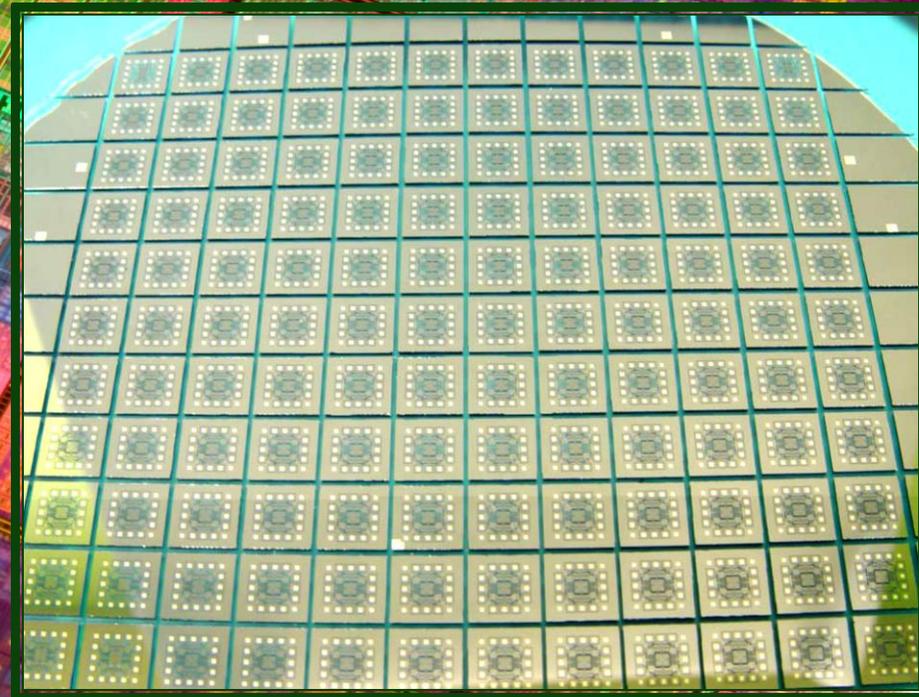


# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ В МИКРО- И НАНОИНЖЕНЕРИИ

Том 1

*В.А. ВАСИН Е.Н. ИВАШОВ  
П.С. КУЗНЕЦОВ В.В. СЛЕТЦОВ*



## СВЕДЕНИЯ ОБ АВТОРАХ УЧЕБНОГО ПОСОБИЯ

**ВАСИН** Владимир Анатольевич, к.т.н., доцент.  
Доцент кафедры «Радиоэлектроника, телекоммуникации и нанотехнологии» «МАТИ» – Российского государственного технологического университета имени К.Э. Циолковского

**ИВАШОВ** Евгений Николаевич, д.т.н., профессор.  
Профессор кафедры «Электроника и наноэлектроника»  
МИЭМ НИУ ВШЭ

**КУЗНЕЦОВ** Павел Сергеевич, к.т.н.  
Начальник сектора ОАО «ГосНИИП»

**СЛЕПЦОВ** Владимир Владимирович, д.т.н., профессор.  
Зав. кафедрой «Радиоэлектроника, телекоммуникации и нанотехнологии» «МАТИ» – Российского государственного технологического университета имени К.Э. Циолковского

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ В МИКРО- И НАНОИНЖЕНЕРИИ

УЧЕБНОЕ ПОСОБИЕ

Том 1

В.А. ВАСИН Е.Н. ИВАШОВ  
П.С. КУЗНЕЦОВ В.В. СЛЕПЦОВ

ISBN 978-5-905872-18-1



НИИП

ИЗДАТЕЛЬСТВО НИИ ПРЕДЕЛЬНЫХ ТЕХНОЛОГИЙ  
2014

УДК 621.01/.03 + 621.315 + 621.37/.39  
ББК 34.44  
И 1

Рецензенты: **Глазунов В.А.**, доктор технических наук и доктор философских наук, профессор, заведующий лабораторией ФГБУН Институт машиноведения имени А.А. Благонравова РАН,

**Шалумов А.С.**, доктор технических наук, профессор, генеральный директор ООО «CALS-технологии»

**И 1** Информационные технологии проектирования в микро- и наноинженерии: Учебное пособие. В 3<sup>х</sup> томах. Том 1. **В.А. Васин, Е.Н. Ивашов, П.С. Кузнецов, В.В. Слепцов.** – Ивanteeвка М.о.: Издательство НИИ предельных технологий, 2014. – 247 с.

**ISBN 978-5-905872-18-1**

Рассмотрены некоторые вопросы информационных технологий проектирования в микро- и наноинженерии, в частности формирование углеродных нанообъектов в тонкоплёночных структурах, автоматизированное проектирование процесса углеродных нанотрубок и фуллеренов, аналитическая процедура Саати в автоматизированном проектировании, моделирование процесса управления качеством электронных устройств, функциональное моделирование системы менеджмента качества на основе методологии IDEFO, математическое моделирование управления качеством электронных устройств на основе углеродных структур, особенности автоматизированного проектирования искусственных нейронных сетей.

Учебное пособие рекомендовано студентам бакалавриата и магистратуры, обучающимся по направлениям 210100 – «Электроника и нанoeлектроника», 210700 – «Инфокоммуникационные технологии и системы связи», 211000 – «Конструирование и технология электронных средств».

**ISBN 978-5-905872-18-1** УДК 621.01/.03 + 621.315 + 621.37/.39

© **В.А. Васин, Е.Н. Ивашов,**  
**П.С. Кузнецов, В.В. Слепцов, 2014**  
© **Издательство НИИ предельных технологий, 2014**  
© **МИЭМ НИУ ВШЭ, 2014**

# **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ В МИКРО- И НАНОИНЖЕНЕРИИ:**

**УЧЕБНОЕ ПОСОБИЕ**

**Том 1**

**ВАСИН** Владимир Анатольевич  
**ИВАШОВ** Евгений Николаевич  
**КУЗНЕЦОВ** Павел Сергеевич  
**СЛЕПЦОВ** Владимир Владимирович

Редактор **И. С. Кравчук**

Технический редактор **П. С. Костомаров**

Рисунок на обложке выполнен художником **Н.Г. Фатьяновой**

Подписано в печать 15.02.2014. Формат 60x84/16.

Бумага типографская. Печать – ризография.

Усл. печ. л. 12,6. Уч.-изд. л. 13,2. Тираж 300 экз.

Заказ 2 Изд. № 3

---

Издательство НИИ предельных технологий  
141282 Мос. обл., г. Ивanteeвка, ул. Ленина, 44.  
Отпечатано в НИИ предельных технологий  
141282 Мос. обл., г. Ивanteeвка, ул. Ленина, 44.

# ПРЕДИСЛОВИЕ

Прогресс современных инженерных и естественных наук на рубеже XX-го и XXI-го столетий привел к рождению новой научно-технической отрасли, получившей название «**нанотехнология**». По мнению ведущих экспертов и специалистов, XXI-й век будет веком нанотехнологий, которые и определяют его облик.

Практически любое техническое решение с использованием нанотехнологии является достаточно сложным сочетанием различных конструктивных подходов с учетом принципиально новых свойств материалов и элементов, появляющихся при переходе в наноразмерную область. Это требует решения инженерных задач, связанных с измерением параметров, привлечением новых технологических процессов и специального технологического оборудования, необходимости сочетания принципиально новых свойств отдельных компонентов и устройств. Во всех этих случаях речь идет о самостоятельном направлении под общим названием «**наноинженерия**», под которой можно понимать научно-практическую деятельность человека по конструированию, изготовлению и применению наноразмерных (наноструктурированных) объектов или структур, созданных методами нанотехнологий.

Термин «наноинженерия» связан с электронными приборами, системами и оборудованием. Это свидетельствует о появлении принципиально новых классов приборов, устройств и оборудования и об их достаточно широком внедрении в современную практику. Необходимо отметить, что микроэлектроника с переходом на проектные нормы 65÷45 нм становится **наноэлектроникой**. Этот термин применяется в соответствии с определениями, предложенными в Российской Концепции развития в области нанотехнологий и в Программе развития наноиндустрии в стране до 2015 года.

Общим направлением наноинженерии следует считать создание

приборов, устройств и оборудования, в состав которых входят компоненты, конструктивные элементы, детали и другие составляющие, построенные на использовании наноструктурных материалов (нанокристаллов, тонких пленок, гетероструктур и др.) и реализующие принципиально новые свойства, обусловленные переходом в наноразмерную область.

Принципиальная особенность функциональных наносистем – необходимость решения проблем, связанных с сочетанием наноразмерных элементов и применением наноинженерии.

Типичные примеры микро- и наносистемной техники – различные наноустройства и, прежде всего, нано- и микроэлектромеханические системы. В наносистемную технику могут быть, в частности, встроены компоненты, разработанные с использованием квантоворазмерных эффектов электронного характера. При этом должны учитываться эффекты, определяющие механические, тепловые, химические свойства не только самих компонентов, но и инженерной системы в целом. В этом случае наноинженерия в качестве самостоятельного направления должна базироваться на специфических подходах, связанных как с необходимостью конструктивного сочетания компонентов, так и с потребностью в разработке новых методов расчета, конструирования и моделирования.

Предлагаемая серия книг охватывает наиболее важные научные и прикладные вопросы развития микро- и наноинженерии в электронном машиностроении, в том числе: наноинженерию туннельных преобразователей, наноинженерию углеродных структур, наноинженерию фуллеренов и нанотрубок, микроинженерию коммутационной аппаратуры, микро- и наноинженерию пьезоэлектрических приводов, микроинженерию магнитных устройств, микро- и наноинженерию электронно-лучевой оптики.

Современное социально-экономическое состояние российского общества характеризуется его направленностью на создание мощной информационной инфраструктуры и формирование единого информационного пространства.

Процесс информатизации и создания информационной среды включает в себя: создание информационной техники и технологий, обеспечивающих производство, обработку и распространение информации; разработку инфраструктуры, обеспечивающей применение и развитие средств и процессов информатизации, производство самой информации, информационных продуктов и услуг.

Информационными ресурсами являются формализованные идеи и знания, различные данные, методы и средства их накопления, хранения и обмена между источниками и потребителями информации. Задача информатизации состоит в сокращении различных видов материальных, энергетических, финансовых и других потоков за счёт их частичной замены и компенсации информационными потоками. Для этих целей применяются информационные системы, реализующие соответствующие информационные технологии.

Информационная технология – это совокупность методов, производственных процессов и программно-технических средств, объединенных в технологическую цепочку, обеспечивающую сбор, обработку, хранение, распространение (транспортировку) и отображение информации с целью снижения трудоёмкости процессов использования информационного ресурса, а также повышения их надёжности и оперативности.

Экономические цели информатизации состоят в получении, обработке и применении информационных средств для повышения эффективности использования всех видов народнохозяйственных ресурсов.

Электронная техника является основой создания и развития информационных технологий и вычислительной техники, на которых в электронной промышленности и её подотрасли – электронном машиностроении, базируются системы автоматизированного проектирования, контроля и диагностики параметров технологического процесса, автоматизированные системы управления и т.д.

Электронное машиностроение – это область науки и техники, занимающаяся изучением физических явлений, исследованием и разработкой технологических процессов, оборудования и систем управления для производства электровакуумных и полупроводниковых приборов, интегральных микросхем, печатных плат, радиокомпонентов и других изделий электронной техники.

Создание современных объектов электронного машиностроения в соответствии с требованиями нормативно-технической документации и в заданные сроки, в общем случае, невозможно без использования информационных технологий проектирования.

Предлагаемая серия книг охватывает наиболее важные научные и прикладные вопросы развития информационных технологий в проектировании объектов электронного машиностроения, в том числе: информационную технологию в проектировании платин плотной записи для накопителей на жёстких магнитных дисках; информационную технологию в проектировании процесса формирования объектов лазерной обработки; информационную технологию в управлении качеством производства материалов на основе углеродных структур; информационную технологию в проектировании пьезосканеров сканирующих зондовых микроскопов; информационную технологию в формировании нейронных сетей на основе твёрдотельных объектов.

Авторы выражают глубокую благодарность рецензентам: А.Н. Антамошкину, доктору технических наук (профессору Сибирского государственного аэрокосмического университета им. академика М.Ф. Решетнева) и Б.Н. Васичеву, доктору физико-математических наук (профессору Российского экономического университета им. Г.В. Плеханова) за тщательный просмотр рукописей, проявленное внимание к рецензируемым книгам, а также высказанные ценные замечания, которые способствовали улучшению содержания предлагаемых читателю изданий.

# ВВЕДЕНИЕ

Интерес к наноструктурам связан с новыми возможностями, которые открывает нанотехнология в создании материалов и изделий из структурных элементов нанометрового диапазона. В последнее время большое внимание привлекают приборы на основе наноразмерных объектов. В таких приборах носители тока могут проявлять как корпускулярные, так и волновые свойства. В этом случае появляется возможность осуществлять контроль фазовой когерентности баллистического транспорта или, иными словами, контролировать фазу волновой функции электрона в структурах, размер которых меньше длины неупругого рассеяния носителей тока. По указанным причинам такие устройства характеризуются рядом уникальных свойств: могут усиливать и преобразовывать очень слабые сигналы, имеют малые размеры и крайне низкое энергопотребление. Для стабильной работы при нормальной температуре они должны иметь размеры порядка десятков нанометров. Для вертикальных размеров эта проблема успешно решается методами молекулярно-лучевой и газофазной эпитаксии. Однако при формировании горизонтальных структур в планарной технологии возникают серьезные трудности.

Например, для фотолитографии необходимо использовать заранее изготовленные фотошаблоны, наносить фоторезист, использовать химическое травление, приводящее к загрязнению поверхности. Модификация поверхности с использованием электронных или ионных пучков, в отличие от фотолитографии, может осуществляться более локально, однако ее использование осложняется трудностью контроля получаемой поверхности по глубине и дороговизной оборудования. Стоимость литографического оборудования составляет, как правило, миллионы долларов только за установку (не считая вспомогательного оборудования), что может себе позволить далеко не всякая исследовательская организация. Кроме того, возможности каждого из перечисленных методов имеют свои ограничения,

вследствие чего, при создании экспериментальных наноразмерных объектов и структур становится целесообразным использование комбинированных методик формирования рисунка.

В то же время, имеется достаточно «дешевая» и активно развивающаяся в последние десятилетия альтернатива микролитографии – сканирующая зондовая литография (СЗЛ), основанная на наномодификации поверхности твердотельным зондом сканирующего зондового микроскопа (СЗМ).

Зарождение сканирующей зондовой литографии неразрывно связано с изобретением сканирующего туннельного микроскопа – СТМ (Рорер, Биннинг – 1981г.). Именно благодаря СТМ, впервые появилась возможность не только наблюдать и исследовать поверхность различных веществ с атомарным разрешением, но и активно воздействовать на нее, то есть получать объекты из конечного числа молекул, удаляя, перемещая или замещая молекулы одного вещества другим. Появилась перспектива работать с отдельными атомами или с их небольшими группами.

Эта возможность, называемая в современных источниках информации нанотехнологией, позволяет значительно расширить диапазон геометрических параметров искусственно созданных объектов применительно к микроэлектронике, которую в этом случае целесообразно называть нанoeлектроникой.

При СЗЛ в области наноконтакта происходит концентрация токов большой плотности, высоких электрических полей и механических давлений. В этих условиях активизируется целый ряд явлений, таких как локальный разогрев, пластическая деформация, поляризация, полевое испарение, массоперенос за счет ионных токов в электролите и электрохимические реакции. Раздельное или совместное действие этих факторов используется для наномодификации материалов в различных режимах СЗЛ.

В данном учебном пособии описывается работа с углеродными (алмазоподобными) тонкими пленками, полученными плазменным осаждением

ем из газовой фазы в вакууме. Целесообразность выбора углерода в качестве подложки и объекта модификации определяется множественностью его аллотропных форм, соединений и широкого диапазона электрофизических свойств – от диэлектрических до проводниковых. Формирование электронных нанообъектов на основе модифицированных углеродных структур является задачей актуальной и своевременной, так как при использовании подобных наноструктур появляется возможность получать как изолирующие, полупроводниковые, так и электропроводящие объекты.

Разработка методов литографии углеродных слоев особенно актуальна в связи с перспективностью использования углеродных материалов в нанoeлектронике. В отличие от кремния, технология формирования элементов интегральных схем на основе углеродных материалов в настоящее время отсутствует. Остаются открытыми вопросы о механизмах локального воздействия острия зонда на углеродные слои.

В процессе развития науки и техники создаваемые технические системы и устройства становятся все более сложными. Широкое внедрение компьютеризации в условиях научно-технического прогресса обеспечивает рост производительности труда в различных областях производства.

Развитие микроэлектроники, переход на нанотехнологии требует использования различных углеродных соединений. Внедрение таких соединений сдерживается сложностью проектирования процесса производства, отсутствием моделей и способов его оценки на отдельных стадиях разработки. Поэтому актуальна разработка элементов автоматизированных систем проектирования, включающая комплекс методических средств и моделей, а также способов оценки качества проекта в целом.

Одним из наиболее мощных средств для исследования и проектирования технических систем является моделирование. Использование моделирования, начиная с ранних стадий проектирования, 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.

Микроминиатюризация изделий и повышение их качества во многих областях современной техники поставили задачу обеспечения 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 аналитическом оборудовании (от 100 мс до 0,1 мс и менее) предъявляет жёсткие требования к динамическим параметрам модулей линейных перемещений.

Во многих случаях, эти модули должны работать в вакуумной среде, которая создает проблемы даже при обычных диапазонах позиционирования. Использование пьезопроводов в вакууме позволяет обходиться без дополнительных устройств передачи движения через герметичную стенку, т.е. проблема создания ввода в вакуумный объём отпадает сама собой.

В нанотехнологии перемещение и позиционирование связано с малыми размерами. В современном технологическом и аналитическом оборудовании используются различные типы устройств прецизионных перемещений, в частности, устройства на пьезопреобразователях, т.н. пьезосканеры, отличительной особенностью которых является их безынерционность. Принцип их действия основан на преобразовании электрической энергии в механическую.

Интерес к субмикронным структурам связан с возможностью существенной модификации свойств известных веществ, а также новыми возможностями, которые открывает технология создания материалов и изделий из структурных элементов микро- и нанометрового диапазона. Данные про-

цессы требуют изменений основных принципов управления процессами производства на отечественных предприятиях, создания систем управления качеством, аналогичных передовым зарубежным разработкам.

Особенно актуальны проблемы повышения качества продукции в отечественной электронной промышленности. Увеличение объёма производства на российских предприятиях качественных электронных компонентов, стоимость которых будет существенно ниже зарубежных аналогов, позволит сэкономить крупные финансовые средства, а также обеспечить независимость ряда важнейших отраслей отечественной промышленности от зарубежных поставщиков. Изделия электронной техники, создаваемые на базе перспективных материалов, используются в настоящее время во многих важнейших производственных сферах: авиационной, космической, оборонной, атомной, энергетической и др. Также областью использования данных материалов в последнее время стала отрасль новых информационных и телекоммуникационных технологий, применение в которых высокотехнологичных проводящих материалов позволяет создавать высокопроизводительные кластерные информационно-управляющие системы. Развитие технологии изготовления приборов микро- и нанoeлектроники связано, прежде всего, с уменьшением геометрических размеров микросхем, элементов электронной памяти и микродатчиков различного назначения. Современные средства традиционных технологических операций позволяют получать размеры элементов в субмикронной области. Дальнейшее уменьшение геометрических размеров функциональных элементов на подложке неизбежно приведёт к физическому пределу, определяемому длиной волны ультрафиолетового излучения, применяемого в традиционной фотолитографии. С изобретением сканирующего туннельного микроскопа (Рорер, Биннинг – 1981г.) появилась возможность не только наблюдать и исследовать поверхность различных веществ с атомарным разрешением, но и активно воздействовать на неё, то есть манипулировать веществом на уровне

отдельных молекул и получать объекты из конечного их числа, удаляя, перемещая или замещая молекулы одного вещества другим. Появилась перспектива работать с отдельными атомами.

Учебное пособие рекомендовано студентам бакалавриата и магистратуры, обучающимся по направлениям 210100 – «Электроника и нанoeлектроника», 210700 – «Инфокоммуникационные технологии и системы связи».

# ГЛАВА 1. ФОРМИРОВАНИЕ УГЛЕРОДНЫХ НАНООБЪЕКТОВ В ТОНКОПЛЕНОЧНЫХ СТРУКТУРАХ

## 1.1. Основная задача формирования углеродных нанообъектов

Применение эффекта туннелирования в зазоре между зондом и поверхностью подложки позволило достичь высокой пространственной разрешающей способности без применения высокоэнергетичных частиц и создать на их основе методы сканирующей зондовой микроскопии. За последнее десятилетие сканирующие туннельные микроскопы были в значительной степени усовершенствованы и дооборудованы функциональными элементами, позволяющими осуществлять в рабочей камере различные по природе воздействия на поверхность образца, что позволило создавать на его поверхности различные нанометровые структуры. Так появились «нанотехнологические установки», основная задача которых, в отличие от обычных СТМ, заключается не просто в исследовании поверхности образцов, а в проведении различных нанотехнологических процессов между технологической средой и подложкой и исследовании физических и химических характеристик создаваемых при этом нанообъектов. На основе этой техники стали развиваться методы нанотехнологии, использующие частицы с величинами энергий, определяемыми оптимальной энергией активации нанотехнологических процессов, а не энергией необходимой для их фокусировки.

В зависимости от фазового состояния используемых сред нанотехнологические реакции можно разделить на три типа: реакции, происходящие в вакууме, в газе и в жидкости.

Однако, вследствие сильной неоднородности вещества непосредственно в области между зондом и подложкой, данное разделение не всегда корректно.

Например, при работе в газовой среде, из-за малой величины туннельного зазора, значительно меньшей среднего расстояния между молекулами газа, процессы прохождения электронов практически эквивалентны прохождению электронов в вакууме. При определенных условиях (повышенном давлении, пониженной температуре или при применении газов состоящих из сильно поляризованных молекул), вследствие капиллярного эффекта газ в зазоре зонд-подложка конденсируется в жидкость. В то же время, при работе в жидкости при значительном энерговыделении в области реакции жидкость может перейти в газообразное состояние в виде нано или микропузырька.

Известно, что состояния с повышенным давлением и с повышенной температурой стимулируют прохождение нанотехнологических реакций, однако реализация таких условий традиционными способами не только значительно усложняет конструирование установки, но и приводит к понижению пространственной разрешающей способности технологических процессов. В то же время использование сил межатомного взаимодействия сопровождается возникновением капиллярного эффекта и приводит к конденсации газа в жидкость. С другой стороны, приложение сильных электрических полей дополнительно поляризует конденсированное вещество, приближая его к состоянию жидкого кристалла.

Выполнение требования повышения температуры может быть реализовано другим путем. Для осуществления химических реакций необходимо возбуждение электронов на орбитах, при этом температура ядер не обязательно должна быть высокой. Выполнение данного условия может быть обеспечено при электромагнитном, т.е. фотонном характере возбуждения.

По характеру воздействия нанотехнологические процессы можно разделить на три вида: полевые, кинетические и электромагнитные [10]. К полевым процессам относятся следующие физико-химические явления: поляризация, стимулированная миграция, диссоциация, десорбция, ионизация, полевая деформация, полевая эмиссия. К группе кинетических воздействий относятся такие явления, как рассеяние электронов и других заряженных частиц, рассеяние нейтральных частиц, а также тепловые воздействия. К электромагнитным нанотехнологическим процессам относятся нерезонансные, резонансные в одном или нескольких спектральных диапазонах, спектрально синхронизированные, пространственно-когерентные и межатомные взаимодействия.

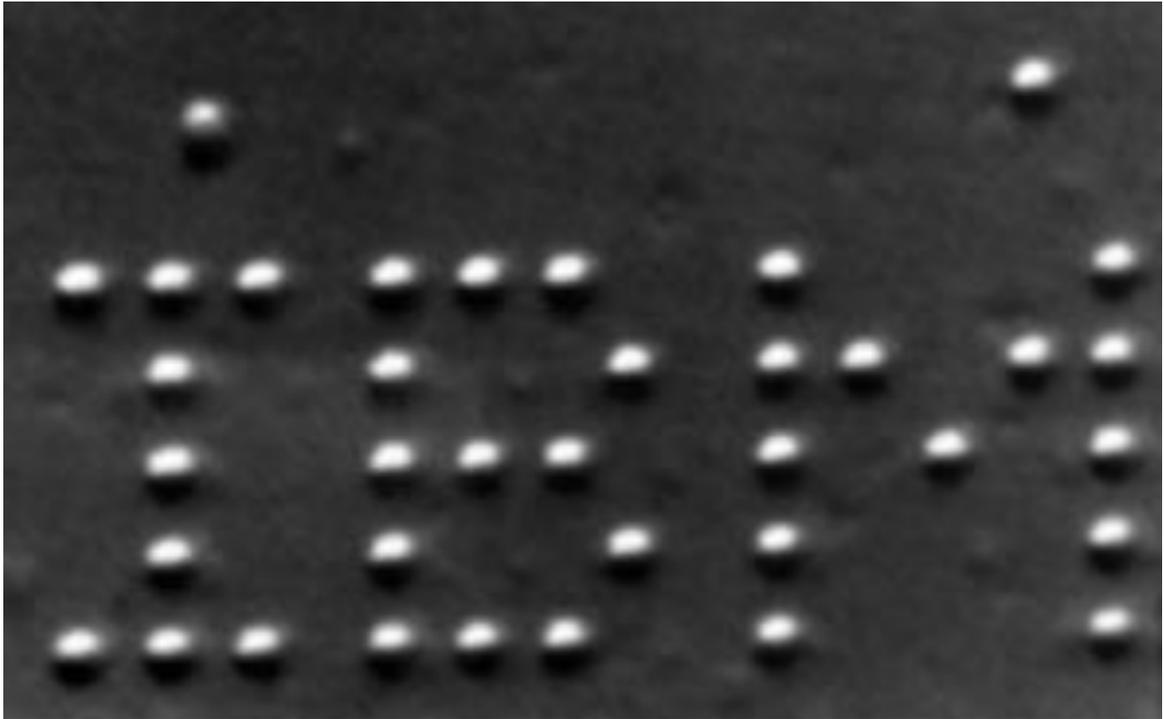
В реальных условиях в технологических установках одновременно реализуется несколько типов нанотехнологических процессов. Кроме процессов воздействия на одиночные атомы и молекулы возможна также стимуляция коллективных процессов, происходящих с веществом, расположенным в зоне активации: нагрев, диффузия, различные перемещения атомов и молекул, термополевая десорбция, инициирование изменения фазового состояния, осуществление локальных химических реакций на подложке с целью изменения ее химического состава, разложение органических молекул, осаждение вещества из газовой фазы.

Возможны также сочетания нескольких процессов. Практически ни один из процессов не протекает исключительно единственным образом.

Как известно, путем изменения структурных и размерных параметров нановключений можно в широких пределах менять энергетические характеристики создаваемой системы, а значит получать необходимую комбинацию различных свойств. Появляется реальная возможность изменять свойства материала, придавая ему характеристики, недостижимые для объемных материалов.

История появления и развития направления в научно-технологической деятельности человечества под названием «нанотехнология» следует отнести к середине XX века. В 1959 году в своей Нобелевской речи выдающийся физик современности Ричард Фейнман произнес фразу, которая стала основой совершенно нового направления в науке и инженерии: «Принципы физики, с моей точки зрения, не отрицают возможности манипулирования атомами, как предметами» [11]. Однако, в период активного развития полупроводниковой микроминиатюризации (начавшегося с 60-х годов) о высказывании Фейнмана забыли. Впервые термин «нанотехнология» произнес в 1974 г. японский инженер-исследователь Танигучи на конференции Японского общества точного машиностроения [12]. А когда в 1990 г. увидела свет статья двух исследователей из лабораторий IBM (Эйглер, Швейцер, журнал "Nature", Vol. 344) под названием «Позиционирование отдельных атомов с помощью сканирующего туннельного микроскопа» [13], стало ясно, что пророчество Фейнмана сбылось – весь мир обошла фотография удивительной мозаики, образующей слово "IBM", "выложенное" отдельными атомами ксенона с атомарной точностью, расположенных при температуре в 4К на подложке из монокристалла Ni (рис.1.1.1). Толщина штриха составляла 1,27 нм. На эту операцию ушло 22 часа. Когда температура поднялась до -228,89°С, буквы испарились. Несколько ранее (1981 г.) Эрик Дрекслер в своей диссертации [14] впервые сформулировал идею робота, способного «собирать» из отдельных атомов сложные структуры. С именем Дрекслера связаны очень многие термины и идеи, существующие в сегодняшней нанотехнологии, - именно этот ученый первым понял всю значимость и справедливость высказывания Фейнмана.

Результат, аналогичный тому, который продемонстрировали Эйглер и Швейцер, был получен и при создании структур из окиси углерода. В дальнейшем были получены успешные результаты по осаждению класте-



**Рис. 1.1.1.** 37 атомов ксенона на грани монокристалла Ni (110)

ров золота с вершины зонда на поверхность подложки [15]. Данный нанотехнологический процесс был реализован благодаря высокой миграционной способности золота. Однако именно это свойство привело к временной нестабильности создаваемых структур.

Наиболее перспективными технологическими процессами являются процессы, использующие локальное осаждение материалов из газовой фазы.

К настоящему времени большое развитие получили различные способы создания наноструктур и закрепления их на разнообразных поверхностях. Уже более десяти лет назад были созданы действующие прототипы нанотранзисторов [16]. Вместе с тем, для практического использования таких структур необходимо знание их физических и физико-химических свойств. Невозможность применения “зонной теории” для описания свойств наноструктур и сложности с экспериментальным измерением физико-химических свойств таких объектов привели к различным моделям оценки их параметров. Так, например, в работе [17] теоретически рассчи-

таны прочностные характеристики нанопластинок и нанобрусков Si и Al в зависимости от размера. Показано, что при размерах менее 100 нм отличие прочностных характеристик наноструктур от аналогичных параметров при макроразмерах превышают 10%. Авторы работы [18] отмечают, что термическая стабильность наноустройств обеспечивается при их линейных размерах, больших 10 нм.

Несмотря на множество публикаций посвященных нанотехнологиям большинство из них носят или обзорный или узко направленный характер. Поэтому исследователям бывает непросто выбрать направление разработок, которое может позволить осуществить создание наноструктур с заданными свойствами. Возможно предложить следующий алгоритм принятия решений:

1. Выбирается категория материалов, которые по физическим и химическим свойствам, условиям эксплуатации (электрофизическим характеристикам, диапазону рабочих температур, механическим напряжениям, стойкости к радиационным воздействиям) могут соответствовать техническим требованиям, предъявляемым к изделию.
2. Разрабатывается технологический процесс создания наноструктур из предварительно выбранных типов материалов.

К процессам, отличающимся высокой разрешающей способностью относится термополевая эмиссия, основанная на сочетании различной степени нагрева зонда и подложки, а также создании между ними электрического поля, стимулирующего процессы переноса атомов и молекул. Данный метод характеризуется разрешающей способностью вплоть до атомарной и позволяет осуществить направленный перенос атомов и молекул независимо от типа проводимости подложки.

Для прецизионного управления нанотехнологическими реакциями возможно использование оптических процессов. Сочетание излучений с различными частотами позволяет провести технологические процессы в

нужном направлении. Для этого необходимо направить в зону реакции излучения, стимулирующие распад соответствующих молекул. Таким образом, можно увеличить вероятность прохождения реакции в нужном направлении и значительно уменьшить протекание реакций в другом, нежелательном направлении.

Следует обратить особое внимание именно на данный метод, т.к. он позволяет при увеличении интенсивности стимулирующего излучения значительно увеличить скорость проведения реакций. Создание электромагнитных полей с различными частотами позволяет активировать различные фрагменты молекул, что приведет к изменению направления хода химических реакций. Метод одновременного воздействия излучением с различными частотами, соответствующими энергиям связи «нежелательных» соединений, позволяет «разрешить» синтез только необходимых веществ. Данный метод характеризуется исключительно малым изменением кинетической энергии молекулы, что позволяет обеспечить высокую локальность проводимых химических процессов.

## **1.2. Физические аспекты формирования наноструктур на поверхности алмазоподобных тонких пленок**

В электронных приборах с размерами активных областей менее нескольких десятков нанометров начинают играть значительную роль квантовые явления, так как размеры приборов становятся сравнимы с характерными размерами, определяющими поведение электронов проводимости.

Поэтому интерес к наноструктурам и связан, прежде всего, с ожиданием различных размерных эффектов на наночастицах или структурах, размеры которых соизмеримы (или меньше) с характерным корреляционным масштабом того или иного физического явления или же с характерной

длиной, фигурирующей в теоретическом описании какого-либо свойства или процесса (например, длина волны де Бройля, длина свободного пробега электронов, длина когерентности, длина сдвоя фазы и радиус экранирования, размер магнитного домена в ферромагнетиках и др.) [19]. Принципиально меняются и физические явления, лежащие в основе работы элементов – осуществляется переход от принципов классической физики к квантовой, т.е. переход от дискретных фаз к атомам (квантовым точкам), цепочкам атомов (одномерным структурам), тонким пленкам (двумерным структурам), а также трехмерным объемным сверхрешеткам.

Прикладной интерес к наноструктурам обусловлен возможностью значительной модификации или даже принципиального изменения свойств известных материалов, а также новыми возможностями, которые открывает нанотехнология в создании материалов и изделий из структурных элементов нанометрового размера. Управление фундаментальными свойствами твердых тел (полупроводники, металлы, полимеры и т.д.), основанное на формировании на поверхности наноразмерных пленок и структур, в настоящее время составляет одну из главных проблем ведущих научных центров мира, работающих в области нанотехнологий [1,20÷25]. Основой развития ряда современных перспективных направлений (микросистемная техника и наноэлектроника, оптика, энергетика, биотехнология и др.) уже сейчас являются новые функциональные материалы и структуры, созданные методами микроэлектроники, нанотехнологии и рядом других способов, обеспечивающие обработку на микро- и нано - уровне. При этом резко расширяется ассортимент используемых материалов, и повышаются требования к их параметрам, в частности, необходимым условием становится обеспечение совокупности функциональных характеристик используемых материалов: кристаллохимической и термохимической совместимости, механической, тепловой и электрической стойкости, биостойкости, низкой механической и термохимической усталостности и электрической

деградации. Решение этих задач требует принципиально новых подходов, как в области материаловедения, так и в области технологии синтеза и формообразования. Переход на субмикронный и нанометровый диапазон размеров элементов требует учета факторов масштабирования, отражающих влияние геометрических размеров на свойства материала. Уменьшение геометрических размеров объектов, составляющих наноструктуры, приводит к тому, что на объемные свойства материала начинают влиять его поверхностные характеристики, а с определенных размеров ( $100\div 1000 \text{ \AA}$ ), поверхностные свойства начинают доминировать над объемными и преимущественно определять свойства системы в целом. Становится возможным путем изменения структурных и размерных параметров (толщины пленок, размеров нановключений, структуры границ раздела и пр.) в широких пределах менять энергетические характеристики создаваемой системы, а значит получать необходимую комбинацию различных свойств (механических, электрофизических, оптических и многих других). Появляется реальная возможность изменять свойства материала, придавая ему характеристики, недостижимые для объемных материалов.

Известно, что поверхность всегда имеет большую энергию, чем объем, и при этом стабильно сохраняет свои характеристики [26]. Затруднительно точно определить границу раздела поверхности и объема. Считается, что поверхностные свойства твердого тела начинаются на границе раздела двух сред и кончаются в глубине материала  $\sim 1000 \text{ \AA}$ . Ширина границ раздела для различных компактных материалов начинается от  $4\div 15 \text{ \AA}$  [27÷29]. Пленочные материалы в диапазоне  $5\div 1000 \text{ \AA}$  также имеют сильную зависимость свойств от толщины [30÷37]. Поэтому поверхность можно рассматривать как некий физический объект, свойства которого определяются совокупностью физических явлений трехмерного объема, двумерной поверхности, а в некоторых случаях и одномерной составляющей. Такой подход позволяет рассмотреть и выявить целый класс структур,

имеющих дробную размерность в приблизительном диапазоне от 0,5 до 3,3, свойства которых существенно превышают свойства массивных материалов.

В связи с этим естественно возникает вопрос о зависимости свойств твердого тела (твердости, механической прочности и др.) от энергетических и электрофизических характеристик поверхностных слоев. Так, важнейшей характеристикой поверхности является поверхностная энергия, величина которой связана с механическими характеристиками.

Обобщенное уравнение первого и второго законов термодинамики для поверхности выглядит следующим образом:

$$dU = TdS + \sigma dF + \varphi dq + \sum \mu_i dn_i, \quad (1.2.1)$$

где  $U$  – внутренняя энергия,  
 $S$  – энтропия,  
 $\sigma$  – удельная поверхностная энергия (поверхностное натяжение),  
 $F$  – площадь,  
 $\varphi$  – потенциал,  
 $q$  – заряд,  
 $\mu_i$  – химический потенциал,  
 $n_i$  – мольная доля вещества.

Из написанного можно сделать вывод, что чем больше внутренняя энергия твердого тела, тем выше его механические характеристики (микротвердость, износостойкость и др.)

Следует обратить внимание на член  $\varphi dq$ , который однозначно свидетельствует об увеличении внутренней энергии системы при возникновении электрического потенциала. Если учесть, что на границе раздела двух сред возникает двойной электрический слой, то естественно будет расти внутренняя энергия нанокompозита. Это приводит к тому, что меняется модуляция электрического поля и возникает зависимость ширины запрещенной зоны от структуры материала, а значит и от толщины покрытия.

Эти факты были подтверждены экспериментально при исследовании зависимости свойств алмазоподобных покрытий от толщины.

Структурные изменения на поверхности влекут за собой изменение величины поверхностной энергии, поэтому развитие и модификация поверхности, влекущая за собой увеличение поверхностной энергии, ведет к увеличению механических характеристик материала. Примером такой модификации может служить ионная бомбардировка растущей алмазоподобной углеродной пленки [38÷41], легирование поверхности твердого тела, а также термообработка, приводящая к изменению структуры и состава поверхности. Исходя из уравнения (1.2.1), можно сделать вывод, что рост члена  $\sum \mu_i dn_i$  приводит к увеличению внутренней энергии системы.

Зависимость электрофизических и механических свойств твердого тела от величины внутренней энергии позволяет сделать следующие выводы:

1. Изменение структуры поверхности в широких пределах позволяет менять энергию поверхности, а значит механические и электрофизические свойства тонких пленок. Диапазон изменения свойств может быть существенно выше, чем у объемных материалов.
2. Изменение толщины пленки позволяет также существенно менять величину поверхностной энергии, а значит менять свойства пленки. Таким образом, появляется реальная технологическая возможность изменять свойства материала, придавая ему характеристики, недостижимые для объемных материалов. Практически увеличение внутренней энергии для повышения механических свойств поверхности решается путем увеличения  $U_F$  за счет деформации связи и создания точечных дефектов за счет легирования методами имплантации или диффузии, или, например, нанесения покрытия из материала, которое априори имеет более высокую внутреннюю энергию, а значит и механические свойства.

Заслуживают рассмотрения также динамические характеристики на-

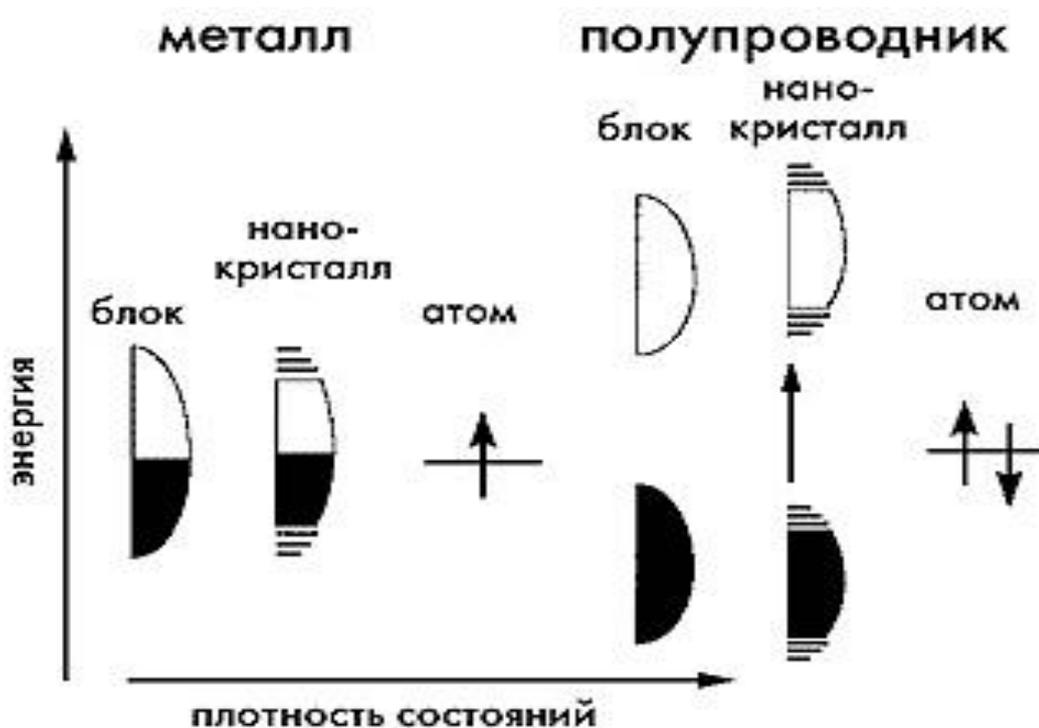
нокомпозитных материалов, которые также существенно отличаются от параметров объемных материалов. Зависимость долговечности от температуры и неравномерного распределения энергии между атомами твердого тела позволяет говорить об увеличении динамической прочности пленочных материалов. Это связано с тем, что уровень накопления энергии, которая определяет скорость деградационных процессов, ограничивается толщиной пленки, а диссипация энергии в поверхности имеет существенно большую скорость, чем в объеме. При рассмотрении характерной функциональной зависимости долговечности  $t = t_0 \exp\left(\frac{U_0 - \delta}{kT}\right)$  от действия растягивающего напряжения и температуры тела [40], видно, что долговечность пропорциональна множителю  $\exp\left(\frac{U_0 - \delta}{kT}\right)$  (где  $U_0$  – энергия начала процесса разрушения (энергия активации), эВ;  $\delta$  – напряжение, ГПа). Опираясь на характерный больцмановский вид формулы для долговечности, можно приписать процессу разрушения термоактивационную природу, подобную природе ряда других кинетических процессов [42]. Это означает, что усталость нанокompозитных структур будет существенно меньше, чем объемных материалов, т.к. у нанокompозитных материалов скорость диссипации избыточной энергии, возникающей в динамическом процессе, существенно выше.

Исходя из вышеизложенного, можно сделать вывод, что с определенного геометрического размера  $\sim 1000 \text{ \AA}$  и менее все материалы имеют существенно более высокую энергию, чем в объеме. Поэтому наноразмерные материалы и структуры являются самоорганизующимися устойчивыми системами, обеспечивающими кооперативное поведение электронных, фононных и других подсистем и обладают недостижимыми для массивных структур механическими и динамическими характеристиками.

Размеры наночастиц  $1 \div 10 \text{ нм}$  соответствуют огромной удельной поверхности. Подобные наночастицы отличаются по свойствам (ширине полосы поглощения, спектральным характеристикам, электронному пере-

носу и др.) как от блочного материала, так и индивидуального атома или молекулы, особо заметны отличия для полупроводниковых частиц. Так, при переходе от нанокристалла CdS к макрокристаллу ширина запрещенной зоны уменьшается от 4,5 до 2,5 эВ, время жизни на нижнем возбужденном уровне увеличивается от пикосекунд до нескольких наносекунд, повышается температура плавления от 400°C до 1600°C [43].

Это определяется тем, что электрическая структура наночастицы принципиально отличается, с одной стороны, от отдельного атома, а с другой стороны, она также отличается от объемного твердого тела. Нанокристалл сочетает в себе свойства отдельных атомов, имея дискретные электронные уровни, а также свойства объемного твердого тела, имея валентную зону и зону проводимости (рис. 1.2.1).



**Рис. 1.2.1.** Заполнение электронных уровней в металле и полупроводнике [43]

Энергетическая зона металла, независимо от размера его частиц, заполнена не вся, поэтому электроны могут переходить на более высокие уровни. У полупроводника же валентная зона заполнена целиком и отделена от зоны проводимости на  $2\div 3$  эВ.

Из-за малых размеров полупроводниковых нанокристаллов эти зоны расщепляются, что приводит к эффективному увеличению ширины запрещенной зоны (до 4,5 эВ) [44]. При изменении размеров наноструктурных элементов изменяются также электрофизические свойства. Так, исследование электропроводности алмазоподобных пленок показало зависимость удельного сопротивления от их толщины  $d$  (рис. 1.2.2).

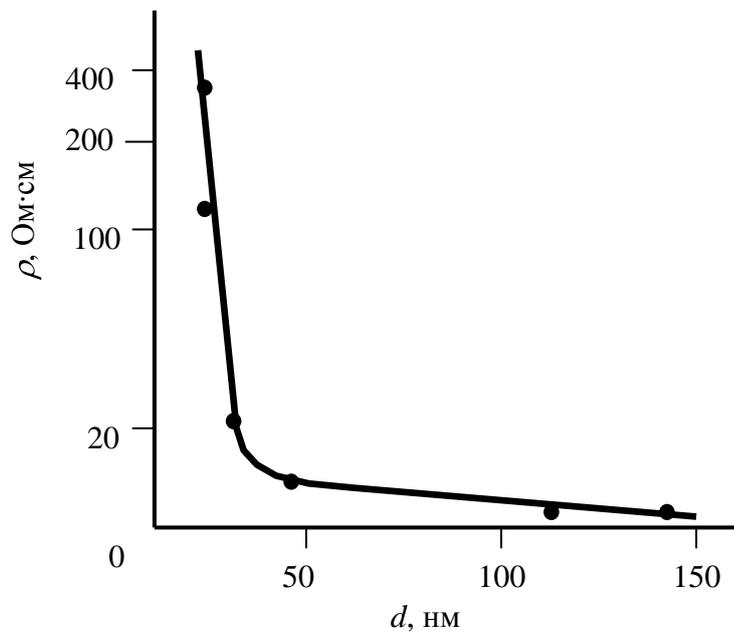
Были получены пленки  $\alpha$ -C:H различной толщины и структуры с помощью магнетронного распыления в атмосфере Ar. Для пленок толщиной больше 50 нм удельное сопротивление слабо зависит от  $d$ , значительное же увеличение удельного сопротивления наблюдается при толщине пленок менее  $20\div 30$  нм.

Для объяснения полученных результатов были проанализированы температурные зависимости проводимости для пленок различной толщины в интервале температур  $77\div 500$  К [32,35].

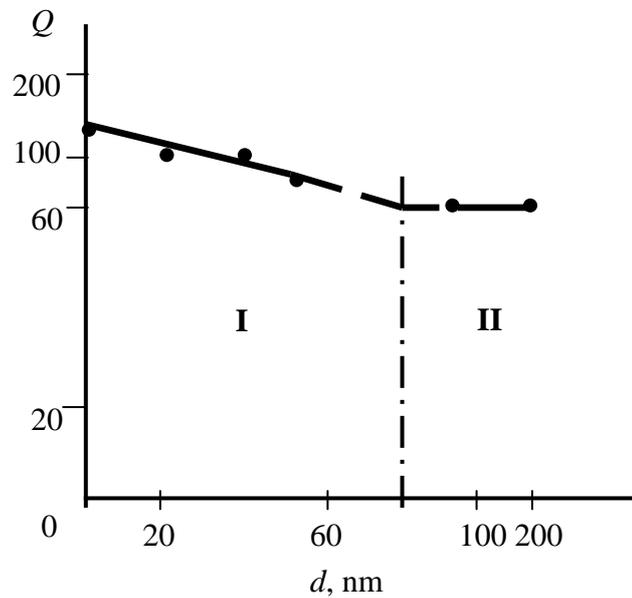
Было установлено, что в температурном диапазоне  $77\div 270$  К для пленок толщиной более 100 нм логарифм проводимости линейно зависит от  $T^{-1/4}$ , кроме того, наклон прямых не зависит от толщины пленок (рис. 1.2.3).

Такой характер зависимости соответствует прыжковому механизму проводимости по локализованным состояниям вблизи уровня Ферми. С уменьшением толщины пленок наблюдается переход от трехмерной прыжковой проводимости к двумерной.

Полученные экспериментальные результаты позволили оценить критическую толщину, при которой этот переход механизмов проводимости ( $\sim 60$  нм).



**Рис. 1.2.2.** Зависимость удельного сопротивления пленок  $\alpha$ -C:H от толщины



**Рис. 1.2.3.** Влияние толщины пленки на тангенс угла наклона прямых ( $Q$ ) в координатах  $\ln(\sigma) = f(T^{1/4})$  (область I) and  $\ln(\sigma) = f(T^{1/3})$  (область II)

### 1.3. Технологические особенности получения тонких алмазоподобных пленок

Интенсивные исследования физики и технологии осаждения алмазоподобных тонких пленок за последние 30 лет привели к их практическому применению, например, для упрочнения режущего инструмента, прецизионных деталей в приборостроении, в качестве просветляющих и защитных покрытий элементов оптических устройств, химической посуды, что обусловлено исключительной химической стойкостью, износостойкостью таких пленок и их прозрачностью в видимом диапазоне длин волн (см. Табл. 1.3.1). Немаловажно применение алмазоподобных углеродных покрытий в качестве защитного слоя дисков для сверхплотной магнитной записи информации, там, где необходима защита этой информации от внешних разрушительных воздействий.

В последнее время особое внимание уделяется поликристаллическим алмазным пленкам на плоских подложках [46]. Некоторые их физические свойства, такие как твердость, теплопроводность, прозрачность, близки к свойствам естественного алмаза. Кроме того, изменяя условия их получения, можно в широких пределах управлять шириной запрещенной зоны, концентрацией примесных атомов и проводимостью слоя, что позволяет использовать алмазные покрытия при изготовлении мощных полупроводниковых приборов.

Еще одна область применения алмазоподобных пленок связана с открытием отрицательного электронного сродства поверхности [111] естественного полупроводникового алмаза типа Шв [47]. Эта особенность алмазной поверхности в принципе позволяет изготовить автоэмиссионные катоды на основе алмазных пленок, работающие с хорошей стабильностью в условиях технического вакуума при сравнительно низких электрических полях. В настоящее время наиболее интенсивно развиваются направления

**Таблица 1.3.1**

Основные физические свойства монокристаллического алмаза [45]

Параметр решетки при температуре 300 К, нм	0,356683 0,35672
Линейный коэффициент термического расширения при температуре 300 К	$0,8 \cdot 10^{-6}$
Показатель преломления при $\lambda = 590$ нм	2,417
Диэлектрическая проницаемость $\varepsilon$	5,7
Ширина запрещенной зоны $E_g$ при 300 К, эВ	5,48
Дрейфовая подвижность электронов $\mu_e$ , $\text{см}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$	2500
Дрейфовая подвижность дырок $\mu_h$ , $\text{см}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$	2100
Плотность, $\text{г} \cdot \text{см}^{-3}$	3,51
Теплоемкость при 300 К, $\text{Дж} \cdot \text{моль}^{-1} \cdot \text{К}^{-1}$	6,12

исследования алмазоподобных покрытий, связанные с изготовлением полупроводниковых приборов и повышением стабильности полевого тока "холодных" электронных источников [48].

Для синтеза алмазных плёнок интерес представляют плазменные "тонкоплёночные" технологии, которые позволяют получать плёнки высокого качества. Как правило, это вакуумные способы.

Первая технология основана на использовании катодного распыления графита при низком давлении. Получаемые алмазоподобные пленки имеют высокие физико-механические свойства: микротвердость

4000÷6000 кг/мм<sup>2</sup>, хорошую адгезию к различным материалам, высокую износостойкость. Эти пленки обладают также высокой химической стойкостью, не вступают в реакцию с кислотами и щелочами.

Вторая технология получения алмазоподобных пленок основана на разложении углеродосодержащей газовой среды в электрическом разряде и осаждении ионов углерода на специальную подложку. Физико-механические свойства алмазоподобных пленок, полученных по этой технологии, близки к свойствам пленок, созданных по первой технологии. В то же время скорость нанесения покрытий по второй технологии значительно выше – 5÷10 мкм/час, что позволяет получать покрытия толщиной более 10 мкм. Имеется возможность плавной регулировки процесса для получения пленок с различными свойствами, начиная от очень твердых (более 4000 кг/мм<sup>2</sup>) до пластичных алмаз полимерных пленок. Технология не требует получения сверхвысокого вакуума и экономична (потребление электроэнергии около 200 Вт/час для нанесения покрытия толщиной 10 мкм на площадь 100 см<sup>2</sup>).

Ростовые процессы третьего типа позволяют выращивать тонкие пленки со структурой алмаза или алмазоподобного углерода осаждением ионов из ионных пучков. Основным достоинством метода является регулирование, существенно влияющей на свойства пленок, энергии падающих ионов в широком диапазоне значений, использование различных, содержащих углеводороды, плазмообразующих смесей газов, существенно расширяющих условия синтеза, и строгое задание массового состава ионного пучка. Типичная упрощенная схема устройства обычно содержит плазменный источник ионов и подложку, находящуюся под отрицательным потенциалом относительно плазмы и служащей извлекающим и ускоряющим электродом, принимающим плазменные ионы, из которых и наращивается пленка.

Осаждение ведется из активируемой газовой фазы, состоящей из ле-

тучего углеродсодержащего вещества и водорода и находящейся при субатмосферном давлении. Существенно, что в используемой для роста пленок области температур и давлений алмаз метастабилен; стабильной формой углерода является другая его аллотропная модификация, а именно, графит.

Активация газовой фазы преследует двоякую цель:

- 1) создать достаточно высокую концентрацию активных углеродсодержащих частиц, которые, сталкиваясь с поверхностью нагретой подложки, распадаются с образованием атомов углерода, из которых и формируется кристаллическая решетка алмаза;
- 2) создать реагент, подавляющий осаждение неалмазных форм углерода – графита, аморфного углерода. Дело в том, что основным препятствием для роста алмазных пленок является не сама по себе метастабильность алмаза при низких давлениях, а именно параллельное образование и рост неалмазного углерода, блокирующего поверхность подложки. Таким реагентом служит атомарный водород, в атмосфере которого графит, условно говоря, «сгорает», а алмаз остается невредимым.

Основные методы активации реакционной газовой смеси – термические и электрические; намного реже применяется химическая и фотохимическая активация. При термической активации рядом с подложкой, на которую предполагается наращивать алмаз, располагают проволоку из тугоплавкого металла (W, Ta). Ее раскаляют до температуры, при которой диссоциация молекул  $H_2$  происходит с достаточно большой скоростью ( $\sim 2000^\circ C$ ). Газовая фаза представляет собой смесь углеродсодержащего газа (метан, пары ацетона, метанола), обычно в концентрации до нескольких процентов, с водородом. При контакте газа с поверхностью активатора образуются, наряду с атомами водорода, возбужденные углеродсодержащие молекулы и радикалы, которые переносятся к поверхности подложки, где и происходит процесс осаждения.

Для электрической активации используют различные формы электрического разряда (дуговой, тлеющий) с применением как постоянного, так и переменного тока (СВЧ-импульсный режим). Каждый из методов активации имеет свои достоинства и недостатки. Например, в дуговом разряде удастся реализовать высокие скорости осаждения: десятки и сотни микрометров в час. При СВЧ-активации возникающая плазма обеспечивает однородные условия осаждения на большой площади поверхности подложки.

Одним из основных достоинств метода осаждения из ионных пучков является возможность получения веществ, не существующих в природе или находящихся при нормальных условиях в метастабильном состоянии.

Рост пленки при ионно-лучевом осаждении позволяет получать плотные кристаллические модификации, не прибегая к традиционным методам.

Существует большое количество веществ имеющих ряд устойчивых кристаллических модификаций. В каждом состоянии решетка будет обладать своим минимумом свободной энергии. Так для углерода существует несколько устойчивых модификаций соответствующих графиту, алмазу, металлическому углероду и другим плотным алмазоподобным структурам. В таком случае обеспечить ионам энергию, необходимого для преодоления потенциального барьера, отделяющего одну кристаллическую фазу от другой, можно простым регулированием потенциала подложки. Причем разброс ионов по энергии не должен превышать разность в высоте потенциальных барьеров, разделяющих две близких кристаллических модификации. Максимальная энергия падающих ионов определяется энергетическим порогом дефектообразования (для алмаза  $60\div 80$  эВ). С учетом возможной потери энергии падающих ионов и диапазон их энергетического распределения является важнейшим, но не единственным условием, т.к. механизм взаимодействия при синтезе материалов из энергетических ион-

ных пучков сложен. Осаждение ионов на поверхность сопровождается релаксационными колебаниями, разогревом поверхности за счет выделения энергии. Перечисленные эффекты, безусловно, не охватывают весь комплекс явлений сопровождающих процесс конденсации. Действие некоторых из них будет отрицательным для синтеза. Воздействовать на степень того или иного эффекта можно различными путями:

- изменять температуру подложки;
- управлять условиями подлета ионов к поверхности подложки, например, изменением электростатического потенциала на последней;
- одновременно с осаждением ионов облучать поверхность электронными или ионными пучками;
- стимулировать или замедлять рост пленки ультрафиолетовым или лазерным излучением.

При подаче между анодом и мишенью постоянного напряжения от блока питания, в области мишени загорается газовый разряд. Со стороны анода в ускоритель поступает поток ионов, которые создаются вблизи анода путем ионизации нейтральных атомов. Металлические стенки ускорительного канала находятся под катодным потенциалом. Магнитная система создает в кольцевом зазоре ускорителя радиальное магнитное поле, постепенно спадающее в области анода и у среза ускорителя. Магнитное поле в основном имеет поперечную компоненту, а электрическое - продольную, поэтому такие системы называют еще ускорителями в скрещенных электрических и магнитных полях или «В-Е»-разрядом. В скрещенных магнитном и электрическом полях на электроны действует сила Лоренца, вызывая их дрейф в азимутальном направлении. В результате, в кольцевом зазоре формируется двойной азимутально-однородный электрический слой. В этих условиях подвижность электронов поперек магнитного поля резко

ограничена, и внешнее электрическое поле совершает работу главным образом над ионами, ускоряя их вдоль оси системы.

Плазма разряда ограничена магнитным полем и имеет форму кольца, расположенного над мишенью. Ионы рабочего газа из плазмы ускоряются в сторону мишени, бомбардирует её, выбивает атомы материала с поверхности мишени, часть которых, осаждаясь на подложке, образует покрытие. Напряжение мишени зависит от типа распыляемого материала, т.к. коэффициент вторичной эмиссии электронов, поддерживающих разряд, у различных веществ различен.

Ток мишени можно в небольшом диапазоне регулировать давлением рабочего газа.

Скорость роста пленок в подобных экспериментах малы. Помещая подложку в область с максимальной плотностью падающих частиц, увеличивается скорость осаждения. Зная геометрическое распределение потока можно оценить неоднородность толщины напыленной пленки или принять специальные меры для ликвидации этой неоднородности.

Алмазоподобные пленки, наряду с высокой твердостью имеют следующий уникальный набор физико-механических свойств:

- сверхнизкий коэффициент трения;
- высокая износостойкость;
- химическая стойкость;
- биологическая совместимость;
- яркая интерференционная окраска и оптическая прозрачность;
- прочное сцепление с основой –  $2\div 12$  кг/мм<sup>2</sup> (адгезия к стальной подложке сравнима с прочностью основы);
- микротвердость  $8000\div 10000$  HV<sub>0,2</sub>.
- плотность 2,92 г/см<sup>3</sup>;
- коэффициент преломления в пределах  $2,2\div 2,4$ .

Получаемые такими способами аморфные тонкие пленки являются, как правило, полупроводниковыми или диэлектрическими.

Путем легирования акцепторной примесью (бором) получают пленки с весьма хорошей проводимостью, которые являются при высоком уровне легирования квазиметаллическими.

Для того чтобы придать электропроводность хотя бы тонкому внешнему слою кристалла и таким образом получить возможность пропускания электрического тока, можно имплантировать в кристалл ионы. Но аргон вообще не является легирующей примесью в алмазе; примеси же кислорода и азота если и придают алмазу проводимость, то лишь при температурах намного выше комнатной. Основным результатом ионной имплантации является, очевидно, аморфизация внешнего слоя алмаза. В результате разрушения кристаллической решетки и появления большего числа атомов углерода с  $sp^2$ -гибридизацией C-C орбиталей в слое аморфизованного алмаза возникает заметная электропроводность.

## **1.4. Формирование нанобъектов фрактального типа**

В результате воздействия на поверхность тонких алмазоподобных пленок в газовой-паровой смеси при определенных соотношениях компонентов  $C_2HCl_3:Ar:H_2O$  происходит самопроизвольное образование объектов, отличающихся специфической формой. Подобные структуры появляются в результате процессов самосборки – формирование химических связей, по-видимому, осуществляется по энергетически наиболее выгодному сценарию в тех местах подложки, где реализована максимальная концентрация свободных водородных связей. Некоторые элементы таких структур самопроизвольно повторяют друг друга в различных масштабах.

В реальности происходит переход от стационарных равновесных состояний к самоорганизующейся диссипативной структуре, подчиняющейся законам синергетики и неравновесной термодинамике. Переход на нанометровый диапазон размеров элементов требует учета факторов масштабирования, отражающих влияние геометрических размеров на свойства материала. Уменьшение геометрических размеров объектов, составляющих наноструктуры, приводит к тому, что на объемные свойства материала начинают влиять его поверхностные характеристики, а с определенных размеров ( $\sim 100\text{\AA}$ ), поверхностные свойства начинают доминировать над объемными и преимущественно определять свойства системы в целом. Поверхность всегда имеет большую энергию, чем объем, и при этом стабильно сохраняет свои характеристики [27]. Попытка формирования абсолютно ровных и гладких линий или каких-либо симметричных структур нанометрового диапазона на реальных поверхностях различных веществ часто приводит к образованию объектов, форма которых с большой натяжкой может быть названа симметричной. Однако как раз в этих формах прослеживается симметрия другого рода. Эти объекты состоят из геометрических фрагментов различного размера и ориентации, но аналогичных по форме.

В литературе объекты такого рода называются фрактальными [49]. Фрактальность подразумевает наличие внутреннего свойства подобия на разных уровнях [50].

Фракталы проявляют хаотическое поведение, благодаря которому они кажутся такими беспорядочными и случайными. Но если взглянуть достаточно близко, можно увидеть много аспектов самоподобия внутри фрактала. Его хаотический вид, в действительности, достаточно упорядочен и подчиняется определенным законам. Проблема состоит в том, что отыскание этих законов может быть очень сложным. Цель изучения фрак-

талов — предсказать закономерность в системах, которые могут казаться непредсказуемыми и абсолютно хаотическими.

Пионером в этой новой области познания был Бенуа Б. Мандельброт. В середине 1960-х годов он разработал то, что впоследствии назвал фрактальная геометрия или геометрия природы [51]. Целью фрактальной геометрии был анализ сломанных, морщинистых и нечетких форм. Мандельброт использовал слово фракталы, предполагая осколочность и фракционность этих форм.

Фракталы находят все большее и большее применение в науке. Основная причина этого заключается в том, что они описывают реальный мир иногда даже лучше, чем традиционные физико-математические описания. Одной из идей, выросших из открытия фрактальной геометрии была идея нецелых значений для количества измерений в пространстве.

В физике поверхности фракталы используются для описания кривизны поверхностей. Неровная поверхность характеризуется комбинацией, по крайней мере, двух разных фракталов. Дело в том, что, поскольку свойства любой реальной поверхности характеризуются поведением элементарных частиц, находящейся на ней, - атомов, электронов, ионов, а также наличием свободных или оборванных связей, дефектов и пр., — то поверхность можно рассматривать как некий физический объект, свойства которого определяются совокупностью физических явлений двумерной поверхности, трехмерного объема, а в некоторых случаях и одномерной составляющей. Такой подход позволяет рассмотреть и выявить целый класс структур, имеющих дробную размерность в приблизительном диапазоне от 0,5 до 3,3, свойства которых существенно превышают свойства массивных материалов.

Реальные поверхности подложек, с которыми приходится работать экспериментатору, не являются совершенно плоскими, а объекты, находящиеся на ней или искусственно образованные, не абсолютно симметрич-

ные, то есть они в той или иной степени фрактальны. Следовательно, объект, имеющий точно 3 измерения невозможен. Вот почему концепция фрактального измерения была нужна для измерения степени неровности вещей. Например, фрактальная размерность листа бумаги приблизительно равна 2,5. Но его нормальная размерность, так же называемая Евклидовой размерностью будет равна 3. Все фракталы, особенно фрактальные кривые, имеют фрактальные размерности.

Существуют различные виды и типы фракталов, описанные в соответствующей литературе [51]. Первыми открытыми фракталами были т.н. детерминированные фракталы. Их отличительной чертой является свойство самоподобия, обусловленное особенностями метода их генерации.

Эти фракталы обычно формируются, начиная с инициатора – фигуры, к которой применяется определенный основной рисунок. Во всех детерминированных фракталах, самоподобие проявляется на всех уровнях. Это значит, что независимо от степени приближения наблюдается все тот же узор.

Детерминистские фракталы образуются в процессе, называемом итерацией, которая применяет основной рисунок к инициатору, после чего применяет его к результату и так далее. Эти фракталы линейны, так как при каждой итерации, что-то убирается либо прибавляется в форме прямых линий. Таковы фрактал Минковского и кривая Коха (рис. 1.4.1 и рис. 1.4.3).

Попытка создать линию на поверхности приводит к образованию объекта, подчиняющегося способностью к реакции данных составляющих систему элементов (молекул и/или атомов) на подложке. Подобные попытки представлены на рис. 1.4.2 и рис. 1.4.3.

Одной из центральных проблем в описании поведения поверхности является невозможность точного предсказания состояния системы. Квантовые частицы, определяющие поведение системы на реальных поверхно-

стях, строго детерминированы лишь в том случае, если мы постоянно измеряем, в каком состоянии они находятся. Но стоит прекратить наблюдение, как квантовые частицы тут же переходят из вполне определенного состояния сразу в несколько различных: их поведение носит вероятностный характер, что определяется суперпозицией квантовых состояний и описывается соответствующей волновой функцией.

Случайное Броуновское движение производит частотную диаграмму, которая может быть использована для предсказания вещей, включающих большие количества данных и статистики.[52].

Если исходные данные изменить даже на очень маленькие величины, скажем, эти величины малы настолько, что соизмеримы с колебаниями числа Авогадро (очень маленькое число порядка  $10^{24}$ ), проверка состояния аттрактора покажет абсолютно другие числа.

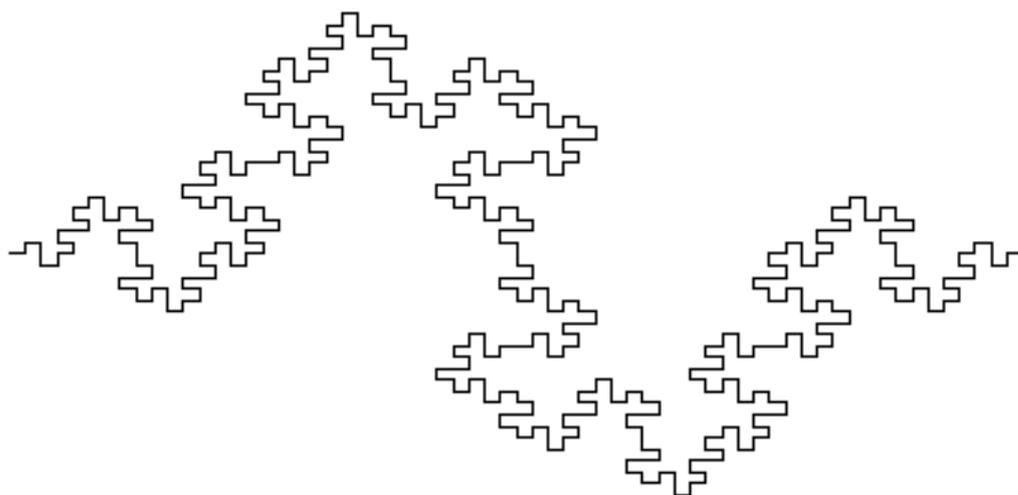
Это происходит потому, что маленькие различия увеличиваются в результате рекурсии.

Фрактальный анализ является одним из основных методов численной характеристики шероховатости поверхностей тонких нанокристаллических пленок. Определение фрактальной размерности позволяет получить меру для количественной оценки рельефа поверхности материала, развивающегося в процессе роста, взаимодействия с внешней средой, разрушения и т.п.

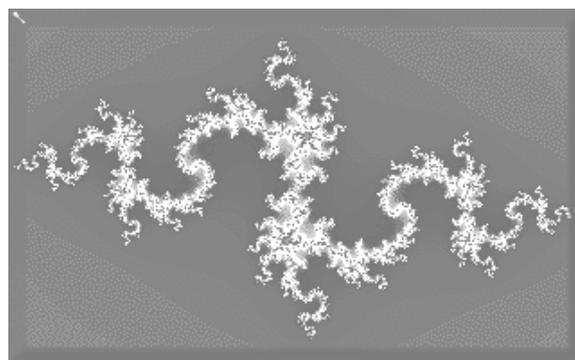
В сочетании с использованием сканирующей зондовой микроскопии, применение фрактального анализа является оптимальным методом контроля состояния поверхности многослойных структур на нанометровом уровне.

Нелинейные динамические системы обладают несколькими устойчивыми состояниями. То состояние, в котором оказалась динамическая система после некоторого числа итераций, зависит от ее начального состояния.

Поэтому каждое устойчивое состояние (аттрактор) обладает некоторой областью начальных состояний, из которых система обязательно попадет в рассматриваемые конечные состояния. Таким образом, фазовое



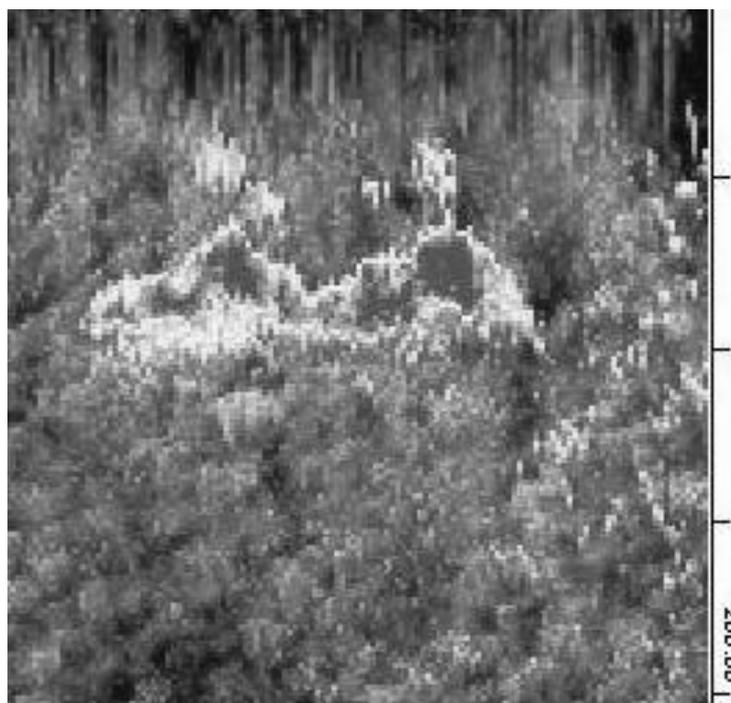
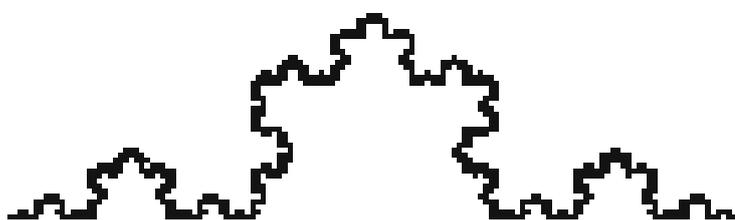
**Рис. 1.4.1.** Фрактал Минковского. Фрактальная размерность  $\ln 8 / \ln 4 = 1,5$



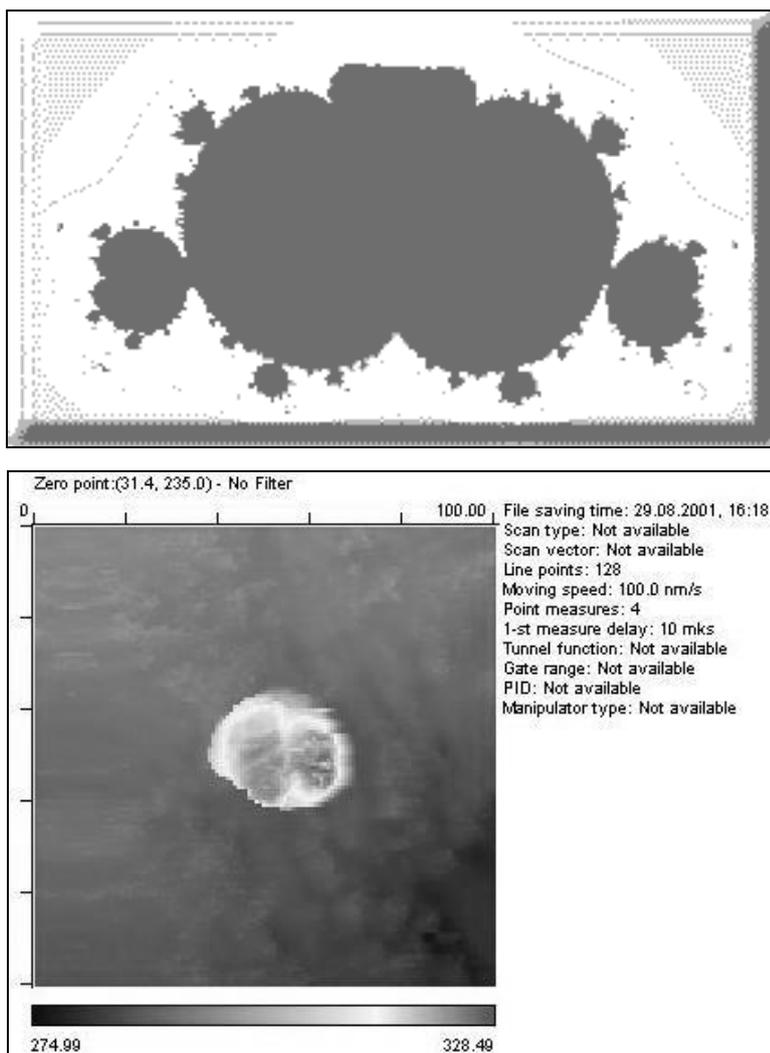
**Рис. 1.4.2.** Множество Жулиа и его экспериментальный аналог

пространство системы разбивается на области притяжения аттракторов.

В качестве примера можно рассмотреть множество Мандельброта (рис.1.4.4). Неожиданностью для математиков стала возможность с помощью примитивных алгоритмов порождать очень сложные нетривиальные структуры. Алгоритм его построения достаточно прост и основан на простом итеративном выражении:  $Z[i+1] = Z[i] * Z[i] + C$ , где  $Z[i]$  и  $C$  – комплексные переменные. Итерации выполняются для каждой стартовой точ-



**Рис. 1.4.3.** Теоретическая Кривая Коха и соответствующая экспериментальная линия на поверхности тонкой пленки  $\alpha$ -СН



**Рис. 1.4.4.** Множество Мандельброта размерностью 3,5 и его экспериментальный аналог

ки  $S$  прямоугольной или квадратной области – подмножество комплексной плоскости. Итерационный процесс продолжается до тех пор, пока  $Z[i]$  не выйдет за пределы окружности радиуса 2.

Существуют и другие классификации фракталов, например деление фракталов на детерминированные (алгебраические и геометрические) и недетерминированные (стохастические).

Известным классом фракталов являются стохастические фракталы, которые получаются в том случае, если в итерационном процессе случай-

ным образом менять какие-либо его параметры.

При этом получаются объекты очень похожие на природные – несимметричные деревья, изрезанные береговые линии и т.д. Двумерные стохастические фракталы используются при моделировании рельефа различных поверхностях [54].

Для описания процесса получения фрактальных структур оптимальным представляется использование метода под названием "Систем Итерированных Функций" (Iterated Functions System – IFS), который представляет собой систему функций из некоторого фиксированного класса функций, отображающих одно многомерное множество на другое.

Наиболее простая IFS состоит из аффинных преобразований плоскости:

$$\begin{aligned} X' &= A \cdot X + B \cdot Y + C \\ Y' &= D \cdot X + E \cdot Y + F \end{aligned} \quad (1.4.1)$$

Теоретическое обоснование метода изложено в [53]. Вкратце метод можно описать следующим образом. Изображение кодируется несколькими простыми преобразованиями (в нашем случае аффинными), т.е. коэффициентами этих преобразований. Например, закодировав какое-то изображение двумя аффинными преобразованиями, мы однозначно определяем его с помощью 12-ти коэффициентов. Если теперь задаться какой-либо начальной точкой (например,  $X=0$   $Y=0$ ) и запустить итерационный процесс, то мы после первой итерации получим две точки, после второй – четыре, после третьей – восемь и т.д. Через несколько десятков итераций совокупность полученных точек будет описывать закодированное изображение. Но проблема состоит в том, что очень трудно найти коэффициенты IFS, которая кодировала бы произвольное изображение.

Для построения IFS применяют кроме аффинных и другие классы простых геометрических преобразований на плоскости, которые задаются небольшим числом параметров. Например, проективные:

$$\begin{aligned} X' &= (A_1 \cdot X + B_1 \cdot Y + C_1) / (D_1 \cdot X + E_1 \cdot Y + F_1) \\ Y' &= (A_2 \cdot X + B_2 \cdot Y + C_2) / (D_2 \cdot X + E_2 \cdot Y + F_2) \end{aligned} \quad (1.4.2)$$

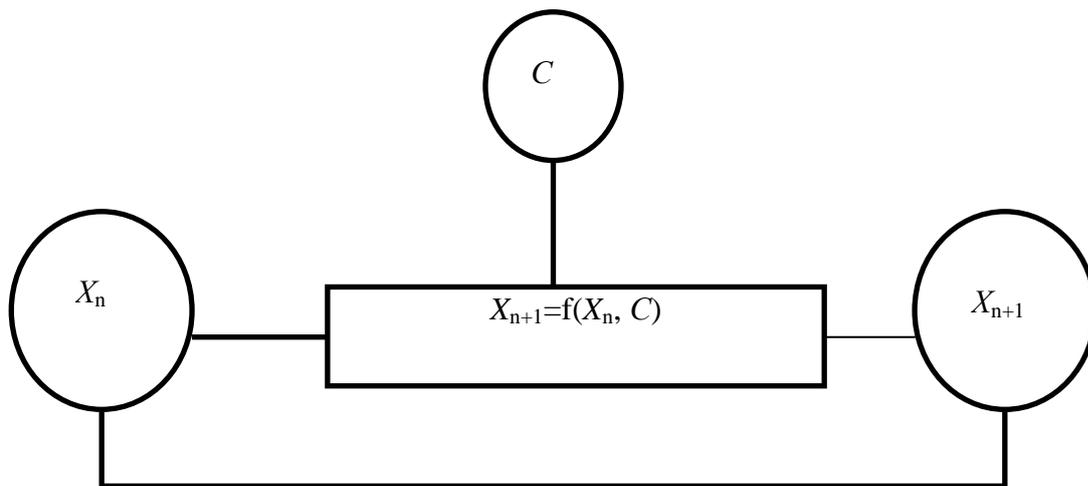
или квадратичные:

$$\begin{aligned} X' &= A_1 \cdot X \cdot X + B_1 \cdot X \cdot Y + C_1 \cdot Y \cdot Y + D_1 \cdot X + E_1 \cdot Y + F_1 \\ Y' &= A_2 \cdot X \cdot X + B_2 \cdot X \cdot Y + C_2 \cdot Y \cdot Y + D_2 \cdot X + E_2 \cdot Y + F_2. \end{aligned} \quad (1.4.3)$$

Схематично подобная процедура показана на рис. 1.4.5.

Здесь  $X_n$  – структура после  $n$  преобразований,  $C$  – алгоритм, по которому начальная структура  $X_n$  преобразуется в последующую  $X_{n+1}$ .

Фрактальная геометрия незаменима при генерации искусственных структур. Фактически – это способ представления сложных неевклидовых объектов, образы которых весьма похожи на природные.



**Рис. 1.4.5.** Математическая модель процесса формирования фрактальной структуры

## ЛИТЕРАТУРА К ГЛАВЕ 1

1. Нанотехнология в ближайшем десятилетии. Прогноз направления исследований. Под ред. М.К.Роко, Р.С.Уильямса и П.Аливасатоса. Пер. с англ. к.ф.-м.н. А.В.Хачояна под ред. д.т.н., проф. Р.А.Андриевского, М.: Мир, 2002.
2. К.А.Валиев, А.А.Орликовский. От микро- и наноэлектроники к твердотельным квантовым компьютерам. Успехи современной радиоэлектроники, № 5-6, 2004, стр. 106-117.
3. Драгунов В.П., Неизвестный И.Г., Гридчин В.А. Основы наноэлектроники. М.: Логос. 2006. - 496с.
4. Ed. by Kazuaki Suzuki, Bruce W. Smith Microlithography: Science and Technology / CRC Press; 2 edition, 2007.
5. J. Orloff, M. Utlaut and L. Swanson High Resolution Focused Ion Beams: FIB and Its Applications. — Springer Press, 2003.
6. Eung-sug Lee, Jun-ho Jeong, Young-suk Sim, Ki-don Kim, Dae-geun Choi, Jun-hyuk Choi. High-throughput step-and-repeat UV-nanoimprint lithography // Current Applied Physics. 6S1 (2006) e92-e98.
7. R. Fabian Pease, Stephen Y. Chou. Lithography and Other Patterning Techniques for Future Electronics // Proceedings of the IEEE. Vol. 96, No. 2, February 2008 P. 624-627.
8. Неволин В.К. Зондовые нанотехнологии в электронике. - М.: Техносфера, 2006. - 160 с.
9. Binning G., Rohrer H. Scanning tunneling microscopy. IBM J.Res.Develop. V.30. №4. P.355. 1986.
10. Лускинович П.Н., Ананян М.А., Дадан Е.В., Кузькин В.И. и др. Нанотехнологические процессы и установки. Труды научного

- семинара «Математическое моделирование нанотехнологических процессов и наноструктур», 2001, вып.1, с.30-47.
11. Feynman R. There's plenty of room at the bottom: an invitation to enter a new field of physics. Eng. Sci. 23, 22 (1960).
  12. Taniguchi N. On the basic concept of nanotechnology. Proc. Int. Conf. Prog. Eng. Part II. Tokyo: Jap. Soc. of Prec. Eng., 1974.
  13. Eigler D.M., Schweizer E.K. Positioning single atoms with a scanning tunneling microscope. Nature (1990) v.344, No. Pt. p.524-526.
  14. Drexler K.E. Molecular engineering: an approach to the development of general capabilities for molecular manipulation. Proc. Natnl. Acad. Sci. USA 78, 5275 (1981).
  15. New Scientist 1992, 133 (1811), pp. 42-46.
  16. Richter S., Cachen D., Cohen S.R., Gartsman K., Lyakhovistskaya V., Nanassen V. // Appl. Phys. Lett. 1998. V 73. P. 1868.
  17. Miller R.E., Shenoy V.B. // Nanotechnology. 2000. V. 11. № 3. P. 139-147
  18. Dryakhlushin V.F., Klimov A. Yu., Rogov V.V., Shashkin V.L, Suchodoev L.V., Volgunov D.G., Vostokov N.V. // Nanotechnology. 2000. V. 11. №3. P. 188-191.
  19. Гусев А.И., Ремпель А.А. Нанокристаллические материалы. М.: Физматлит, 2000.
  20. Workshop Report: SRC/NASA Ames Workshop on Emerging Opportunities and Issues in Nanotubes and Nanoelectronics; Workshop: November 12-13, 1998; (SRC/NASA), D. Herr, M. Meyyappan and V. Zhirnov
  21. «National Nanotechnology Initiative: Leading to the Next Industrial Revolution», Internal government report, supplement to the President's FY 2001 Budget, February 2000.

22. "Nanostructure Science and Technology", NSTC Report, 336 pages; Worldwide status and trends; Also: hard copy publication by Kluwer Academic Publ., October 1999.
23. "Nanotechnology: Shaping the World Atom by Atom", NSTC Report, 12 pages; Color brochure for the public.
24. "Russian Research and Development Activities on Nanoparticles and Nanostructured Materials" (proceedings of an August 1997 workshop held in St. Petersburg, Russia)
25. "R&D Status and Trends in Nanoparticles, Nanostructured Materials, and Nanodevices in the United States" (proceedings of a May 1997 workshop held in Arlington, VA)
26. Ормонт Б.Ф., Введение в физическую химию и кристаллохимию полупроводников. Издание второе, переработанное и дополненное. М; Высшая школа. 1973. с. 389.
27. Петров Ю.И. Кластеры и малые частицы. М. : Наука. 1986.
28. Abraham F.F. CRC Crit. Revs Solid State and Mater. Sci., 1981, May, 169.
29. Wunderlich W., Ishida Y., Maurer R. Scripta Metall. Mater. 1990, 24, No. 2, P. 403.
30. Thomas G.J., Siegel R.W., Eastmen J.A. Scripta Metall. Mater. 1990, 24, No. 1, P. 201.
31. Mutschele T., Kirchheim R. Scripta Met. 1987, 21, No. 2, P. 135.
32. Sleptsov V.V., Elinson V.M., Baranov A.M., Tereshin S.A. Optical and Electrical properties of quantum-dimensional Multilayer Structures Based on Carbon Films. Wide Band Gap Electronic Materials NATO ASI Series. 3 High Technology, 1995,1, P. 257-264.
33. Polyakov V.I., Perov P.I., Ermakov M.G., Ermakova O.N., Sleptsov V.V., Elinson V.M. Thin Solid Films, 1992, 212, P. 226-231.

34. Баранов А.М., Елинсон В.М., Кондратов П.Е., Слепцов В.В. "Влияние толщины пленок а-С на механизм переноса носителей заряда". Письма в ЖТФ, 1989,16(3), С. 36-39.
35. Sleptsov V.V., Kyzin A.A., Baranov A.M., Elinson V.M. Optical absorption in ct-C:H multilayer periodic structures, *Diamond and Related Materials*, 1992,1, P. 570-571.
36. Sleptsov V.V., Kyzin A.A., Baranov A.M., Elinson V.M. Electrical and optical properties of carbon films. In Book "Physics and Technology of Diamond Materials", Poland Publishers, Moscow, 1994, p.80-87.
37. Meng F., Schlup J.R., Fan L.T. *Chem Mater.*, 1997, 9, p. 2459.
38. Смирнов Б.М. Физика фрактальных кластеров. М.: Наука, 1991.; Усп. физ. наук, 1991. 161. N 6. 141; 1993. 163. N 7. 51.
39. Hou J.G., Wu Z.Q. *Phys. Rev., B*, 1990, 42, 3271
40. Dan Nir, Intrinsic stress in diamond-like carbon films and its depends on deposition parameters. *Thin Solid Films*, 1987,146, P. 27-43.
41. Rossi E., Andre B. Low-temperature synthesis of diamond films in termoassisted RF-plasma CVD. *Japaness J, Appl. Phys.*, 1992, 31(3), С. 872-879.
42. Регель В.Р. и др. Кинетическая природа прочности твердых тел. М., «Наука», 1974, стр. 110.
43. Vossmeier T., Katsikas L., Giersig M., Popovic I., *J. Chem. Phys.*, 1994, 98, P.7665—7673.
44. Чвалун С.Н. Полимерные нанокompозиты. *Природа*, 2000, №7
45. Ю.В. Плесков, "Электрохимия алмаза". УРСС. Москва, 2003.
46. Вавилов В.С. // УФН. 1997. Т. 167. № 1.
47. Himpsel F.J., Knapp J.A., Van Vechten J.A., Eastman D.E. // *Phys. Rev. B*. 1979. Vol. 20. P. 624-627.

48. Пшеничнюк С.А., Юмагузин Ю. М., Бахтизин Р.З. // ПЖТФ. 1999. Т. 25. Вып. 15. С. 46-52.
49. Леонард М. Сандер, Фрактальный рост, пер. с англ., В мире науки, 1987, т.3, с.62.
50. Львов Г.А. Фрактальные среды. СПб.: Изд-во СПбГТУ, 2001.
51. В.В.Mandelbrot, The Fractal Geometry of nature. San Francisco: Freeman 1982.
52. <http://www.ghcube.com/fractals/>
53. Бондаренко В.А., Дольников В.Л. Фрактальное сжатие изображений по Барнсли-Слоану. Автоматика и телемеханика.-1994. №15. с.12-20.
54. Витолин Д. Применение фракталов в машинной графике. Computerworld-Россия.-1995. №15. с. 11.

## **ГЛАВА 2. ПРОИЗВОДСТВО И ПРИМЕНЕНИЕ ФУЛЛЕРЕНОВ И УГЛЕРОДНЫХ НАНОТРУБОК**

### **2.1. Автоматизированное проектирование процесса производства углеродных нанотрубок и фуллеренов**

Проектирование как инвариантная по отношению к конкретным объектам форма инженерной деятельности имеет ряд характерных особенностей [1]:

1. Процесс имеет итерационный характер.
2. Решения принимаются на отдельных этапах в условиях неполной или недостаточной информации, которая в эти случаях поступает или из внешней среды, или вырабатывается проектировщиком в процессе творческой деятельности.
3. В процессе проектирования сочетаются процедуры алгоритмического и эвристического характера.
4. В проектной деятельности используются различные ресурсы, среди которых одним из наиболее важных являются знания проектировщика.
5. Цель проектирования устанавливается вне процесса проектирования и остается неизменной в течении этого процесса.
6. Процесс проектирования производит информацию, которая может быть использована в производстве.

Информационная система поддержки принятия решений при проектировании оборудования играет роль мощного средства, эффективное применение которого невозможно без разработки комплекса методических указаний и инструкций, регламентирующих последовательность этапов и используемых на каждом этапе. Поскольку на каждом этапе автоматизиро-

ванного проектирования осуществляются различные операции с материальными и нематериальными (информационными) объектами, а также возникает проблема наиболее эффективного распределения этих операций во времени и оптимального соотношения в пространстве с целью экономии трудовых и материальных ресурсов, то представляется целесообразной необходимость разработки и отработки технологии автоматизированного проектирования [1].

Процесс проектирования оборудования начинается со сбора информации о спроектированных разновидностях объекта, результатах выполненных научно-исследовательских работ, сбора данных об испытаниях аналогов, условиях снабжения материалами и т.д. На основании анализа требований к техническим характеристикам, условий изготовления и эксплуатации, выявления тенденций развития оборудования для производства наноматериалов составляется и корректируется исходное техническое задание на проектирование [1].

Реализация технологии автоматизированного проектирования оборудования для производства наноматериалов предъявляет к разрабатываемой системе комплекс следующих требований:

- Наличие автоматизированного рабочего места проектировщика;
- наличие математической модели;
- возможность формулировать ТЗ на проектирования в понятной форме для проектировщика;
- обеспечения необходимой проектировщикам ясности и однозначности формулировок цели проектирования (лучше всего описать в терминах характеристик системы);
- наличие средств, для эффективной корректировки задания на проектирование;
- отсутствие жестких ограничений на структуру и объем входных данных и формы носителей информации, на которых они хранятся;

- возможность оперативного подключения к программному обеспечению системы новых модулей и исключение устаревших;
- представление возможностей проектировщику на основе промежуточных результатов принимать решение о выборе методов для продолжения проектной задачи, а также изменений значений отдельных параметров в используемом методе решения;
- возможность в ходе выполнения проектных операций проследить значения основных показателей процесса, свидетельствующих о его эффективности, и в зависимости от их значений корректировать вычислительный процесс;
- наличие актуальной базы знаний, базы материалов;
- допустимость включения обучающих программ для повышения квалификации проектировщика.

## **2.2. Применение углеродных наноматериалов и технологий на их основе**

Нанoeлектроника является новой областью науки и техники, формирующейся сегодня на основе последних достижений физики твердого тела, физической химии и технологии полупроводниковой электроники. Ее содержание определяется необходимостью установления фундаментальных закономерностей, определяющих физико-химические особенности формирования наноразмерных структур (структур с размером от единиц до десятков нанометров,  $1\text{ нм} = 10^{-9}\text{ м}$ ), их электронные и оптические свойства. Исследования в области нанoeлектроники важны для разработки новых принципов, а вместе с ними и нового поколения сверхминиатюрных быстродействующих систем обработки информации [2].

Можно выделить три направления развития нанотехнологий, тесно связанные между собой:

- изготовление электронных схем (в том числе объемных) с активными элементами, чьи размеры сравнимы с размерами единичных молекул или атомов;
- разработка и изготовление наномашин, т.е. механизмов роботов величиной с молекулу, использование которых открывает перед человечеством невиданные перспективы;
- непосредственная манипуляция атомами и молекулами и «сборка» из них всевозможных материалов. Эта задача в свою очередь распадается на две концепции. Первая – перестройка имеющихся структур (например, перестроив порядок атомов в графите, можно «изготовить» алмаз). Вторая – сборка большего из меньшего (так, используя молекулы воды и углекислого газа, можно «изготовить» из них сахар или крахмал, как это делают растения) [3].

На основе наноструктур возможно создание новых материалов с заранее заданными свойствами – более прочных, лёгких, твёрдых и экологических по сравнению с известными. Наноструктурированные углеродные и керамические материалы в десятки раз более прочные, чем сталь, могут быть использованы в промышленности и транспорте, полимерные материалы в три раза более прочные и примерно на 100°C более тугоплавкие, чем существующие в настоящее время, найдут свое применение как в наземных транспортных средствах, так и в самолётостроении.

Уже достаточно давно ведутся работы по применению нанотрубок в компьютерной индустрии, а именно, в области средств отображения информации. Например, созданы и опробованы прототипы тонких плоских дисплеев, работающих на матрице из нанотрубок (т.н. «автоэмиссионные дисплеи»). Под действием электрического поля нанотрубки испускают электроны, которые попадают на фосфоресцирующий экран и вызывают свечение пикселей. Получающееся при этом зерно изображения будет

чрезвычайно малым — порядка микрона [6], что дает возможность производить высокоразрешающие микродисплеи.

Одно из возможных революционных решений задачи записи и хранения информации связано с применением эндоэдральных комплексов. Эндоэдраны – это сложные объекты, представляющие собой фуллерен  $C_N$ , состоящий из  $N$  атомов углерода, и «начинённый» внутри почти любым элементом таблицы Менделеева или даже простой молекулой. Эндоэдраны с редкоземельными элементами (гадолиний, диспрозий) обладают большим магнитным моментом и представляют собой магнитные диполи. Различные ориентации такого диполя могут служить состояниями, соответствующими "1" и "0". То есть эндоэдрал может рассматриваться в качестве элемента хранения одного бита информации. Величина магнитного момента позволяет управлять ориентацией диполя с помощью внешнего (записывающего) магнитного поля. Если принять размер диполя с прилегающим буферным пространством порядка 10 нм, то плотность записи информации на носитель, состоящий из ячеек-эндоэдралов, может достигать  $10^{10}$  бит/мм<sup>2</sup>. В оптических дисках на сегодняшний день этот параметр не превосходит  $5 \cdot 10^5$  бит/мм [4].

Делаются попытки создания компьютерной памяти на основе нанотрубок. Элементы памяти, в этом варианте, должны состоять из коротких отрезков закрытых нанотрубок диаметром порядка 1,4 нм. Внутри каждой нанотрубки находится молекула  $C_{60}$ , содержащая ион щелочного металла. Под действием электрического поля такая молекула сможет перемещаться в одну (бит 0) или в другую (бит 1) сторону вдоль нанотрубки. Оценки показывают, что для перемещения молекулы, то есть для переключения состояния "0"- "1", потребуется время около 4 пс, что на порядок меньше, чем у обычных переключающих элементов [4].

Международная группа ученых показала, что нанотрубки можно использовать для создания «искусственных мускулов», которые при одина-

ковом объеме могут быть втрое сильнее биологических, не боятся высоких температур, и многих химических реагентов [6].

Нелинейные оптические свойства фуллеренов, введенных в жидкое или твердофазное вещество (матрицу), позволяют создать высокоэффективные устройства для защиты глаз и приборов от поражения оптическим излучением. К таким устройствам предъявляются сложные требования:

- для того, чтобы не ухудшать условия наблюдения при умеренных уровнях освещенности, фильтр должен обладать высокой прозрачностью ( $0,9 \div 0,98$ ) при низкой интенсивности оптического потока в широкой полосе спектра ( $400 \div 800$  нм),
- чтобы надежно защитить глаза или оптические приборы от внезапной вспышки света, время реакции на изменение интенсивности облучающего потока должно быть минимально (не более нескольких наносекунд),
- коэффициент поглощения должен меняться в широком диапазоне (вплоть до нескольких порядков),
- высокая лучевая прочность (до  $100$  кДж/м), позволит фильтру сохранять работоспособность при интенсивном облучении,
- малые габариты и низкое энергопотребление сделают устройство удобным в эксплуатации.

Область применения оптических фильтров более чем широка:

- защитные очки для сварщиков;
- защитные очки для водителей, управляющих автомобилем в ночное время;
- защитное покрытие «фонаря» кабины летчика боевого самолета;
- защитные шлемы, предохраняющие солдат от воздействия лазерного оружия;

- нанесение защитных пленок на окна домов, расположенных в субтропиках и тропиках;
- прочие применения в технологических и исследовательских процессах.

Имеющиеся в настоящее время другие типы фильтров (их еще называют оптическими модуляторами, поглотителями), не отвечают в полной мере вышеприведенным требованиям.

Оптические свойства наноматериалов позволяют создать на их основе литографические резисты нового поколения. Эти резисты должны удовлетворять следующим двум требованиям:

- реализуемое разрешение, при заданном уровне чувствительности, должно быть не хуже десятков нм;
- возможность использования материала для работы одновременно с ультрафиолетовым, рентгеновским и электроннолучевым потоками.

Пленки  $C_{60}$  и  $C_{70}$  проявляют высокую чувствительность в указанных диапазонах длин волн актиничного излучения. При облучении пленки фуллеренов имеют место два фотохимических эффекта. Во-первых, это образование C—C связей между молекулами фуллеренов в отсутствие кислорода. Во-вторых – это так называемая фотоиндуцированная диффузия кислорода в решетку фуллерена. Оба эти эффекта ведут к потере растворимости пленок фуллерена в ароматических растворителях, например, в толуоле. Применение пленки фуллерена в качестве литографической маски позволило достигнуть нанометровых разрешений [4].

Другой пример использования нанотрубки – «насаживание» ее на острие сканирующего туннельного или атомного силового микроскопа. Обычно такое острие представляет собой остро заточенную вольфрамовую иглу (для заточки используются электрохимическое травление и бомбардировка тяжелыми ионами), однако, по атомным меркам подобная заточка все равно достаточно «грубая». Нанотрубка же представляет собой идеальную иглу диаметром порядка нескольких атомов. Приклады-

вая определенное напряжение, можно подхватывать атомы и целые молекулы, находящиеся на подложке непосредственно под иглой, и переносить их с места на место [5].

Целесообразно использовать зонды из нанотрубок с целью повышения величины межатомного взаимодействия между вершиной зонда и подложкой, что увеличивает размеры области взаимодействия. Наиболее эффективным способом локализации является использование трехмерного потенциального барьера, создаваемого на конце нанотрубки [5].

Применение нанотрубок в качестве острых зондов позволяет повысить достоверность получаемой информации благодаря большому отношению диаметр/длина трубки [4].

Исследователи из Гарвардского университета P. Kim и Ch. Lieber оснастили сканирующий зондовый микроскоп зондом в форме пинцета. Зонд-пинцет сделан из углеродных нанотрубок диаметром  $1 \div 2$  нм. Таким «двухзубым» зондом, если приложить к нему электрическое поле, можно захватывать и двигать молекулы и атомы; использовать его в качестве электрических контактов к исследуемому образцу или в качестве зондов иммерсионного типа для исследования жидкости.

Другой пример применения нанотрубок в сверхминиатюрных механических устройствах – создание элементов наноробототехники. Адсорбированные на поверхности нанотрубок частицы могут выполнять роль молекулярных зубцов, превращая нанотрубку в зубчатую передачу. На рис. 2.1.1 представлено схематическое изображение такого узла [4].

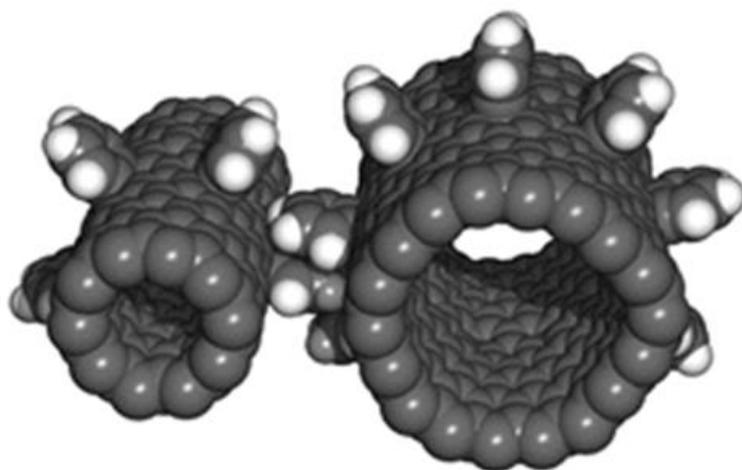
Моделирование молекулярно-зубчатых передач позволяет предсказать, что с помощью данных устройств может быть достигнута скорость передачи углового момента вплоть до 100 ГГц. Валами шестеренок в такой «нанокоробке передач» являются углеродные нанотрубки, а зубцами могут служить молекулы бензола. Характерные частоты вращения шестеренок составляют несколько десятков гигагерц. Устройства «работают» либо в

глубоком вакууме, либо в инертной среде при комнатной температуре. Для «охлаждения» устройства предлагается использовать инертные газы.

Применение фуллеренов в триботехнике позволяет оптимизировать характеристики смазок и продлить срок службы узлов, содержащих пары трения. Установлено, что добавки фуллеренов даже в незначительных количествах в смазочные масла заметно улучшают условия работы узлов трения механизмов [4].

На сегодняшний день объяснение снижения трения в присутствии фуллеренов сводится к следующей простой модели: молекула фуллерена имеет шаровидную форму и выполняет роль аналогичную шарик в шарикоподшипнике. Другими словами, молекулы фуллеренов обеспечивают частичный переход от трения скольжения к трению качения и препятствуют массопереносу вещества под действием сил трения. Высокая механическая прочность молекул снижает вероятность возникновения задиров и уменьшает т.н. «выработку контактирующих поверхностей» [4].

Использование фуллеренов в смазочных маслах позволяет увеличить давление на трущиеся поверхности в  $1,5 \div 2$  раза, снизить работу сил трения в  $2 \div 2,5$  раза, более чем вдвое уменьшить объемный износ трущихся поверхностей [4].



**Рис. 2.1.1.** Зубчатая передача на основе нанотрубок [8]

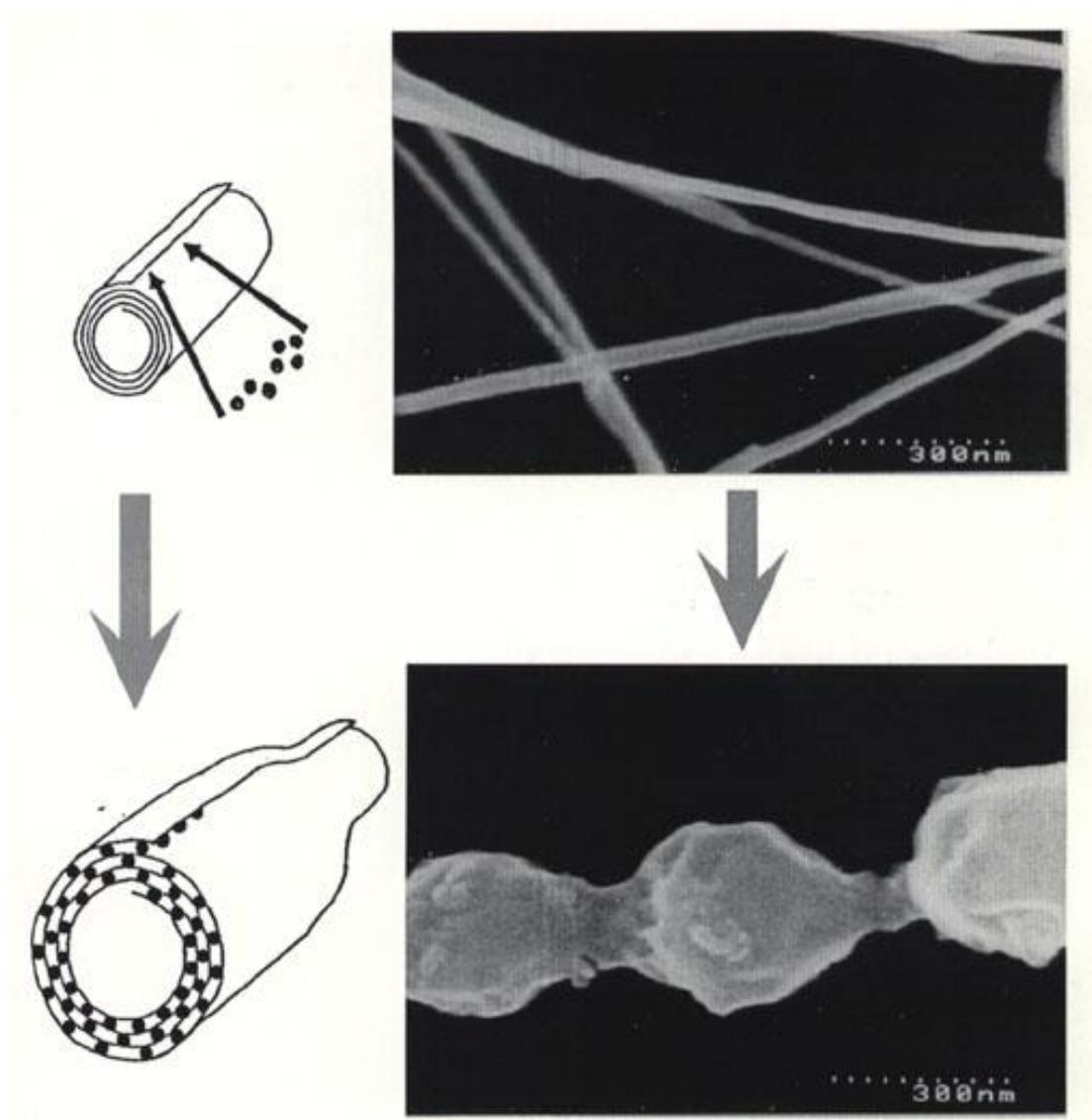
Применение фуллеренсодержащих присадок к твердым смазкам приводит к еще более ощутимым положительным результатам. Так например, добавление фуллеренсодержащей сажи в твердую смазку тяжело нагруженных подшипников полуосей электропоездов метрополитена увеличило срок эксплуатации подшипников с 6 месяцев до 2 лет.

Большое количество проектов посвящено капиллярным явлениям в нанотрубках. В самом деле, заполнение нанотрубки атомами металла создало бы металлический провод диаметром десятков ангстрем, готовая деталь сверхминиатюрной электроники. Уже удалось частично (пока только на 1%) провести капиллярное заполнение многослойных нанотрубок жидким свинцом, висмутом [11].

Зато межслоевые пространства рулонных нанотрубок охотно принимают в себя моноатомные слои щелочных металлов и мономолекулярные слои различных летучих хлоридов (реакция внедрения из горячей газовой фазы). Схема этой реакции и электронные микрофотографии нанотрубок до и после внедрения показаны на рис. 2.2.2.

Развитие и реализация методов заполнения нанотрубок различными материалами в настоящее время является бурно развивающимся направлением исследований. Нанотрубка может использоваться в качестве как носителя заполняющего ее материала, так и изолирующей оболочки, предохраняющей данный материал от электрического контакта либо от химического взаимодействия с окружающими объектами. Один из примеров использования нанотрубок – хранилище для газообразного водорода. Практическая реализация этой разработки может способствовать созданию высокоэффективного экологически безопасного автомобильного транспорта на водородном топливе [12].

Существует возможность применения углеродных нанотрубок для изоляции и хранения радиоактивных отходов. Преимущества такого спо-



**Рис. 2.2.2.** Внедрение частиц в нанотрубки-рулоны

способа хранения связаны с тем обстоятельством, что радиоактивный материал, капсулированный внутри нанотрубки, в химическом отношении инертен и не должен вымываться подземными водами из хранилища. Тем самым можно рассчитывать на повышенную безопасность такого способа хранения материала по сравнению с традиционными, сопровождаемыми значительными утечками радиоактивных материалов.

Первые исследования капиллярных явлений показали, что имеется связь между величиной поверхностного натяжения жидкости и возможно-

стью ее втягивания внутрь канала нанотрубки. Оказалось, что жидкость проникает внутрь канала нанотрубки, если ее поверхностное натяжение не выше 200 мН/м. Поэтому для ввода каких-либо веществ внутрь нанотрубок используют растворители, имеющие низкое поверхностное натяжение. Так, например, для ввода в канал нанотрубки некоторых металлов используют концентрированную азотную кислоту, поверхностное натяжение которой невелико (43 мН/м). Затем проводят отжиг при 400°C в течение 4 ч в атмосфере водорода, что приводит к восстановлению металла. Таким образом, были получены нанотрубки, содержащие никель, кобальт и железо [13].

Высокая удельная поверхность нанотрубок открывает широкие перспективы по их промышленному применению. Нанотрубки могут, переплетаясь между собой случайным образом, образовывать спиралевидные структуры. Они имеют огромное число полостей нанометрового размера, доступных для проникновения извне жидкостей и газов. В результате удельная поверхность материала, составленного из нанотрубок, оказывается близкой к соответствующей величине для индивидуальной нанотрубки. Это значение в случае однослойной нанотрубки составляет около 600 м<sup>2</sup>г<sup>-1</sup>. Столь высокое значение удельной поверхности нанотрубок открывает возможность их использования в качестве сорбентов в фильтрах, в аппаратах химической технологии и др. [4].

Весьма разнообразны варианты использования нанотрубок в электронике. Такие свойства нанотрубки как малые размеры, меняющаяся в значительных пределах (в зависимости от условий синтеза) электропроводность, механическая прочность и химическая стабильность позволяют рассматривать нанотрубку в качестве основы будущих элементов нанoeлектроники. Возможная схема одного из таких элементов связана с программируемым введением дефектов в структуры нанотрубок. Например, внедрение в идеальную структуру однослойной нанотрубки в качестве дефекта пары пятиугольник-семиугольник (рис. 2.2.3) изменяет ее хиральность и, как следствие, ее электронные свойства. Мо-

делирование свойств нанотрубок со структурой хиральностей (8,0)/(7,1) показало следующее: трубка с хиральностью (8,0) представляет собой полупроводник с шириной запрещенной зоны 1,2 эВ, в то время как трубка с хиральностью (7,1) является полуметаллом, для которого ширина запрещенной зоны равна нулю. То есть, такая нанотрубка представляет собой аналог барьера Шоттки (металл-полупроводник). Аналогичным образом в результате внедрения дефекта могут быть получены гетеропереходы полупроводник-полупроводник с различными значениями ширины запрещенной зоны. Тем самым нанотрубка с внедренными в нее дефектами может рассматриваться как вентильный элемент рекордно малых размеров. Задача внедрения дефекта в идеальную структуру однослойной нанотрубки представляет определенные технические трудности, однако, не является непреодолимой проблемой [4].

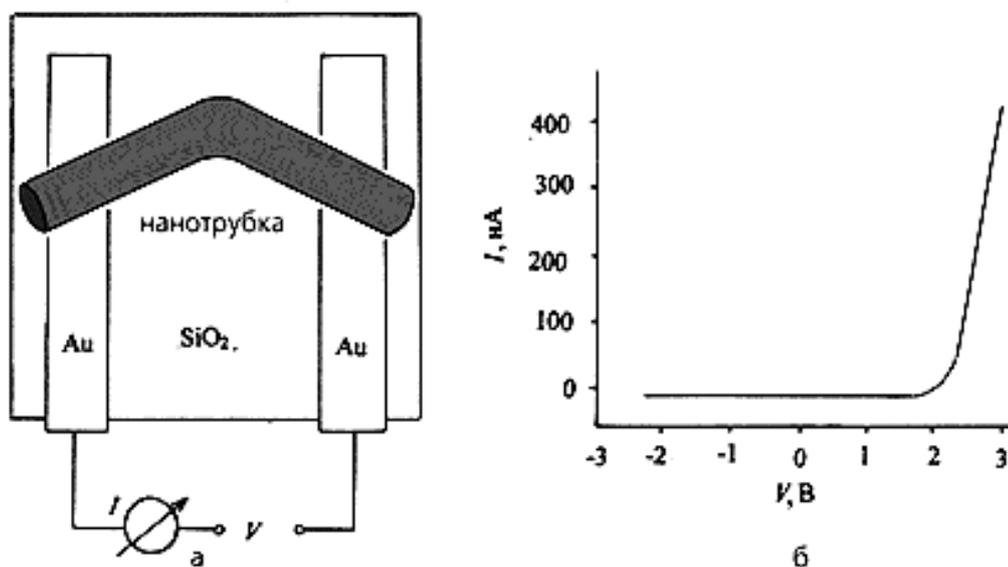
Группа концентрических нанотрубок, вложенная внутрь другого набора трубок, может скользить назад и вперед миллиард раз в секунду. Такой гигагерцовый осциллятор мог бы стать большим прорывом в области нанотехнологий, послужив основой для таких устройств, как сверхбыстрые оптические фильтры и наноантенны [15].

Если вытянуть внутреннюю часть концентрического набора углеродных нанотрубок и отпустить, то такой "поршень" (рис.2.2.4) начнет колебаться назад и вперед на гигагерцовых частотах - быстрее, чем любой другой механический осциллятор.

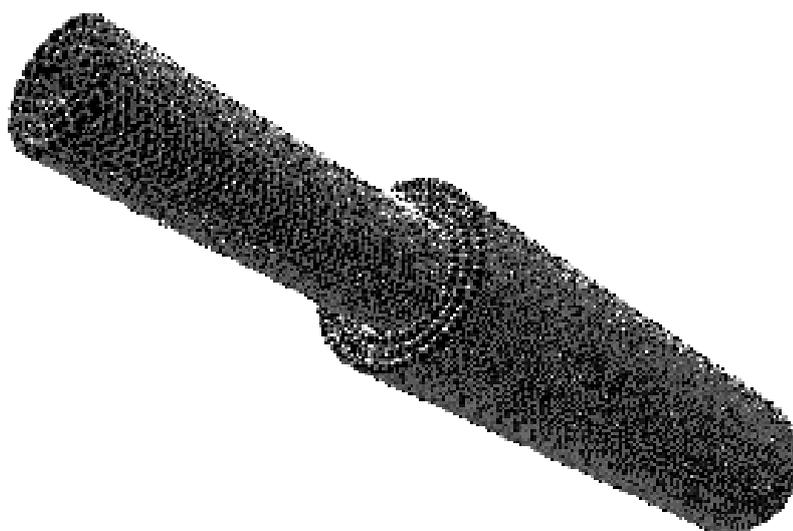
Большие надежды исследователи возлагают на создание технологии получения нанопроводников, заключенных внутрь однослойных нанотрубок. Решение этой проблемы, а также родственных ей проблем создания на основе таких проводников электронных схем может привести к уменьшению минимальных размеров элементов микроэлектронных устройств на один -два порядка и, в конечном счете, к очередному скачку в технологиях обработки и передачи информации [4].

Весьма обнадеживающими представляются перспективы использования нанотрубок в химической технологии, что связано, с одной, стороны, с их высокой удельной поверхностью и химической стабильностью, а с другой стороны – с возможностью присоединения к поверхности нанотрубок разнообразных радикалов, которые могут служить в дальнейшем либо каталитическими центрами, либо зародышами для осуществления разнообразных химических превращений [4].

Ученые Bell Labs, научно-исследовательского центра Lucent Technologies, изготовили индивидуально адресуемый транзистор, канал которого состоит всего из одной молекулы. Канал транзистора располагается между его электродами, и именно там происходит процесс электронной коммутации и усиления [16].



**Рис. 2.2.3.** Выпрямляющий диод на изогнутой нанотрубке. Нанотрубка лежит на непроводящей (кварцевой) подложке в контакте с двумя сверхтонкими проводами (а); вольт-амперная характеристика для такой системы (б) [14]



**Рис. 2.2.4.** Телескопическая трубка

Новые транзисторы изготовлены на базе тиолов, органических полупроводниковых веществ, в состав которых входят углерод, водород и сера. Была использована техника «самосборки». В кремниевой подложке было сделано углубление, покрытое слоем золота – это был один из трех электродов транзистора. Затем подложка погружалась в раствор тиола и инертного органического вещества, использовавшегося для получения нужной концентрации тиола. После высушивания пластины на золотом электроде оставалась пленка толщиной ровно в одну молекулу. Подбирая соотношение количества молекул тиола и инертного вещества, обеспечивалось наличие только одной активной молекулы над золотым электродом. Затем на пленку наносился еще один слой золота, который и служил третьим электродом [16].

Одним из показательных примеров, демонстрирующих возможность эффективного применения нанотрубок в химической технологии, может служить использование материала нанотрубок в качестве несущей под-

ложки для химических катализаторов. Для производства метана используется, протекающая в присутствии катализатора, реакция газофазной гидрогенизации СО:  $\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$ .

Реакция идет на поверхности катализатора, а эффективность катализатора при прочих равных условиях оценивается глубиной конверсии реагентов. Глубина конверсии измеряется в процентах и показывает, сколько процентов исходного сырья участвовало в реакции [4].

Для увеличения глубины конверсии необходимо увеличивать площадь поверхности катализатора. Катализатором для рассматриваемой реакции является Ni или Ru. Один из этих металлов наносят тонким слоем на подложку, обычно графитовую и полученный «сендвич» размещают в реакционной камере. Для увеличения удельной площади катализатора, металл осаждается не на графит, а на промежуточный слой, состоящий из нанотрубок. Каталитическая эффективность нанотрубок заметно превышает соответствующий параметр для графита. Причем максимальная эффективность достигается с использованием открытых нанотрубок.

Повышенная каталитическая активность открытых нанотрубок обусловлена наличием металлических частиц внутри нанотрубок, что благоприятствует протеканию каталитической реакции.

Среди известных материалов нанотрубки отличаются максимальным соотношением прочность/размер. Модуль Юнга нанотрубки в аксиальном направлении составляет порядка 7000 ГПа, в то время как для стали и иридия, традиционно используемых для изготовления таких зондов, значение этого параметра составляет 200 и 520 ГПа соответственно. Благодаря этому обстоятельству нанотрубки могут служить основой тончайших измерительных инструментов, используемых для контроля неоднородностей поверхности. Например, наноиндентор на базе нанотрубок позволит измерять твердость поверхности в нанометровом масштабе [4].

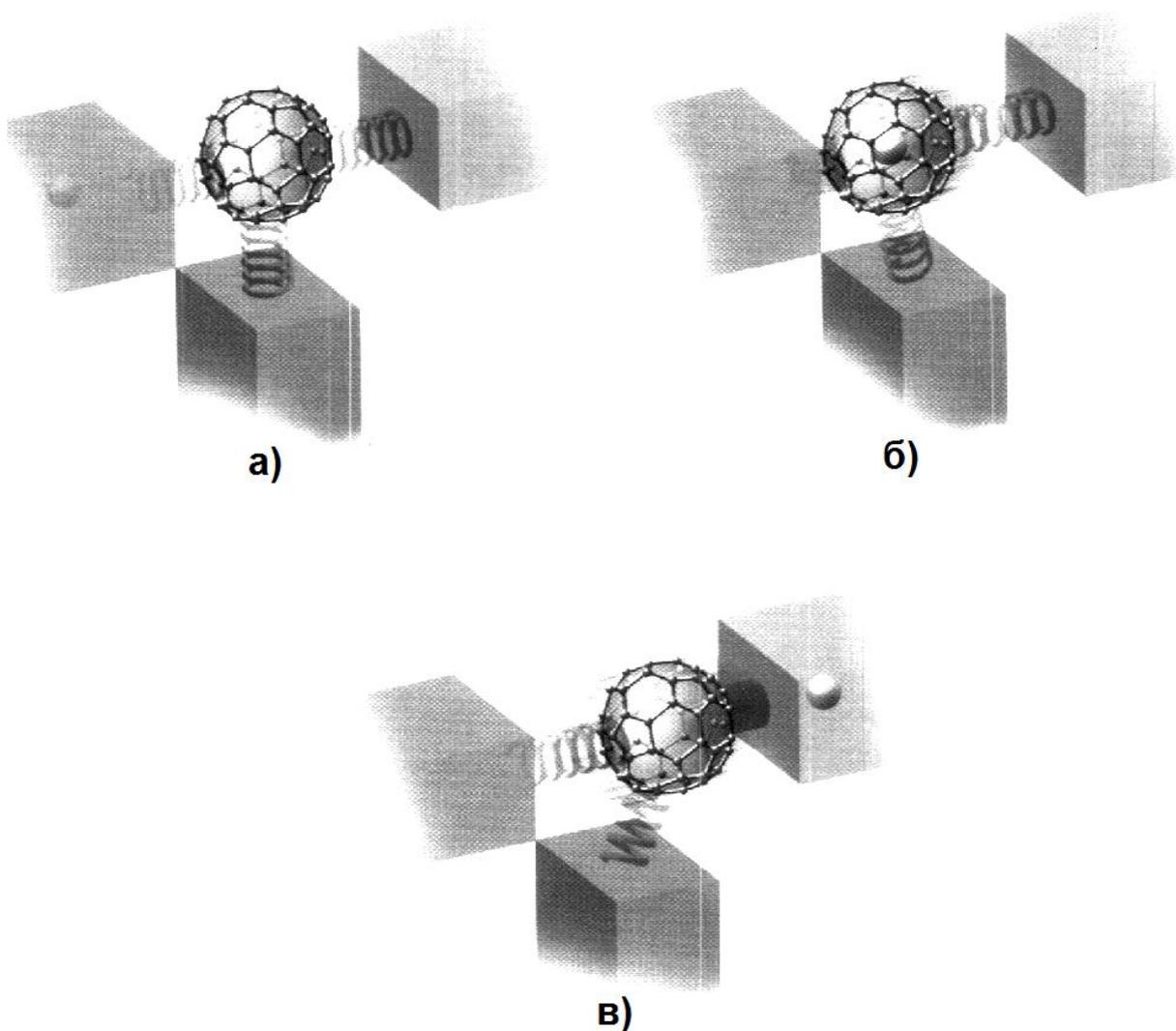
Фуллерены на порядки повышают адгезионную способность и устойчивость к внешним воздействиям полимерных защитных материалов. Эта способность фуллереновых присадок может принести эффект в ряде направлений:

- внедрение новых типов электроизоляционных покрытий с высокими эксплуатационными характеристиками (повышенный ресурс, увеличенные межремонтные сроки, сниженная стоимость обслуживания);
- внедрение новых типов мастик и герметиков (высокая адгезия, прочность, теплопроводность);
- разработка и внедрение лакокрасочных покрытий и материалов с улучшенными эксплуатационными характеристиками, низкой токсичностью, повышенной климатической стойкостью;
- применение новых композиционных материалов на основе стекла и керамики с усиленными эксплуатационными характеристиками;
- создание и внедрение лакирующих материалов для трубопроводов.

Известен транзистор на основе молекулы углерода  $C_{60}$  (рис.2.2.5): фуллерен  $C_{60}$  размещен в центральной части транзистора, или, ключа, который управляет током в цепи. На пластину из двуокиси кремния нанесены тончайшие золотые нити. Пластина покрыта раствором толуола, содержащего фуллерены [17].

По каждой из нитей пропускается электрический ток, который прожигает зазор шириной 1 нанометр в том месте, где проводник был особенно тонким.

Далее подается напряжение на кремниевую прослойку внутри пластины из двуокиси кремния. Она выполняет функции «затвора» – ключевого электрода транзистора. Если напряжение затвора будет равно нулю, электроны проникнут в фуллерен из золотого электрода, или «истока», а выходя из него, попадают в другой электрод – «сток». Однако напряжение



**Рис. 2.2.5.** Транзисор на основе фуллерена:

- а) фуллерен зависает между электродами, удерживаемый электрическими силами;
- б) как только электрон попадает внутрь фуллерена, происходит смещение электрических сил, и молекула сдвигается к одному из электродов;
- в) электрон вылетает в соседний электрод, или "сток", в результате происходит новое смещение сил, и фуллерен "отскакивает" назад - к "истоку", минуя свое первоначальное положение. Затем снова вперед, и так далее.

затвора препятствует переносу зарядов, изменяя направление электрического поля. Чтобы электроны могли преодолеть «затвор», напряжение между «истокком» и «стокком» или напряжение смещения должно быть выше порогового уровня [17].

Увеличение напряжения смещения выше порогового уровня приводит к тому, что ток в цепи начинает быстро расти. Этот рост скачкообразный. Через каждые 5 милливольт проходящий через транзистор ток перескакивает на более высокий уровень.

Энергетические уровни электронов в фуллеренах распределены неравномерно, поэтому электроны на этих уровнях не могут вызвать равномерное скачкообразное увеличение силы тока.

Отскоки могут происходить только в том случае, если электроны обладают необходимой энергией. Соответственно увеличение тока будет происходить скачкообразно по мере того, как электроны получают новый способ попасть в молекулу. Величина энергетических уровней скачущей молекулы соответствует увеличению напряжения смещения на каждые 5 милливольт.

Фуллерен отскакивает недалеко – всего на несколько тысячных нанометра. Но за одну секунду он успевает совершить более триллиона скачков. Такой эффект молекула демонстрирует благодаря одному удару электрона.

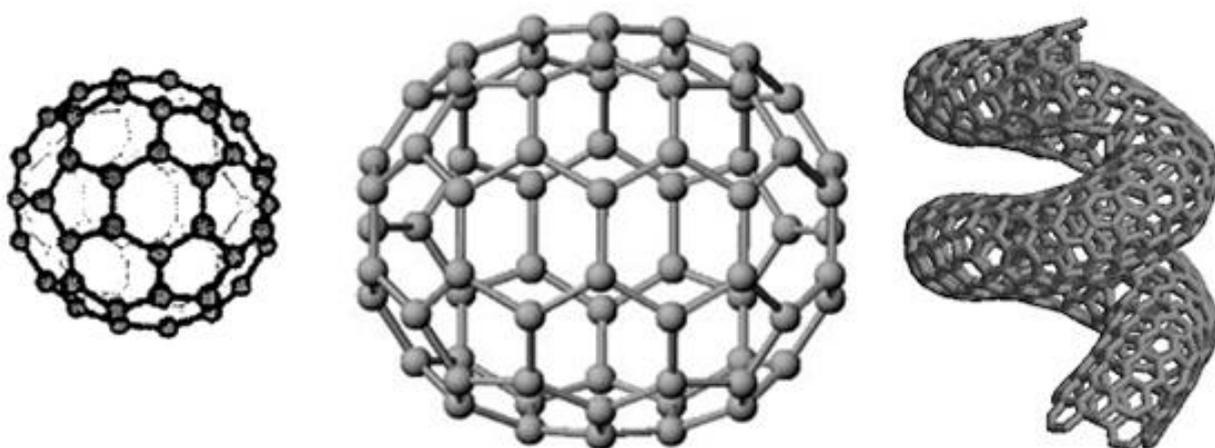
Привести фуллерен в движение может и фотон. Такой транзистор можно использовать в качестве чувствительного датчика излучения. Если закрепить его на каком-либо предмете, то получится сверхчувствительный датчик силы. Задрожавшая даже при ничтожном воздействии молекула вызовет изменение тока, которое можно измерить. Устройство способно также регистрировать и слабейший электрический заряд, находящийся поблизости, так как электрическое поле будет воздействовать на молекулу, изменяя частоту ее колебаний.

## 2.3. Определение основных признаков и свойств углеродных нанотрубок и фуллеренов

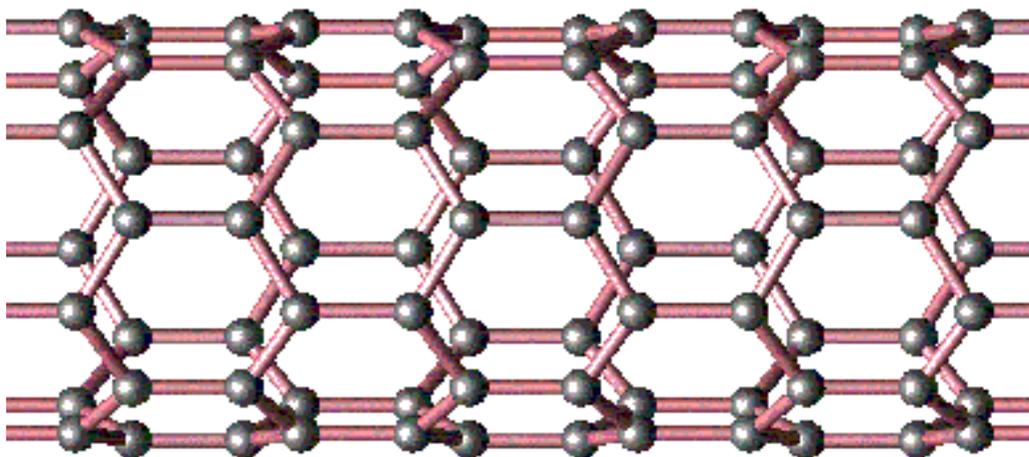
Углерод может существовать в виде объемных замкнутых структур, содержащих  $n$  атомов углерода. Такие структуры обозначают, как  $C_n$ . Замкнутые соединения атомов углерода, имеющие близкую к сферической или сфероидальной форму, получили название фуллерены (рис. 2.3.1). Поверхность фуллеренов образована пяти- и шестиугольниками, в углах которых находится по одному атому [4].

Взаимодействие между атомами разных фуллеренов слабее, чем взаимодействие между ближайшими атомами, принадлежащими одному фуллерену. Взаимодействие двух молекул фуллерена не приведет к их объединению с потерей индивидуальных свойств, образуется система двух слабосвязанных молекул фуллерена, где каждая молекула сохраняет свою индивидуальность.

В процессе образования фуллеренов из графита формируются также и другие структуры, составленные, как и фуллерены, из шести- и пятиугольных колец углерода. Среди них выделяются нанотрубки (рис. 2.3.2).



**Рис. 2.3.1.** Фуллерены  $C_{60}$  [бакминстерфуллерен],  $C_{70}$  и наноспираль [18]



**Рис. 2.3.2.** Однослойная нанотрубка

Нанотрубки представляют собой свернутые в однослойную или многослойную трубку графитовые слои. Длина таких образований может достигать десятков микрометров и на несколько порядков превышает их диаметр, составляющий обычно от одного до несколько нанометров.

Однослойные нанотрубки обычно синтезируют разрядно-дуговым и лазерным методами, получают их всегда считанные миллиграммы, нанотрубки, получаемые этими методами, как правило, искривленные и обычно перепутаны друг с другом (рис. 2.3.3.).

Существует две разновидности многослойных нанотрубок – трубки-матрешки и рулонные (рис. 2.3.4) [11]. Матрешки составлены из нескольких вложенных друг в друга однослойных трубок, примерно до двадцати слоев. Рулонные нанотрубки – это рулоны, свернутые из одного графитового листа.

Многослойные нанотрубки гораздо крупнее однослойных, их даже можно увидеть в обычный растровый электронный микроскоп (индивидуальные однослойные нанотрубки наблюдаются только в просвечивающий электронный микроскоп высокого разрешения). Различить матрешку и рулон крайне сложно, для этого надо получить четкое изображение попе-

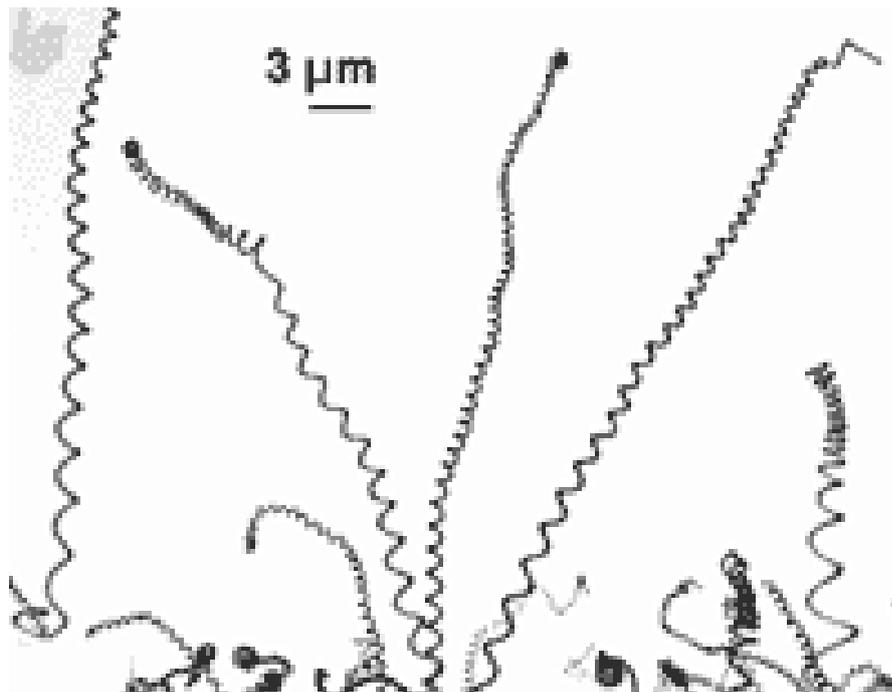


Рис. 2.3.3. Нанотрубки [18]

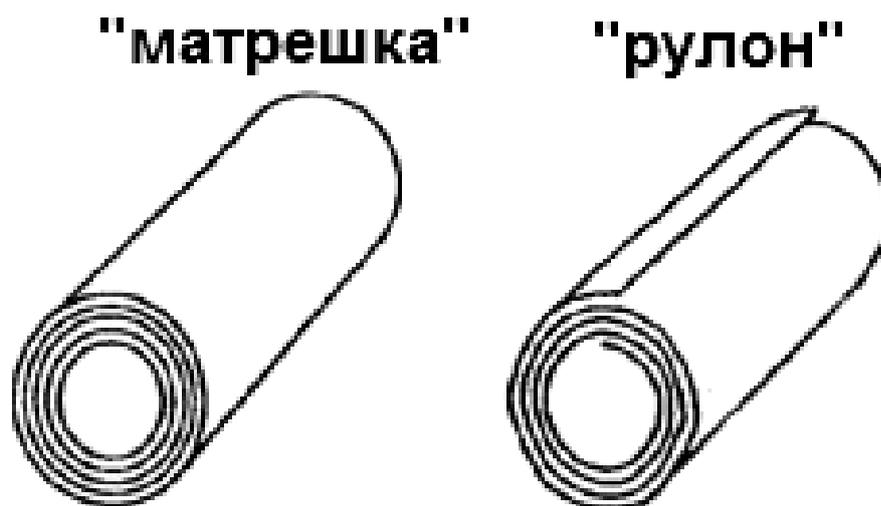


Рис. 2.3.4. Схематическое изображение нанотрубок типов «рулон» и «матрешка»

речного сечения трубки. Химические свойства рулона и матрешки очень разные. При заполнении рулонной нанотрубки моноатомными слоями щелочных металлов и мономолекулярными слоями различных летучих хлоридов, рулон при этом слегка разматывается и вздувается. Трубки «матрешки» не могут вступать в эту реакцию [11].

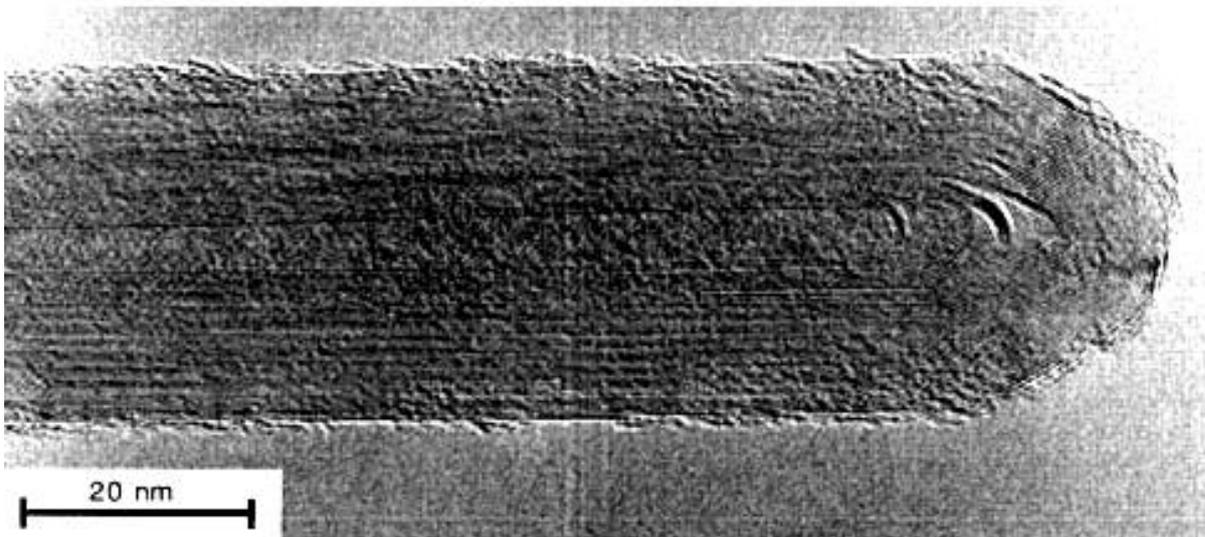
Многослойные нанотрубки могут быть прямыми (рис. 2.3.5). Прямые нанотрубки получают только разрядно-дуговым методом, а это очень важно для будущих применений. Высокопроизводительный ХОП-метод дает всегда только искривленные нанотрубки, спутанные в клубок (рис. 2.3.6).

### **Электрические свойства углеродных нанотрубок**

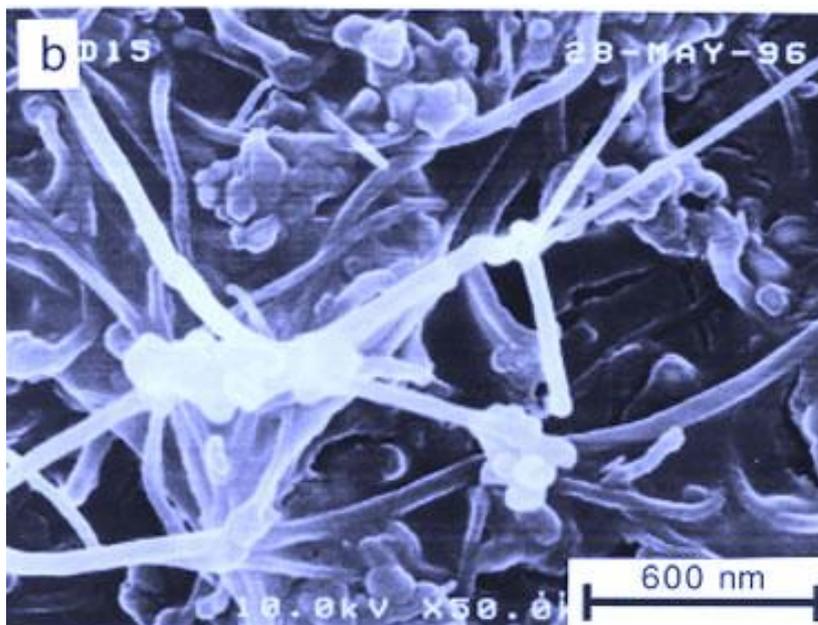
Электрические свойства углеродных нанотрубок (полупроводник, полуметалл, металл) зависят от их угла хиральности (рис. 2.3.7, 2.3.8), т.е. ориентации сторон правильного шестиугольника относительно оси трубки [4].

Все атомы углерода в нанотрубках трехкратно координированы, а это значит, что нанотрубки являются сопряженными ароматическими системами, в которых три из четырех валентных электронов каждого атома углерода образуют локализованные  $\sigma$ -связи, а четвертый участвует в образовании делокализованной  $\pi$ -системы.  $\pi$ -электроны слабо связаны со своими атомами, поэтому с участием именно этих электронов может осуществляться перенос заряда в системе. Высокой (металлической) проводимостью должны обладать те нанотрубки, в которых занятые  $\pi$ -состояния не отделены от вакантных  $\pi^*$ -состояний. В противном случае нанотрубка – полупроводник.

Цилиндрические неизогнутые нанотрубки образуются из повторяющихся углеродных шестиугольников. Если углеродный шестиугольник заменить, например, на пятиугольник, семиугольник или на два таких де-



**Рис. 2.3.5.** Прямая нанотрубка – микрофотография сделана просвечивающим электронным микроскопом



**Рис. 2.3.6.** Искривленные нанотрубки – микрофотография сделана растровым электронным микроскопом

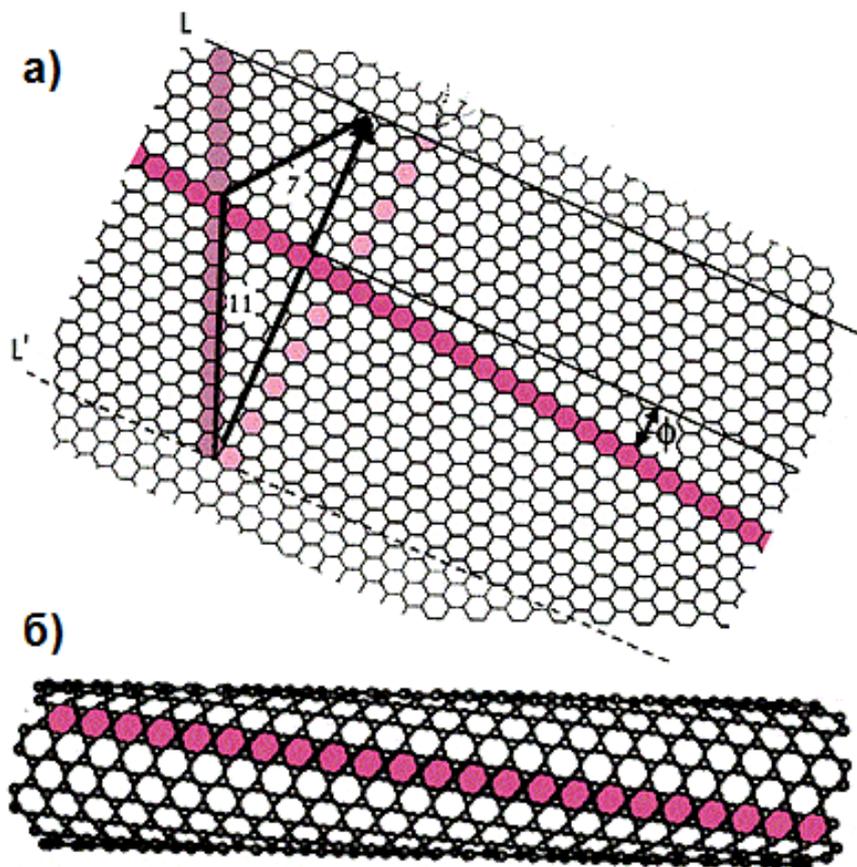
фекта, как показано на рис.2.3.9, нанотрубка изогнется. С разных сторон относительно изгиба ориентация углеродных шестиугольников оказывается различной. Но с изменением ориентации шестиугольников по отношению к оси нанотрубки меняется ее электронный спектр, положение уровня Ферми, ширина оптической щели и т.п. В частности, для приведенного нарис.2.3.9 случая, слева относительно изгиба нанотрубка должна быть металлической, а справа - полупроводниковой. Таким образом, эта изогнутая нанотрубка должна представлять собой молекулярный переход металл-полупроводник.

Если рассматривать данные куски нанотрубки изолированно, с разных сторон относительно изгиба электроны на уровне Ферми обладают разной энергией. В единой системе выигрыш в энергии приводит к перетеканию заряда и образованию потенциального барьера. Электрический ток в таком переходе течет только в том случае, если электроны перемещаются из области нанотрубки с большей энергией Ферми в область с меньшей. Иначе говоря, ток может течь только в одном направлении. «Одностороннее» прохождение тока через нанотрубку с изгибом используется для создания выпрямляющего диода – одного из основных элементов электронных схем.

### ***Электромеханические свойства углеродных нанотрубок***

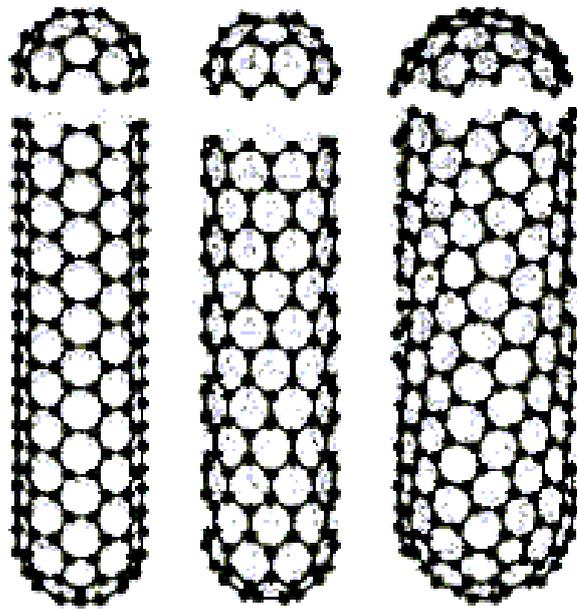
Несмотря на уже ставшее привычным открытие новых необычных свойств нанотрубок, недавно обнаруженный электромеханический эффект в индивидуальных нанотрубках следует отнести к сенсации. Сущность его состоит в значительном изменении электрической проводимости нанотрубки даже при небольшом изгибе. Детальные исследования электромеханических характеристик нанотрубок выполнены группой исследователей из ряда лабораторий Университета Атланты (Джорджия, США) [14]. В эксперименте использовали индивидуальные многослойные углеродные нанотрубки, полученные стандартным электроразрядным методом. К на-

нотрубке прикрепляли тонкую золотую проволоку, на расстоянии от 5 до 20 мкм от конца нанотрубки размещали контрэлектрод, ориентированный под некоторым углом к ней. При подаче на нанотрубку внешнего потенциала она испытывала притяжение к контрэлектроду и изгибалась на некоторый угол, т. е. нанотрубка оказалась способна преобразовывать механическую энергию в электрическую и обратно.

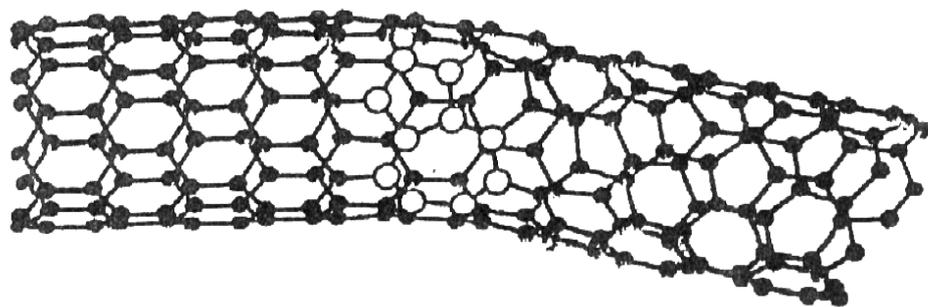


**Рис. 2.3.7.** Построение модели нанотрубки:

а) графитовый слой и лента (11, 7); б) нанотрубка (11, 7). Кроме индексов  $(n, m)$ , геометрию нанотрубки можно охарактеризовать длиной окружности цилиндра  $C$  и углом хиральности  $\phi$ . Если вектор  $C$  совпадает с вертикальной или наклонной “разрезанной” линией шестиугольников, получаются нехиральные трубки  $(n, 0)$  и  $(n, n)$  [14]



**Рис. 2.3.8.** Примеры однослойных нанотрубок с различной хиральностью



**Рис. 2.3.9.** Диод

## ***Теплопроводность нанотрубок***

Согласно проведенным в Пенсильванском университете исследованиям, углеродные нанотрубки, возможно, являются лучшими среди известных проводников тепла. Как известно, тепло переносится звуковыми волнами (фононами). Ранее считалось, что звуковые волны в нанотрубках рассеиваются во всех направлениях и поэтому теплопроводность нанотрубок невелика. Однако, Дж. Е. Фишер и А.Т. Джонсон обнаружили, что звук распространяется строго вдоль нанотрубок со скоростью 10 км/с и очень эффективно переносит тепло. Особенно замечательно то, что тепло хорошо передается даже вдоль многих соединенных между собой нанотрубок. Данное свойство можно использовать для эффективного отвода тепла от элементов микросхем. П. Шенг и его коллеги из Гонконгского Университета наук и технологий впервые зарегистрировали собственную сверхпроводимость углеродных нанотрубок при температуре ниже 20 К.

## ***Механические свойства***

Уникальные механические свойства углеродных нанотрубок (наряду с их электрическими свойствами) привлекают к себе внимание исследователей с момента открытия нанотрубок в 1991 году. Нанотрубки демонстрируют удивительное сочетание прочности и эластичности, обладая способностью обратимым образом реагировать даже на очень сильные внешние воздействия. По этой причине нанотрубки рассматриваются как наиболее перспективный кандидат для нового поколения чрезвычайно лёгких и суперпрочных волокон.

Эксперименты по изучению жёсткости нанотрубок очень сложны. Согласно теоретическим расчетам, критический уровень деформации однослойных нанотрубок составляет 5÷6 %. В принципе, это согласуется с немногочисленными имеющимися на сегодня экспериментальными дан-

ными. Однако есть все основания ожидать, что истинный предел прочности нанотрубок на самом деле гораздо выше. Дело в том, что, с одной стороны, в экспериментах качество образцов не контролировалось должным образом. С другой стороны, в теоретических расчётах основное внимание уделялось равновесным (в термодинамическом смысле) характеристикам деформированных нанотрубок, то есть выяснялось, при каком уровне деформации становится термодинамически выгодным образование топологических дефектов (энергия недеформированной нанотрубки минимальна в отсутствие дефектов, тогда как при деформации свыше критического уровня бездефектное состояние становится метастабильным). А при появлении дефектов нанотрубка под внешним воздействием разрушается быстрее.

Но наряду с равновесными характеристиками дефектных и бездефектных нанотрубок очень важна и динамика образования дефектов. Для возникновения дефекта не достаточно, чтобы энергия содержащей дефект нанотрубки была ниже, чем у бездефектной. Необходимо ещё и преодолеть активационный барьер, отделяющий состояние без дефекта от состояния с дефектом. Если величина активационного барьера достаточно велика, то даже при превышении критического уровня деформации нанотрубка будет сохранять свою структуру, хотя и окажется при этом в термодинамически метастабильном состоянии.

Основные характеристики нанотрубок приведены в таблице 2.3.1.

Интенсивные работы по созданию и совершенствованию процессов производства наноматериалов позволили на сегодняшний день отработать несколько технологий их получения. Установить оптимальные режимы получения тех или иных наноразмерных материалов. Рассмотрим схемы реализации наиболее часто используемых процессов.

**Таблица 2.3.1.**

**Основные характеристики нанотрубок**

Свойства	Однослойные нанотрубки	Сравнение с известными данными
Характерный размер	Диаметр от 0,6 до 1,8 нм	Предел электронной литографии 7 нм
Плотность	1,33 - 1,4 г/см <sup>3</sup>	Плотность алюминия 2,7 г/см <sup>3</sup>
Прочность	45 ГПа	Самый прочный сплав стали Разламывается при 2 ГПа
Упругость	Упруго изгибается под любым углом	Металлы и волокна из углерода ломаются по границам зерен
Плотность тока	Оценки дают до 1 ГА/см <sup>2</sup>	Медные провода выгорают при 1 МА/см <sup>2</sup>
Автоэмиссия	Активируются при 1-3 В при расстоянии 1 мкм	Молибденовые иглы требуют 50-100 В и недолговечны
Теплопроводность	До 6000 Вт/м*К	Чистый алмаз имеет 3320 Вт/м*К
Стабильность по температуре	До 2800°С в вакууме и 750°С на воздухе	Металлизация в схемах плавится при 600-1000°С
Цена	500 \$/г	Золото 10 \$/г

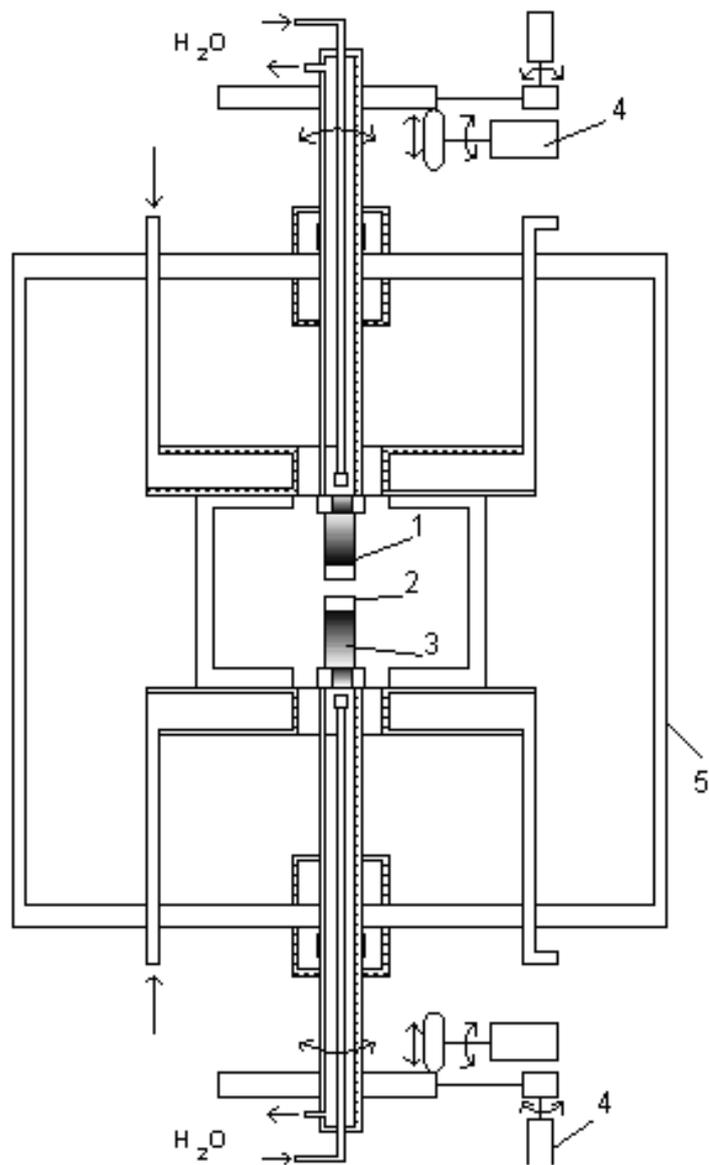
## 2.4. Методы получения углеродных нанотрубок и фуллеренов

### *Метод, основанный на термическом распылении графитового электрода в гелии*

Широко распространен метод получения углеродных нанотрубок, основанный на термическом распылении графитового электрода в плазме дугового разряда, горящей в атмосфере He. Этот метод обеспечивает самую высокую производительность и является основой промышленного производства фуллеренов.

На рис.2.4.1 приведена схема установки получения фуллеренсодержащих материалов. В дуговом разряде постоянного тока с графитовыми электродами при напряжении 15÷25 В, токе в несколько десятков ампер, межэлектродном расстоянии в несколько миллиметров и давлении He в несколько сот мм рт. ст. происходит интенсивное термическое распыление материала анода.

Катод представляет собой графитовый стержень длиной 40 мм и диаметром 16 мм, а анод – графитовый стержень длиной 100 мм и диаметром 6 мм с высверленным отверстием глубиной 40 мм и диаметром 3,2 мм, заполненным смесью металлического катализатора и графитового порошка. Падение напряжения между электродами составляет 30 В при токе дуги 100 А и зазором между электродами 3 мм. В качестве катализатора используются смеси: Ni-Co, Co-Y или Ni-Y с различным процентным содержанием атомов металла по отношению к углероду. Найдено оптимальное соотношение концентрации металлов (1 ат.%Y и 4ат.%Ni) при котором осаждение продуктов конденсации наблюдается в виде: сажи на стенках реакционной камеры; паутины между катодом и стенками реакционной камеры; депозита на катоде; пористого ободка вокруг катода [4].



**Рис. 2.4.1** Схема электродуговой установки для производства наноразмерных углеродных материалов:  
 1 - графитовый анод; 2 - дуга; 3 - графитовый катод;  
 4 - устройство регулировки межэлектродного зазора;  
 5 - охлаждаемая стенка рабочей камеры

Продукты распыления (сажа), содержат до 12÷14% фуллеренов и осаждаются на охлаждаемых стенках разрядной камеры, а также на поверхности катода, более холодного по сравнению с анодом. Катодный осадок (депозит) содержит преимущественно нанотрубки [4].

Во всех продуктах конденсации были обнаружены протяженные структуры, концентрация которых зависела от места осаждения.

Сажа со стенок реакционной камеры, также как и паутина между катодом и стенками камеры, содержала одностенные нанотрубки, частицы катализатора окруженные аморфным углеродом или графеновыми слоями, сажевые частицы различной степени структуризации. Одностенные нанотрубки организованы в пучки с диаметром от 5 до 20 нм. Межосевое расстояние в пучке для одностенных нанотрубок диаметром 1,4 нм составило 1,7 нм. Распределение диаметров нанотрубок, относительная концентрация и структура нанотрубного материала определяются типом используемого катализатора, а также параметрами дугового разряда. Так, например, использование металлов группы железа (Co, Ni, Fe) способствует получению одностенных нанотрубок с разбросом диаметров 0,7÷2,0 нм, с максимумами 0,84 нм; 1,05 нм и 1,4 нм соответственно [4].

Добавление к металлическому катализатору серы или использование сульфидов металлов подгруппы железа приводит к уширению распределения диаметров одностенных нанотрубок до 1÷6 нм. Позднее было показано, что висмут и свинец обладают схожим с серой действием, т.е. их добавление к катализатору приводит к формированию одностенных нанотрубок большего диаметра [4].

Модифицирование метода, заключавшееся в нахождении оптимальных параметров давления He, тока дуги, напряжения и зазора между электродами, позволило увеличить выход нанотрубок. Установка подключается к вакуумной линии с диффузионным насосом и к источнику газообраз-

ного He. Непрерывный поток He при постоянном давлении является более предпочтительным, чем статическая газовая атмосфера.

В дуговом разряде между анодом катодом при напряжении 20÷25 В, стабилизированном постоянном токе дуги 50÷100 А, межэлектродном расстоянии 0,5÷2 мм и давлении He 10÷60 кПа происходит интенсивное распыление материала анода. Часть продуктов распыления, содержащая графит, сажу, и фуллерены осаждается на охлаждаемых стенках камеры, часть, содержащая графит и многослойные углеродные нанотрубки, осаждается на поверхности катода.

На выход нанотрубок влияет множество факторов. Наиболее важным является давление He в реакционной камере, которое в оптимальных с точки зрения производства нанотрубок условиях составляет 60 кПа, а не 10÷60 кПа, как в случае фуллеренов. Другим не менее важным фактором является ток дуги: максимальный выход нанотрубок наблюдается при минимально возможном токе дуги, необходимым для ее стабильного горения. Эффективное охлаждение стенок камеры и электродов также важно для избежания растрескивания анода и его равномерного испарения, что влияет на содержание нанотрубок в катодном депозите [4].

Наиболее высокая эффективность получения одностенных нанотрубок достигается при использовании смешанных катализаторов, в состав которых входят два или три металла группы железа. Кроме металлов группы железа, эффективной для получения одностенных нанотрубок оказалась и группа платины. В установке, при разряде постоянного тока 70 А в гелиевой атмосфере, давлении 10÷80 кПа, с композитным анодом, заполненным смесью металлического (Ru, Rh, Pd, Os, Ir, Pt) и графитового порошков в массовом соотношении 5:1, продукты термического распыления осаждались на стенках камеры, боковой и торцевой поверхностях катода. Межэлектродное расстояние поддерживалось на уровне 3÷2 мм, а напряжение – 25 В [4].

Материал катодного осадка, по данным ПЭМ, содержал заполненные металлом многослойные полиэдрические наночастицы от 20 до 200 нм. Сажка, собранная со стенок разрядной камеры и боковой поверхности катода, содержала сферические частицы металлического катализатора, размером  $5\div 30$  нм, окруженные аморфным углеродом. Образование одностенных нанотрубок, диаметром  $1,3\div 1,7$  нм, длиной  $10\div 200$  нм наблюдалось при использовании в качестве катализатора Rh, Pd и Pt. В случае Rh наблюдались каталитические частицы с растущими на них одностенными нанотрубками образующими структуру похожую на морского ежа. Материал, собранный с боковой поверхности катода оказался более богатым одностенными нанотрубками, чем материал со стенок разрядной камеры. Вдобавок к перечисленным выше рост одностенных нанотрубок наблюдался на катализаторах Ce, Gd, La, Mn, Sc, V, Zr и не наблюдался на Co/Ru, Ni/B, Cu, Ti [4].

Использование автоматического устройства поддержания межэлектродного расстояния на фиксированном уровне способствует увеличению стабильности параметров дугового разряда и обогащению нанотрубками материала катодного депозита. Добавление в систему атомов металла вещественным образом меняет морфологию продуктов конденсации.

Изменение параметров процесса и конструкции установки ведет к изменению эффективности процесса и состава продукта. Качество продукта подтверждается как масс-спектрометрическими измерениями, так и другими методами (ядерный магнитный резонанс, электронный парамагнитный резонанс, ИК-спектроскопия и др.) [20].

### ***Лазерное распыление***

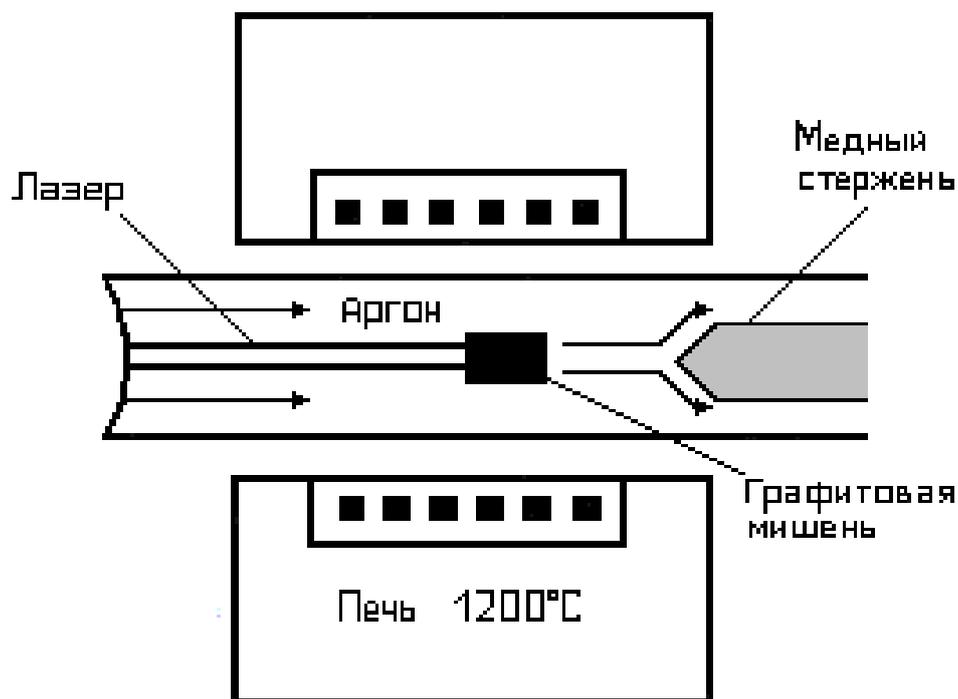
В 1995 году появилось сообщение о синтезировании углеродных НТ методом распыления графитовой мишени под воздействием импульсного лазерного излучения в атмосфере инертного (He или Ag) газа. В данном случае источником высококонцентрированной энергии служит импульсное

излучение лазера. Графитовый стержень (мишень) диаметром 10÷20 мм находится внутри кварцевой трубки диаметром 25÷40 мм. Сама трубка помещается в ИК-печи. Мишень находится в атмосфере аргона при давлении 70 кПа и облучается импульсами длительностью в десятки или сотни нс и энергией в доли Дж. При этом частота следования импульсов не превышает нескольких десятков Гц. Продукт лазерного распыления, содержащий наноматериалы, частично оседает на стенках кварцевой трубки, а частично уносится потоком аргона на водоохлаждаемый сборник [4].

Фокусирующийся системой линз лазерный пучок (рис.2.4.2) сканирует поверхность графитовой мишени для обеспечения равномерного испарения материала мишени. Получающийся в результате лазерного испарения пар попадает в поток инертного газа и выносится из высокотемпературной области в низкотемпературную, где осаждается на охлаждаемой водой медной подложке. Сажа, содержащая нанотрубки, собирается с медной подложки, стенок кварцевой трубки и обратной стороны мишени.

Также как и в дуговом методе получается несколько видов конечного материала:

1. В экспериментах, где в качестве мишени использовался чистый графит, получались многостенные нанотрубки, которые имели длину до 300 нм и состояли из 4÷24 графеновых цилиндров. Структура и концентрация таких нанотрубок в исходном материале главным образом определялись температурой. При 1200°C все наблюдаемые нанотрубки не содержали дефектов и имели шапочки на окончаниях.
2. При понижении температуры синтеза до 900°C в нанотрубках появлялись дефекты, число которых увеличивалось с дальнейшим понижением температуры, и при 200°C образование нанотрубок не наблюдалось.



**Рис. 2.4.2.** Схема установки лазерного испарения графита

3. При добавлении в мишень небольшого количества переходных металлов, в продуктах конденсации наблюдались одностенные нанотрубки. Однако в процессе испарения мишень обогащалась металлом, и выход одностенных нанотрубок снижался [4]. Для решения этой проблемы стали использовать две облучаемые одновременно мишени, одна из которых является чистым графитом, а другая состоит из сплавов металлов.

Процентный выход нанотрубок резко меняется в зависимости от катализатора. Так, например, высокий выход нанотрубок получается на катализаторах Ni, Co смеси Ni и Co с другими элементами. Получаемые одностенные нанотрубки имели одинаковый диаметр и были объединены в пучки диаметром 5÷20 нм. Смеси Ni/Pt и Co/Pt дают высокий выход нанотрубок, тогда как использование чистой платины приводит к низкому выходу одно-

стенных нанотрубок. Смесь Со/Си дает низкий выход одностенных нанотрубок, а использование чистой меди вообще не приводит к формированию одностенных нанотрубок. На окончаниях одностенных нанотрубок, свободных от частиц катализатора, наблюдались сферические шапочки [4].

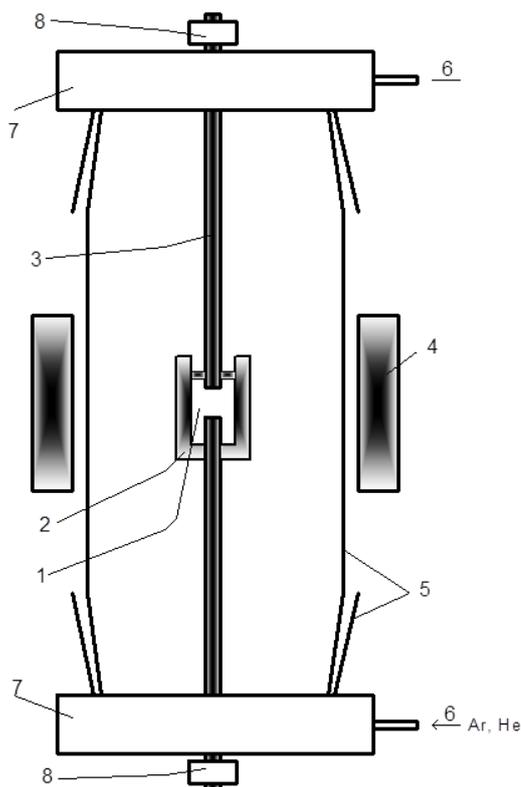
Метод лазерного испарения графита имеет большие перспективы в промышленности, как основа технологии при селективном получении нанотрубок с заданными характеристиками [3].

### **Электролитический синтез**

Данный процесс не получил распространения для промышленного производства, однако с успехом может применяться в лабораторной практике.

Нанотрубки эффективно образуются не только в условиях приповерхностной низкотемпературной плазмы, но также при электролизе, когда все компоненты процесса находятся в конденсированном состоянии.

Основная идея этого метода состоит в том, чтобы получить углеродные нанотрубки, пропуская электрический ток между графитовыми электродами, находящимися в расплавленной ионной соли. Графитовый катод расходуется в процессе реакции и служит источником атомов углерода. В результате формируется широкий спектр наноматериалов. Анод представляет собой лодочку, сделанную из высоко чистого графита и заполненную хлоридом лития. Лодочка нагревается до температуры плавления хлорида лития ( $604^{\circ}\text{C}$ ) на воздухе или в атмосфере инертного газа (аргона). В расплавленный хлорид лития погружается катод и в течение одной минуты между электродами пропускается ток  $1\div 30$  А. За время пропускания тока погруженная в расплав часть катода эродирует. Далее расплав электролита, содержащий частицы углерода, охлаждается до комнатной температуры [4].



**Рис. 2.4.3.** Схема экспериментальной установки для электролитического синтеза нанотрубок: 1 - жидкий электролит; 2 - графитовый анод с отверстием; 3 - графитовый катод; 4 - печь; 5 - кварцевая трубка; 6 - трубка для прокачивания газа; 7 - медные фланцы; 8 - уплотнительные кольца

Впервые электролитическим путем нанотрубки получены в работе, в которой в качестве анода использовался графитовый куб размером  $50 \times 50 \times 50$  мм с отверстием диаметром 2,5 и глубиной 3 мм, заполненным 1 г соли LiCl. Расплавленная соль, нагретая до температуры  $600^\circ\text{C}$ , служила электролитом. Графитовый катод, погружаемый в расплав, имел диаметр 3 мм. В результате пропускания через электролит тока 30 А в течение 1 мин на поверхности катода происходило образование нанотрубок и других наночастиц углерода, которые наблюдались с помощью электронного микроскопа.

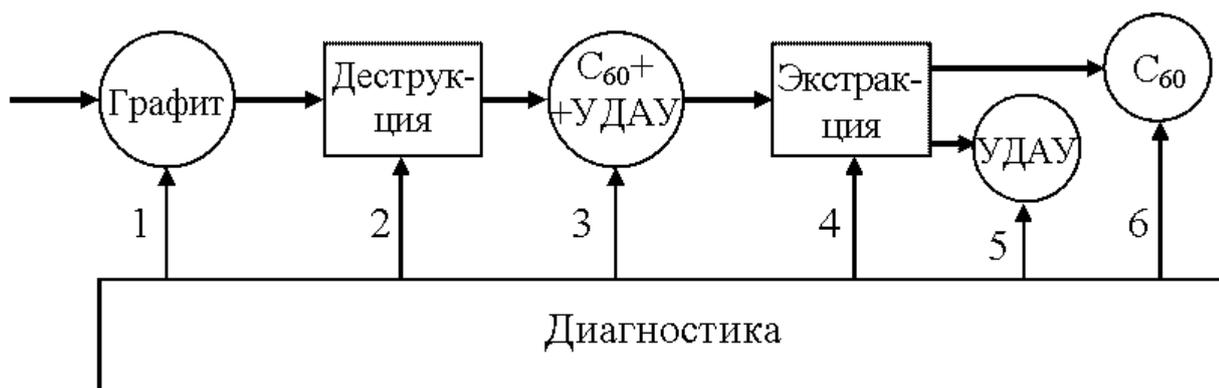
Вариант установки для электролитического синтеза нанотрубок представлен на рис.2.4.3.

Для того чтобы выделить частицы углерода, получившиеся вследствие эрозии катода, соль растворяется в воде. Осадок выделяется, растворяется в толуоле и диспергируется в ультразвуковой ванне. Продукты электролитического синтеза состоят из закапсулированных частиц металла, лукович и углеродных нанотрубок различной морфологии, включая спиральные и сильно изогнутые. В зависимости от условий эксперимента диаметр нанотрубок образованных цилиндрическими графеновыми слоями колеблется от 2 до 20 нм. Длина многослойных нанотрубок достигает 5 мкм. Найдены оптимальные условия по току – 3÷5 А. При высоком значении тока (10÷30 А) образуются только закапсулированные частицы и аморфный углерод. При низких значениях тока (< 1А) образуется только аморфный углерод.

### ***Синтез фуллеренов и УДАУ***

*Ультрадисперсные агрегаты углерода (УДАУ)* — ассоциаты углеродных кластеров. УДАУ выделяются многообразием строения. Среди элементов ассоциатов могут быть самые разные структуры, в том числе цепочечные, однослойные, многослойные ("луковичные").

В Республике Беларусь создан автоматизированный технологический комплекс по производству фуллеренов, позволяющий производить десятки граммов фуллеренов в смену. Создание комплекса потребовало разработки новых технических решений по всей технологической цепочке — от выбора сырья, получения фуллеренсодержащей сажи, разделения фуллеренов и УДАУ до упаковки, хранения и диагностики полученных продуктов (рис.2.4.2). Оптимизация технологии на стадии получения фуллеренсодержащей сажи обеспечивается выбором и автоматической стабилизацией режима дугового разряда, расстояния между электродами, давле-



**Рис. 2.4.2.** Структурная схема технологического комплекса по производству фуллеренов: 1, 3, 5, 6 — операции контроля качества материалов; 2, 4 — операции контроля технологических режимов [21]

ния инертного газа и режима его циркуляции. Выход фуллеренов зависит от температуры синтеза, частоты столкновений малых углеродных кластеров и условий охлаждения синтезированных молекул [21].

### ***Каталитическое разложение углеводородов***

Широко используемый способ получения нанотрубок основан на использовании процесса разложения ацетилена в присутствии катализаторов. В качестве катализаторов использовались частицы металлов Ni, Co, Cu и Fe размером несколько нанометров. В кварцевую трубку длиной 60 см, внутренним диаметром 4 мм, помещается керамическая лодочка с 20÷50 мг катализатора. Смесь ацетилена  $C_2H_2$  (2,5÷10%) и азота прокачивается через трубку в течение нескольких часов при температуре 500÷1100°C. После чего система охлаждается до комнатной температуры. На эксперименте с кобальтовым катализатором наблюдались четыре типа структур: аморфные слои углерода на частицах катализатора; закапсулирован-

ные графеновыми слоями частицы металлического катализатора; нити, образованные аморфным углеродом; многостенные нанотрубки [4].

Наименьшее значение внутреннего диаметра этих многостенных нанотрубок составляло 10 нм. Наружный диаметр свободных от аморфного углерода нанотрубки находился в пределах 25÷30 нм, а для нанотрубок, покрытых аморфным углеродом – до 130 нм. Длина нанотрубок определялась временем протекания реакции и изменялась от 100 нм до 10 мкм.

Выход и структура нанотрубок зависит от типа катализатора – замена Co на Fe дает меньшую концентрацию нанотрубок и количество бездефектных нанотрубок сокращается. При использовании никелевого катализатора большинство нитей имело аморфную структуру, иногда встречались нанотрубки с графитизированной бездефектной структурой. На медном катализаторе формируются нити с нерегулярной формой и аморфной структурой. В образце наблюдаются закапсулированные в графеновые слои частицы металла. Получаемые нанотрубки и нити принимают различные формы - прямые; изогнутые, состоящие из прямых участков; зигзагообразные; спиральные. В некоторых случаях шаг спирали имеет псевдопостоянную величину.

В настоящее время возникла необходимость получить массив ориентированных нанотрубок, что продиктовано использованием таких структур в качестве эмиттеров. Существует два пути получения массивов ориентированных нанотрубок: ориентация уже выросших нанотрубок и рост ориентированных нанотрубок, используя каталитические методы [4].

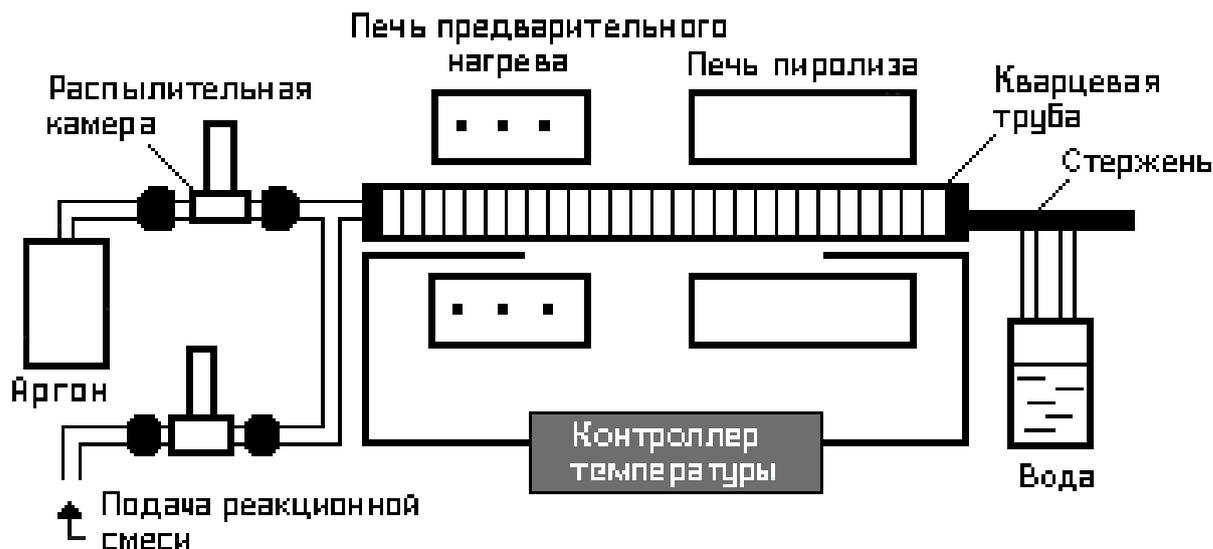
Можно использовать в качестве подложки для роста нанотрубок пористый кремний, поры которого заполнены наночастицами железа. Подложка помещается в среду буферного газа и ацетилена, находящихся при температуре 700°C, где железо катализировало процесс термического распада ацетилена. В результате, на площадях в несколько мм<sup>2</sup>, перпендикулярно подложке, формируются ориентированные многослойные нанотрубки.

Полученные нанотрубки обладают высокой степенью графитизации, внутренним и внешним диаметром  $2.5 \div 4$  нм и  $10 \div 12$  нм соответственно. Исследовано влияние на процесс различных видов катализаторов, подложек, температуры, времени протекания реакции и расхода газов. Установлено, что оптимальная температура реакции сильно зависит от типа катализатора [4].

Вариант установки каталитического пиролиза углеводородов, в которой можно реализовать гибкое и отдельное управление условиями образования нанотрубок изображен на рис. 2.4.5.

Используемый здесь способ получения нанотрубок основан на процессе разложения ацетилена ( $C_2H_2$ ) в присутствии катализаторов. В качестве катализатора обычно используется железо, которое образуется в восстановительной среде из различных соединений железа (хлорид железа (III), салицилат железа (III) или пентакарбонил железа). Смесь солей железа с углеводородом (ацетиленом) распыляется в реакционную камеру либо направленным потоком аргона, либо с использованием ультразвукового распылителя. Полученный аэрозоль с потоком аргона поступает в кварцевый реактор. В зоне печи предварительного нагрева аэрозольный поток прогревается до температуры  $250^\circ C$ , происходит испарение углеводорода и начинается процесс разложения металлсодержащей соли. Далее аэрозоль попадает в зону печи пиролиза, температура в которой составляет  $900^\circ C$ . При этой температуре происходит процесс образования микро- и наноразмерных частиц катализатора, пиролиз углеводорода, образование на частицах металла и стенках реактора различных углеродных структур, в том числе нанотрубок. Затем газовый поток, двигаясь по реакционной трубе, поступает в зону охлаждения. Продукты пиролиза осаждаются в конце зоны пиролиза на охлаждаемом водой медном стержне [21].

На эксперименте с никелевым катализатором наблюдались четыре типа структур:



**Рис. 2.4.5.** Схема установки каталитического пиролиза углеводородов

1. аморфные слои углерода на частицах катализатора;
2. закапсулированные графеновыми слоями частицы металлического катализатора;
3. нити, образованные аморфным углеродом;
4. многослойные нанотрубки.

Наименьшее значение внутреннего диаметра этих многослойных нанотрубок составляло 10 нм. Наружный диаметр свободных от аморфного углерода нанотрубок находился в пределах 25÷30 нм, а для нанотрубок, покрытых аморфным углеродом – до 130 нм. Длина нанотрубок определялась временем протекания реакции и изменялась от 100 нм до 10 мкм.

### ***Метод с использованием анодированного алюминия в качестве подложки***

Аналогичный метод – использование в качестве подложки анодированного алюминия. Поры анодированного алюминия заполняются

кобальтом. Подложка помещается в проточную смесь ацетилена и азота при температуре 800°C. Получаемые ориентированные нанотрубки имеют средний диаметр  $50,1 \pm 0,7$  нм с расстоянием между трубками  $104,2 \pm 2,3$  нм. Средняя плотность определяется на уровне  $1,1 \cdot 10^{10}$  нанотрубок /см<sup>2</sup>. ПЭМ нанотрубок выявила хорошо графитизированную структуру с расстоянием между графеновыми слоями 0,34 нм. Изменяя параметры и время обработки алюминиевой подложки можно менять как диаметр нанотрубок, так и расстояние между ними.

### ***Низкотемпературные методы (~600°C)***

Метод, протекающий при более низких температурах (ниже 650°C) также описан в статьях. Низкие температуры в процессе синтеза позволяют использовать в качестве подложки стекло с нанесенной пленкой никеля. Никелевая пленка служила катализатором для роста нанотрубок методом осаждения из газовой фазы в активированной плазме с горячей нитью. В качестве источника углерода использовался ацетилен. Меняя условия эксперимента можно менять диаметр трубок от 20 до 400 нм и их длину в пределах 0,1÷50 мкм. Получаемые нанотрубки большого диаметра (>100 нм) прямые и их оси направлены строго перпендикулярно подложке. Наблюдаемая плотность нанотрубок по данным растровой электронной микроскопии составляет 10 нанотрубок /мм. Когда диаметр нанотрубок становится меньше 100 нм преимущественная ориентация, перпендикулярная плоскости подложки, исчезает. Ориентированные массивы нанотрубок могут создаваться на площадях в несколько квадратных сантиметров.

### ***Метод квазисвободной конденсации пара***

В методе квазисвободной конденсации пара, углеродной пар образуется в результате резистивного нагрева графитовой ленты и конденсирует-

ся на подложку из высокоупорядоченного пиролиитического графита, охлаждаемую до температуры 30°C в вакууме  $133 \cdot 10^8$  Па. Исследования, полученных пленок толщиной  $2 \div 6$  нм показывают, что они содержат углеродные нанотрубки диаметром  $1 \div 7$  нм, длиной до 200 нм, большинство из которых заканчивается сферическими окончаниями. Содержание нанотрубок в осадке превышает 50%. Для многослойных нанотрубок расстояние между образующими их графеновыми слоями составляет 0,34 нм. Трубки располагаются на подложке практически горизонтально [4].

### ***Метод получения нанотрубок в пламени газов***

Метод получения нанотрубок в пламени ацетилена, бензола или этилена, предварительно смешанных с кислородом и буферным газом. Давление газа варьировалось в диапазоне  $20 \div 100$  мм.рт.ст.; атомное соотношение C/O составляло 1,06 в случае  $C_2H_2$ ,  $0,86 \div 1,0$  – в случае  $C_6H_6$  и 1,07 – в случае  $C_2H_6$  при скорости потока газа в горелке  $25 \div 50$  см/сек, максимальная температура пламени была около 2000 К. Образцы сажи отбирались из пламени с использованием водоохлаждаемого всасывающего зонда, а также со стенок камеры сгорания, после чего диспергировались в толуоле с помощью ультразвука. В образцах сажи находились многостенные нанотрубки диаметром  $2 \div 15$  нм, с расстоянием между графеновыми слоями 0,34 нм. Многостенные нанотрубки имели сферические окончания [4].

### ***Лазерно-порошковый метод синтеза одностенных углеродных нанотрубок***

Лазерно-порошковый метод непрерывного синтеза одностенных углеродных нанотрубок. Метод базируется на лазерной абляции смеси графитового и каталитических порошков (Ni/Co=1/1) лазером мощностью 2 кВт в потоке аргона или азота. Основной идеей предложенного метода является более эф-

эффективное использование лазерной мощности, поглощаемой на частицах микронного размера, для испарения твердого материала по сравнению с абляцией массивных мишеней в известных лазерных методах за счет меньших (на  $1\div 2$  порядка величины) теплопроводностных потерь в газе. В результате стало возможным испарять до 5 грамм графитового порошка в час [22].

Полученные образцы сажи были исследованы с помощью спектроскопии комбинационного рассеяния и просвечивающей электронной микроскопии высокого разрешения. На основе этих исследований содержание нанотрубок в сажистом осадке оценивается  $20\div 40\%$ . Разработан метод непосредственного обогащения сажи нанотрубками в процессе эксперимента. Предложены пути оптимизации параметров процесса для увеличения процентного содержания нанотрубок в саже [22].

### ***Холодная деструкция графита***

В качестве исходного графитсодержащего сырья используется или природный чешуйчатый графит, или графит в виде порошка.

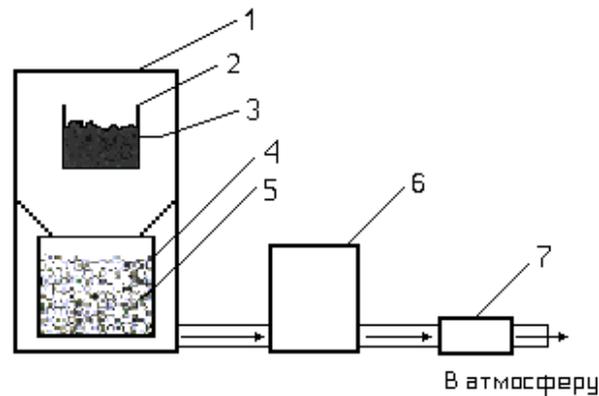
Смесь представляет собой углеродный материал, обладающий огромной реакционной способностью по отношению к любым углеводородным соединениям. Реакционная способность углеродной смеси обусловлена тем, что при её изготовлении из природного чешуйчатого графита, графита в виде порошка или другого графитсодержащего сырья происходит не только расслаивание кристаллитов на отдельные пакеты базисных плоскостей, как при известных способах изготовления расширенного графита, но и разрыв межгексагональных ковалентных связей. Это приводит к образованию энергетически напряженных атомарных соединений углерода [20].

Для обеспечения оптимального режима изготовления смеси весовое соотношение исходного графитсодержащего сырья и галоген-кислородного соединения равно 2:1.

Изготовление углеродной смеси производят путем химической обработки исходного графитсодержащего сырья по крайней мере одним галоген-кислородным соединением общей формулы  $\text{MXO}_n$ , где М – одно из химических веществ ряда: Н,  $\text{NH}_4$ , Na, К; X – одно из химических веществ ряда: Cl, Br, J; а  $n = 1 \div 4$ , с образованием иницирующих комплексов, способных в результате фотохимического, сонохимического или прямого химического воздействия к экзотермическому взрывообразному разложению с последующим иницированием автокаталитического процесса распада соединения. Иницирующие комплексы вводятся в межслоевые пространства графита, инициируется их взрывообразное разложение и происходит разрыв не только Ван-дер-ваальсовых, но и ковалентных связей, что приводит к образованию углеродной смеси, имеющей тенденцию к высокой реакционности. Процесс осуществляется в любой емкости, в том числе и без доступа кислорода.

Процесс преобразования графита (разрыв Ван-дер-ваальсовых связей) осуществляется под воздействием микровзрывов, вводимых в межслойные пространства графита взрывчатых веществ, в данном случае, названных иницирующими комплексами. Взрывчатое вещество находится в межслоевом пространстве на молекулярном уровне и химическим путем инициируется до взрыва. В результате энергий, высвобождаемых микровзрывом, происходят разрывы не только Ван-дер-ваальсовых, но и межатомарных связей с образованием не только свободных радикалов С,  $\text{C}_2$ ,  $\text{C}_3$  и т.д., но и радикалов в виде гексагонов (одного или нескольких) с присоединенными к ним радикалами вида С,  $\text{C}_2$ ,  $\text{C}_3$ ,  $\text{C}_4$  и  $\text{C}_5$ , обеспечивающими в совокупности высокую реакционную способность получаемой углеродной смеси.

После загрузки исходного сырья в установку (рис. 2.4.6), любым из вышеуказанных способов инициируется взрывообразный процесс. В результате преобразования графита и образования углеродной смеси нанобъектов объем исходного вещества увеличивается в несколько сотен раз,



**Рис. 2.4.6.** Схема установки для проведения холодной деструкции графита: 1 – наружный корпус; 2 – загрузочная емкость; 3 – сырье (графит); 4 – сетчатый бункер; 5 – готовый продукт; 6 – поглотитель паров; 7 – компрессор лазерного испарения графита

и готовая смесь, поднимаясь, высыпается через край загрузочной ёмкости (2) и попадает в приёмный сетчатый бункер (4). Образующиеся внутри устройства пары и газы через поглотитель (6) откачиваются компрессором (7). Поглотитель паров и газов представляет собой молекулярное сито для улавливания вредных составляющих газовой смеси, например, паров соляной кислоты и хлора, и в виде уже безвредных паров воды, углекислого газа и др., которые выбрасывается в атмосферу. Для выхода газа из корпуса в его нижней части предусмотрен патрубок, снабженный клапаном. После завершения процесса наружный корпус (1) раскрывается, и готовый продукт (5) выгружается из приёмного сетчатого бункера (4).

## 2.5. Аналитическая процедура Саати

### в автоматизированном проектировании

В начале 1970 года американский математик Томас Саати разработал процедуру поддержки принятия решений, которую назвал "Analytic hierarchy

process" (АНР). Этот метод относится к классу критериальных и занимает особое место, благодаря тому, что он получил исключительно широкое распространение и активно применяется по сей день, особенно в США. На основе этого метода разработаны достаточно серьезные системы поддержки принятия решений, например «Expert choice» [23].

$K_i^R$  - значение критерия  $K^R$  по  $i$ -ому столбцу ( $i$ -ой альтернативе).

$K_j^R$  - значение критерия  $K^R$  по  $j$ -ой строке ( $j$ -ой альтернативе);

$\frac{K_i^R}{K_j^R}$  - парное соотношение критериев  $K^R$ , расположенных в  $i$ -ом столбце и  $j$ -ой строке;

$\sum_{j=1}^G \frac{K_i^R}{K_j^R}$  - суммарное построчное значение парных соотношений альтернатив;

$\sum_{i=1}^J \left( \sum_{j=1}^G \frac{K_i^R}{K_j^R} \right)$  - общая столбцовая сумма значений парных соотношений альтернатив.

$J = G$

$a_i^R = \frac{\sum_{j=1}^G \frac{K_i^R}{K_j^R}}{\sum_{i=1}^J \left( \sum_{j=1}^G \frac{K_i^R}{K_j^R} \right)}$  - нормированное значение альтернатив по критерию  $R$  или оценка

альтернатив. Причем  $\sum_{i=1}^J a_i^R = 1$ , для критерия  $R$ ,  $\sum_{R=1}^N a_i^R, i = 1 \dots J \Rightarrow \max$

Описание метода выполним на примере выбора альтернатив.

**Критерии:**

- A1
- A2
- A3.

## Критерии:

- К1
- К2
- К3
- К4

В основе АНР все та же линейная свертка, но оценки альтернатив и веса критериев получаются особым образом. В модели АНР вместо критериальной таблицы принята иерархия. Представим ее следующим образом:

**Уровень 0:** Цель - выбрать одну из альтернатив.

### Уровень 1:

Критерии -

-К1

-К2

-К3

- К4

Уровней может быть сколько угодно. Например, критерий 1-го уровня "К1" можно раскрыть уровнем 2 и т.д. Ограничимся Уровнем 1 [23].

Теперь нужно получить оценки каждой альтернативы по каждому критерию. Если существуют объективные оценки, то они просто выписываются и нормируются таким образом, чтобы их сумма была равна единице. Например, если бы нас интересовал критерий "процентный выход наноматериалов" и имелись бы соответствующие данные: по каждому методу, то нужно было бы составить следующую таблицу [23].

Альтернативы	Максимальный вы- ход	Нормирование зна- чения
A1	94	0,277
A2	95	0,280
A3	60	0,177
A4	90	0,265
<i>Сумма</i>		<i>1,000</i>

Иногда процедура Саати рекомендует использовать парные сравнения. Для фиксации результата сравнения пары альтернатив может использоваться, например, шкала следующего типа:

- 1 - равноценность
- 3 - умеренное превосходство
- 5 - сильное превосходство
- 7 - очень сильное превосходство
- 9 - высшее (крайнее) превосходство

Лицо, принимающее решение (ЛПР), просят попарно сравнить альтернативы. Результат парных сравнений альтернатив для критерия записывается в виде таблицы:

	A1	A2	A3	A4
A1	1/1	1/2	1/4	1/6
A2	2/1	1/1	1/2	1/5
A3	4/1	2/1	1/1	1/2
A4	6/1	5/1	2/1	1/1

Простые дроби в клетках трактуются следующим образом. Например, на пересечении строки A2 и столбца A1 записана дробь 4/1. Это выражает мнение ЛПР о том, что критерий A2 в 4 раза выше, чем критерий A1. Далее простые дроби переводятся в десятичные. Получается такая таблица:

	A1	A2	A3	A4
A1	1.000	0,5	0,25	0,17
A2	2	1.000	0,5	0,2
A3	4	2	1.000	0,5
A4	6	5	2	1.000

Эта таблица есть не что иное, как таблица результатов парных сравнений. Теперь посчитаем строчные суммы:

	A1	A2	A3	A4	Сумма по строке
A1	1.000	0,5	0,25	0,17	1,92
A2	2	1.000	0,5	0,2	3,7
A3	4	2	1.000	0,5	7,5
A4	6	5	2	1.000	14
				Сумма	27,12

Теперь нормируем суммы таким образом, чтобы их сумма в свою очередь была равна 1. Для этого просто разделим сумму каждой строки на 32,37 (сумма последнего столбца, т.е. сумма самих строчных сумм).

Получим:

	A1	A2	A3	A4	Сумма
A1	1.000	0,5	0,25	0,17	0,07
A2	2	1.000	0,5	0,2	0,14
A3	4	2	1.000	0,5	0,28
A4	6	5	2	1.000	0,52
				Сумма	1,00

В методе Саати полученные таким образом нормированные суммы принимаются в качестве оценок альтернатив по критерию "наиболее эффективный метод, для получения фуллеренов" [23].

Аналогичным образом получаются веса критериев. Предположим, конкретное ЛПР сравнило попарно критерии с точки зрения их сравнительной важности. Запишем результаты сравнений в виде таблицы [23].

	K1	K2	K3
K1	1/1	K12	K13
K2	K21	1/1	K23
K3	K31	K32	1/1

	K1	K2	K3	Сумма
K1	1/1	K12	K13	K11+K12+K13
K2	K21	1/1	K23	K21+K22+K23
K3	K31	K32	1/1	K31+K32+K33
				$\sum_{i,j=1}^n K_{ij}$

	K1	K2	K3	Сумма
K1	1/1	K12	K13	$(K11+K12+K13) / \sum_{i,j=1}^n K_{ij}$
K2	K21	1/1	K23	$(K21+K22+K23) / \sum_{i,j=1}^n K_{ij}$
K3	K31	K32	1/1	$(K31+K32+K33) / \sum_{i,j=1}^n K_{ij}$
				1

Применяя к этой таблице описанную выше процедуру, получим веса критериев:

$$W_i = \frac{\sum_{j=1}^n K_{ij}}{n \sum_{ij=1} K_{ij}}.$$

Таким образом, мы можем получить как веса критериев, так и оценки альтернатив по критериям [23].

Далее, применяя линейную свертку (взвешенную сумму), получим интегральные оценки альтернатив (функция полезности).

## ЛИТЕРАТУРА К ГЛАВЕ 2

1. <http://elmech.mpei.ac.ru/homes/Petrichenko/bac2.html>
2. Борисенко В.Е. Нанoeлектроника – основа информационных систем XXI века. // Соросовский образовательный журнал. №5. 1997.
3. Чумаченко Б., Лавров К. Нанотехнологии – ключевой приоритет обозримого будущего. // Международный журнал «Проблемы теории и практики управления». №5. 2001.
4. Рыбалко В.В. Наноразмерные углеродсодержащие материалы. – М.:МГИЭМ. 2003г. – 50с., ил.
5. Нанотехнологические процессы и установки. Лускинович П.Н. и др. // сб. докладов. Математическое моделирование нанотехнологических процессов и наноструктур. Труды научных семинаров. – М.:МИФИ. 2001. – 116с.
6. Елецкий А. В. Углеродные нанотрубки. // УФН, сентябрь 1997г, т. 167, № 9
7. Раков Э.Г. Удивительные нанотрубки: свойства. // Химия. №9. 2001.
8. Чудеса нанотехники. Часть первая: наноначало. membrana (<http://www.membrana.ru/articles/simply/2002/01/29/163800.html>)
9. Jie Hany, Al Globus, Richard Jaffe and Glenn Deardorff Molecular dynamics simulations of carbon nanotube-based gears // Nanotechnology 8 (1997) 95–102.
10. Нанотехнологии и вычислительная математика. Г.Г. Еленин // сб. докладов. Математическое моделирование нанотехнологических процессов и наноструктур. Труды научных семинаров. – М.:МИФИ. 2001. – 116с.
11. Мордкович В.З. Соломинки для микробов или об углеродных нанотрубках. // Химия и жизнь. №7. 1999.
12. Неволин В.К. Электронные устройства с элементами нанометровых размеров. // Электронная промышленность, 10, 20.

13. Золотухин И.В. Углеродные нанотрубки. // Соросовский образовательный журнал. №3. 1999.
14. Дьячков. Углеродные нанотрубки: строение, свойства, применения. – М.: Бином. Лаборатория знаний, 2006 г.
15. Посецельский А.А., Крашенинников А.П. Скоростные нанотрубки. (<http://nature.web.ru/db/msg.html?mid=1183580&s=120500000>)
16. William F. Brinkman and Mark R. Pinto. The Future of Solid-State Electronics // Bell Labs Technical Journal. Autumn 1997.
17. NewScientist Игра в бакибол. Ломоносов №5 2003г.
18. <http://news.fsu.edu/More-FSU-News/Decades-old-mystery-of-buckyballs-cracked-by-team-led-by-FSU-MagLab-researchers>
19. Елецкий А.В. «Экзотические» объекты атомной физики. // Соросовский образовательный журнал. №4. 1999.
20. Соколов В. И., Станкевич И. В. Фуллерены - новые аллотропные формы углерода: структура, электронное строение и химические свойства//Успехи химии, т.62 (5), с.455, 1993.
21. Shpilevskii M. E., Shpilevskii E. M., Stel'makh V. F. Fullerenes and fullerene-like structures: the basis for promising materials. Фуллерены и фуллереноподобные структуры — основа перспективных материалов // Journal of Engineering Physics and Thermophysics. — 2001. — Т. 74. — № 6. — С. 1499—1508.
22. Новый лазерно-порошковый метод синтеза одностенных углеродных нанотрубок. Углов С.А., Большаков А.П., Савельев А.В., Конов В.И., Горбунов А.А., Помпе В., Граф А. // XIII симпозиум «Современная химическая физика» (25 сентября - 6 октября 2001 года, Туапсе).
23. Саати Т. принятие решений. Метод анализа иерархий. – М.: Радиосвязь, 1993г.

# ГЛАВА 3. АЛГОРИТМЫ ПОИСКА ОПТИМАЛЬНЫХ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В ОБЛАСТИ КАЧЕСТВА ПРОИЗВОДСТВА МАТЕРИАЛОВ НА ОСНОВЕ УГЛЕРОДНЫХ СТРУКТУР

## 3.1. Адаптивное управление качеством производства наноматериалов

Адаптивное управление качеством производства наноматериалов – это оптимальное управление недостаточно хорошо определенными производственными процессами, результаты которых зависят от входных, управляющих и неконтролируемых переменных [3]. Из-за воздействия неконтролируемых переменных технологические операции теряют такое свойство, как установившееся состояние, и процесс может дрейфовать во времени [4].

Предположим, что математическая модель процесса производства электронных нанобъектов имеет вид:

$$Y_{вых} = f(Y_{тр.вых}, X_{упр}, Y_{тек}, X_{тек}, V_{вх}), \quad (3.1.1)$$

где  $Y_{вых}$  – вектор значений выходных переменных;  
 $Y_{тр.вых}$  – вектор требуемых значений выходных переменных;  
 $X_{упр}$  – вектор требуемых значений управляющих переменных;  
 $Y_{тек}$  – вектор текущих значений выходных переменных;  
 $X_{тек}$  – вектор текущих значений управляющих переменных;  
 $V_{вх}$  – вектор входных переменных, учитывающий влияние параметров исходных материалов и полуфабрикатов и/или результатов предшествующих операций.

Для упрощения выкладок исключим вектор  $V_{вх}$  из дальнейшего рассмотрения, включив его в качестве компоненты  $X_{зад}$  в вектор  $X_{упр} = (X_{упр}, X_{зад})$ . Тогда математическая модель операции преобразуется к виду:

$$Y_{вых} = f(Y_{тр.вых}, X_{упр}, Y_{тек}, X_{тек}). \quad (3.1.2)$$

В общем случае математическая модель (3.1.2) представляет собой систему нелинейных уравнений [7]. От степени нелинейности зависит сложность алгоритмов управления. Нелинейные модели требуют итерационных методов вычисления значений выходных переменных на каждом шаге оптимизации, что увеличивает сложность и время расчётов. Поэтому целесообразно использовать математические модели минимальной степени нелинейности, которые обеспечивают компромисс между точностью и затратами времени на вычисления [8]. Учитывая, что на управляющие и выходные переменные накладываются двухсторонние ограничения, область работоспособности математической модели ограничивается диапазоном изменения этих переменных, поэтому для практических целей можно использовать полиномиальные (как правило, линейные или квадратичные) аппроксимации математических моделей:

$$Y_{вых} = a_0 + \sum_{i=1}^n a_i X_i; \quad (3.1.3)$$

$$Y_{вых} = a_0 + \sum_{i=1}^n a_i X_i + \sum_{i=1}^n a_i X_i^2 + \sum_{i=1}^n a_i X_i Y_i. \quad (3.1.4)$$

По существу, оптимальное управление связано с решением следующих задач: выбором целевой функции, проверкой выполнения ограничений (попаданием в область работоспособности) и нахождением значений переменных управления  $X_{упр}$ , обеспечивающих экстремум целевой функции в области работоспособности [9]. Для широкого круга задач управления в

качестве целевой функции  $F = F(Y_{тр.вых}, X_{упр}, Y_{тек}, X_{тек})$  можно использовать функцию качества в виде взвешенной суммы квадратов:

$$F = \psi_1(Y_{тр.вых} - Y_{тек})^2 + \psi_2(X_{упр} - X_{тек})^2, \quad (3.1.5)$$

где  $\psi_1$  и  $\psi_2$  – векторы весовых коэффициентов выходных и управляющих переменных соответственно.

Если на величину управляющих переменных  $X_{упр}$  не накладывается никаких ограничений, то целевая функция преобразуется к виду

$$F = \psi(Y_{тр.вых} - Y_{тек})^2. \quad (3.1.6)$$

Подставим в уравнение (3.1.6) вместо текущего значения вектора выходных переменных  $Y_{тек}$  его значение, вычисляемое по математической модели (3.1.3)

$$Y_{тек} = a_0 + \sum_{i=1}^n a_i X_{тек\ i}. \quad (3.1.7)$$

В результате получим целевую функцию в форме  $G$ , зависящую только от управляющих переменных:

$$G(X_{тек}) = \psi(Y_{тр.вых} - a_0 - \sum_{i=1}^n a_i X_{тек\ i})^2. \quad (3.1.8)$$

Для нахождения оптимальных значений переменных управления  $X_{онт}$  воспользуемся необходимыми условиями существования экстремума. Для этого продифференцируем целевую функцию в форме  $G(X_{тек})$  по компонентам вектора  $X_{тек}$  и приравняем результаты дифференцирования нулю. Получим систему уравнений (в общем случае нелинейную), из которой можно найти искомый вектор  $X_{онт}$ . Так, в одномерном случае, при использовании линейной математической модели имеем [10]:

$$\frac{G(X_{тек})}{\partial X_{тек}} = 2(Y_{тр. вых} - a_0 - a_1 X_{тек})(-a_1) = 0,$$

$$Y_{тр. вых} - a_0 - a_1 X_{тек} = 0, \quad (3.1.9)$$

$$X_{онт} = X_{тек} = \frac{Y_{тр. вых} - a_0}{a_1}.$$

Если коэффициенты математической модели (3.1.3) заданы с погрешностью (будем называть такие коэффициенты оценками и обозначать их, соответственно, через  $a'_0 \dots a'_i, (i=1..n)$ ), то найденное управленческое решение (3.1.9) преобразуется к виду

$$X'_{онт} = \frac{Y_{тр. вых} - a'_0}{a'_1}. \quad (3.1.10)$$

Для прогнозирования результатов реального процесса производства по найденному значению переменной управления подставим полученное значение  $X'_{онт}$  в математическую модель с истинными значениями коэффициентов [11]. В результате получим:

$$(3.1.11)$$

$$Y_{вых} = a_0 + a_1 \left( \frac{Y_{тр. вых} - a'_0}{a'_1} \right) = a_0 + \frac{a_1 Y_{тр. вых}}{a'_1} - \frac{a_1 a'_0}{a'_1} = \frac{a_1}{a'_1} \left( Y_{тр. вых} + a_0 \left( \frac{a'_1}{a_1} - \frac{a'_0}{a_0} \right) \right)$$

Если коэффициенты математической модели (3.1.3) заданы точно, т.е.  $a_0 = a'_0$  и  $a_1 = a'_1$ , то значение выходной переменной будет равно требуемому значению:  $Y_{вых} = Y_{тр. вых}$ . В случае, когда оценки не равны реальным значениям коэффициентов математической модели ( $a_0 \neq a'_0, \dots, a_1 \neq a'_1$ ), величина выходной переменной  $Y_{вых}$  будет отличаться от  $Y_{тр. вых}$ , причём тем больше, чем больше отклонение оценок от их истинных значений. В этом случае для оптимального управления качеством требуется настройка оценок параметров математических моделей под их действительные значения [12].

На реально действующем производстве отклонение оценок коэффициентов модели может быть обусловлено действием меняющих своё значение неконтролируемых переменных, вызывающих дрейф параметров математических моделей. Для оптимального управления такими процессами требуются адаптивные алгоритмы. Рассмотрим один из таких методов.

Допустим, что в двумерном случае математическая модель технического объекта описывается уравнениями

$$\begin{cases} Y_1 = a_{01} + a_{11}x_1 + a_{12}x_2 \\ Y_2 = a_{02} + a_{21}x_1 + a_{22}x_2 \end{cases} \quad (3.1.12)$$

В качестве целевой функции используем функцию качества в виде взвешенной суммы квадратов:

$$F = \psi_1(\tilde{Y}_1 - Y_1)^2 + \psi_2(\tilde{Y}_2 - Y_2)^2 + \psi_3(\tilde{X}_1 - X_1)^2 + \psi_4(\tilde{X}_2 - X_2)^2, \quad (3.1.13)$$

где  $\tilde{Y}_1, \tilde{Y}_2, \tilde{X}_1, \tilde{X}_2$  – требуемые значения выходных и управляющих переменных.

Будем наблюдать за технологическим процессом в течение  $q$  моментов времени  $t_1, t_2, \dots, t_q$ , а все результаты наблюдений заносить в табл. № 3.1.1. Значения выходных и управляющих переменных для  $i$ -го момента обозначим через  $Y_1^i, Y_2^i, X_1^i, X_2^i$  соответственно.

Так как таблица рассчитана на  $q$  строк (наблюдения за  $q$  моментами времени), занесение информации, соответствующей  $t_{q+1}$  моменту времени осуществляется следующим образом:

1. стирается устаревшая информация из последней строки таблицы (соответствующая моменту времени  $t_1$ , наиболее удалённому от текущего момента);
2. все строки таблицы смещаются вниз на единицу;
3. освободившаяся таким образом первая строка таблицы используется

для записи и хранения информации, соответствующей текущему моменту времени  $t_{q+1}$ .

Под действием меняющихся во времени неконтролируемых переменных коэффициенты математической модели могут менять свои значения, и может наступить такой момент, когда модель перестанет адекватно описывать процесс. При этом возникает необходимость уточнения параметров модели под их реальные значения. Для определения момента времени, когда математическая модель перестанет адекватно описывать процесс, используется следующий алгоритм [13].

Значения целевой функции для каждого момента времени  $t_i$  рассчитываем дважды. Первый раз путём подстановки в целевую функцию (3.1.8) вместо текущих значений  $Y_{тек}$  величин  $Y_1^i$  и  $Y_2^i$ , измеряемых на выходе технологического процесса. Второй раз – подстановкой значений выходных переменных  $Y_1^i$  и  $Y_2^i$ , рассчитанных по математической модели (3.1.12). Полученные таким образом значения целевой функции обозначим через

**Таблица 3.1.1.**  
Результаты наблюдений

Момент времени	Значения выходных переменных		Значения управляющих переменных		Значения целевой функции	
	$Y_1^i$	$Y_2^i$	$X_1^i$	$X_2^i$	$F_{действ}^i$	$F_{расч}^i$
$t_i$	$Y_1^i$	$Y_2^i$	$X_1^i$	$X_2^i$	$F_{действ}^i$	$F_{расч}^i$
$t_q$	$Y_1^q$	$Y_2^q$	$X_1^q$	$X_2^q$	$F_{действ}^q$	$F_{расч}^q$
$t_{q-1}$	$Y_1^{q-1}$	$Y_2^{q-1}$	$X_1^{q-1}$	$X_2^{q-1}$	$F_{действ}^{q-1}$	$F_{расч}^{q-1}$
...	...	...	...	...	...	...
$t_2$	$Y_1^2$	$Y_2^2$	$X_1^2$	$X_2^2$	$F_{действ}^2$	$F_{расч}^2$
$t_1$	$Y_1^1$	$Y_1^2$	$X_1^1$	$X_2^1$	$F_{действ}^1$	$F_{расч}^1$

$F_{действ}^i$  и  $F_{расч}^i$ , соответственно [14]. Затем рассчитывается модуль разности действительного и расчётного значений целевой функции:

$$\left| F_{действ}^i - F_{расч}^i \right| \leq \Delta_{доп} . \quad (3.1.14)$$

При этом, если модуль разности не превышает некоторого допустимого значения, считается, что математическая модель адекватно описывает реальный процесс и в дальнейшем она может использоваться для управления процессом. И наоборот, нарушение условия (3.1.14) свидетельствует о неадекватности модели реальному процессу, при этом возникает необходимость уточнения параметров модели под данные наблюдений [15].

Для подстройки параметров модели под фактическое состояние процесса (предполагается регулярное по времени изменение величины неконтролируемых переменных) используется регрессионный анализ, в основе которого лежит экспериментально-статистический подход [16]. В качестве экспериментальных данных для регрессионного анализа используются данные табл. 3.1.1.

В целом алгоритм вышеописанного подхода, соответствующий моменту времени  $t_i$ , описывается следующей блок-схемой (рис. 3.3.1):

## 3.2. Регрессионный анализ

Для построения экспериментально-статистических моделей используются статистические методы, такие, как корреляционный, регрессионный, конъюнктный и др. Рассмотрим метод регрессионного анализа на примере нахождения оценок линейного регрессионного уравнения.

При использовании регрессионного анализа на переменные накладываются следующие ограничения:

- независимые переменные  $X$  измеряются без ошибок;
- зависимая переменная  $Y$  измеряется с ошибкой, распределенной по



**Рис.3.1.1.** Алгоритм адаптивного управления качеством с подстройкой параметров модели

нормальному закону, что позволяет учесть влияние случайно распределённых неконтролируемых переменных [17].

### **Математическая модель линейной регрессии**

Линейная регрессионная модель имеет вид

$$Y = A \cdot X + B + \sigma, \quad (3.2.1)$$

где  $A, B$  – оценки коэффициентов линейной регрессии;

$\sigma$  – погрешность измерения выходной переменной  $Y$ . Считается постоянной при изменении независимых переменных  $X$ , т.е.  $\sigma^2 = \text{const}$ .

При выполнении вышеописанных условий для получения несмещённых и состоятельных оценок коэффициентов регрессии можно воспользоваться методом наименьших квадратов. Суть метода состоит в том, чтобы найти такие неизвестные оценки параметров регрессии, при которых сумма квадратов отклонений экспериментальных точек  $Y_i$  от линии регрессии была бы наименьшей [18]. Это условие можно записать в виде

$$S = \min \sum_{i=1}^N (Y_i - (A \cdot X_i + B))^2, \quad (3.2.2)$$

где  $S$  – остаточная сумма квадратов;

$N$  – число экспериментальных точек (в нашем случае  $N = q$ ).

Воспользуемся необходимыми условиями существования экстремума. Для этого продифференцируем уравнение (3.2.2) по коэффициенту регрессии  $A$  и приравняем полученную производную нулю:

$$\frac{\partial S}{\partial A} = 2 \cdot \sum_{i=1}^N (\bar{Y}_i - (A \cdot X_i + B)) \cdot (-X_i) = 0. \quad (3.2.3)$$

Раскроем скобки и решим уравнение относительно искомым оценок:

$$\sum_{i=1}^N Y_i X_i - \sum_{i=1}^N X_i^2 \cdot A - \sum_{i=1}^N X_i \cdot B = 0,$$

или

(3.2.4)

$$A \cdot \sum_{i=1}^N X_i^2 + B \cdot \sum_{i=1}^N X_i = \sum_{i=1}^N X_i Y_i$$

Аналогично для коэффициента  $B$ :

$$\frac{\partial S}{\partial B} = 2 \cdot \sum_{i=1}^N (\bar{Y}_i - (A \cdot X_i + B)) \cdot (-1) = 0 .$$
(3.2.5)

Раскроем скобки и решим уравнение относительно искомых оценок:

$$\sum_{i=1}^N Y_i - \sum_{i=1}^N X_i \cdot A - \sum_{i=1}^N B = 0 ,$$

или

(3.2.6)

$$A \cdot \sum_{i=1}^N X_i + B \cdot N = \sum_{i=1}^N Y_i .$$

Объединяя уравнения (3.2.4) и (3.2.6), получим систему из двух уравнений с двумя неизвестными (в  $n$ -мерном случае систему  $n+1$  уравнений с  $n+1$  неизвестным):

$$\begin{cases} A \cdot \sum_{i=1}^N X_i^2 + B \cdot \sum_{i=1}^N X_i = \sum_{i=1}^N X_i Y_i, \\ A \cdot \sum_{i=1}^N X_i + B \cdot N = \sum_{i=1}^N Y_i. \end{cases}$$
(3.2.7)

Вводя обозначение  $\sum = \sum_{i=1}^N$  перепишем уравнение (3.2.7) в матричной форме:

$$\begin{bmatrix} \sum X_i^2 & \sum X_i \\ \sum X_i & N \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{bmatrix} \sum X_i Y_i \\ \sum Y_i \end{bmatrix} .$$
(3.2.8)

Для нахождения коэффициентов регрессии рассчитаем определители:

$$\det_0 = \begin{bmatrix} \sum X_i^2 & \sum X_i \\ \sum X_i & N \end{bmatrix} = N \cdot \sum X_i^2 - \sum X_i^2;$$

$$\det_A = \begin{bmatrix} \sum X_i Y_i & \sum X_i \\ \sum Y_i & N \end{bmatrix} = N \cdot \sum X_i Y_i - \sum X_i \cdot \sum Y_i; \quad (3.2.9)$$

$$\det_B = \begin{bmatrix} \sum X_i^2 & \sum X_i Y_i \\ \sum X_i & \sum Y_i \end{bmatrix} = \sum X_i^2 \cdot \sum Y_i - \sum X_i \cdot \sum X_i Y_i.$$

Таким образом, оценки коэффициентов регрессии вычисляются:

$$A = \frac{\det_A}{\det_0} = \frac{N \cdot \sum X_i Y_i - \sum X_i \cdot \sum Y_i}{N \cdot \sum X_i^2 - \sum X_i^2}; \quad (3.2.10)$$

$$B = \frac{\det_B}{\det_0} = \frac{\sum X_i^2 \cdot \sum Y_i - \sum X_i \cdot \sum X_i Y_i}{N \cdot \sum X_i^2 - \sum X_i^2}.$$

### **Математическая модель квадратичной регрессии**

Квадратичная регрессионная модель имеет вид:

$$Y = A \cdot X^2 + B \cdot X + C + \sigma, \quad (3.2.11)$$

где  $A$ ,  $B$ ,  $C$  – оценки коэффициентов линейной регрессии;

$\sigma$  – погрешность измерения выходной переменной  $Y$ . Считается постоянной при изменении независимых переменных  $X$ , т.е.  $\sigma^2 = const$ .

При выполнении вышеописанных условий для получения несмещенных и состоятельных оценок коэффициентов регрессии можно воспользоваться методом наименьших квадратов. Суть метода состоит в том, чтобы

найти такие неизвестные оценки параметров регрессии, при которых сумма квадратов отклонений экспериментальных точек  $Y_i$  от линии регрессии была бы наименьшей [98]. Это условие можно записать в виде

$$S = \min \sum_{i=1}^N (Y_i - (A \cdot X_i^2 + B \cdot X_i + C))^2, \quad (3.2.12)$$

где  $S$  – остаточная сумма квадратов;

$N$  – число экспериментальных точек (в нашем случае  $N = q$ ).

Воспользуемся необходимыми условиями существования экстремума. Для этого продифференцируем уравнение (3.2.12) по коэффициенту регрессии  $A$  и приравняем полученную производную нулю:

$$\frac{\partial S}{\partial A} = 2 \cdot \sum_{i=1}^N (\bar{Y}_i - (A \cdot X_i^2 + B \cdot X_i + C)) \cdot (-X_i^2) = 0. \quad (3.2.13)$$

Раскроем скобки и решим уравнение относительно искомых оценок:

$$\sum_{i=1}^N Y_i X_i^2 - \sum_{i=1}^N X_i^4 \cdot A - \sum_{i=1}^N X_i^3 \cdot B - \sum_{i=1}^N X_i^2 \cdot C = 0, \quad (3.2.14)$$

или

$$A \cdot \sum_{i=1}^N X_i^4 + B \cdot \sum_{i=1}^N X_i^3 + C \cdot \sum_{i=1}^N X_i^2 = \sum_{i=1}^N X_i^2 Y_i$$

Аналогично для коэффициента  $B$ :

$$\frac{\partial S}{\partial B} = 2 \cdot \sum_{i=1}^N (\bar{Y}_i - (A \cdot X_i^2 + B \cdot X_i + C)) \cdot (-X_i) = 0. \quad (3.2.15)$$

Раскроем скобки и решим уравнение относительно искомых оценок:

$$\sum_{i=1}^N Y_i X_i - \sum_{i=1}^N X_i^3 \cdot A - \sum_{i=1}^N X_i^2 \cdot B - \sum_{i=1}^N X_i \cdot C = 0, \quad (3.2.16)$$

или

$$A \cdot \sum_{i=1}^N X_i^3 + B \cdot \sum_{i=1}^N X_i^2 + C \cdot \sum_{i=1}^N X_i = \sum_{i=1}^N X_i Y_i$$

И для коэффициента  $C$ :

$$\frac{\partial S}{\partial C} = 2 \cdot \sum_{i=1}^N (\bar{Y}_i - (A \cdot X_i^2 + B \cdot X_i + C)) \cdot (-1) = 0. \quad (3.2.17)$$

Раскроем скобки и решим уравнение относительно искомым оценок:

$$\sum_{i=1}^N Y_i - \sum_{i=1}^N X_i^2 \cdot A - \sum_{i=1}^N X_i \cdot B - C \cdot N = 0, \quad (3.2.18)$$

или

$$A \cdot \sum_{i=1}^N X_i^2 + B \cdot \sum_{i=1}^N X_i + C \cdot N = \sum_{i=1}^N Y_i$$

Объединяя уравнения (3.2.14), (3.2.16) и (3.2.18), получим систему из трёх уравнений с тремя неизвестными (в  $n$ -мерном случае систему  $n+1$  уравнений с  $n+1$  неизвестным):

$$\begin{cases} A \cdot \sum_{i=1}^N X_i^4 + B \cdot \sum_{i=1}^N X_i^3 + C \cdot \sum_{i=1}^N X_i^2 = \sum_{i=1}^N X_i^2 Y_i, \\ A \cdot \sum_{i=1}^N X_i^3 + B \cdot \sum_{i=1}^N X_i^2 + C \cdot \sum_{i=1}^N X_i = \sum_{i=1}^N X_i Y_i, \\ A \cdot \sum_{i=1}^N X_i^2 + B \cdot \sum_{i=1}^N X_i + C \cdot N = \sum_{i=1}^N Y_i. \end{cases} \quad (3.2.19)$$

Вводя обозначение  $\sum = \sum_{i=1}^N$  перепишем уравнение (3.2.19) в матричной форме:

$$\begin{bmatrix} \sum X_i^4 & \sum X_i^3 & \sum X_i^2 \\ \sum X_i^3 & \sum X_i^2 & \sum X_i \\ \sum X_i^2 & \sum X_i & N \end{bmatrix} \begin{Bmatrix} A \\ B \\ C \end{Bmatrix} = \begin{bmatrix} \sum X_i^2 Y_i \\ \sum X_i Y_i \\ \sum Y_i \end{bmatrix}. \quad (3.2.20)$$

Для нахождения коэффициентов регрессии рассчитаем определители:

$$\det_0 = \begin{bmatrix} \sum X_i^4 & \sum X_i^3 & \sum X_i^2 \\ \sum X_i^3 & \sum X_i^2 & \sum X_i \\ \sum X_i^2 & \sum X_i & N \end{bmatrix} =$$

$$= \begin{pmatrix} N \sum X_i^4 \sum X_i^2 + 2 \sum X_i \sum X_i^2 \sum X_i^3 - \\ - \sum X_i^2 \sum X_i^3 - N \sum X_i^3 \sum X_i^2 - \sum X_i^2 \sum X_i^4 \end{pmatrix};$$

$$\det_A = \begin{bmatrix} \sum X_i^2 Y_i & \sum X_i^3 & \sum X_i^2 \\ \sum X_i Y_i & \sum X_i^2 & \sum X_i \\ \sum Y_i & \sum X_i & N \end{bmatrix} = \begin{pmatrix} N \sum X_i^2 \sum X_i^2 Y_i + \sum X_i \sum X_i^3 \sum Y_i + \\ + \sum X_i \sum X_i^2 \sum X_i Y_i - \sum X_i^2 \sum Y_i - \\ - N \sum X_i^3 \sum X_i Y_i - \sum X_i^2 \sum X_i^2 Y_i \end{pmatrix};$$

$$\det_B = \begin{bmatrix} \sum X_i^4 & \sum X_i^2 Y_i & \sum X_i^2 \\ \sum X_i^3 & \sum X_i Y_i & \sum X_i \\ \sum X_i^2 & \sum Y_i & N \end{bmatrix} = \begin{pmatrix} N \sum X_i^4 \sum X_i Y_i + \sum X_i \sum X_i^2 \sum X_i^2 Y_i + \\ + \sum X_i^2 \sum X_i^3 \sum Y_i - \sum X_i^2 \sum X_i Y_i - \\ - N \sum X_i^3 \sum X_i^2 Y_i - \sum X_i \sum X_i^4 \sum Y_i \end{pmatrix};$$

$$\det_C = \begin{bmatrix} \sum X_i^4 & \sum X_i^3 & \sum X_i^2 Y_i \\ \sum X_i^3 & \sum X_i^2 & \sum X_i Y_i \\ \sum X_i^2 & \sum X_i & \sum Y_i \end{bmatrix} =$$

$$= \begin{pmatrix} \sum X_i^2 \sum X_i^4 \sum Y_i + \sum X_i^2 \sum X_i^3 \sum X_i Y_i + \\ + \sum X_i \sum X_i^3 \sum X_i^2 Y_i - \sum X_i^2 \sum X_i^2 Y_i - \\ - \sum X_i^3 \sum Y_i - \sum X_i \sum X_i^4 \sum X_i Y_i \end{pmatrix}.$$

Таким образом, оценки коэффициентов регрессии вычисляются:

$$\begin{aligned}
 A &= \frac{\det A}{\det_0}; \\
 B &= \frac{\det B}{\det_0}; \\
 C &= \frac{\det C}{\det_0}.
 \end{aligned}
 \tag{3.2.21}$$

### 3.3. Моделирование процесса адаптивного управления с подстройкой параметров модели

Поскольку изучение алгоритмов адаптивного управления качеством производится на ЭВМ, а не на реальном технологическом процессе, воздействия режимов обработки, а также регулярной и случайной составляющих неконтролируемых переменных на величину выходных переменных операции моделируется на ЭВМ [9]. Для этого используются две модели. Первая – описывающая реальный технологический процесс с фактическими (точными) значениями параметров, учитывающая влияние регулярной и случайной составляющих неконтролируемых переменных. И вторая, используемая для управления технологическим процессом, параметры которой известны пользователю с некоторой погрешностью. Как правило, это аппроксимация реальной математической модели [10].

Предположим, что модель реального технологического процесса описывается уравнением

$$Y = A + Bx. \tag{3.3.1}$$

Будем полагать, что суммарное влияние систематически действующих неконтролируемых переменных  $\bar{\xi}$  на результаты процесса проявляется аддитивно с воздействием основного процесса:

$$Y = A + Bx + \bar{\xi}. \tag{3.3.2}$$

Кроме того, результаты процесса зависят от случайных неконтролируемых переменных и от ошибки измерения выходной переменной  $Y$ . Характер их влияния примем случайным, распределенным по нормальному закону со средним арифметическим  $Y = 0$  и средним квадратичным отклонением  $y = \text{const} > 0$ . Тогда реальная выходная переменная  $Y$  может быть рассчитана следующим образом:

$$Y = \bar{Y} + \bar{\xi} \pm \sigma, \quad (3.3.3)$$

где  $\bar{Y}$  – значение выходной переменной, рассчитанное по математической модели (3.3.1).

Влияние неконтролируемых переменных  $\bar{\xi}$  можно аппроксимировать одним из следующих уравнений:

$$\left\{ \begin{array}{l} 1. \text{ линейным: } \xi = A \cdot t \cdot K_{oe} + B; \\ 2. \text{ параболическим: } \xi = A \cdot t^2 \cdot K_{oe} + B; \\ 3. \text{ гиперболическим: } \xi = \frac{A}{t} \cdot K_{oe} + B; \\ 4. \text{ степенным: } \xi = A \cdot t^A \cdot K_{oe}; \\ 5. \text{ показательным: } \xi = A^t \cdot K_{oe}; \\ 6. \text{ логарифмическим: } \xi = A \cdot \lg t \cdot K_{oe} + B; \\ 7. \text{ синусоидальным: } \xi = A \cdot K_{oe} \cdot \sin t + B, \text{ и т.д.,} \end{array} \right. \quad (3.3.4)$$

где  $K_{oe}$  – коэффициент, регулирующий скорость роста регулярной составляющей погрешности (подбирается экспериментально).

Вследствие регулярного, меняющегося во времени, воздействия неконтролируемых переменных наблюдается дрейф выходных переменных технологической операции [11]. Для учёта влияния неконтролируемых переменных и ошибок измерения на выходные переменные составляется табл. 3.3.1, рассчитанная на  $q$  моментов времени (строк), предшествую-

**Таблица 3.3.1.**

Результаты наблюдений

Расчетный момент времени $q$	Значение управляющей переменной $X$	Значение $\bar{Y}$ , рассчитанное по математической модели	Реальное значение выходной переменной $Y$	Регулярная составляющая погрешности $\eta$	Случайная составляющая погрешности $\sigma$
1	$X_1$	$\bar{Y}_1$	$Y_1$	$\eta_1$	$\sigma_1$
...	...	...	...	...	...
$i$	$X_i$	$\bar{Y}_i$	$Y_i$	$\eta_i$	$\sigma_i$
...	...	...	...	...	...
$q$	$X_q$	$\bar{Y}_q$	$Y_q$	$\eta_q$	$\sigma_q$

-щих началу управления.

В первый столбец таблицы заносятся номера моментов времени. Причём в первой строке, соответствующей моменту времени  $t = 1$ , содержится информация, непосредственно предшествующая началу управления, а  $q$ -ая строка соответствует моменту времени, наиболее удалённому от текущего момента, поэтому в ней содержится наиболее устаревшая информация.

В третий столбец таблицы заносятся значения выходной переменной  $\bar{Y}_i$ , рассчитанные по математической модели (3.3.1). В качестве управляющей переменной  $X_i$  в уравнение (3.3.1) подставляются значения, получаемые с помощью датчика случайных чисел с равномерным распределением случайной величины из допустимого диапазона изменения  $X$ :  $(X_{\min}, X_{\max})$ . Полученные таким образом значения управляющей переменной заносятся во второй столбец таблицы. В четвертый столбец таблицы заносятся реальные значения выходной переменной  $Y$ , рассчитанные по модели (3.3.3), с учётом влияния регулярной и случайной составляющих, взятых, соответ-

ственно, из пятого и шестого столбцов таблицы. Моделирование регулярной составляющей осуществляется по одному из уравнений системы (3.3.4), а расчёт случайной составляющей – с помощью датчика случайных чисел с нормальным законом распределения и параметрами распределения: средним арифметическим  $Y = 0$ , средним квадратичным отклонением  $\sigma_0 = K_x \cdot \bar{Y}$ , где  $K_x$  – коэффициент разброса, выбираемый из диапазона 1–20% (0,01–0,2) от  $\bar{Y}$ , в зависимости от предполагаемой точности измерения выходной переменной [12].

Для управления технологическим процессом используются концептуальные математические модели, известные оператору или технологу, ведущему процесс, с некоторой погрешностью. В качестве такой модели используется полиномиальная аппроксимация, получаемая по данным табл. 3.3.1 с использованием программы регрессионного анализа.

### **Запуск системы управления**

Управление технологическим процессом начинается с момента времени  $t_{q+1}$ , что соответствует моменту включения технологического оборудования. Все данные, получаемые на этом этапе, заносятся в таблицу в соответствии с алгоритмом, описанным выше (при заполнении табл. 3.1.1). Предварительно производится сдвиг всех строк, начиная с первой, вниз на одну строку. А затем получаемые в процессе управления данные заносятся в первую строку. При этом предполагается, что в начальный момент времени параметры математической модели известны нам с идеальной точностью (рассчитаны по данным табл. 3.3.1). Пусть эти коэффициенты равны:  $A=5$  и  $B=1,5$ . Тогда, используя уравнение (3.1.9), рассчитывается значение управляющей переменной и реализуется выбранное управление. Работа технологического оборудования моделируется на компьютере с помощью математических моделей (3.3.1) и (3.3.3). Ожидаемое значение

выходной переменной  $\bar{Y}_i$  определяется по формуле (3.3.1), а реальное значение  $Y_i$  – рассчитывается по формуле (3.3.3) с учётом действия регулярной и случайной составляющих [13]. Полученные значения управляющей  $X_1$ , ожидаемой  $\bar{Y}_1$  и реальной  $Y_1$  значений переменных заносятся в первую строку таблицы.

Затем рассчитываются действительное  $F_{действ}^i$  и расчетное  $F_{расч}^i$  значения критерия качества управления. Далее эти значения сравниваются. Если  $\left| F_{расч}^i - F_{действ}^i \right| \leq \Delta_{дон}$ , то модель (3.3.1) считается адекватной процессу и в дальнейшем может использоваться для расчёта управляющих воздействий очередной партии изделий. В противном случае требуется коррекция модели по данным табл. 3.3.1. И так по всему временному диапазону от  $t_{q+1}$  до  $t_{q+n}$ . Результаты управления выводятся на экран дисплея диск.

### **3.4. Метод выбора недоминируемого решения из ряда недоминируемых альтернатив**

Метод выбора недоминируемого решения из ряда недоминируемых альтернатив – это метод принятия решения на множестве альтернатив, соответствующих различным процессам производства углеродных наноструктур, в котором учитывается не только неопределённость в оценках лица, принимающего решение (ЛПР), но и его интуитивные склонности относительно реальных характеристик критериев качества получаемых наноматериалов и их специфики [43].

Предлагаемый метод позволяет выбрать в качестве оптимальной альтернативу, которая минимизирует взвешенное расстояние от назначенной «идеальной точки» для размытой и интуитивной составляющих. Таким образом, данный метод комбинирует оценки каждого значения критерия

при отсутствии полной информации о предпочтениях и субъективные оценки относительно важности всех критериев [45, 46].

Допустим, имеется  $m$  альтернатив, представляющих собой различные АТО, из которых необходимо сделать выбор. Каждому выбору соответствует  $n$  значений критериев. Для  $k$ -той альтернативы существует множество критериев  $(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ , где  $x_i^{(k)}$  – значение  $i$ -го критерия, если  $k$ -ая альтернатива выбрана. Будем обозначать через  $A^{(k)}$ ,  $k$ -тую альтернативу, однозначно определяющую множество  $(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$  и соответствующую одному из вариантов состава АТО [20].

Согласно концепции «идеальной точки» обозначим через  $A^*$   $(x_1^*, x_2^*, \dots, x_n^*)$  «идеальную» альтернативу, для которой  $x_i^*$  предпочтительнее  $x_n^{(k)}$  для любых  $k$  и  $i$ . Отметим, что некоторые значения  $x_i^*$  могут и не соответствовать значениям альтернатив  $A^{(k)}$  рассматриваемого множества.

Обозначим субъективную оценку относительной важности каждого  $i$ -го критерия с помощью нечёткого множества  $H_i$ .

$$X_i = \{x_i^{(k)} : 1 \leq k \leq m\} \cup \{x_i^*\}, \quad (3.4.1)$$

где  $x_i^*$ , в общем случае, не является частью множества  $\{x_i^{(k)} : 1 \leq k \leq m\}$ .

$H_i$  описывается своей характеристической функцией вида

$$f_{H_i} : X_i \rightarrow [0;1]. \quad (3.4.2)$$

Таким образом,  $f_{H_i}(x_i^{(k)})$  характеризует предпочтительность  $A^{(k)}$  по  $i$ -му критерию. Очевидно, что  $f_{H_i}(x_i^*) = 1$ , т.к.  $x_i^*$  – наилучшее возможное значение  $i$ -го критерия. В некоторых случаях может быть использовано предварительно определённое соотношение для нахождения значений  $f_{H_i}(x_i^{(k)})$ , например:

если  $0 < x_i^{(k)} \leq x_i^* \forall k$ , то  $f_{H_i}(x_i^{(k)}) = \frac{x_i^{(k)}}{x_i^*}$ ;

если  $0 < x_i^* \leq x_i^{(k)} \forall k$ , то  $f_{H_i}(x_i^{(k)}) = \frac{x_i^*}{x_i^{(k)}}$ .

Как правило, специалист хотел бы использовать свои собственные субъективные знания на этапе назначения «удовлетворяющих» коэффициентов для  $i$ -го значения критерия, обеспеченного  $A^{(k)}$  альтернативой.

Выбор альтернативы представляет собой попытку минимизировать «расстояние» между  $A^{(k)}$  и  $A^*$ , которое, несомненно, является функцией его субъективных предпочтений, характеризующих квалификацию, опыт практической работы, уровень знаний о предметной области, а также зависит от уровня информированности о представленном для выбора реальном множестве альтернатив в конкретной ситуации. Другими словами, расстояние между двумя альтернативами, принимая во внимание относительную важность (вес), назначенную для каждого критерия, должно быть минимально. При этом благодаря сгенерированным коэффициентам в процессе выбора учитывается неопределённость, присущая данной задаче. Очевидно, что если для всех  $x_i^{(k)}$  ( $1 \leq k \leq m$ ) на этапе принятия решения информация о предпочтительности недоступна (скрыта), то неопределённость по  $i$ -му критерию становится относительно высока. Однако, высокий уровень для  $x_i^{(k)}$  и низкий для  $x_i^{(l)}$ , где  $k \neq l$ , снижает неопределённость по  $i$ -му критерию.

Разрабатывая решающий алгоритм, применим наиболее часто используемую меру неопределённости – энтропию, обозначив через  $p_i$  вероятностную функцию распределения:

$$-\sum_{i=1}^m p_i \ln p_i . \quad (3.4.3)$$

Чтобы выразить нечёткое множество  $H_i$  через распределение вероятностей, установим:

$$p_i^{(k)} = \frac{f_{H_i}(x_i^{(k)})}{\sum_{l=1}^m f_{H_i}(x_i^{(l)})}, \quad (3.4.4)$$

где  $p_i^{(k)}$  – оценка степени принадлежности  $x_i^{(k)}$  нечёткому множеству  $H_i$ .

Теперь можно определить неопределённость по  $i$ -му критерию как функцию энтропии:

$$e_i = -M \sum_{k=1}^m p_i^{(k)} \ln p_i^{(k)}, \quad (3.4.5)$$

где  $M$  – нормирующая константа, допускающая максимальное значение  $e_i$ , равное 1. Фактически,  $e_i$  максимально, если  $p_i^{(k)} = 1/m$ , следовательно,  $M = 1/\ln m$  и  $0 \leq e_i \leq 1$ .

Общая неопределённость получается путём суммирования неопределённостей по каждому критерию:

$$E = \sum_{i=1}^n e_i = -\frac{1}{\ln m} \sum_{i=1}^n \sum_{k=1}^m p_i^{(k)} \ln p_i^{(k)}. \quad (3.4.6)$$

В расстоянии между двумя альтернативами весовые коэффициенты, зависящие от неопределённости, будут малы, если неопределённость высока. Высокое значение неопределённости относительно некоторого  $i$ -го критерия означает, что довольно большое количество альтернатив имеет близкие по значению «удовлетворяющие» коэффициенты относительно этого  $i$ -го критерия, и поэтому решение о выборе некоторой альтернативы не будет существенно зависеть от этого критерия [21]. Другими словами,  $i$ -й критерий, не отличающийся по большинству альтернатив, имеет более существенное значение для принятия решения. Определим «удовлетворя-

ющий» коэффициент как

$$\bar{S}_i = \frac{1 - e_i}{\sum_{i=1}^n (1 - e_i)} = \frac{1 - e_i}{n - E}, \quad 0 < i \leq n. \quad (3.4.7)$$

Отметим, что когда неопределённость, соответствующая  $i$ -му критерию, высока, то  $e_i$  близко к 1 и коэффициент  $\bar{S}_i$  мал.

Предлагаемая ниже процедура комбинирует нечёткость, характеризующую полученные «удовлетворяющие» коэффициенты и субъективную составляющую, присущую конкретному специалисту на этапе принятия решения о составе АТО.

Пусть  $w_i$  – относительные веса по важности, назначенные ЛПР  $i$ -му критерию априори. Тогда комбинированный весовой коэффициент можно определить как

$$S_i = \frac{S_i w_i}{\sum_{i=1}^n \bar{S}_i w_i}, \quad 0 < i \leq n. \quad (3.4.8)$$

Мера того, насколько удалена  $A^{(k)}$  от идеальной альтернативы  $A^*$ , соответствует тому, насколько значения атрибутов  $A^{(k)}$  удовлетворяют субъективным целевым представлениям ЛПР. Введём понятие *дополнения* нечёткого множества  $H_i$ , которое будем обозначать  $U_i$  и для которого выполняется  $f_{U_i}(x_i^{(k)}) = 1 - f_{H_i}(x_i^{(k)})$ .

Итак, если необходимо определить, насколько удалены друг от друга альтернативы  $A^{(k)}$  и  $A^*$  по  $i$ -му критерию, то можно увидеть это по степени неудовлетворённости ЛПР значением  $i$ -го критерия для  $A^{(k)}$ , т.е. определим

$$Dis^{(i)}[A^*, A^{(k)}] = S_i f_{U_i}(x_i^{(k)}), \quad (3.4.9)$$

и тогда

$$\sigma[A^*, A^{(k)}] = \sqrt{\sum_{i=1}^n (Dis^{(i)}[A^*, A^{(k)}])^2} . \quad (3.4.10)$$

Здесь значение  $Dis^{(i)}$  характеризует *дисперсию* по каждому отдельному  $i$ -му критерию, а  $\sigma[A^*, A^{(k)}]$  соответствует среднеквадратичному отклонению  $Dis^{(i)}$ . Таким образом, целью является выбор альтернативы  $A^{(k_0)}$ , которая максимально близка к  $A^*$  (чем меньше различие между  $A^*$  и  $A^{(k)}$ , тем лучше):

$$\sigma[A^*, A^{(k_0)}] = \min_k \sigma[A^*, A^{(k)}]. \quad (3.4.11)$$

### **Алгоритм выбора недоминируемого решения, основанный на субъективных предпочтениях ЛПР**

На предыдущем этапе решения задачи было сформировано множество недоминируемых альтернатив. Для осуществления окончательного выбора на этом множестве применим метод, описанный ранее. Алгоритм, реализующий данный метод, выглядит следующим образом (рис. 3.4.1) [22].

**Шаг 0.** Назначение идеальной альтернативы. Представление пользователю всех имеющихся недоминируемых альтернатив.

**Шаг 1.** Назначение пользователем  $f_{H_i}(x_i^{(k)})$ , характеризующих предпочтительность  $k$ -й альтернативы по  $i$ -му критерию для  $\forall k, i$ . А также назначение весовых коэффициентов  $w_i$ , определяющих относительную важность  $i$ -го критерия.

**Шаг 2.** Определение наилучшей альтернативы путём выполнения следующих вычислительных процедур:

1. Расчёт частных энтропий  $e_i$  по каждому критерию  $\forall i$ .
2. Расчёт «удовлетворяющих» коэффициентов  $\bar{S}_i$  по каждому кри-

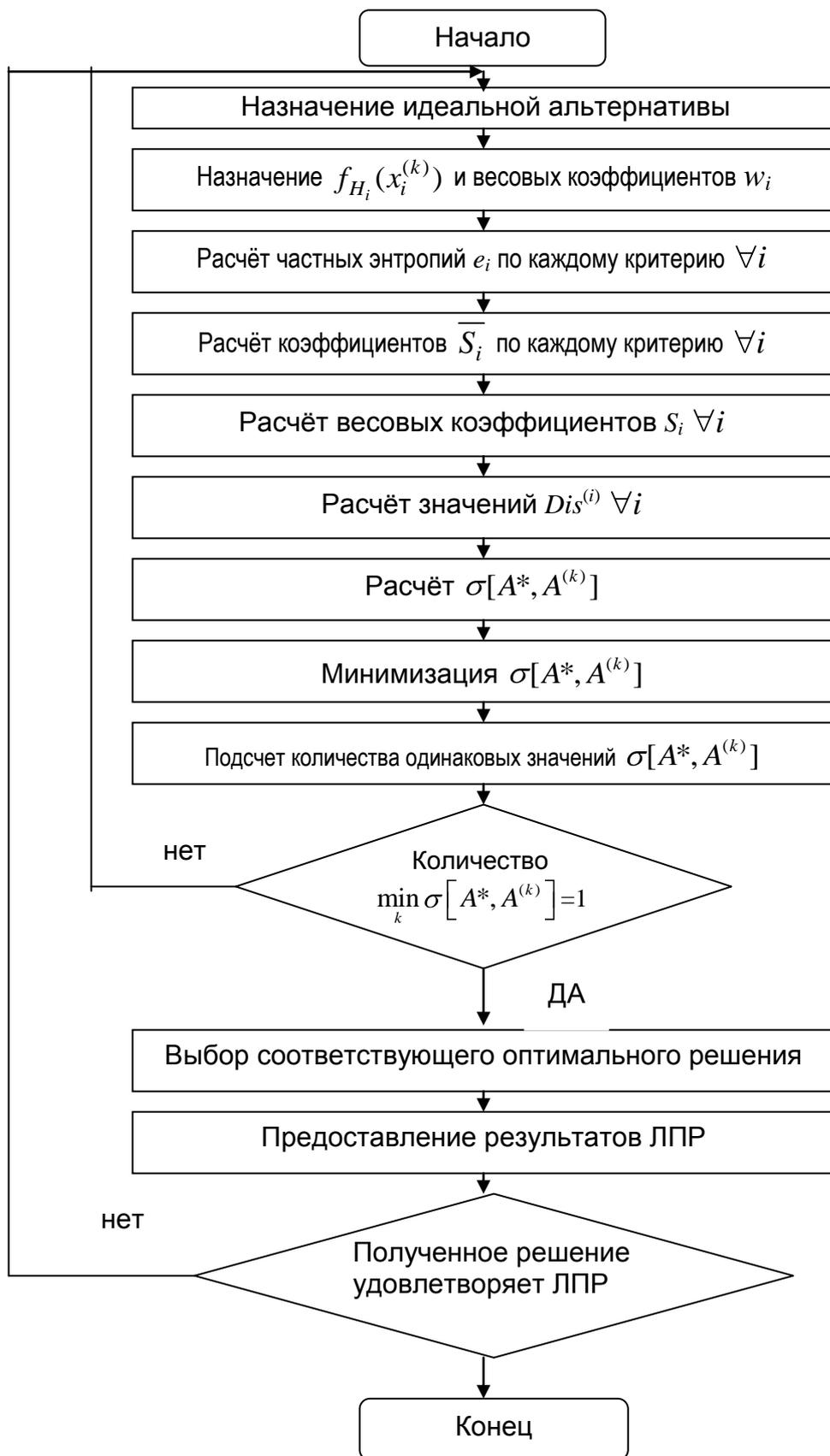
- терию  $\forall i$ .
3. Расчёт весовых коэффициентов  $S_i \forall i$ , комбинирующих нечёткость и субъективную оценку.
  4. Расчёт значений  $Dis^{(i)} \forall i$ , определяющих расстояние между  $k$ -й альтернативой и идеальной альтернативой по  $i$ -му критерию.
  5. Расчёт  $Dis[A^*, A^{(k)}]$ .
  6. Определение наилучшей альтернативы путём минимизации  $Dis[A^*, A^{(k)}]$ .

**Шаг 3.** Предоставление ЛПР значения  $Dis [A^*, A^{(k)}]$  и соответствующего ему решения. Если полученное решение удовлетворяет ЛПР, то завершение, иначе – переход на **Шаг 1**.

### ***Метод упорядоченного предпочтения через сходство с идеальным решением***

Метод упорядоченного предпочтения через сходство с идеальным решением базируется на концепции, что выбранная альтернатива должна иметь наименьшее расстояние до идеального решения и наибольшее расстояние до идеального негативного решения [43].

Принимая, что каждый критерий имеет монотонно возрастающую (убывающую) целевую функцию, легко найти «идеальное» решение, которое состоит из лучших достижимых значений всех критериев и «идеальное негативное» решение, состоящее из худших допустимых значений всех критериев. Один из подходов – взять альтернативу, которая имеет минимальное (взвешенное) евклидово расстояние до идеального решения в геометрическом смысле. Однако этот подход противоречит утверждению о наибольшем расстоянии до идеального негативного решения [23]. Иногда выбранная альтернатива, имеющая наименьшее евклидово расстояние до



**Рис. 3.4.1.** Алгоритм выбора недоминируемого решения, основанный на субъективных предпочтениях ЛПР



**Рис. 3.4.2.** Евклидово расстояние до идеального и идеального негативного решения в двумерном пространстве

идеального решения, ближе всех других альтернатив к идеальному негативному решению. Например, на рис. 3.4.2 альтернатива  $A_1$  имеет меньшие расстояния (и до идеального решения  $A^*$ , и до идеального негативного решения  $A^H$ ), чем другая альтернатива  $A_2$ . Сложно оправдать выбор  $A_1$ . Предлагаемый метод одновременно рассматривает оба расстояния до идеального и идеального негативного решения, учитывая относительную близость к идеальному решению. Этот метод прост и даёт бесспорный порядок предпочтения решения.

### ***Алгоритм работы метода упорядоченного предпочтения через сходство с идеальным решением***

Метод упорядоченного предпочтения через сходство с идеальным решением оценивает следующую матрицу решения, которая содержит  $m$

альтернатив, каждой из которых соответствует  $n$  критериев. Предполагается, что каждый критерий матрицы решения имеет либо монотонно возрастающую, либо монотонно убывающую целевую функцию [38]. Другими словами, чем больше значение каждого критерия, тем больше предпочтения для «выгоды» и меньше предпочтения для «затрат». Кроме того, все критерии, имеющие нечисловой вид, должны быть приведены к числовому виду. Поскольку все критерии не могут иметь равную важность, ЛПР назначает веса их важности.

$$D =$$

	$X_1$	$X_2$	...	$X_j$	...	$X_n$
$A_1$	$x_{1,1}$	$x_{1,2}$	...	$x_{1j}$	...	$x_{1n}$
$A_2$	$x_{2,1}$	$x_{2,2}$	...	$x_{2j}$	...	$x_{2n}$
...	...	...	...	...	...	...
$A_i$	$x_{i,1}$	$x_{i,2}$	...	$x_{ij}$	...	$x_{in}$
...	...	...	...	...	...	...
$A_m$	$x_{m,1}$	$x_{m,2}$	...	$x_{mj}$	...	$x_{mn}$

где  $A_i$  –  $i$ -ая альтернатива;  $x_{ij}$  – значение  $j$ -го критерия  $i$ -ой альтернативы.

Ниже приведена последовательность шагов решения метода.

**Шаг 1.** Нормализация матрицы решения. На этом шаге критерии, имеющие различные измерения, преобразуются в безразмерные критерии, что позволит в дальнейшем их сравнение. Один из подходов – разделить значение каждого критерия на норму вектора суммы значений критериев. Элемент  $r_{ij}$  нормализованной матрицы  $R$  вычисляется как

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}$$

**Шаг 2.** Построение взвешенной нормализованной матрицы решения.

Здесь к нормализованной матрице решения применяются веса

$w = (w_1, w_2, \dots, w_n)$ ,  $\sum_{j=1}^n w_j = 1$ , заданные ЛПР. Матрица получается

умножением каждого столбца матрицы  $R$  на соответствующий весовой коэффициент  $w_j$ . Следовательно, взвешенная нормализованная матрица решения  $V$  равна

$$V = \begin{bmatrix} w_1 r_{1,1} & w_2 r_{1,2} & \dots & w_j r_{1j} & \dots & w_n r_{1n} \\ w_1 r_{2,1} & w_2 r_{2,2} & \dots & w_j r_{2j} & \dots & w_n r_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ w_1 r_{i1} & w_2 r_{i2} & \dots & w_j r_{ij} & \dots & w_n r_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ w_1 r_{m1} & w_2 r_{m2} & \dots & w_j r_{mj} & \dots & w_n r_{mn} \end{bmatrix}.$$

**Шаг 3.** Определение идеального и идеального негативного решения. Определим две искусственные альтернативы  $A^*$  и  $A^H$  как

$$A^* = (\max_i v_{ij} | j \in J), (\min_i v_{ij} | j \in J') | i = 1, 2, \dots, m = v_1^*, v_2^*, \dots, v_n^*,$$

$$A^H = (\min_i v_{ij} | j \in J), (\max_i v_{ij} | j \in J') | i = 1, 2, \dots, m = v_1^H, v_2^H, \dots, v_n^H,$$

где  $J = j = 1, 2, \dots, n | j$  – связанные с критерием выгоды ,

$J' = j = 1, 2, \dots, n | j$  – связанные с критерием затраты .

Эти две искусственно созданные альтернативы  $A^*$  и  $A^H$  определённо будут являться наиболее предпочтительной (идеальное решение) и наименее предпочтительной (идеальное негативное решение) альтернативами соответственно.

**Шаг 4.** Вычисление степени удалённости. Удалённость одной альтернативы от другой может быть измерена  $n$ -мерным евклидовым расстоянием.

Удалённость каждой альтернативы от идеальной вычисляется как

$$S_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2}, \quad i = 1, 2, \dots, m.$$

Аналогично вычисляется удалённость от идеального негативного решения:

$$S_i^H = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^H)^2}, \quad i = 1, 2, \dots, m.$$

**Шаг 5.** Вычисление относительной близости к идеальному решению. Относительная близость альтернативы  $A_i$  к  $A^*$  определяется как

$$C_i^* = \frac{S_i^*}{S_i^* + S_i^H}, \quad 0 < C_i^* < 1, \quad i = 1, 2, \dots, m.$$

Ясно, что  $C_i^* = 1$ , если  $A_i = A^*$  и  $C_i^* = 0$ , если  $A_i = A^H$ . Альтернатива  $A_i$  тем ближе к  $A^*$ , чем  $C_i^*$  ближе к 1.

**Шаг 6.** Ранжирование порядка предпочтений. Набор альтернатив теперь может быть ранжирован в соответствии со значениями  $C_i^*$ . Наилучшим решением будет альтернатива с наибольшим значением  $C_i^*$ .

### **3.5. Метод упорядоченного предпочтения через сходство с идеальным решением для задач многоцелевого принятия решений**

Большинство задач принятия решений имеют множество целевых функций, которые не могут быть оптимизированы одновременно из-за присущей им несоразмерности и конфликту между этими целями. Поэтому основной целью этих задач становится получение «лучшего компромисс-

ного» решения. Для решения задач многоцелевого принятия решений (multiple objectives decision making) было предложено большое число методологий [24].

Математически задачу многоцелевого принятия решений (МЦПР) можно представить в виде:

$$\max / \min [f_1(x), f_2(x), \dots, f_k(x)], \quad (3.5.1)$$

где  $x \in X = \{x \mid g_h(x) \geq, =, < 0, \quad h = 1, 2, \dots, m\}$ ,

$f_j(x)$ : цель (выгода) для максимизации,  $j \in J$ ,

$f_i(x)$ : цель (затраты) для минимизации,  $i \in I$ .

Задача состоит из  $n$  переменных,  $m$  условий и  $k$  целей. Функции  $f_i(x), f_j(x)$  и  $g_h(x) \forall i, j, h$  могут быть линейными и нелинейными. Для решения формулы (1) был разработан новый подход – метод упорядоченного предпочтения через сходство с идеальным решением для решения задач МЦПР (method of ordered preference through likeness with ideal decision – TOPSIS).

Сначала этот метод был предложен для решения задач принятия многоатрибутивных решений (multiple attribute decision making, MADM). Данный подход обеспечивает принцип компромисса, т.е. выбранная альтернатива должна иметь «наикратчайшее расстояние до позитивного идеального решения (ПИР)» и «наибольшее расстояние до негативного идеального решения (НИР)». На практике TOPSIS для MADM используется при решении задач оценки/выбора с конечным числом альтернатив. Однако концепция TOPSIS никогда не использовалась для разработки методологии решения задач МЦПР, имеющих неограниченное число альтернатив [18].

Используя принцип компромисса, TOPSIS для МЦПР вначале уменьшает  $k$ -мерное пространство целей до двумерного (расстояние до ПИР и расстояние до НИР), и избавляется от несоразмерности между исходными целями. Так как обычно существует конфликт между расстоянием до ПИР

и расстоянием до НИР, то одновременно достигнуть компромисса между наикратчайшим расстоянием до ПИР и наибольшим расстоянием до НИР мы не можем. Поэтому оба критерия – «наикратчайшее расстояние до позитивного идеального решения» и «наибольшее расстояние до негативного идеального решения» – заменяются на «как можно ближе к ПИР» и «как можно дальше от НИР». Термины «как можно ближе» и «как можно дальше» – нечеткие, и, чтобы их смоделировать, применяются функции принадлежности из теории нечетких множеств. Исходная задача становится теперь нечеткой двучелевой задачей программирования, и для достижения компромисса при её решении используется оператор минимума-максимума Беллмана-Заде. Это компромиссное решение и становится решением исходной  $k$ -целевой задачи.

Нужно заметить, что критерии «как можно ближе к ПИР» и «как можно дальше от НИР» похожи на «как можно больше выгоды» и «как можно меньше риска». Ниже приводится описание самого метода упорядоченного предпочтения через сходство с идеальным решением.

Многокритериальное принятие решений (МПР) широко изучалось с 1970-х г.г. В отличие от других методологий, первым шагом при решении задач МПР обычно является определение контрольных точек в многоцелевом/многокритериальном пространстве. С учётом этих точек оцениваются альтернативы и принимается компромиссное решение/альтернатива [19].

Принцип компромисса состоит в том, что выбранная альтернатива должна иметь «наикратчайшее расстояние до позитивного идеального решения» и «наибольшее расстояние до негативного идеального решения». Для задач с бесконечным множеством критериев этот принцип формулируется так – минимизировать расстояние до ПИР и максимизировать расстояние до НИР. Т.к. эти два критерия обычно конфликтуют между собой, они оба могут достигнуть только *части* индивидуального оптимума. Поэтому принцип компромисса становится таким – выбранное решение должно

находиться «как можно ближе к ПИР» и «как можно дальше от НИР». Для того, чтобы математически сформулировать принцип компромисса, сначала определим контрольные точки ПИР и НИР для формулы как

$$f_t^+ = \max_{x \in X} f_j(x) \quad \forall j \in J = \min_{x \in X} f_i(x) \quad \forall i \in I ; \quad (3.5.2)$$

$$f_t^- = \min_{x \in X} f_j(x) \quad \forall j \in J = \max_{x \in X} f_i(x) \quad \forall i \in I , \quad (3.5.3)$$

где  $t=1, 2, \dots, k$ .

Пусть  $f^+ = f_1^+, f_2^+, \dots, f_k^+$  является вектором решения формулы (3.5.2).

Этот вектор состоит из индивидуальных наилучших возможных решений для всех целей и называется ПИР. Подобно ему,  $f^- = f_1^-, f_2^-, \dots, f_k^-$  является вектором решения формулы (3.5.3) и состоит из наихудших возможных решений для всех целей. Заметим, что  $f^+$  и  $f^-$  всегда находятся в пределах допустимой области значений формулы (1).

Кроме ПИР и НИР для измерения расстояния используется  $L_p$  – метрика Миньковского. Из-за несоизмерности между целями необходимо сначала нормализовать компонентное расстояние (от ПИР и НИР) для каждой цели. Получаем следующие функции расстояния:

$$d_p^{ПИР} = \left\{ \sum_{j \in J} w_j^p \left[ \frac{f_j^+ - f_j(x)}{f_j^+ - f_j^-} \right]^p + \sum_{i \in I} w_i^p \left[ \frac{f_i(x) - f_i^+}{f_i^- - f_i^+} \right]^p \right\}^{1/p} ; \quad (3.5.4)$$

$$d_p^{НИР} = \left\{ \sum_{j \in J} w_j^p \left[ \frac{f_j(x) - f_j^-}{f_j^+ - f_j^-} \right]^p + \sum_{i \in I} w_i^p \left[ \frac{f_i^- - f_i(x)}{f_i^- - f_i^+} \right]^p \right\}^{1/p} , \quad (3.5.5)$$

где  $w_t, t=1, 2, \dots, k$  – относительная важность (веса) целей;

$p=1, 2, \dots, \infty$  – параметр функций расстояния;

$d_p^{ПИР}$  (уступки) и  $d_p^{НИР}$  (поощрения) – расстояния до ПИР и НИР соответственно.

Подчеркнем, что  $w_t$  показывает степень важности  $t$ -ой цели. С другой стороны, свойство параметра расстояния  $p$  таково, что, когда  $p$  увеличивается, расстояние  $d_p$  уменьшается, т.е.  $d_1 > d_2 > \dots > d_\infty$  и в формировании суммарного расстояния упор делается на наибольшее отклонение;  $p=1$  подразумевает равную важность при формировании функции расстояния  $d_t$  для всех индивидуальных отклонений;  $p=2$  подразумевает, что наибольшая важность дается наибольшему отклонению, пропорционально; при  $p=\infty$  наибольшее отклонение полностью доминирует при определении расстояния.  $d_1$  (расстояние Манхэттена) и  $d_2$  (расстояние Евклида) – наидлиннейшее и наикратчайшее в геометрическом смысле, а  $d_\infty$  (расстояние Чебышева) – наикратчайшее в числовом смысле. Среди всех значений  $p$  случаи  $p=1, 2, \infty$  важны практически и даже являются хорошо известным стандартом в областях теории управления и МЦПР.

Теперь вместо  $k$  исходных целей имеем цели «минимизировать расстояние до ПИР  $d_p^{ПИР}$ » и «максимизировать расстояние от НИР  $d_p^{НИР}$ ».

Имеем двучелевую задачу программирования:

$$\begin{aligned} \min d_p^{ПИР}(x) \\ \max d_p^{НИР}(x) , \\ x \in X \end{aligned} \quad (3.5.6)$$

где  $p=1, 2, \dots, \infty$ .

Так как эти две цели обычно конфликтуют друг с другом, то невозможно одновременно достичь их индивидуального оптимума.

# ГЛАВА 4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ УПРАВЛЕНИЯ КАЧЕСТВОМ ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ УСТРОЙСТВ НА ОСНОВЕ УГЛЕРОДНЫХ СТРУКТУР

## 4.1. Классификация задач оптимизации

В общем виде задача оптимизации может быть записана следующим образом:

$$F(X) \rightarrow \underset{X \in D \subseteq S}{opt}, \quad (4.1.1)$$

где:  $F = [f_1, f_2, \dots, f_k]^T$  – оптимизируемая функция, которая обычно называется целевой функцией, целевым функционалом, критерием качества и т.п., и численно выражает степень достижения целей функционирования оптимизируемого объекта;

$X = (x_1, x_2, \dots, x_n)^T$  – называется вектором независимых переменных, его компоненты – неизвестными в задаче оптимизации. Они являются управляемыми входами объекта оптимизации;

$D$  – множество допустимых значений неизвестных, определяемое налагаемыми на неизвестные ограничениями и по-другому называется допустимой областью или допустимым множеством;

$S$  – пространство оптимизации.

Классификация задач оптимизации может быть проведена на основе обсуждения того, какими могут быть объекты, вошедшие в данную математическую формулировку [36].

Прежде всего отметим, что выражение "opt" в формулировке задачи означает минимум или максимум. Так как принципиальной разницы здесь

нет [  $\min F(X) = -\max (-F(X))$  ], то обычно в теории оптимизации заменяют "opt" на "min" или "max" в зависимости от целей задач. Единственное требование – это соблюдение последовательности в терминологии (необходимо придерживаться выражений одного и того же типа при изложении материала).

Таким образом, решать задачи максимизации функции  $F(X)$  можно методами минимизации, но применёнными к целевой функции  $(-F(X))$ . Решением будет та же самая точка, а оптимальное значение только сменит знак.

Будем классифицировать задачи оптимизации, обсуждая по очереди термины, вошедшие в общую формулировку.

### **Основные подходы в многокритериальной оптимизации**

При выборе рациональных вариантов сложных систем большинство возникающих задач оптимизации являются многокритериальными, так как функционирование сложной системы должно удовлетворять многим критериям эффективности одновременно. Основной концепцией, используемой при многокритериальной оптимизации, является концепция недоминируемых точек в пространстве решений и в критериальном пространстве (множества Парето) совместно с методикой последовательного сужения множества таких точек [37]. Рассмотрим паретовскую концепцию применительно к задачам дискретной и комбинаторной оптимизации.

Пусть функционирование системы оценивается по  $p$  критериям качества  $f_1, f_2, \dots, f_p$ . Тогда задача оптимизации примет вид:

$$F(X) = f_1(X), f_2(X), \dots, f_p(X) \rightarrow \min_{X \in D \subset R^n}, \quad (4.1.2)$$

где  $D$  – область допустимых решений (альтернатив) – является конечным или счётным.

Критерии могут быть согласованными, нейтральными или противоречивыми. В первом случае оптимизация одного из критериев приводит к улучшению других. Во втором случае оптимизация одного критерия никак не влияет на другие. Интерес представляет случай конфликтующих (противоречивых) критериев, когда попытка улучшить один из них приводит к ухудшению других. В таком случае решение возможно только на основе компромисса. Математическая модель компромисса в оптимизации обычно строится на основе понятия множества Парето, названного так в честь итальянского экономиста, первым исследовавшего такие модели в начале XX века.

Решение  $X^0 \in D$  называется эффективным (недоминируемым, паретовским, неулучшаемым), если в множестве допустимых альтернатив  $D$  не существует решения, которое по целевым функциям было бы не хуже, чем  $X^0$ , и, по крайней мере по одной целевой функции было бы строго лучше, чем  $X^0$ . (Паретовская точка не может быть улучшена по совокупности всех целевых функций).

(4.1.3)

$$X^0 \in D_p \stackrel{def}{\Leftrightarrow} \neg \exists X \in D ((f_i(X) \leq f_i(X^0), i = \overline{1, p}) \wedge \exists i_0 (f_{i_0}(X) < f_{i_0}(X^0))).$$

Множество  $D_{\Pi}$  всех эффективных точек называется множеством Парето в пространстве переменных (альтернатив).

Множество Парето в критериальном пространстве (в пространстве критериальных функций) – это множество, определяемое как

$$\Pi = f(D_{\Pi}) = \{ (f_1(X), f_2(X), \dots, f_p(X)) \in R^p, X \in D_{\Pi} \}. \quad (4.1.4)$$

Для любой допустимой точки, лежащей вне множества Парето, найдётся точка в множестве Парето, дающая по всем целевым функциям значения не хуже, чем в этой точке, и хотя бы по одной целевой функции – строго лучше. Из определения следует, что решение многокритериальной

задачи оптимизации целесообразно выбирать из множества Парето, так как любое другое, очевидно, может быть улучшено некоторой точкой Парето как минимум по одному критерию без ухудшения других критериев. С позиции математики решения из множества Парето не могут быть предпочтены друг другу, поэтому после формирования множества Парето задача может считаться математически решенной [2]. Очевидно, что с точки зрения практики такой подход не может быть удовлетворительным и, следовательно, требуются какие-то методы, позволяющие облегчить для лица, принимающего решения выбор эффективного варианта из множества предложенных после формального решения задачи многокритериальной оптимизации. Такие подходы будут рассмотрены в данном разделе. Однако прежде всего рассмотрим подходы к построению множества Парето, которое само по себе является довольно-таки сложной процедурой, особенно для задач дискретной и комбинаторной оптимизации.

Сформулируем теоремы о способах получения недоминируемых точек многокритериальной задачи дискретной оптимизации.

Если существует единственное решение задачи

$$\sum_{i=1}^p \lambda_i \cdot f_i(X) \rightarrow \min, \quad \lambda_i \geq 0, \quad \sum_{i=1}^p \lambda_i = 1, \quad (4.1.5)$$

причём, если все  $\lambda_i > 0$ , то данное решение является недоминируемым (точка Парето) для задачи (4.1.4).

Если существует несколько решений задачи (4.1.5) и существуют некоторые  $\lambda_i = 0$ , то решение принадлежит множеству Парето задачи (4.1.4), если и только если оно не доминируется по совокупности критериев точками из множества

$$\text{Arg min} \sum_{i=1}^p \lambda_i \cdot f_i(X).$$

Пусть все  $f_i(X) > 0$  на множестве допустимых решений. Тогда, если

существует единственное решение задачи

$$\phi(X, \lambda) = \max_{i \in 1, \dots, p} \lambda_i \cdot f_i(X) \rightarrow \min, \quad \lambda_i > 0, \quad \sum \lambda_i = 1, \quad (4.1.6)$$

то данное решение является недоминируемым (точка Парето) для задачи (4.1.4).

Если существует несколько решений задачи (4.1.6), то решение принадлежит множеству Парето задачи (4.1.4) тогда и только тогда, когда оно не доминируется по совокупности критериев точками из  $Arg \min \phi(X, \lambda)$ . [5].

Очевидно, что требование неотрицательности критериев на допустимом множестве не сужает общности задачи, так как в противном случае мы можем получить условие теоремы заменой соответствующего критерия на функцию

$$f_i(X) + i \min f_i(X_i).$$

Множество Парето задачи (4.1.4) при условии конечности множества допустимых решений  $D$  и неотрицательности целевых функций на  $D$  совпадает с объединением множеств решений задач

$$\phi_z(X, \bar{\lambda}) = \sum_{i=1}^p \lambda_i \cdot f_i(X)^{z_i} \rightarrow \min_{X \in D}, \quad (4.1.7)$$

при всевозможных

$$\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p) \in \Lambda = \left\{ \bar{\lambda} \in R^p : \lambda_i > 0, \quad \sum_{i=1}^p \lambda_i = 1 \right\},$$

где  $z_1 = 1$ , а  $z_i (i = 2, 3, \dots, p)$  – произвольные фиксированные числа, не меньшие некоторых пороговых величин.

$$D_{II} = \bigcup_{\bar{\lambda} \in \Lambda} Arg \min \sum_{i=1}^p \lambda_i \cdot f_i(X)^{z_i}. \quad (4.1.8)$$

К сожалению, прямое использование результатов этой теоремы чрез-

мерно затруднено из-за сложности вычисления пороговых величин, для которых неизвестны оценки снизу, а оценки сверху очень сложно получить.

Если допустимое множество задачи (4.1.4) конечно и все целевые функции принимают значения в интервале  $[c, d]$ ,  $c > 0$ , то множество Парето задачи (4.1.4) совпадает с множеством решений задач

$$\phi_{\alpha}(X, \bar{\lambda}) = \left[ \sum_{i=1}^p (\lambda_i \cdot f_i(X))^{\alpha} \right]^{1/\alpha} \rightarrow \min_{X \in D}, \quad (4.1.9)$$

при всевозможных

$$\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p) \in \Lambda^{\sigma} = \left\{ \bar{\lambda} \in R^p : \lambda_i \geq \sigma, \sum_{i=1}^p \lambda_i = 1, \sigma = \frac{c}{dp} \right\},$$

где  $\alpha$  – произвольное число, не меньшее некоторого порогового значения  $\bar{\alpha} \geq 1$ .

$$D_{\Pi} = \bigcup_{\bar{\lambda} \in \Lambda^{\sigma}} \text{Arg min}_{X \in D} \left[ \sum_{i=1}^p (\lambda_i \cdot f_i(X))^{\alpha} \right]^{1/\alpha}. \quad (4.1.10)$$

На практике пороговое значение  $\bar{\alpha} \geq 1$  определяется на тестовых задачах последовательным увеличением до тех пор, пока будут выявляться новые точки Парето. Очевидно, что это существенно затрудняет использование теоремы в реальных задачах.

Методы получения точек Парето с помощью свёрток наименее трудоёмки в смысле получения решения отдельных задач оптимизации, но при этом весьма неудобны, так как для нахождения всех эффективных точек требуют изменять коэффициенты  $\lambda_i$  в континуальном пространстве. Кроме того, существует возможность, что при изменении коэффициентов  $\lambda_i$  во всем диапазоне будет получаться одна и та же точка, хотя мощность множества Парето значительна.

Поэтому были сделаны попытки предложить и другие подходы к по-

строению множества Парето, в частности основанные на методе последовательного сужения множества альтернатив. Другим подходом является так называемый метод последовательной генерации точек Парето. Он требует минимального числа итераций, равного мощности множества Парето, однако трудоёмкость итераций резко возрастает для каждой следующей эффективной точки. К сожалению, эти методы ориентированы только на задачи целочисленного программирования, в частности метод предназначен для задач ЦЛП с целыми коэффициентами.

Особое место занимает метод варьирования ограничений, который обеспечивают следующие теоремы.

**Теорема 1.**

Пусть  $X^{*i}$  является решением задачи оптимизации

$$f_i(X) \rightarrow \min_{X \in D_i}; \quad (4.1.11)$$

$$D_i = \{X \in D : f_j(X) \leq \xi_j, j = \overline{1, p}, j \neq i\},$$

где  $\xi_j$  – произвольные числа, такие, что  $D_i \neq \emptyset$ , а  $i = 1, 2, \dots, p$ .

Тогда, в случае единственности решения задачи (4.1.11)  $X^{*i}$  принадлежит множеству Парето задачи (4.1.4). Если же решением задачи (4.1.11) является конечное множество точек, то среди них найдётся хотя бы одна недоминируемая по совокупности критериев задачи (4.1.4).

**Теорема 2.**

Пусть  $X^0 \in D_{\Pi}$  – решение задачи (4.1.4). Тогда, для любого  $i = \overline{1, p}$  найдутся некоторые числа  $\xi_j, j = \overline{1, p}, j \neq i$ , такие, что  $X^0$  является решением задачи (4.1.11).

В многокритериальных задачах дискретной оптимизации с конеч-

ным множеством допустимых альтернатив методом варьирования ограничений могут быть получены все точки Парето за конечное число шагов.

### **Основные подходы к построению множества Парето**

На основе приведённых теорем можно построить несколько возможных подходов и предложить для многокритериальной оптимизации, с учётом указанных особенностей и ограничений, несколько способов построения эффективных точек. В каждом из них задача многокритериальной оптимизации сводится к одной или нескольким задачам однокритериальной оптимизации. Методы должны работать многократно, чтобы сгенерировать множество точек Парето с тем, чтобы дальше выполнить оценку решения.

### **Свёртка критериев**

Задача многокритериальной оптимизации преобразуется в задачу однокритериальной оптимизации с критерием, равным взвешенной сумме исходных критериев.

**Аддитивная свёртка:**

$$f(X) = w_1 f_1(X) + \dots + w_k f_k(X), \quad w_1 + \dots + w_k = 1, \quad w_i > 0. \quad (4.1.12)$$

Пользователь должен указать весовые коэффициенты и выбрать алгоритм оптимизации.

Иногда используется также **мультипликативная свёртка**, где вместо суммы применяется произведение частных критериев, возведённых каждый в свою степень.

В результате решения оптимизационной задачи по свёртке обычно получают Парето–оптимальную точку (для выпуклых функций доказаны соответствующие необходимые и достаточные условия) [38].

Метод свёрток применялся и применяется наиболее часто, но имеет труднопреодолимые недостатки:

- не всегда потеря качества по одному критерию компенсируется приращением по другому. «Оптимальное» по свёртке решение может характеризоваться низким качеством некоторых частных критериев и в связи с этим будет неприемлемым;
- не всегда можно задать весовые коэффициенты критериев. Зачастую известна лишь сопоставимая важность критериев, иногда нет никакой информации о важности;
- величина функции цели, полученная по свёртке, не имеет никакого физического смысла;
- многократный запуск алгоритма по свёртке может давать только несколько различных точек Парето (или одну и ту же) даже в случае, когда в действительности этих точек очень много.

### ***Иерархическая оптимизация (метод уступок)***

Сначала решаются задачи оптимизации по каждому из критериев в отдельности при исходной системе ограничений и отбрасывании остальных критериев и получаются соответствующие минимальные значения функции  $f_k^*$ . Затем критерии упорядочиваются по важности и решается последовательность однокритериальных задач оптимизации по каждому из критериев при дополнительных ограничениях на предыдущие по важности критерии, объявляющих допустимые отклонения (уступки) от найденных ранее минимумов:

$$\begin{aligned} f_i(X) &\rightarrow \min_{X \in D}, \\ f_j(X) &\leq \left(1 + \frac{\Delta_j}{100}\right) f_k^*, \quad j = 1, \dots, i-1. \end{aligned} \tag{4.1.13}$$

Здесь  $\Delta_j$  – допустимая уступка по  $j$ -му критерию. Пользователь должен упорядочить критерии, указать допустимые уступки по каждому из них и выбрать оптимизационные алгоритмы для каждого из критериев.

Недостатками подхода является то, что величины уступок не соизмеримы друг с другом (что исключительно затрудняет их выбор), и то, что в общем случае могут быть получены точки, не являющиеся Парето-оптимальными [55].

### ***Дискриминационный подход***

Пользователь выбирает один из критериев в качестве целевой функции, а остальные критерии учитываются как ограничения по отношению к их минимумам:

$$\begin{aligned} f_i(X) &\rightarrow \min_{X \in D}, \\ f_j(X) &\leq \Delta_j, \quad j = 1, \dots, K, \quad i \neq j. \end{aligned} \tag{4.1.14}$$

Пользователь должен указать главный критерий и величины ограничений на остальные, а также выбрать алгоритм решения получаемой задачи оптимизации.

Дискриминационный подход используется также для сокращения количества частных критериев. В этом случае несколько частных критериев остаются целевыми функциями, а остальные преобразуются в ограничения.

Проблемой здесь является выбор главного критерия и назначение ограничений на остальные. При слишком жёстких ограничениях можно получить пустое допустимое множество, а при слишком слабых – неудовлетворительное решение по некоторым частным критериям [39].

Модификацией этого подхода является метод, при котором все критерии по очереди назначаются ведущими, а полученные затем результаты сравниваются.

## **Целевое программирование**

Предполагается наличие по каждому из частных критериев некоторой цели, к которой должно стремиться.

### ***Минимизация расстояния до целевой точки в $l_1$ -метрике***

Многокритериальная задача заменяется однокритериальной, где в качестве целевой функции выбрана сумма абсолютных значений отклонений критериев от выбранных заранее целей  $y_i$ :

$$f(X) = |f_1(X) - y_1| + \dots + |f_k(X) - y_k|. \quad (4.1.15)$$

Пользователь указывает цели (ими могут быть и индивидуальные минимумы частных критериев) и выбирает алгоритм оптимизации.

### ***Минимизация расстояния до целевой точки в $l_2$ -метрике***

Многокритериальная задача заменяется однокритериальной, где в качестве целевой функции выбрана сумма квадратов отклонений критериев от выбранных заранее целей  $y_i$ :

$$f(X) = (f_1(X) - y_1)^2 + \dots + (f_k(X) - y_k)^2. \quad (4.1.16)$$

Пользователь указывает цели (ими могут быть и индивидуальные минимумы частных критериев) и выбирает алгоритм оптимизации.

При этом подходе проблемой является несопоставимость величин отклонений от цели по каждому из частных критериев [40].

## **Управление по идеалу**

Выполняется минимизация суммарного относительного отклонения целевых функций от их минимумов, полученных при однокритериальной оптимизации в  $l_1$  и  $l_2$ -нормах.

***Метод глобального критерия в  $l_1$ -норме.*** Минимизируется скалярная функция, представляющая собой сумму относительных расстояний

частных критериев от их минимальных значений:

$$f(X) = \frac{f_1(X) - f_1^*}{|f_1^*|} + \dots + \frac{f_k(X) - f_k^*}{|f_k^*|}, \quad (4.1.17)$$

где  $f_i^*$  – оптимальные значения частных критериев, полученные путём их оптимизации при исходных ограничениях и отбрасывании других критериев. Пользователь должен выбрать алгоритмы оптимизации для каждого частного критерия и для обобщённого.

**Метод глобального критерия в  $l_2$ -метрике.** Минимизируется скалярная функция, представляющая собой сумму квадратов относительных отклонений частных критериев от их минимальных значений:

$$f(X) = \left( \frac{f_1(X) - f_1^*}{f_1^*} \right)^2 + \dots + \left( \frac{f_k(X) - f_k^*}{f_k^*} \right)^2. \quad (4.1.18)$$

Полностью аналогичен предыдущему методу.

Здесь неявно предполагается, что критерии равноценны, что редко бывает на практике [41].

### **Минимаксные методы (методы компромисса)**

Минимизируются наибольшие значения, а не суммы целевых функций или их относительных отклонений, в том числе взвешенных, от целевого или идеального значения.

**Минимаксный метод № 1.** Минимизируется максимум абсолютных значений всех частных критериев:

$$\max \left\{ |f_i(X)|, i = 1, \dots, k \right\} \rightarrow \min. \quad (4.1.19)$$

Пользователь выбирает алгоритм оптимизации для обобщённого критерия.

**Минимаксный метод № 2.** Минимизируется максимум значений всех частных критериев:

$$\max f_i(X), i = 1, \dots, k \rightarrow \min. \quad (4.1.20)$$

Пользователь выбирает алгоритм оптимизации для обобщённого критерия.

**Минимаксный метод № 3.** Минимизируется максимум абсолютных расстояний значений всех частных критериев от заданных целей:

$$\max |f_i(X) - y_i|, i = 1, \dots, k \rightarrow \min. \quad (4.1.21)$$

Пользователь определяет целевые значения  $y_i$  для каждого частного критерия и выбирает алгоритм оптимизации.

Общая проблема этих трёх подходов – критерии измерены в различных единицах, следовательно, невозможно сравнивать критерии между собой и вести их совместную оптимизацию.

**Минимаксный метод № 4.** Минимизируется максимум относительных расстояний значений частных критериев от их индивидуальных минимумов:

$$\max \left\{ \frac{f_i(X) - f_i^*}{|f_i^*|}, i = 1, \dots, k \right\} \rightarrow \min, \quad (4.1.22)$$

где  $f_i^*$  – оптимальные значения частных критериев, полученные путём их оптимизации при исходных ограничениях и отбрасывании других критериев. Пользователь должен выбрать алгоритмы оптимизации для каждого частного критерия и для обобщённого. Критерии неявно предполагаются равнозначными [60].

**Минимаксный метод № 5.** Минимизируется максимум взвешенных относительных расстояний значений частных критериев от их индивидуальных минимумов:

$$\max \left\{ w_i \left( \frac{f_i(X) - f_i^*}{|f_i^*|} \right), i = 1, \dots, k \right\} \rightarrow \min, \quad (4.1.23)$$

где  $f_i^*$  – оптимальные значения частных критериев, полученные путём их оптимизации при исходных ограничениях и отбрасывании других критериев. Пользователь должен выбрать алгоритмы оптимизации для каждого частного критерия и для обобщённого и указать весовые коэффициенты (что и является его главной проблемой при выборе этого подхода).

**Минимаксный метод № 6.** Минимизируется максимум взвешенных значений частных критериев:

$$\max w_i f_i(X), i = 1, \dots, k \rightarrow \min. \quad (4.1.24)$$

Пользователь должен выбрать алгоритм оптимизации для обобщённого критерия и указать весовые коэффициенты. Проблемами будут как выбор весовых коэффициентов, так и несопоставимость величин критериев.

Все минимаксные методы порождают проблему выбора метода оптимизации обобщённого критерия. Решение таких задач требует разработки специальных процедур оптимизации. Чаще всего используются методы теории игр.

Применение указанных способов позволяет строить одну или несколько (или все), в зависимости от выбранного подхода, точек множества Парето.

### **Применение математического моделирования для оценки качества производимых нанообъектов**

При моделировании для исследования какого-либо объекта (оригинала) используют другой объект (модель), заменяющий в этом случае первый.

Оценку и прогнозирование качества функционирования можно осуществлять по данным реальной эксплуатации объектов, по результатам лабораторных стендовых испытаний, моделирования процессов функционирования объектов или в результате расчётов. Расчёты и моделирование удобны тем, что дают возможность при минимальных затратах предсказать качество функционирования объектов задолго до их изготовления. Однако для проведения расчётов необходимы математические модели, основанные обычно на довольно грубых допущениях. Поэтому результаты расчётов недостаточно точны. Моделирование, занимая промежуточное положение между расчётами и испытаниями натуральных образцов, позволяет отказаться от введения ряда грубых допущений.

При оценке качества функционирования систем с помощью вероятностного моделирования воспроизводят на модели случайные процессы изменения состояний элементов, находят характеристики случайного процесса изменения определяющего (выходного) параметра системы и приписывают эти характеристики исследуемой системе – оригиналу.

В зависимости от степени физического сходства между оригиналом и моделью различают моделирование по аналогии с использованием математических моделей или при малоформализованных моделях и моделирование на основе подобия.

### ***Дискриминационный метод оптимизации***

Семейство дискриминационных методов предполагает наложение ограничений на критерии на основе некоторой системы требований, предъявляемой к оптимальному решению.

Предлагаемый здесь метод является методом выделения главного критерия. Основная идея этого метода – максимизация наиболее важного (главного) критерия  $f_1(x)$  при условии, что значения других критериев  $f_i(x)$ ,

$i = 2, \bar{k}$  не превышают пороговых значений  $f_i^0$ :

$$\min_{x \in X'} f_1(x), \quad (4.1.25)$$

$$X' = X \cap \{x | f_i(x) \leq f_i^0, i = 2, \bar{k}\}.$$

Этот метод рационально применять в тех случаях, когда имеются соображения о примерных значениях пороговых величин (или довольно узких пределах для них), позволяющие ограничиться рассмотрением сравнительно небольшой части всего множества решений.

## 4.2. Оптимизация обобщённого критерия качества дискриминационным методом

Семейство дискриминационных методов предполагает наложение ограничений на критерии на основе некоторой системы требований, предъявляемой к оптимальному решению [42].

Дискриминационный метод применим к оптимизации сложной системы (в данном случае обобщённого критерия качества (рис. 4.2.1)), которая состоит из отдельных подсистем [43]. Каждая из подсистем может быть реализована  $u_j$  способами, характеризующимися различными значениями параметров, например, функциональными параметрами, стоимостью, габаритами и т.п. Требуется определить вариант обобщенного критерия качества (то есть выбрать наиболее оптимальный вариант каждого параметра), который доставляет экстремум целевой функции  $P(v_i)$  и обеспечивает наиболее высокое качество изготавливаемых нанообъектов, с вероятностями не ниже заданных, при этом затраты не должны превосходить заданной границы [44].

Этот метод рационально применять в тех случаях, когда имеются соображения о примерных значениях пороговых величин (или довольно узких пределах для них), позволяющие ограничиться рассмотрением сравнительно небольшой части всего множества решений [45]. Математическая



**Рис. 4.2.1.** Обобщенный критерий качества

модель этой задачи имеет следующий вид: определить вариант  $v_i$ , доставляющий максимум целевой функции:

$$P(v_i) = \max \left\{ \sum_{p=1}^q g_p(v_i) \right\} \quad (4.2.1)$$

при наличии ограничений:

$$g_p(v_i) = \sum_{j=1}^n g_p(v_j); \quad p = 1, \dots, q; \quad k = 1, \dots, k; \quad j = 1, \dots, n; \quad (4.2.2)$$

$$g_p(v_i) \geq g_p^*, \quad (4.2.3)$$

$$v_i \in V,$$

$$u_i \in U$$

где  $u_j$  – отдельные параметры качества;

$U = \{u_1, \dots, u_j\}, (j = 1, \dots, n)$  – совокупность всех возможных отдельных параметров качества, которые могут быть использованы для оптимизации обобщенного критерия качества;

$v_i$  – видовые критерии качества (совокупности отдельных видов параметров качества);

$V = \{v_1, \dots, v_i\}, (i = 1, \dots, k)$  – совокупность всех возможных видовых критериев качества;

$g_p(u_j)$  – значения параметров качества;

$g_p(v_i)$  – значения видовых критериев качества;

$g_p^*$  – введенные приближительные ограничения значений параметров качества;

$P(v_i)$  – целевая функция – обобщенный критерий качества.

Одним из подходов к решению проблемы выбора является следующая итерационная процедура. Поочередно производится отсев по каждому из критериев. Для этого, рассматривая каждый критерий, для каждой из  $n$  подсистем упорядочим все  $\xi_j$  типов элементов по возрастанию, согласно значениям рассматриваемого критерия [46]. Сумма первых значений упорядоченных типов элементов представляет необходимый минимум ресурсов для формирования системы. Очевидно, что эта сумма должна быть меньше ограничения на текущий критерий, как необходимое условие существования допустимых решений.

Следующим шагом является определение «допуска» для каждой подсистемы, он вычисляется по формуле:

$$\Delta g_p^j = g_p^* - g_p^S,$$

где  $g_p^S$  – сумма первых элементов.

Все элементы  $j$ -ой подсистемы, превосходящие значение «допуска», отбрасываются и в дальнейшем уже не рассматриваются. Рассмотрев все критерии, выполняем следующую итерацию. Процедура заканчивается,

когда отсев элементов больше не происходит [47].

Может, однако, случиться и так, что все элементы будут отброшены, в этом случае необходимо выполнить один или оба из следующих пунктов:

- 1) расширить наложенные на систему ограничения;
- 2) использовать менее ресурсоёмкие элементы подсистем.

Если результат решения задачи оптимального проектирования неудовлетворителен (то есть полученная надёжность системы мала), но используемые ресурсы исчерпаны не полностью, то необходимо прибегнуть к резервированию элементов подсистемы, понимая под этим следующее: если произведено резервирование в  $j$ -ой подсистеме, и она содержит  $\lambda_j + 1$  элементов ( $\lambda_j$  резервных и 1 основной), то выход её из строя происходит при выходе из строя всех  $\lambda_j + 1$  элементов (так называемое «параллельное резервирование»). Максимальная кратность резервирования определяется по формуле

$$\lambda_{j(lj)}^* = \min_{p=1, \dots, q} \left[ \frac{\Delta g_p^j}{g_p(u_{j(lj)})} - 1 \right], \quad (4.2.4)$$

где  $\Delta g_p^j$  – допуски, вычисленные на последней итерации.

Определив по формуле (4.2.4) максимальную кратность резервирования оставшихся после отсева по типам элементов подсистем, можно их продублировать. Завершающий этап предлагаемого метода – поиск оптимального решения [1]. При нахождении решения используются элементы подсистем, с учётом кратности резервирования и целевой функции. При решении используется итерационная процедура, предложенная выше. Однако прежде рассмотрения допустимости по критериям, рассматривается ограничение на целевую функцию [48]. Это ограничение задано не жестким образом, а находится на каждой итерации по формуле

$$f(v) \geq f_i^* = \frac{f_{\max} + f_{i-1}^*}{2}, \quad (4.2.5)$$

для первой итерации  $f_{i-1}^* = f_{\min}$ ,

где  $f_{\min}$  – соответственно сумма максимальных и минимальных в каждой подсистеме значений целевой функции.

Вычисления заканчиваются, когда прекращается отсев, то есть найдено оптимальное решение, или же когда все элементы отброшены. Во втором случае приходится расширить ограничение на целевую функцию. Решением же будет результат последней итерации [49].

Алгоритм работы метода выделения главного критерия качества представлен на рис. 4.2.2.

### **Аналитическая процедура Саати взвешенных критериев**

Для того, чтобы выбрать наилучшее техническое решение, нужно вначале дать оценку полученным вариантам, затем эти оценки сравнить, и только после этого выбрать вариант с наивысшей оценкой [50]. Сравнить варианты зачастую приходится по многим признакам, т. е. пользоваться несколькими шкалами. Причем по одной шкале лучшими оказываются одни варианты, по другой – другие. Признаки, используемые для сравнения, неравнозначны: одни более важны, другие – менее. Некоторые противоречивы. В качестве критериев выбираются признаки, отвечающие наиболее важным целям проектирования.

Одновременно достижение оптимума по всем критериям практически невозможно. Приходится идти на компромисс. В одних случаях это уравнивание критериев, а в других – выбор главного [51].

В начале 1970 года американский математик Томас Саати разработал процедуру поддержки принятия решений. Этот метод относится к классу критериальных и занимает особое место благодаря тому, что он получил исключительно широкое распространение. На основе этого метода разработаны достаточно серьезные системы поддержки принятия решений [52].



**Рис. 4.2.2.** Алгоритм работы метода выделения главного критерия качества

Суть этой процедуры заключается в следующем – вычисляются:

$K_i^R$  – значение критерия  $K^R$  по  $i$ -ому столбцу ( $i$ -ой альтернативе);

$K_j^R$  – значение критерия  $K^R$  по  $j$ -ой строке ( $j$ -ой альтернативе);

$\frac{K_i^R}{K_j^R}$  – парное соотношение критериев  $K^R$ , расположенных в  $i$ -ом столбце и  $j$ -ой строке;

$\sum_{j=1}^G \frac{K_i^R}{K_j^R}$  – суммарное построчное значение парных соотношений альтернатив;

$\sum_{i=1}^J \left( \sum_{j=1}^G \frac{K_i^R}{K_j^R} \right)$  – общая столбцовая сумма значений парных соотношений альтернатив,  $J = G$ ;

$$a_i^R = \frac{\sum_{j=1}^G \frac{K_i^R}{K_j^R}}{\sum_{i=1}^J \left( \sum_{j=1}^G \frac{K_i^R}{K_j^R} \right)} \quad (4.2.8)$$

– нормированное значение альтернатив по критерию  $R$  или оценка альтернатив.

Причём  $\sum_{i=1}^J a_i^R = 1$  для критерия  $R$ , а  $\sum_{R=1}^N a_i$  при  $i = 1 \dots J \Rightarrow \max$ .

Далее нормированные (взвешенные) значения критериев заносятся в таблицу, как исходные данные для дальнейшей оптимизации [53].

### **Пример расчёта обобщенного критерия качества**

Выбрать оптимальный вариант технологического решения с учётом себестоимости получаемой продукции. Рассчитать обобщенный критерий качества углеродных нанотрубок для альтернативных производственных процессов. При расчёте будем использовать данные, полученные в предыдущей задаче и экономических расчётах (табл. 4.2.1).

Обобщённый критерий качества рассчитывается по следующей формуле

$$K_{\text{сум}} = \frac{1}{\sum_{i=1}^N \frac{a_i}{k_i}},$$

**Таблица 4.2.1.**

Расчёт экономического критерия для каждой альтернативы

Альтернатива	Себестоимость наноматериала	Удельный вес в обратной пропорции	Нормированное значение
A1	9174,74	5,878	0,323
A2	9666,07	5,579	0,307
A3	22605,26	2,385	0,133
A4	12485,7	4,319	0,237
Сумма	53931,77	18,16307097	1

где  $a_i$  – весовые коэффициенты (из метода Саати);

$k_i = 0,2$  – возьмем в качестве примера среднее значение выхода годных изделий при производстве электронных объектов в традиционной кремниевой микроэлектронике (20%).

**1.** Произведем расчет обобщённого критерия  $K_{\text{сум1}}$  для первой альтернативы A1 (т. е. метода термического разложения графита):

1.1. Процентный выход однослойных нанотрубок:

$$n_{1;1} = \frac{a_i}{k_i} = \frac{0,348}{0,2} = 1,74 .$$

1.2. Металлический тип проводимости:

$$n_{1;2} = \frac{0,391}{0,2} = 1,955 .$$

1.3. Средняя длина нанотрубок:

$$n_{1;3} = \frac{0,435}{0,2} = 2,175 .$$

1.4. Полупроводниковый тип проводимости:

$$n_{1;4} = \frac{0,058}{0,2} = 0,29 .$$

1.5. Себестоимость:

$$n_{1;5} = \frac{0,323}{0,2} = 1,615 .$$

1.6. Обобщённый критерий:

$$K_{сум1} = \frac{1}{n_{1;1} + n_{1;2} + n_{1;3} + n_{1;4} + n_{1;5}} = \frac{1}{1,74 + 1,955 + 2,175 + 0,29 + 1,615} = 0,13 .$$

2. Рассчитаем обобщённый критерий  $K_{сум2}$  для второй альтернативы А2 (метода плазмохимического осаждения):

2.1. Процентный выход однослойных нанотрубок:

$$n_{2;1} = \frac{0,131}{0,2} = 0,655 .$$

2.2. Металлический тип проводимости:

$$n_{2;2} = \frac{0,352}{0,2} = 1,76 .$$

2.3. Средняя длина нанотрубок:

$$n_{2;3} = \frac{0,058}{0,2} = 0,29 .$$

2.4. Полупроводниковый тип проводимости:

$$n_{2;4} = \frac{0,118}{0,2} = 0,59 .$$

2.5. Себестоимость:

$$n_{2;5} = \frac{0,307}{0,2} = 1,535 .$$

2.6. Обобщённый критерий:

$$K_{\text{сум}2} = \frac{1}{n_{2;1} + n_{2;2} + n_{2;3} + n_{2;4} + n_{2;5}} = \frac{1}{0,655 + 1,76 + 0,29 + 0,59 + 1,535} = 0,21 .$$

3. Рассчитаем обобщённый критерий  $K_{\text{сум}3}$  для третьей альтернативы А3 (метода лазерного испарения):

3.1. Процентный выход однослойных нанотрубок:

$$n_{3;1} = \frac{0,304}{0,2} = 1,52 .$$

3.2. Металлический тип проводимости:

$$n_{3;2} = \frac{0,043}{0,2} = 0,215 .$$

3.3. Средняя длина нанотрубок:

$$n_{3;3} = \frac{0,362}{0,2} = 1,81 .$$

3.4. Полупроводниковый тип проводимости:

$$n_{3;4} = \frac{0,529}{0,2} = 2,645 .$$

3.5. Себестоимость:

$$n_{3;5} = \frac{0,133}{0,2} = 0,665$$

3.6. Обобщённый критерий:

$$K_{\text{сум}3} = \frac{1}{n_{3;1} + n_{3;2} + n_{3;3} + n_{3;4} + n_{3;5}} = \frac{1}{1,52 + 0,215 + 1,81 + 2,645 + 0,665} = 0,15 .$$

4. Рассчитаем обобщённый критерий  $K_{\text{сум}4}$  для четвертой альтернативы А4 (метода холодной деструкции графита):

4.1. Процентный выход однослойных нанотрубок:

$$n_{4;1} = \frac{0,217}{0,2} = 1,085 .$$

4.2. Металлический тип проводимости:

$$n_{4;2} = \frac{0,223}{0,2} = 1,115 .$$

4.3. Средняя длина нанотрубок:

$$n_{4;3} = \frac{0,145}{0,2} = 0,725 .$$

4.4. Полупроводниковый тип проводимости:

$$n_{4;4} = \frac{0,295}{0,2} = 1,475 .$$

4.5. Себестоимость:

$$n_{4;5} = \frac{0,237}{0,2} = 1,185 .$$

4.6. Обобщённый критерий:

$$K_{\text{сум}4} = \frac{1}{n_{4;1} + n_{4;2} + n_{4;3} + n_{4;4} + n_{4;5}} = \frac{1}{1,085 + 1,115 + 0,725 + 1,475 + 1,185} = 0,18$$

Сведем полученные данные в общую таблицу 4.2.2:

Расчёты суммарных критериев качества с учётом себестоимости получаемых наноматериалов показали, что наиболее эффективными методами производства углеродных нанотрубок можно считать метод терми-

**Таблица 4.2.2.**

Общая таблица суммарных критериев качества

Суммарные критерии качества	Сравнительные значения
К1	0,13
К2	0,21
К3	0,15
К4	0,18

ческого разложения и метод лазерного испарения графита;

Расчёты подтвердили, что наилучшим способом производства однослойных углеродных нанотрубок, обладающих полупроводниковыми свойствами и достаточной длиной с целью их дальнейшего позиционирования на микросхемах, является метод лазерного испарения графита, хотя себестоимость получаемого наноматериала выше, чем в других методах производства.

### **4.3. Вероятностное моделирование**

#### ***Моделирование по аналогии***

#### ***с использованием математических моделей***

Этот вид моделирования наиболее близок к расчетам по уровню формализации моделей. Под аналогией понимают частичное сходство, то есть сходство в каком-либо отношении между предметами и явлениями в остальном различными. При моделировании по аналогии с использованием математических моделей, которое за последние годы нашло широкое при-

менение для исследования функционирования систем, имитируются на ЭВМ изменения состояний системы («имитационное моделирование») [37].

Для применения метода первичное описание системы преобразуется в математическую модель процесса функционирования системы.

Для системы с дискретными состояниями математическая модель учитывает все возможные состояния системы, пути и интенсивности переходов из одного состояния в другое. Предполагают, что элементы могут находиться в нескольких несовместимых состояниях, например, в работоспособном или неработоспособном. Функциональные связи между элементами заменяют логическими, характеризующими состояние системы в зависимости от состояний элементов, например условия работоспособности системы при отказах элементов записывают с помощью логических соотношений.

Для системы, имеющей элементы, параметры которых являются непрерывными случайными функциями времени функционирования, математическая модель содержит зависимость определяющего параметра системы от определяющих параметров  $X_j(t)$  элементов:

$$Y(t) = \phi X_1(t), \dots, X_n(t) .$$

При оценке и прогнозировании параметрической (по постепенным отказам) надежности для  $Y(t)$  должна быть задана рабочая область  $\Omega(t)$ , выход за границы которой расценивается как отказ системы.

В математических моделях для вероятностного моделирования допущения могут быть менее жёсткими, чем в расчётных моделях, например можно учесть корреляцию между различными случайными величинами. Однако эти принципиальные возможности не всегда можно реализовать, так как это связано со значительным усложнением процесса моделирования.

Математическая модель лежит в основе моделирующих алгоритмов для ЭВМ или для специализированных установок, моделирующих системы

массового обслуживания. Поскольку при составлении математической модели приходится деформировать, исказить реальные системы, искомый результат содержит методическую ошибку, которую трудно (часто невозможно) оценить и учесть.

Преимуществом метода моделирования по аналогии с использованием математических моделей является возможность применения стандартных технических средств. Кроме того, такое моделирование часто рассматривают как способ избавления от трудностей, встретившихся при попытке оценки и прогнозирования качества функционирования системы расчётным путём. Расчётную математическую модель незначительно изменяют и используют при моделировании [5].

Однако сильная формализация процессов функционирования реальных систем при составлении математических моделей, пригодных для имитации этих процессов на современных технических средствах, приводит к трудностям и ограничениям в применении этого вида моделирования [59]. Кроме того, даже для сравнительно простых систем моделирующие алгоритмы часто бывают очень громоздкими и необходимы большие трудозатраты на разработку и отладку программ. Поскольку математическая модель должна учитывать все возможные состояния системы, моделирование систем, которые могут находиться во многих дискретных состояниях, практически неосуществимо.

Попытки преодоления этих трудностей развиваются в двух направлениях:

- 1) отбрасывание маловероятных состояний и состояний с малыми значениями показателей качества функционирования;
- 2) объединение состояний, то есть упрощение математической модели системы.

Аналогично для систем с непрерывными параметрами нахождение реализаций случайного процесса  $Y(t)$  по характеристикам случайных про-

цессов изменения свойств элементов  $X_1(t), \dots, X_n(t)$  (обычно нестационарных) связано с очень громоздкими вычислениями даже для сравнительно простых моделей. Эти трудности можно обойти, используя метод моделирования при малоформализованных моделях, который можно также условно назвать вероятностным физическим моделированием (ВФМ).

**Моделирование по аналогии  
при малоформализованных моделях  
(вероятностное физическое моделирование)**

При таком моделировании функционирование системы-оригинала имитируют физическим процессом, происходящим обычно в другом масштабе времени, но с сохранением однозначного соответствия между параметрами модели и оригинала.

При осуществлении ВФМ процесса функционирования системы имитируют изменения состояний ее элементов изменениями параметров физических (обычно электрических) элементов, соединение которых соответствует принципиальной схеме системы. Проводят статистические испытания полученной физической модели, многократно повторяя процесс её функционирования в течение заданного времени. Результаты испытаний подвергают статистической обработке.

Для реализации этого простого на первый взгляд алгоритма необходимо иметь систему аналогий для процессов, происходящих в оригинале и модели, и разработать приёмы перехода от оригинала к модели и обратно с определёнными практическими целями. Разработка таких аналогий и приёмов перехода (то есть способов моделирования) является главной частью проблемы и встречает известные трудности, так как связана с выявлением физических или психофизиологических особенностей процессов в различных системах [2].

Применение ВФМ в задачах оценки и прогнозирования качества

функционирования систем позволяет повысить точность результатов за счёт снятия ограничений по сложности исследуемой системы и уменьшения методической ошибки. Кроме того, появляется возможность обсуждать ход исследований вариантов системы-оригинала в физических или психофизиологических терминах и категориях рассматриваемой области науки, не переходя каждый раз к абстрактной математической модели и от неё опять к исследуемой системе [54]. При этом достигается важный психологический эффект: к обсуждению хода исследования в полной мере приобщаются специалисты по исследуемой системе. В ряде случаев можно наблюдать за динамикой функционирования объектов. Одновременно сокращаются трудозатраты на проведение моделирования и его стоимость.

Степень формализации процесса функционирования системы при вероятностном физическом моделировании может быть различной. Наибольшая формализация может быть достигнута при оценке и прогнозировании качества функционирования «направленных» систем при грубых (внезапных) отказах элементов [6]. В этом случае не возникает необходимости имитации физических параметров элементов. При оценке и прогнозировании качества функционирования системы при непрерывном изменении её свойств приходится воспроизводить случайные процессы изменения физических параметров элементов.

Необходимо отметить, что вероятностное физическое моделирование появилось и получило развитие в нашей стране как метод исследования надёжности и качества функционирования систем.

### ***Моделирование на основе подобия***

Системы или явления называют подобными, если все количественные характеристики одного из них можно получить пропорциональным преобразованием характеристик другого. Теория подобия изучает соотношения

между оригиналом и моделью при одной и той же форме движения матери, в частности при механическом движении [5].

При моделировании на основе подобия создают подобную оригиналу материальную систему – физическую модель, с помощью которой изучают процессы, протекающие в оригинале. Такое моделирование можно использовать для оценки качества функционирования уникальных объектов, для которых трудно провести аналогии с другими объектами [55]. При этом возможны два подхода к решению задачи:

1. Для оценки качества функционирования уникальных объектов, изготовленных по существующей технологии, можно использовать опыт эксплуатации подобных объектов, которые в рассматриваемом случае играют роль моделей. С помощью теории подобия показатели качества функционирования этих моделей пересчитывают на оригинал – исследуемый уникальный объект.
2. Для оценки качества функционирования уникальных объектов, изготовленных по новой технологии, можно использовать испытания малых моделей. При этом необходимо изготовить по той же технологии значительное количество малых моделей рассматриваемых объектов и провести их испытания. Затем результаты испытаний пересчитывают на оригинал. Испытания моделей целесообразно проводить в форсированном режиме и осуществлять двойной пересчёт – на нормальный режим и затем на оригинал.

Достоинства моделирования на основе подобия: большая уверенность в правильности результатов исследований, возможность выявления физических причин явлений. Однако часто возможности такого моделирования ограничены из-за трудностей изготовления моделей и проведения экспериментов, высокой стоимости экспериментов, недостаточной точности методов и средств измерений.

## **Достаточные условия в задачах оптимального управления**

Рассмотрим постановку задачи оптимального управления и соответствующие достаточные условия Кротова для случая управляемых процессов без последствия с полной информацией. Начнём с дискретных систем.

Элементами множества  $D$  будем считать пары функций  $(x(t), u(t))$ ,  $x = (x^1, x^2, \dots, x^n)$ ,  $u = (u^1, u^2, \dots, u^r)$ , удовлетворяющие условиям:  $(x(t), u(t)) \in V(t)$  – ограничения на состояние и управляющие органы.

Минимизируемый функционал задаётся как сумма:

$$I = \sum_{t=t_n+1}^{t_k-1} f^0(t, x(t), u(t)) + F(x(t_n), x(t_k), u(t_n), u(t_k)), \quad (4.3.1)$$

где  $f^0(\cdot)$ ,  $F(\cdot)$  – скалярные функции (функционалы). Они отражают представления о качестве процесса (точность, стоимость, затраченная энергия и т.п.) [40].

### **4.4. Задача о замене оборудования**

На производственном участке работает машина, расходы на эксплуатацию которой непрерывно растут с течением времени в результате износа, так что слишком долго её эксплуатировать невыгодно, а целесообразно в некоторый момент заменить новой машиной [26]. Задача состоит в нахождении таких моментов замены в течение рассматриваемого периода, при которых суммарные расходы на закупку и эксплуатацию минимальны.

Пусть  $N$  – число шагов планового периода;  $K$  – начальная цена одной машины;  $p(z) = a + bz$  – расходы на эксплуатацию машины за время  $z$ ;  $a, b > 0$ ;

$x^1(t)$  – суммарные расходы на шаге  $t$ ;  $x^2(t)$  – время эксплуатации одной машины на шаге  $t$ ;  $u$  – признак замены:

$$u = \begin{cases} 1 - \text{если замена,} \\ 0 - \text{в противном случае.} \end{cases}$$

Тогда

$$\begin{aligned} x^1(t+1) &= x^1(t) + uK + (1-u)(a + bx^2(t)), \\ x^2(t+1) &= (1-u)(x^2(t) + 1). \end{aligned} \quad (4.4.1)$$

В начальный момент планового периода

$$x^1(0) = x^2(0) = 0.$$

Требуется найти такую программу, при которой  $x^1(N)$  – минимально. Очевидно, те значения  $t$ , при которых  $\bar{u}(t) = 1$ , и будут моментами замены.

Принцип расширения в данном случае реализуется следующим образом. Игнорируется связь  $x(t+1) = f(t, x(t), u(t))$ . Тем самым определяется множество  $E \supset D$ . Строится функционал на  $E$ .

$$L = G(x(t_n), x(t_k), u(t_n), u(t_k)) - \sum_{t_n}^{t_k-1} R(t, x(t), u(t)), \quad (4.4.2)$$

где

$$G = F(\cdot) + \phi(t_k, x(t_k)) - \phi(t_n, x(t_n)), \quad (4.4.3)$$

$$R = \phi(t+1, f(t, x(t), u(t))) - f^0(t, x(t), u(t)) - \phi(t, x(t)), \quad (4.4.4)$$

$\phi(t, x(t))$  – вспомогательная неопределённая пока функция (функционал),

$$f^0(t_n, \cdot) = 0.$$

Если  $(x(t), u(t)) \in D$ , то, учитывая, что при этом  $f(t, x(t), u(t)) = x(t+1)$ , будем иметь:

$$L = F(\cdot) + \sum_{t=t_H}^{t_k-1} (\phi(t, x(t)) - \phi(t, x(t))) + \sum_{t=t_H}^{t_k-1} f^0(t, x(t), u(t)) = I, \quad (2.4.5)$$

то есть  $L$  есть доопределение  $I$  на более широкое множество.

### Теорема 1.

Пусть  $x_s(t), u_s(t) \subset D$  и найдётся такая функция  $\phi(t, x(t))$ , что

$$R(t, x_s(t), u_s(t)) \rightarrow \mu(t) = \sup_{(x,u) \in V(t)} R, \quad (2.4.6)$$

$$G(x_s(t_H), x_s(t_k), u_s(t_H), u_s(t_k)) \rightarrow m = \inf_{(x,u) \in V(t)} G.$$

Тогда  $x_s(t), u_s(t)$  – минимизирующая последовательность  $T$  на  $D$ .

Здесь  $\sup_{(x,u) \in V(t)} R$  (supremum) – точная верхняя граница функции  $R$  на

$V(t)$ , т.е. наименьшая из всех верхних границ.

Очевидно,  $L \geq l = m - \sum_{t_H}^{t_k-1} \mu(t)$ , но по условиям теоремы  $L_s = I_s \rightarrow l$  на

рассматриваемой последовательности из  $D$ . Следовательно,  $l = d = \inf_D I$ .

Теорема 1 сводит задачу оптимального управления к более простым задачам об экстремуме функций многих переменных.

Проблема сводится теперь к тому, чтобы удачно задать неопределённый элемент – функцию  $\phi$ , которая и доопределяет функционал  $L$  на  $E$ , т.е. так, чтобы минимум принадлежал  $D$  или допускал неограниченное приближение из  $D$ . Тогда задача разрешится. Оказывается, что очень важно, функций  $\phi$  существует много, и это создаёт возможность приспособливаться к специфике задачи. В настоящее время разработаны по крайней мере три характерных способа задания  $\phi$ , два из которых ведут к процедурам Понтрягина и Беллмана, а третий аналогов не имеет и хорошо приспособ-

лен к решению широко распространённых на практике так называемых вырожденных задач [56]. Рассмотрим этот способ на наглядном примере.

Пусть  $x(T) = x_T$  – задано. Если оно свободно, то, решив задачу для различных  $x_T$ , можно затем проварьировать  $x_T$  и выбрать наилучшее. Зададим  $\varphi$  так, чтобы  $R$  не зависела от  $u$  (пользуясь тем, что  $R$  от  $u$  зависит линейно). Тем самым опять-таки достигается неединственность максимумов  $R$ , которая может быть использована потом для удовлетворения дифференциальному уравнению. Имеем  $\varphi_x = 1$ ,  $R = ax$ . Максимум  $R$  достигается на верхней границе  $x$ . Она существует и строится как решение уравнения  $x = ax - u$  при  $u = 0$  до  $x = x_T$  и при  $u = ax$  далее. Очевидно, эти значения дают максимум и минимум  $x$  для данных ограничений, так что выше нельзя подняться из начальной точки  $(0, x, (0))$ , либо из любой точки выше  $x = x_T$  нельзя попасть в точку  $x = x_T$ . Отсюда получается искомое решение.

## ЛИТЕРАТУРА К ГЛАВАМ 3 и 4

1. Дружинин Г.В. Методы оценки и прогнозирования качества. – М.: Радио и связь, 1982. – 160 с. – (Межиздательская серия: Надежность и качество).
2. Варакута С.А. Управление качеством продукции: Учебное пособие. – М.: ИНФРА-М, 2001. – 207 с. – (Серия «Вопрос – ответ»).
3. Кравчук И.С., Тихоглаз Ю.С., Ву Тхи Туэт Ланг. Эволюционная стратегия управления в задачах распознавания образов // Системы управления и информационные технологии, 2008, 2.3(32). – С. 358-360.
4. Ивашов Е.Н., Кравчук И.С. Адаптивное управление качеством производства электронных нанообъектов: Труды X Международной конференции «Опто-, наноэлектроника, нанотехно-логии и микросистемы». – Ульяновск: УлГУ, 2008.
5. Аристов О.В. Управление качеством: Учебник. – М.: ИНФРА-М, 2008. – 240 с. – (Серия «Высшее образование»).
6. Шестопал Ю.Т., Дорофеев В.Д., Шестопал Н.Ю., Андреева Э.А. Управление качеством: Учебное пособие. – М.: ИНФРА-М, 2008. – 331 с. – (Серия «Высшее образование»).
7. Ивашов Е.Н., Кравчук И.С. Математическая модель адаптивного управления качеством производства наноматериалов. / INTERMATIC-2008 // Материалы Международной научно-технической конференции «Фундаментальные проблемы радиоэлектронного приборостроения». – М.: МИРЭА, 2008.
8. Ивашов Е.Н., Кравчук И.С., Тихоглаз Ю.С. Математические модели и алгоритмы управления качеством распознавания образов в кластерных системах обработки информации. – М.: МИЭМ, 2008.
9. Автоматизированные системы управления оборудованием. Адаптивное управление: Методические указания по проведению лабораторных ра-

- бот // Московский государственный институт электроники и математики. Сост. В.П. Чулков. – М., 2006. – 36 с.
10. Ли Т.Т., Адамс Т.Э., Гейнз У.М. Управление процессами с помощью ЭВМ. Моделирование и оптимизация. – М.: Советское радио, 1972.
  11. Фаронов В.В. ТУРБО Паскаль 7.0. Начальный курс: Учебное пособие. – М.: Норидж, 1997. – 616 с.
  12. Чулков В.П. САПР оборудования и технологии производства СБИС: Учебное пособие // Московский государственный институт электроники и математики. – М., 2003. – 180 с.
  13. В.П. Чулков. Моделирование технологических операций производства СБИС: Учебное пособие // Московский государственный институт электроники и математики. – М., 1994. – 73 с.
  14. Вальков В.М., Вершин В.Е. Автоматизированные системы управления технологическими процессами. – М.: Машиностроение, 1977.
  15. Норенков И.П. Введение в автоматизированное проектирование технических устройств и систем. – М.: Высшая школа, 1980. – 311 с.
  16. Чулков В.П. Системы машинного проектирования технологических процессов производства БИС: Учеб. пособие. – М.: МИЭМ, 1990. – 77 с.
  17. Практикум по высшей математике для студентов экономических специальностей: Методические указания к самостоятельной работе студентов / Под ред. Сагитова Р.В. // РЭА им. Г.В. Плеханова. – М.: Издательство «Менеджер», 2008. – 208 с.
  18. Клименко Ю.И. Высшая математика для экономистов: теория, примеры, задачи: Учебник для вузов. – М.: Издательство «Экзамен», 2005. – 736 с. (Серия «Учебник для вузов»).
  19. Пискунов Н.С. Дифференциальное и интегральное исчисления. Т.2. – М.: Наука, 1976. – 576 с.
  20. Кравченко В.А., Цидилин С.М., Федосеева Т.Л. Алгоритмы решения задач многокритериальной оптимизации: Учебное пособие. – М.:

МИЭМ, 1988. – 74 с.

21. Рапопорт Э.Я. Анализ и синтез систем автоматического управления с распределенными параметрами: Учеб. пособие. – М.: Высш. шк., 2005. – 292 с.
22. Оптимальное управление. Сборник. – М.: «Знание», 1978. – 144 с. (Нар. ун-т. Естественнонаучный фак. Издается с 1961 г.).
23. Семенкин Е.С., Семенкина О.Э., Коробейников С.П. Оптимизация технических систем: Учебное пособие. – Красноярск: СИБУП, 1996. – 284 с.
24. Выявление экспертных знаний (процедуры и реализации) / О.И. Ларичев, А.И. Мечитов, Е.М. Мошкович, Е.М. Фуремс. – М.: Наука, 1989. – 128 с.
25. Ивашов Е.Н., Пак М.М. Туннельный метод измерения нанорельефа поверхности. Деп. рук. ВИНТИ № 1201 В. 2004: 13.07.04. – 10 с.
26. Быков В.П. Методическое обеспечение САПР в машиностроении. – Л.: Машиностроение. Ленингр. отд-ние, 1989. – 255 с.
27. Норенков И.П., Маничев В.Б. Основы теории и проектирования САПР: Учеб. для втузов по спец. «Вычислительные машины, комплексы, системы и сети». – М.: Высш. шк., 1990. – 335 с.
28. Лускинович П.Н. и др. Нанотехнологические процессы и установки // Сб. докладов «Математическое моделирование нанотехнологических процессов и наноструктур»: Труды научных семинаров. – М.: МИФИ, 2001.
29. Норенков И.П. Введение в автоматизированное проектирование технических устройств и систем: Учеб. пособие для втузов. – 2-е изд., перераб. и доп. – М.: Высшая. школа, 1986. – 304 с.
30. Савостьянов В.П., Филатова Г.А., Филатов В.В. Расчёт и конструирование деталей аппаратуры САУ: Учеб. пособие. – М.: Машиностроение, 1982. – 328 с.
31. Ивашов Е.Н., Реутова М.В. Технологические устройства для получения наноструктур с использованием углеродных нанотрубок // Сборник докладов научно-технической конференции «Вакуумная наука и техника». – М.: МИЭМ, 2003.

32. Ивашов Е.Н., Пак М.М. Оптоволоконная нанотехнология в электронике и методы ее реализации. Деп. рук. ВИНТИ № 1203 В. 2004: 13.07.04. – 11 с.
33. Черняев В.Н. Технология производства интегральных микросхем и микропроцессоров: Учебник для вузов. – 2-е изд., перераб. и доп. – М.: Радио и связь, 1987. – 464 с.
34. Основы системотехники: Учеб. пособие для вузов / Г.П. Беляков, В.А. Сарычев, В.А. Сорокин, В.О. Чернышев. Под ред. В.О. Чернышева. – Томск: МГП «РАСКО», 1992. – 312 с.
35. Виброзащита радиоэлектронной аппаратуры полимерными копчаундами // Ю.В. Зеленева, А.А. Кирилин, Э.Б. Слободник, Е.Н. Талицкий; Под ред. Ю.В. Зеленева. – М.: Радио и связь, 1984. – 120 с.
36. Неволин В.К. Электронные устройства с элементами нанометровых размеров // Электронная промышленность. – М., 2002.
37. Слепов Н.Н. Современные технологии цифровых оптоволоконных сетей связи. – М.: Радио и связь, 2000. – 468 с.
38. Чуличков А.И. Математические модели нелинейной динамики. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 296 с.
39. Советов Б.Я., Яковлев С.А. Моделирование систем. Практикум: Учеб. пособие для вузов. – 3-е изд., стер. – М.: Высш. шк., 2005. – 295 с.
40. Еленин Г.Г. Нанотехнологии и вычислительная математика // Сб. докладов «Математическое моделирование нанотехнологических процессов и наноструктур»: Труды научных семинаров. – М.: МИФИ, 2001.
41. Солодовников И.В., Зародов А.Ф. Использование методов имитационного моделирования в исследовании экологических систем // Труды семинара «Новые информационные технологии». – М.: МИЭМ, 2001.
42. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных. Докл. АН СССР, 1956. Т. 108. – С. 179-182.
43. Киндлер Е. Языки моделирования. – М.: Энергоатомиздат, 1985. – 288 с.

44. Ивашов Е.Н., Кравчук И.С. Математическая модель оценки обобщённого критерия качества производимых наноструктур: Труды IX Международной конференции «Опто-, наноэлектроника, нанотехнологии и микросистемы». – Ульяновск: УлГУ, 2007.
45. Ивашов Е.Н., Кравчук И.С. Математическое моделирование в области управления качеством производства наноматериалов: Труды Всероссийской конференции с Международным интернет-участием «От наноструктур, наноматериалов и нанотехнологий к наноиндустрии». – Ижевск: ИжГТУ, 2007.
46. Ивашов Е.Н., Кравчук И.С. Математическая модель оценки обобщённого критерия качества для альтернативных производственных процессов // Химия твёрдого тела и современные микро- и нанотехнологии; VII Международная конференция. Кисловодск – Ставрополь: СевКавГТУ, 2007.
47. Ивашов Е.Н., Кравчук И.С. Математическая модель и алгоритм выбора оптимального процесса производства углеродных наноструктур / INTERMATIC-2007 // Материалы Международной научно-технической конференции «Фундаментальные проблемы радиоэлектронного приборостроения». – М.: МИРЭА, 2007, часть 3. – С. 162-165.
48. 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.
49. Рапопорт Э.Я. Анализ и синтез систем автоматического управления с распределенными параметрами: Учеб. пособие. – М.: Высш. шк., 2005. – 292 с.
50. Оптимальное управление. Сборник. – М.: «Знание», 1978. – 144 с. (Нар. ун-т. Естественнонаучный фак. Издается с 1961 г.).
51. Семенкин Е.С., Семенкина О.Э., Коробейников С.П. Оптимизация тех-

- нических систем: Учебное пособие. – Красноярск: СИБУП, 1996. – 284 с.
52. Ивашов Е.Н., Реутова М.В. Применение метода Саати при структурировании множества альтернатив получения углеродных нанотрубок, фуллеренов и кластеров. Деп. рукопись ВИНТИ № 2325 В. 2003: 31.12.2003. – 12 с.
53. Саати Т.Л. Элементы теории массового обслуживания и её приложения. – М.: Советское радио, 1977.
54. Беленов В.К., Блинникова-Вяземская Е.В., Иоаннисянц Т.А. Количественный анализ технологического процесса // Микроэлектроника, 1975, вып. 8.
55. Саати Т. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.
56. Клыков Ю.И., Горьков Л.Н. Банки данных для принятия решений. – М.: Советское радио, 1980. – 208 с.
57. Зародов А.Ф. Представление иерархии моделей при проектировании систем // Аэрокосмические технологии. – М.: МГТУ им Н.Э. Баумана, 2003.
58. Корячко В.П. и др. Теоретические основы САПР: Учебник для вузов // В.П. Корячко, В.М. Курейчик, И.П. Норенков. – М.: Энергоатомиздат, 1987. – 400 с.
59. Алексеев О.В., Головков А.А., Пивоваров И.Ю. и др. Автоматизация проектирования радиоэлектронных средств. – М.: Высш. шк., 2000. – 479 с.
60. Смирнов С.А. Оценка интеллектуальной собственности. – М.: Финансы и статистика, 2002. – 352с.
61. Фаронов В.В. Турбо Паскаль 7.0. Начальный курс: Учеб. пособие. – М.: КНОРУС, 2006. – 576 с.
62. Смирнов С.А. Оценка интеллектуальной собственности. – М.: Финансы и статистика, 2002. – 352с.

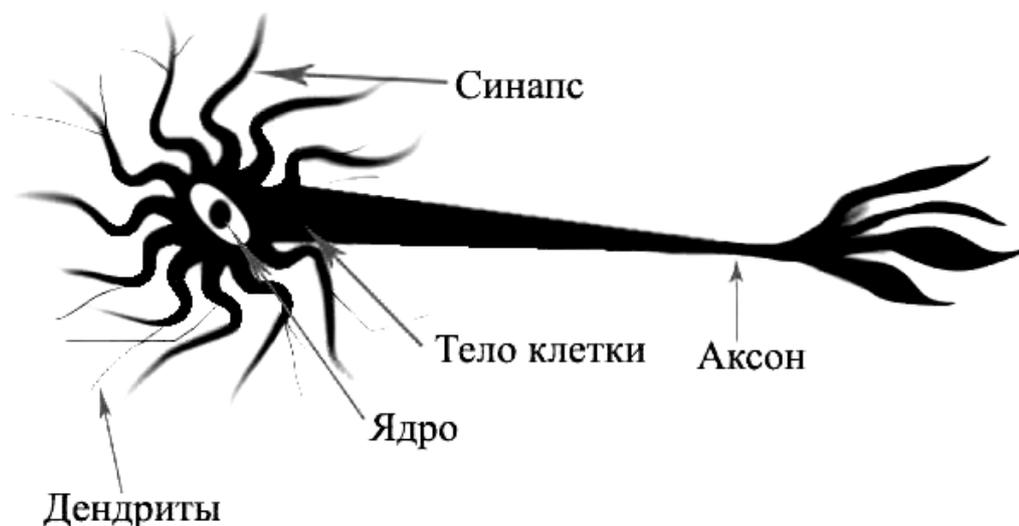
# ГЛАВА 5. ПРОЕКТИРОВАНИЕ ЭЛЕМЕНТОВ НЕЙРОННЫХ СЕТЕЙ

## 5.1. Особенности автоматизированного проектирования искусственных нейронных сетей

Перцептрон, перцептрон (англ. perceptron, нем. Perzeptron, от лат. perceptio – понимание, познавание, восприятие), математическая модель процесса восприятия.

Исследования в области нейронных сетей (НС) пережили три периода активизации. Первый пик в 40-х годах обусловлен пионерской работой МакКаллока и Питтса. Вторым возник в 60-х – благодаря теореме сходимости перцептрона Розенблатта и работе Минского и Пейперта, указавшей ограниченные возможности простейшего перцептрона. Результаты Минского и Пейперта погасили энтузиазм большинства исследователей, особенно тех, кто работал в области вычислительных наук. Возникшее в исследованиях по нейронным сетям затишье продлилось почти 20 лет. С начала 80-х годов НС вновь привлекли интерес исследователей, что связано с энергетическим подходом Хопфилда и алгоритмом обратного распространения для обучения многослойного перцептрона (многослойные сети прямого распространения), впервые предложенного Вербесом и независимо разработанного рядом других авторов. Алгоритм получил известность благодаря Румельхарту в 1986 году. Андерсон и Розенфельд подготовили подробную историческую справку о развитии НС [1÷7].

Нейрон (нервная клетка) является особой биологической клеткой, которая обрабатывает информацию (рис. 5.1.1). Она состоит из тела клетки (cell body), или сомы (soma), и двух типов внешних древоподобных ветвей:



**Рис. 5.1.1.** Схема биологического нейрона

аксона (axon) и дендритов (dendrites). Тело клетки включает ядро (nucleus), которое содержит информацию о наследственных свойствах, и плазму, обладающую молекулярными средствами необходимых нейрону материалов. Нейрон получает сигналы (импульсы) от других нейронов через дендриты (приемники) и передает сигналы, сгенерированные телом клетки, вдоль аксона (передатчик), который в конце разветвляется на волокна (strands). На окончаниях этих волокон находятся синапсы (synapses).

Синапс является элементарной структурой и функциональным узлом между двумя нейронами (волокно аксона одного нейрона и дендрит другого). Когда импульс достигает синаптического окончания, высвобождаются определенные химические вещества, которые называются нейротрансмиттерами.

Нейротрансмиттеры диффундируют через синаптическую щель, возбуждая или затормаживая, в зависимости от типа синапса, способность нейрона-приёмника генерировать электрические импульсы. Результативность синапса может настраиваться проходящими через него сигналами,

так что синапсы могут обучаться в зависимости от активности процессов, в которых они участвуют. Эта зависимость от предыстории действует на память, которая, возможно, ответственна за память человека.

Кора головного мозга человека является протяжённой, образованной нейронами поверхностью толщиной от 2 до 3 мм с площадью около 2200 см<sup>2</sup>, что вдвое превышает площадь поверхности стандартной клавиатуры. Кора головного мозга содержит около 10<sup>11</sup> нейронов, что приблизительно равно числу звезд Млечного пути. Каждый нейрон связан с 10<sup>3</sup> ÷ 10<sup>4</sup> другими нейронами. В целом мозг человека содержит от 10<sup>14</sup> до 10<sup>15</sup> взаимосвязей [8÷11].

Нейроны взаимодействуют посредством короткой серии импульсов, как правило, продолжительностью несколько мс. Сообщение передается посредством частотно-импульсной модуляции. Частота может изменяться от нескольких единиц до сотен Гц, что в миллион раз медленнее, чем самые быстродействующие переключательные электронные схемы. Тем не менее сложные решения по восприятию информации, как например, распознавание лица, человек принимает за несколько сотен мс. Эти решения контролируются сетью нейронов, которые имеют скорость выполнения операций всего несколько мс. Это означает, что вычисления требуют не более 100 последовательных стадий. Другими словами, для таких сложных задач мозг «запускает» параллельные программы, содержащие около 100 шагов. Это известно как правило 100 шагов [12]. Рассуждая аналогичным образом, можно обнаружить, что количество информации, посылаемое от одного нейрона другому, должно быть очень маленьким (несколько бит). Отсюда следует, что основная информация не передается непосредственно, а захватывается и распределяется в связях между нейронами. Этим объясняется такое название, как коннекционистская модель, применяемое к НС.

## ***От биологических сетей к нейронным***

Современные цифровые вычислительные машины превосходят человека по способности производить числовые и символьные вычисления. В свою очередь, НС могут без усилий решать сложные задачи восприятия внешних данных (например, узнавание человека в толпе только по его промелькнувшему лицу) с такой скоростью и точностью, что мощнейший в мире компьютер по сравнению с ними кажется безнадёжным тугодумом. В чем причина столь значительного различия в их производительности? Архитектура биологической нейронной системы совершенно не похожа на архитектуру машины фон Неймана (табл. 5.1.1) и существенно влияет на типы функций, которые более эффективно исполняются каждой из существующих моделей.

Подобно биологической нейронной системе НС является вычислительной системой с огромным числом параллельно функционирующих простых процессоров с множеством связей. Модели НС в некоторой степени воспроизводят «организованные» принципы, свойственные мозгу человека. Моделирование биологической нейронной системы с использованием НС также может способствовать лучшему пониманию биологических функций. Такие технологии производства, как VLSI (сверхвысокий уровень интеграции) и оптические аппаратные средства, делают возможным подобное моделирование [13].

Глубокое изучение НС требует знания нейрофизиологии, науки о познании, психологии, физики (статической механики), теории управления, теории вычислений, проблем искусственного интеллекта, статистики/математики, распознавания образов, компьютерного зрения, параллельных вычислений и аппаратных средств (цифровых / аналоговых / VLSI / оптических). С другой стороны, НС также стимулируют эти дисциплины, обеспечивая их новыми инструментами и представлениями. Этот симбиоз

**Таблица 5.1.1**

Машина фон Неймана по сравнению с  
биологической нейронной системой

	<b>Машина фон Неймана</b>	<b>Биологическая нейронная система</b>
<b>Процессор</b>	Сложный	Простой
	Высокоскоростной	Низкоскоростной
	Один или несколько	Большое количество
<b>Память</b>	Отделена от процессора	Интегрирована в процессор
	Локализована	Распределенная
	Адресация не по содержанию	Адресация по содержанию
<b>Вычисления</b>	Централизованные	Распределенные
	Последовательные	Параллельные
	Хранимые программы	Самообучение
<b>Надежность</b>	Высокая уязвимость	Живучесть
<b>Специализация</b>	Численные и символные операции	Проблемы восприятия
<b>Среда функционирования</b>	Строго определенная	Плохо определенная
	Строго ограниченная	Без ограничений

жизненно необходим для проведения исследований по нейронным сетям.

Представим некоторые проблемы, решаемые в контексте НС и представляющие интерес для учённых и инженеров.

**Классификация образов.** Задача состоит в указании принадлежности входного образа, представленного вектором признаков, одному или нескольким предварительно определённым классом.

**Кластеризация/категоризация.** При решении задачи кластеризации, которая известна также как классификация образов «без учителя», отсутствует обучающая выборка с метками классов. Алгоритм кластеризации основан на подобии образов и размещает близкие образы в один кластер. Известны случаи применения кластеризации для извлечения знаний, сжатия данных и исследования свойств данных.

**Аппроксимация функций.** Предположим что имеется обучающая выборка  $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$  (пары данных вход-выход), которая генерируется неизвестной функцией  $(x)$ , искаженной шумом. Задача аппроксимации состоит в нахождении оценки известной функции  $(x)$ . Аппроксимация функций необходима при решении многочисленных инженерных и научных задач моделирования.

**Предсказание/прогноз.** Пусть заданы  $n$  дискретных отсчетов  $\{y(t_1), y(t_2), \dots, y(t_n)\}$  в последовательные моменты времени  $t_1, t_2, \dots, t_n$ . Задача состоит в предсказании значения  $y(t_{n+1})$  в некоторый будущий момент времени  $t_{n+1}$ . Предсказание/прогноз имеют значительное влияние на принятие решений в бизнесе, науке и технике. Предсказание цен на фондовой бирже и прогноз погоды являются типичными приложениями техники предсказания/прогноза.

**Оптимизация.** Многочисленные проблемы в математике, статике, технике, науке, медицине и экономики могут рассматриваться как проблемы оптимизации. Задачей алгоритма оптимизации является нахождение такого решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию. Задача коммивояжера, относящаяся к классу NP-полных, является классическим примером задачи оптимизации.

**Память адресуемая по содержанию.** В модели вычислений фон Неймана обращение к памяти доступно только посредством адреса, который не зависит от содержания памяти. Более того, если допущена ошибка в вычислении адреса, то может быть найдена совершенно другая информация. Ассоциативная память, или память, адресуемая по содержанию, доступна по указанию заданного содержания. Содержимое памяти может быть вызвано даже по частичному входу или искаженному содержанию. Ассоциативная память желательна при создании мультимедийных баз данных.

**Управление.** Рассмотрим динамическую систему, заданную совокупностью  $\{u(t), y(t)\}$ , где  $u(t)$  является входным управляющим воздействием, а  $y(t)$  – выходом системы в момент времени  $t$ . В системах управления с эталонной моделью целью управления является расчёт такого входного воздействия  $u(t)$ , при котором система следует по желаемой траектории, диктуемой эталонной моделью. Примером является оптимальное управление двигателем [9].

## **5.2. Модель нейрона**

МакКаллок и Питтс 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.

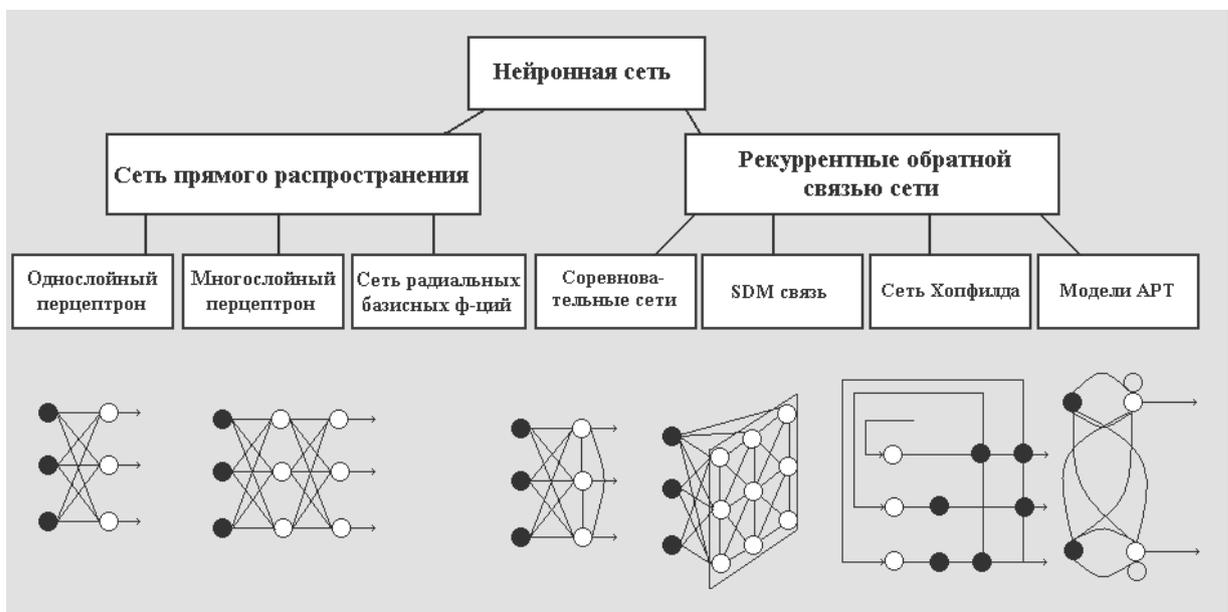
Часто удобно рассматривать  $u$  как весовой коэффициент, связанный с постоянным входом  $x_0 = 1$ . Положительные веса соответствуют возбуждающим связям, а отрицательные – тормозным. МакКаллок и Питтс доказали, что при соответствующим образом подобранных весах, совокупность параллельно функционирующих нейронов подобного типа способна выполнять универсальные вычисления. Здесь наблюдается определенная

аналогия с биологическим нейроном: передачу сигнала и взаимосвязи имитируют аксоны и дендриты, веса связей соответствуют синапсам, а пороговая функция отражает активность сомы [1÷5].

### **Архитектура нейронной сети**

НС может рассматриваться как направленный граф со взвешенными связями, в котором искусственные нейроны являются узлами. По архитектуре связей НС могут быть сгруппированы в два класса: сети прямого распространения, в которых графы не имеют петель, и рекуррентные сети, или сети с обратными связями [1÷10].

В наиболее распространенном семействе сетей первого класса, называемых многослойным персептроном, нейроны расположены слоями и имеют однонаправленные связи между слоями. На рис. 5.2.1 представлены типовые сети каждого класса. Сети прямого распространения являются



**Рис. 5.2.1** Систематизация архитектур сетей прямого распространения и рекуррентных (с обратной связью)

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 НС 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 определённой 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: ёмкость, сложность образцов и вычислительная сложность. Под ёмкостью 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, подчинённых тому же статистическому распределению. Известны четыре основных типа правил обучения: коррекция по ошибке, машина Больцмана, правило Хебба и обучение методом соревнования.

*Правило коррекции по ошибке.* При обучении с учителем для каждого входного примера задан желаемый выход  $d$ . Реальный выход сети  $y$  может не совпадать с желаемым. Принцип коррекции по ошибке при обучении состоит в использовании сигнала  $(d - y)$  для модификации весов, обеспечивающей постепенное уменьшение ошибки. Обучение имеет место только в случае, когда персептрон ошибается. Известны различные модификации этого алгоритма обучения [2].

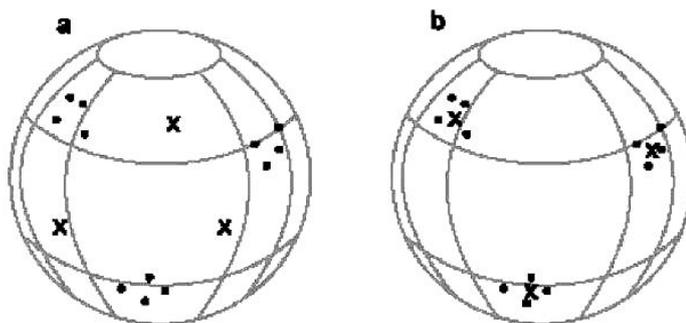
*Обучение Больцмана.* Представляет собой стохастическое правило обучения, которое следует из информационных теоретических и термодинамических принципов [10]. Целью обучения Больцмана является такая

настройка весовых коэффициентов, при которой состояния видимых нейронов удовлетворяют желаемому распределению вероятностей. Обучение Больцмана может рассматриваться как специальный случай коррекции по ошибке, в котором под ошибкой понимается расхождение корреляций состояний в двух режимах.

*Правило Хебба.* Самым старым обучающим правилом является постулат обучения Хебба [13]. Хебб опирался на следующие нейрофизиологические наблюдения: если нейроны с обеих сторон синапса активизируются одновременно и регулярно, то сила синаптической связи возрастает. Важной особенностью этого правила является то, что изменение синаптического веса зависит только от активности нейронов, которые связаны данным синапсом. Это существенно упрощает цепи обучения в реализации VLSI.

*Обучение методом соревнования.* В отличие от обучения Хебба, в котором множество выходных нейронов могут возбуждаться одновременно, при соревновательном обучении выходные нейроны соревнуются между собой за активизацию. Это явление известно как правило "победитель берет все". Подобное обучение имеет место в биологических нейронных сетях. Обучение посредством соревнования позволяет кластеризовать входные данные: подобные примеры группируются сетью в соответствии с корреляциями и представляются одним элементом.

При обучении модифицируются только веса "победившего" нейрона. Эффект этого правила достигается за счёт такого изменения сохраненного в сети образца (вектора весов связей победившего нейрона), при котором он становится чуть ближе ко входному примеру. На рис. 1.3 дана геометрическая иллюстрация обучения методом соревнования. Входные векторы нормализованы и представлены точками на поверхности сферы. Векторы весов для трёх нейронов инициализированы случайными значениями. Их



**Рис. 5.2.2.** Пример обучения методом соревнования:

а – перед обучением;

б – после обучения

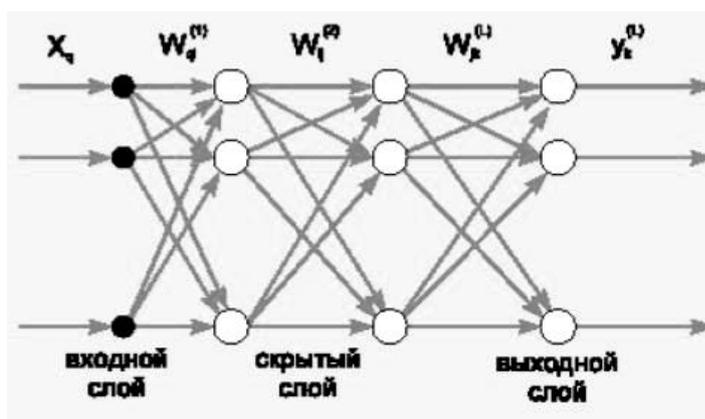
начальные и конечные значения после обучения отмечены X на рис. 1.3,а и 1.3,б соответственно. Каждая из трёх групп примеров обнаружена одним из выходных нейронов, чей весовой вектор настроился на центр тяжести обнаруженной группы.

Можно заметить, что сеть никогда не перестанет обучаться, если параметр скорости обучения не равен 0. Некоторый входной образец может активизировать другой выходной нейрон на последующих итерациях в процессе обучения. Это ставит вопрос об устойчивости обучающей системы. Система считается устойчивой, если ни один из примеров обучающей выборки не изменяет своей принадлежности к категории после конечного числа итераций обучающего процесса. Один из способов достижения стабильности состоит в постепенном уменьшении до 0 параметра скорости обучения. Однако это искусственное торможение обучения вызывает другую проблему, называемую пластичностью и связанную со способностью к адаптации к новым данным. Эти особенности обучения методом соревнования известны под названием дилеммы стабильности-пластичности Гроссберга.

В табл. 5.2.1 представлены различные алгоритмы обучения и связанные с ними архитектуры сетей (список не является исчерпывающим). В последней колонке перечислены задачи, для которых может быть применен каждый алгоритм. Каждый алгоритм обучения ориентирован на сеть определенной архитектуры и предназначен для ограниченного класса задач. Кроме рассмотренных, следует упомянуть некоторые другие алгоритмы: Adaline и Madaline, линейный дискриминантный анализ, проекции Саммона, анализ главных компонент [10÷12].

### 5.3. Многослойные сети прямого распространения

Стандартная  $L$ -слойная сеть прямого распространения состоит из слоя входных узлов (будем придерживаться утверждения, что он не включается в сеть в качестве самостоятельного слоя),  $(L-1)$  скрытых слоёв и выходного слоя, соединённых последовательно в прямом направлении и не содержащих связей между элементами внутри слоя и обратных связей между слоями. На рис. 5.3.1 приведена структура трёхслойной сети.



**Рис. 5.3.1** Типовая архитектура трехслойной сети прямого распространения

**Таблица 5.2.1**

**Известные алгоритмы обучения**

Парадигма	Обучающее правило	Архитектура	Алгоритм обучения	Задача	
С Учителем	Коррекция ошибки	Однослойный и многослойный персептрон	Алгоритмы обучения персептрона Обратное распространение Adaline и Madaline	Классификация образов Аппроксимация функций Предсказание, управление	
	Больцман	Рекуррентная	Алгоритм обучения Больцмана	Классификация образов	
	Хебб	Многослойная прямого распространения	Линейный дискриминантный анализ	Анализ данных Классификация образов	
	Соревнование	Соревнование	Соревнование	Векторное квантование	Категоризация внутри класса Сжатие данных
			Сеть ART	ARTMap	Классификация образов
Без учителя	Коррекция ошибки	Многослойная прямого распространения	Проекция Саммона	Категоризация внутри класса Анализ данных	
	Хебб	Прямого распространения или соревнование	Анализ главных компонентов	Анализ данных Сжатие данных	
		Сеть Хопфилда	Обучение ассоциативной памяти	Ассоциативная память	
Без учителя	Соревнование	Соревнование	Векторное квантование	Категоризация Сжатие данных	
		SOM Кохонена	SOM Кохонена	Категоризация Анализ данных	
		Сети ART	ART1, ART2	Категоризация	
Смешанная	Коррекция ошибки и соревнование	Сеть RBF	Алгоритм обучения RBF	Классификация образов Аппроксимация функций Предсказание, управление	

## ***Многослойный персептрон***

Наиболее популярный класс многослойных сетей прямого распространения образуют многослойные персептроны, в которых каждый вычислительный элемент использует пороговую или сигмоидальную функцию активации. Многослойный персептрон может формировать сколь угодно сложные границы принятия решения и реализовывать произвольные булевы функции [6]. Разработка алгоритма обратного распространения для определения весов в многослойном персептроне сделала эти сети наиболее популярными у исследователей и пользователей нейронных сетей.

Геометрическая интерпретация объясняет роль элементов скрытых слоёв (используется пороговая активационная функция) [13, 14].

## ***RBF-сети***

Сети, использующие радиальные базисные функции (RBF-сети), являются частным случаем двухслойной сети прямого распространения. Каждый элемент скрытого слоя использует в качестве активационной функции радиальную базисную функцию типа гауссовой. Радиальная базисная функция (функция ядра) центрируется в точке, которая определяется весовым вектором, связанным с нейроном. Как позиция, так и ширина функции ядра должны быть обучены по выборочным образцам. Обычно ядер гораздо меньше, чем создаётся обучающих примеров. Каждый выходной элемент вычисляет линейную комбинацию этих радиальных базисных функций. С точки зрения задачи аппроксимации скрытые элементы формируют совокупность функций, которые образуют базисную систему для представления входных примеров в построенном на ней информационном пространстве.

Существуют различные алгоритмы обучения RBF-сетей [3]. Основным алгоритмом используется двушаговую стратегию обучения, или смешанное обучение. Он оценивает позицию и ширину ядра с использованием algo-

ритма кластеризации "без учителя", а затем алгоритм минимизации среднеквадратической ошибки "с учителем" для определения весов связей между скрытым и выходным слоями. Поскольку выходные элементы линейны, применяется неитерационный алгоритм. После получения этого начального приближения используется градиентный спуск для уточнения параметров сети.

Этот смешанный алгоритм обучения RBF-сети сходится гораздо быстрее, чем алгоритм обратного распространения для обучения многослойных персептронов. Однако RBF-сеть часто содержит слишком большое число скрытых элементов. Это влечёт более медленное функционирование RBF-сети, чем многослойного персептрона. Эффективность (ошибка в зависимости от размера сети) RBF-сети и многослойного персептрона зависят от решаемой задачи [13,17,45].

### ***Нерешённые проблемы***

Существует множество спорных вопросов при проектировании сетей прямого распространения – например, сколько слоёв необходимы для данной задачи, сколько следует выбрать элементов в каждом слое, как сеть будет реагировать на данные, не включенные в обучающую выборку (какова способность сети к обобщению), и какой размер обучающей выборки необходим для достижения "хорошей" способности сети к обобщению.

Хотя многослойные сети прямого распространения широко применяются для классификации и аппроксимации функций [2], многие параметры ещё должны быть определены путём проб и ошибок. Существующие теоретические результаты дают лишь слабые ориентиры для выбора этих параметров в практических приложениях.

## **Самоорганизующиеся карты Кохонена**

Самоорганизующиеся карты Кохонена [16] обладают благоприятным свойством сохранения топологии, которое воспроизводит важный аспект карт признаков в коре головного мозга высокоорганизованных животных. В отображении с сохранением топологии близкие входные примеры возбуждают близкие выходные элементы. По существу, основная архитектура сети Кохонена представляет собой двумерный массив элементов, причём каждый элемент связан со всеми  $n$  входными узлами.

Такая сеть является специальным случаем сети, обучающейся методом соревнования, в которой определяется пространственная окрестность для каждого выходного элемента. Локальная окрестность может быть квадратом, прямоугольником или окружностью. Начальный размер окрестности часто устанавливается в пределах от  $1/2$  до  $2/3$  размера сети и сокращается согласно определённому закону (например, по экспоненциально убывающей зависимости). Во время обучения модифицируются все веса, связанные с победителем и его соседними элементами.

Самоорганизующиеся карты Кохонена могут быть использованы для проектирования многомерных данных, аппроксимации плотности и кластеризации. Эта сеть успешно применялась для распознавания речи, обработки изображений, в робототехнике и в задачах управления [2]. Параметры сети включают в себя размерность массива нейронов, число нейронов в каждом измерении, форму окрестности, закон сжатия окрестности и скорость обучения.

## **Модели теории адаптивного резонанса**

Дилемма стабильности-пластичности является важной особенностью обучения методом соревнования. Как обучать новым явлениям (пластичность) и в то же время сохранить стабильность, чтобы существующие знания не были стерты или разрушены?

Карпентер и Гроссберг, разработавшие модели теории адаптивного резонанса (ART1, ART2 и ARTMAP) [17], сделали попытку решить эту дилемму. Сеть имеет достаточное число выходных элементов, но они не используются до тех пор, пока не возникнет в этом необходимость. Будем говорить, что элемент распределен (не распределен), если он используется (не используется). Обучающий алгоритм корректирует имеющийся прототип категории, только если входной вектор в достаточной степени ему подобен.

В этом случае они резонируют. Степень подобия контролируется параметром сходства  $k$ ,  $0 < k < 1$ , который связан также с числом категорий. Когда входной вектор недостаточно подобен ни одному существующему прототипу сети, создается новая категория, и с ней связывается нераспределённый элемент со входным вектором в качестве начального значения прототипа. Если не находится нераспределённого элемента, то новый вектор не вызывает реакции сети.

Чтобы проиллюстрировать модель, рассмотрим сеть ART1, которая рассчитана на бинарный (0/1) вход. Упрощённая схема архитектуры ART1 [2] представлена на рис. 5.3.2. Она содержит два слоя элементов с полными связями.

Направленный сверху вниз весовой вектор  $W_j$  соответствует элементу  $j$  входного слоя, а направленный снизу вверх весовой вектор  $i$  связан с выходным элементом  $i$ .  $i$  является нормализованной версией  $W_i$ . Векторы  $W$ , сохраняют прототипы кластеров. Роль нормализации состоит в том, чтобы предотвратить доминирование векторов с большой длиной над векторами с малой длиной. Сигнал сброса  $R$  генерируется только тогда, когда подобие ниже заданного уровня.

Модель ART1 может создать новые категории и отбросить входные примеры, когда сеть исчерпала свою ёмкость [1÷5,35,36,41,42,45]. Однако число обнаруженных сетью категорий чувствительно к параметру сходства.

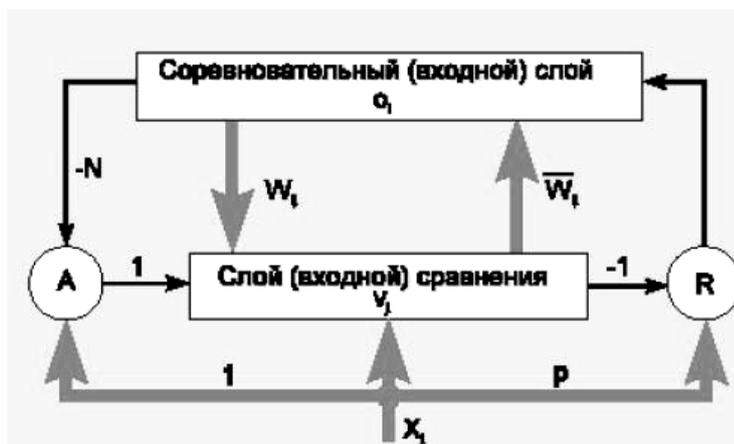


Рис.5.3.2. Сеть ART1

## 5.4. Твёрдые объекты в нейронных сетях

Твёрдые объекты входят во многие наноструктуры в виде основных единиц, формирующих твёрдое тело. Процессы ведущие к образованию таких нанокластеров, весьма распространены в природе, например, кристаллизация из раствора или расплава, спекания, различного рода мартенситные превращения, кристаллизация из аморфных систем, образование магнитных и сегнетоэлектрических доменов, спинодальный распад. Все эти процессы подчиняются законам термодинамики и сопровождаются явлением упорядочения и самоорганизации [49].

Образование и организация кластеров в твёрдотельную наносистему во многом определяются способами их получения. При этом формирование наноструктуры возможно из отдельных кластеров, или путём наноструктурирования массивного твёрдого тела. Все эти способы уже имеют большое значение для создания наноматериалов на основе металлов, сплавов, оксидов, керамик и т.д. кроме формирования наноструктур, важным аспектом является их структурные механические и тепловые свойства,

определяющие качество и назначение многих материалов.

Молекулярные кластеры металлов – это многоядерные комплексные соединения, в основе молекулярной структуры которых находится окруженный лигандами остов из атомов металлов. Кластером считается ядро, включающее более двух атомов. Металлический остов представляет собой цепи различной длины, разветвленные циклы, полиэдры и их комбинации.

Молекулярные лигандные кластеры металлов образуются из металлокомплексных соединений в результате проведения химических реакций в растворе. Наибольшее распространение среди методов синтеза больших кластеров получили методы конденсации многоатомных кластеров и восстановление комплексов металлов. В качестве стабилизирующих лигандов используются органические фосфины, особенно  $PPh_3$ , или фенантролины. Таким путём были синтезированы кластеры палладия, обладающие икосаэдрическим ядром и кластерные анионы молибдена. Безлигандные получают в основном тремя основными способами: с помощью сверхзвукового сопла, с помощью газовой агрегации и с помощью испарения с поверхности твёрдого тела или жидкости. Однако от момента получения кластера до момента его фиксации, когда, так сказать, его можно подержать в руках, путь гораздо более длинный, чем для молекулярных кластеров, синтезированных из раствора. Кластеры генерируются с помощью звукового сопла, проходят через диафрагму, ионизируются с помощью электронных или фоновых столкновений, разделяются по массам (по отношению  $m/e$  на масс-спектрометре) и регистрируются детектором. Такая схема уже даёт основные элементы получения кластеров: это источники кластеров, масс-спектрометры и детекторы [48÷49].

Весьма удобной, можно сказать близкой к модельной, оказалась реакция термического разложения оксалата железа –  $Fe_2(C_2O_4)_3 \cdot 5H_2O$ . При температуре  $T_d = 200$  °C происходит дегидратация и разложение оксалата железа с выделением CO и CO<sub>2</sub>, формируется та активная среда, в которой

начинается нуклеация и образуются нанокластеры оксида железа. Второй минимум температуры  $T_d = 260$  °С связан с дальнейшим выделением СО и СО<sub>2</sub>, началом спекания и образованием наноструктуры, включающей нанокластеры оксида железа. Размеры кластеров увеличиваются с 1 до 6÷7 нм с увеличением температуры разложения и времени выдерживания при данной температуре (увеличение времени способствует гомогенизации кластеров по размерам) [40,41,49].

Термическое разложение оксалатов, цитратов и формиатов железа, кобальта, никеля, меди при температуре 200÷260 °С в вакууме или инертной атмосфере приводит к получению кластеров металлов с размерами 100÷300 нм. Нанокластеры карбидов и нитридов кремния можно синтезировать с помощью высокотемпературного пиролиза при 1300 °С полисилазанов, поликарбосиланов и поликарбосилаксанов.

Нанокластеры боридов переходных металлов получают пиролизом борогидридов при более низких температурах 300÷400 °С, иногда с помощью лазерного воздействия [42÷49].

## **5.5. Схема работы нейрона**

Рассмотрим классический вариант эксперимента Юнга, в котором имеется лишь один барьер между источником частиц и детектором. Допустим, что можно помещать (или не помещать) в пространство до и/или после барьера пластинку с показателем преломления  $n$ . Таким образом, оказывается возможным сформировать четыре варианта входа в систему: (1,1), (1,  $n$ ), ( $n$ ,1), ( $n$ , $n$ ). Значения показателя преломления 1 (пластинка отсутствует) и  $n$  (пластинка присутствует) могут быть использованы для представления бинарных значений входа (например, 1 и 0) [13÷24].

Попытаемся подобрать параметры системы таким образом, чтобы ее выход давал значения функции XOR, а именно

$x_1$	$x_2$	$y$
1	1	0
1	$n$	1
$n$	1	1
$n$	$n$	0

Выберем в качестве значений показателя преломления  $n = 5/3$ , положим  $h_0 = h_1 = h$ ,  $r^{(2)}(D) = 0$ , и предположим, что  $\lambda \gg h$  [14].

Выберем также положения двух щелей таким образом, чтобы длины обоих звеньев траектории частицы в мире 1  $l_1^1, l_1^2$  превосходили длины ее звеньев в мире 2  $l_2^1, l_2^2$  на  $\frac{3}{4}\lambda$  (рис. 5.5.1). Этого можно добиться, полагая [16,18÷20]

$$r_1 = r_2 + \sqrt{1 + \frac{h^2}{r^2}} \cdot \frac{3}{4}\lambda . \quad (5.5.1)$$

Для таких значений параметров системы мы получаем:

$$\langle D|s_1^1 \rangle \langle s_1^1|S \rangle = \frac{1}{l_1^1 l_1^2} \exp\left(\frac{2\pi i}{\lambda} n_1 l_1^1 + n_2 l_1^2\right) ; \quad (5.5.2)$$

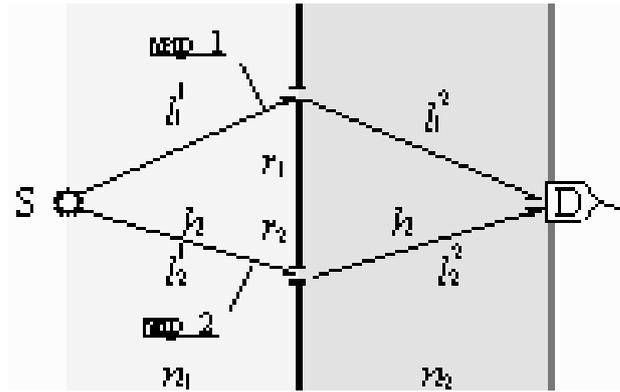
$$\langle D|s_2^1 \rangle \langle s_2^1|S \rangle = \frac{1}{l_2^1 l_2^2} \exp\left(\frac{2\pi i}{\lambda} n_1 l_2^1 + n_2 l_2^2\right) . \quad (5.5.3)$$

Учитывая, что  $\lambda \gg h$  можно приближенно положить, что вероятность обнаружения частицы в детекторе, которая определяется интерференцией выходов нейронов из разных миров, составляет

$$P_{n_1, n_2} = \left| \langle D|s_1^1 \rangle \langle s_1^1|S \rangle + \langle D|s_2^1 \rangle \langle s_2^1|S \rangle \right|^2 \quad (5.5.4)$$

$$\cong \frac{1}{r^2 + h^2} \left| 1 + \cos \frac{2\pi}{\lambda} (n_1 + n_2) \Delta l \right|^2 ,$$

где  $\Delta l = l_1^1 + l_1^2 - l_2^1 - l_2^2 = \frac{3}{2}\lambda . \quad (5.5.5)$



**Рис. 5.5.1.** Экспериментальная реализация функции XOR с помощью интерферирующих нейронов в двух мирах

Тогда следует, что [14]

$$\begin{aligned}
 P_{1,1} &= A|1 + \cos 3\pi|^2 = 0 ; \\
 P_{1,n} &= A|1 + \cos 4\pi|^2 = 4A ; \\
 P_{n,1} &= A|1 + \cos 4\pi|^2 = 4A ; \\
 P_{n,n} &= A|1 + \cos 5\pi|^2 = 0 .
 \end{aligned}
 \tag{5.5.6}$$

Таким образом, выход системы совпадает со значением функции XOR с точностью до нормировочного фактора  $4A = 4(r^2 + h^2)^{-2}$  [15÷17,32].

Заметим, что *единственный квантовый нейрон* оказался способным выполнить функцию XOR, реализация которой на классических нейронах требует построения двухслойной *сети нейронов*.

### Обобщённая модель

Можно обобщить описанную выше схему квантового нейрона. В фейнмановском представлении, эволюция любой квантовой системы может быть определена если мы сможем определить все пути, по которым она может достичь данного состояния. Каждый из этих путей определяет

отдельный мир. Далее необходимо вычислить комплексную амплитуду соответствующего перехода из начального  $r_0$  в конечное положение  $r$  в каждом из этих миров [12÷18]

$$\varphi \propto \exp\left(\frac{i}{\hbar} S_{r_0, r}\right). \quad (5.5.7)$$

Здесь значение действия

$$S_{r_0, r} = \int_{t_0}^t L(r, \dot{r}, t) dt = \int_{t_0}^t T(\dot{r}, t) - U(r) dt, \quad (5.5.8)$$

в котором  $L(r, \dot{r}, t)$  – функция Лагранжа,

$T(\dot{r}, t)$  – кинетическая энергия,

$U(r)$  – потенциальная энергия,

является аддитивным и может быть рассмотрено как аналог аддитивной активации квантового нейрона.

Хотя такое рассмотрение в некотором смысле аналогично данному Э. Берман и др. [21]. Оно также ясно выявляет нейроноподобное нелинейное преобразование активации, даваемое выражением (5.5.7) и отличается тем, что не квантовое состояние определяет вход нейрона, а распределение потенциала  $U(r)$ . Выходом же нейрона является комплекснозначная амплитуда  $\varphi$ . А соотнесённая с решением задачи регрессии вероятность детектирования частицы является результатом интерференции комплексных выходов нейронов (интерференции миров).

Главное отличие предлагаемого рассмотрения от ранее предложенных подходов к построению квантовых нейронных систем заключается в том, что оно по сути указывает на необязательность использования квантовых нейронных сетей как таковых и, возможно, на достаточность создания единственного квантового нейрона для решения достаточно общих проблем нейрокомпьютинга [15÷19].

Фактически, структура квантовой множественной Вселенной (по Эверетту) представляет в наше распоряжение множество нейронов, существующих в различных мирах, которые, кооперируя друг с другом, и дают требуемый ответ в результате интерференции этих миров [17].

Важно подчеркнуть, что понимание предложенного здесь подхода возможно лишь в рамках Эвереттовской интерпретацией квантовой механики. Это также согласуется с мнением Д. Дойча [18], согласно которому такая интерпретация существенна для понимания квантовых вычислений вообще. Аналогичную точку зрения поддерживают А. Нараянан и Т. Меннеер [22].

Другое важное свойство предложенной системы состоит в отсутствии необходимости решать очень сложную проблему удержания когерентности квантового состояния, которая является ключевой для реализации алгоритмических (не нейронных) квантовых вычислений. Причиной этого является то, что квантовый нейрон обрабатывает аналоговые входы. В действительности, классический нейрокомпьютинг имеет дело, в основном, с обработкой *аналоговых* сигналов. Обработка образов, кодируемых векторами с дискретными (в частности, бинарными) компонентами, конечно, возможна, но не она составляет главное поле приложения искусственных нейронных сетей. С другой стороны фон Неймановские компьютеры в общем случае являются *цифровыми*. Это приводит к необходимости оперировать с кубитами (вместо битов) в их квантовых аналогах [23,24].

Поскольку нейрокомпьютеры не обязаны обрабатывать биты, их квантовые аналоги не обязаны работать с кубитами. Поэтому, все трудности связанные с необходимостью обрабатывать кубитовый регистр для них не возникают [25].

Конечно, предложенный подход к построению системы квантовой нейронной обработки данных требует использования ансамбля квантов, поскольку вероятность детектирования частицы (градуальный выход сис-

темы) может быть оценена лишь статистически. Таким образом, эта система по сути реализует стохастические вычисления [22].

Обучение квантового нейрона в общем случае может быть реализовано стандартным методом градиентного спуска. При этом минимизация функции ошибки

$$E(r) = \frac{1}{2} \sum_u^P |y^u - t^u|^2, \quad (5.5.9)$$

где обучающее множество содержит  $P$  примеров, может быть достигнута путём коррекции положений щелей согласно выражению [23]

$$\Delta r_k^j = -\frac{\partial E}{\partial r_k^j}, \quad (5.5.10)$$

описали лишь простейшую физическую реализацию нейронной системы.

Для оценки перспективности её использования, необходимо исследовать такие характеристики как:

- универсальность – возможность аппроксимации с необходимой точностью любой функции многих переменных. Очевидно, что увеличение числа миров (определяемого числом щелей в барьерах) увеличивает гибкость системы [22];
- способность к обобщению – необходимо исследовать способность системы обрабатывать новые образы [29].

Одной из задач процесса проектирования является синтез конструктивного варианта объекта, в наибольшей степени удовлетворяющего требованиям технического задания. На начальных стадиях проектирования требования технического задания конкретизируются в виде системы ограничений, которым должны удовлетворять характеристики объекта проектирования, обеспечивающие успешное решение проектной задачи.

# ГЛАВА 6. ИССЛЕДОВАНИЕ ПРОЦЕССА СОЗДАНИЯ ЭЛЕМЕНТОВ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПРОЕКТИРОВАНИЯ НЕЙРОННЫХ СЕТЕЙ

## 6.1. Элементы нейронных сетей

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 схемотехники – 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 схемотехника 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 применения [35].

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. 6.1.1. Адаптивным называем его из-за наличия вектора настраиваемых параметров  $a$  [36].

Для многих задач полезно иметь линейную неоднородную функцию выходных сигналов. Её вычисление также можно представить с помощью адаптивного сумматора, имеющего  $n + 1$  вход и получающего на 0-й вход постоянный единичный сигнал (рис. 6.1.2).

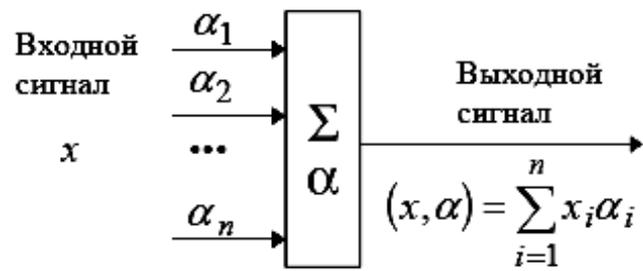


Рис. 6.1.1. Адаптивный сумматор

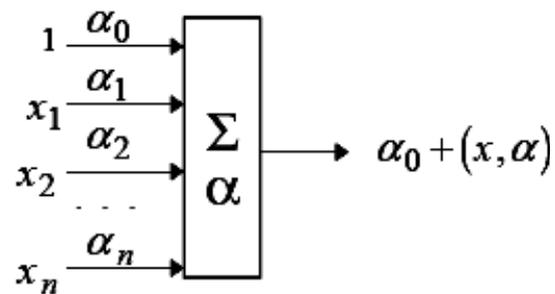


Рис. 6.1.2. Неоднородный адаптивный сумматор

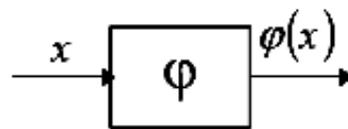


Рис. 6.1.3. Нелинейный преобразователь сигнала

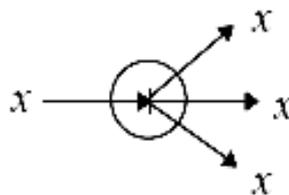
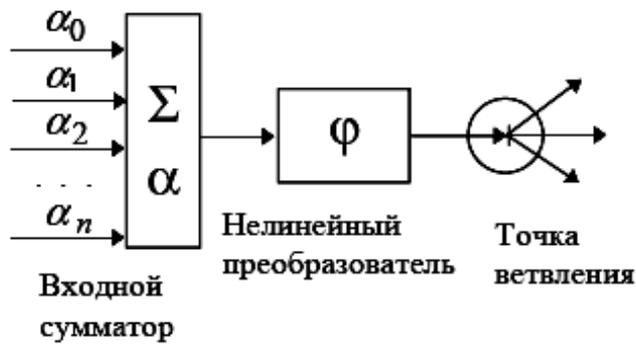
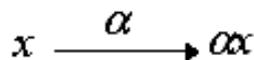


Рис. 6.1.4. Точка ветвления



**Рис. 6.1.5.** Формальный нейрон



**Рис. 6.1.6.** Синапс

*Нелинейный преобразователь сигнала* изображен на рис. 6.1.3. Он получает скалярный входной сигнал  $x$  и переводит его в  $\varphi(x)$ .

*Точка ветвления* служит для рассылки одного сигнала по нескольким адресам (рис. 6.1.4). Она получает скалярный входной сигнал  $x$  и передает его всем своим выходам [34].

*Стандартный формальный нейрон* составлен из входного сумматора, нелинейного преобразователя и точки ветвления на выходе (рис. 6.1.5).

*Линейная связь – синапс* – отдельно от сумматоров не встречается, однако для некоторых рассуждений бывает удобно выделить этот элемент (рис. 6.1.6). Он умножает входной сигнал  $x$  на «вес синапса»  $a$ .

Также бывает полезно «присоединить» связи не ко 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.

Итак, дано описание основных элементов, из которых составляются нейронные сети [35÷45].

## Теория архитектуры нейронной сети

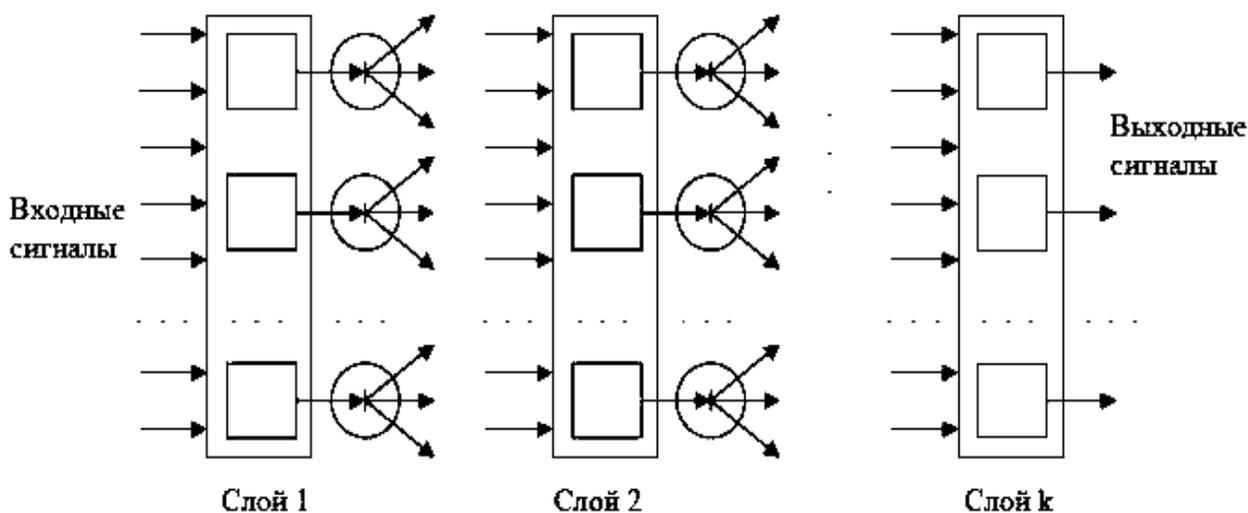
Как можно составлять сети из элементов? Строго говоря, как угодно, лишь бы входы получали какие-нибудь сигналы. Но такой произвол слишком необозрим, поэтому используют несколько стандартных архитектур, из которых путём вырезания лишнего или (реже) добавления строят большинство используемых сетей.

Сначала следует решить вопрос о том, как будет согласована работа различных нейронов во времени – вопрос о синхронности функционирования. Здесь и далее рассматриваются только нейронные сети, синхронно функционирующие в дискретные моменты времени.

Выделяется две базовых архитектуры нейронных сетей – *слоистые и полносвязные сети*.

**Слоистые сети:** нейроны расположены в несколько слоёв (рис. 6.1.7). Нейроны первого слоя получают входные сигналы, преобразуют их и через точки ветвления передают нейронам второго слоя. Далее срабатывает второй слой и т.д. до  $k$ -го слоя, который выдаёт выходные сигналы для интерпретатора и пользователя. Если не оговорено противоположное, то каждый выходной сигнал  $i$ -го слоя подается на вход всех нейронов  $i + 1$ -го. Число нейронов в каждом слое может быть любым и никак заранее не связано с количеством нейронов в других слоях. Стандартный способ подачи входных сигналов: все нейроны первого слоя получают каждый входной сигнал. Особое распространение получили трёхслойные сети, в которых каждый слой имеет своё наименование: первый – входной, второй – скрытый, третий – выходной [41÷46].

**Полносвязные сети:** каждый нейрон передает свой выходной сигнал остальным нейронам, включая самого себя. Выходными сигналами сети могут быть все или некоторые выходные сигналы нейронов после нескольких тактов функционирования сети. Все входные сигналы подаются всем нейронам.



**Рис. 6.1.7.** Слоистая сеть

Элементы слоистых и полносвязных сетей могут выбираться по-разному. Существует, стандартный выбор – нейрон с адаптивным неоднородным линейным сумматором на входе (см. рис. 6.1.5).

Для полносвязной сети входной сумматор нейрона фактически распадается на два: первый вычисляет линейную функцию от входных сигналов сети, второй – линейную функцию от выходных сигналов других нейронов, полученных на предыдущем шаге.

*Функция активации нейронов (характеристическая функция)  $\varphi$*  – нелинейный преобразователь, преобразующий выходной сигнал сумматора (см. рис. 6.1.5) – может быть одной и той же для всех нейронов сети. В этом случае сеть называют *однородной (гомогенной)*. Если же  $\varphi$  зависит ещё от одного или нескольких параметров, значения которых меняются от нейрона к нейрону, то сеть называют *неоднородной (гетерогенной)*.

Составление сети из нейронов стандартного вида (см. рис. 6.1.5) не является обязательным. Слоистая или полносвязная архитектуры не налагают существенных ограничений на участвующие в них элементы. Единственное жесткое требование, предъявляемое архитектурой к элементам сети, это соответствие размерности вектора входных сигналов элемента (она определяется архитектурой) числу его входов.

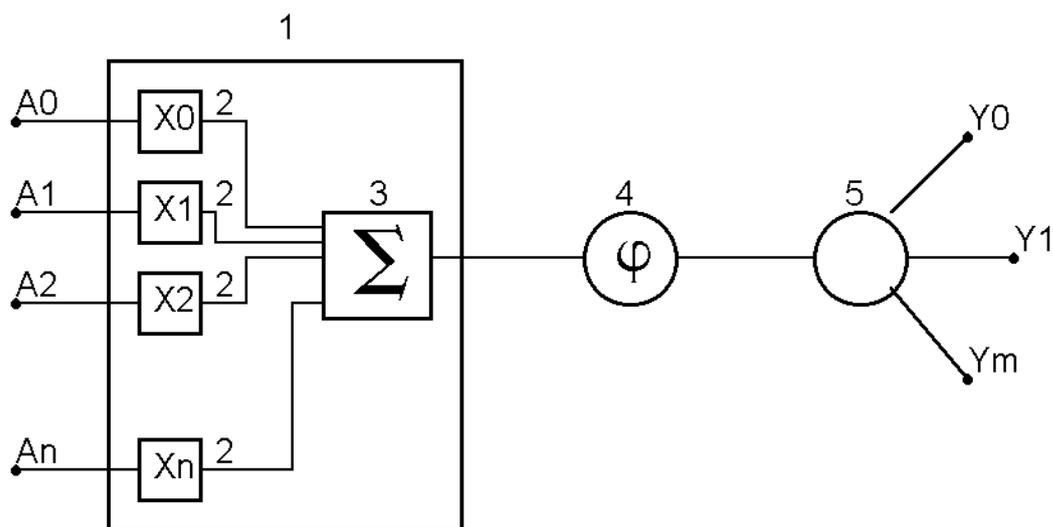
Если полносвязная сеть функционирует до получения ответа (заданное число тактов  $k$ ), то её можно представить как частный случай  $L$ -слойной сети, все слои которой одинаковы и каждый из них соответствует такту функционирования полносвязной сети [47].

Существенное различие между полносвязной и слоистой сетями возникает тогда, когда число тактов функционирования заранее не ограничено – слоистая сеть так работать не может.

Элемент нейронной сети (формальный нейрон) (рис. 6.1.8.) реализуется в составе интегральной микросхемы и в общем случае содержит входной сумматор 1, последовательно связанный с нелинейным преобразователем сигналов 4 и точкой ветвления 5, предназначенной для передачи выходного сигнала по нескольким адресам. Каждый вход  $A_i$  входного сумматора 1 называется синапсом. В состав сумматора входит схема умножения входного сигнала на вес синапса 2 и блок суммирования сигналов 3.

Формальный нейрон работает следующим образом. Входной сумматор 1 получает на входы  $A_i$  вектор входного сигнала  $\{a_0, a_1, \dots, a_n\}$ . Каждая компонента входного вектора умножается в блоке 2 на вес входа (вес синапса)  $x_i$ , а затем результаты суммируются в блоке 3 и передаются на вход нелинейного преобразователя 4. Нелинейный преобразователь 4 на основе выходного сигнала сумматора вырабатывает выходной сигнал нейрона в соответствии с заложенной в него функцией активации  $y = \varphi(x)$ . Выходной сигнал нейрона через точку ветвления 5 передаётся на входы других формальных нейронов.

Не нарушая общности, поставим задачу создания нейронной сети, предназначенной для распознавания графических образов методом сравнения с эталонами. Графические образы образуются, например, в видеокамере или цифровом фотоаппарате. Эталоны также задаются, например, в форме цифровых фотографий. Алгоритм работы нейронной сети при распознавании изображений состоит из двух обобщенных этапов.



**Рис. 6.1.8.** Схема формального нейрона

1. Предварительный этап, состоящий в наведении видеокамеры на объект, автофокусировки изображения, масштабирования и центрирования изображения на фотоприемнике видеокамеры.
- 2 Собственно распознавание изображения методом сравнения с эталонами.

*Первый этап* решается стандартным набором функций видеокамер, предназначенных для работы в системах видеонаблюдения и обеспечения безопасности. Это, например, видеокамера PANASONIC WV-CS570. В результате получается стандартизованное изображение, грубо расположенное в центре экрана и имеющее стандартный масштаб.

*Второй этап*, состоящий собственно в распознавании и идентификации изображения, в системах обеспечения безопасности обычно решается оператором. В работе сделана попытка автоматизировать процесс распознавания графических изображений.

На основе теоретического анализа предложен алгоритм распознавания, состоящий из двух этапов:

1. Для каждого графического объекта формируется множество эталонов, отличающихся масштабом, сдвигом, контрастностью и т.д. То, что все эти

эталон относится к заданному объекту, определяет оператор. На первом этапе система распознавания должна определить эталон, наиболее похожий на графический объект, выбирая его в каждой группе эталонов.

2. На втором этапе из группы отобранных эталонов выбирается один, наиболее похожий на заданный графический объект.

В процессе распознавания каждому эталону назначается число (вес), определяющее близость эталона и графического объекта. В результате анализа выбирается образ (группа эталонов), к которому относится графический объект.

Понятно, что получение однозначного результата при распознавании графических изображений возможно с определенной вероятностью узнавания. На экране видеокамеры может появиться изображение, для которого нет эталона. Тогда результат распознавания может быть неоднозначным, например, система обнаружит два или три похожих эталона.

Рассмотренный алгоритм соответствует алгоритму работы двухслойной нейронной сети с прямым распространением сигналов.

Рассмотрим технологические ограничения на архитектуру нейронной сети. Характерным элементом конструкции ПЛИС является наличие трассировочных матриц, состоящих из параллельных проводников. Две трассировочные матрицы с взаимно пересекающимися проводниками выполняется в двух изолированных друг от друга технологических слоях. Связь проводников с другими элементами микросхемы осуществляется через колодцы, выполненные в слоях изоляции микросхемы.

Эти особенности технологии делают обоснованным выбор регулярной плоской архитектуры нейронной сети, согласованной с технологией её изготовления.

Первый слой нейронной сети работает по следующему алгоритму. Графическое изображение, не уменьшая общности, можно представить в виде вектора  $X_i$ , заданного компонентами  $\{x_0, x_1, x_2, \dots, x_n\}$  в  $n$ -мерном евк-

лидовом пространстве. Каждая координата вектора соответствует пикселю графического изображения. Величина  $x_i$  соответствует яркости пикселя изображения.

Изображение-эталон  $j$  также представлено в виде группы из  $k$ , вектор  $Y_{jk}$ , компоненты которого  $\{y_{1k}, y_{2k}, y_{3k}, \dots, y_{nk}\}$  соответствуют яркости соответствующего пикселя в эталонном изображении.

Результат сравнения изображения и эталона вычисляется как скалярное произведение

$$B_j = (X, Y_j) = (x_1 y_{1j} + \dots + x_n y_{nj}) . \quad (6.1.1)$$

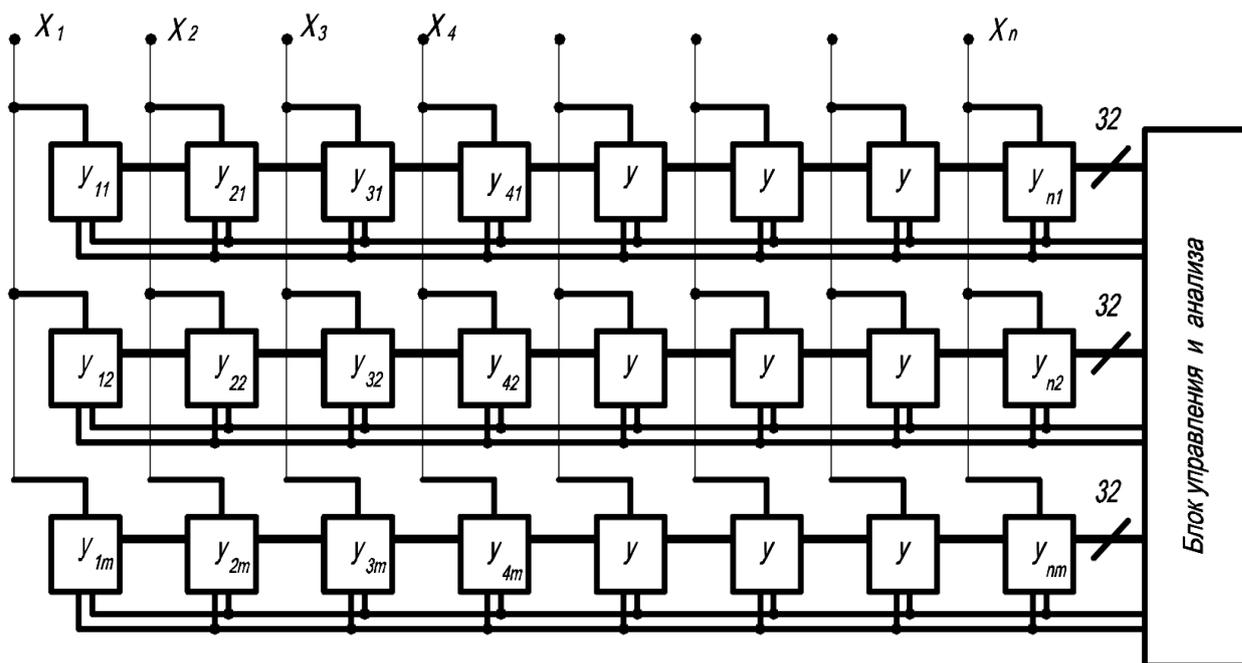
Величина скалярного произведения  $B_j$  зависит от выбранного эталона. Поэтому в качестве меры близости двух векторов необходимо взять нормированное на единицу скалярное произведение (или косинус угла между векторами  $X_i Y_{ij}$ ).

$$b_j = \frac{B_j}{|X||Y_j|} = \frac{x_1 y_{1j} + x_2 y_{2j} + \dots + x_n y_{nj}}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \sqrt{y_{1j}^2 + y_{2j}^2 + \dots + y_{nj}^2}} . \quad (6.1.2)$$

Максимальное значение нормированного скалярного произведения  $b_j$  (6.1.2) равно единице, если на вход системы подан вектор-эталон. Для произвольного вектора  $X$  величина  $b_j$  меньше единицы.

*Нормированное скалярное произведение принимается за меру близости вектора-изображения  $X$  и вектора-эталона  $Y_j$ .*

Вычисление скалярного произведения и нормировка проводится аппаратно в плоском модуле сети (матрице нейронов), представленном на рис. 6.1.9. Вектор-изображение  $X = \{x_i\}$  подается на входы сети. В строках сети расположены формальные нейроны, в которые записаны компоненты векторов  $Y_j = \{y_{ij}\}$  (весовые коэффициенты). Таким образом, число строк матрицы нейронов  $m$  соответствует максимальному числу эталонов для одного изображения.



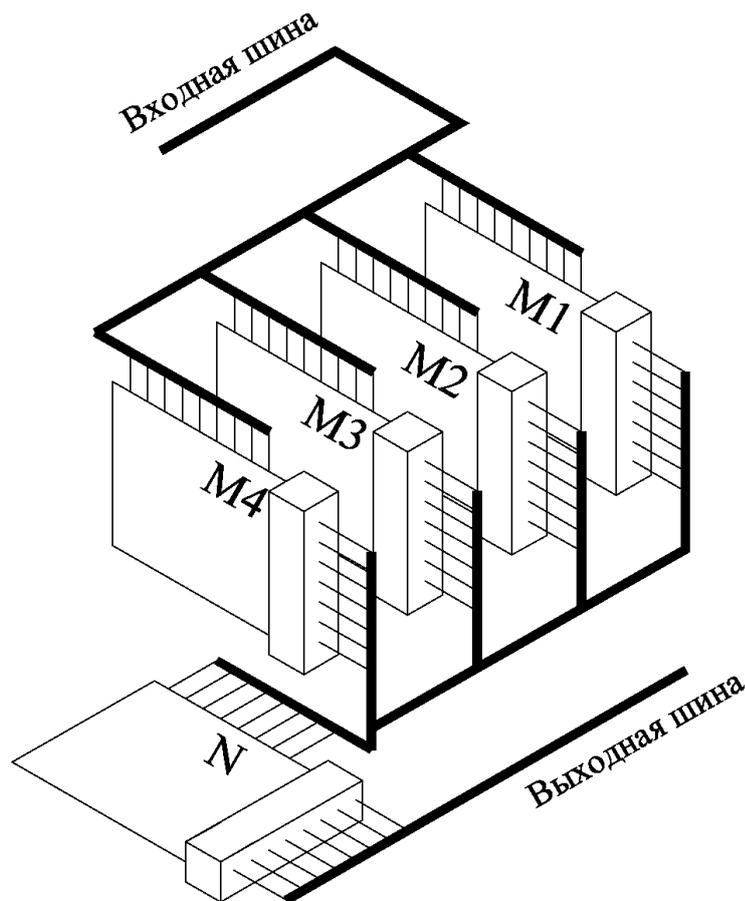
**Рис. 6.1.9.** Матрица нейронов – модуль плоской нейронной сети

Каждый формальный нейрон проводит умножение  $x_i y_{ij}$ , складывает полученную величину с значением  $x_{i-1} y_{(i-1)j}$ , полученным слева и передает сумму нейрону, расположенному справа. Таким образом, строка нейронов параллельно вычисляет скалярное произведение  $B_j$  (6.1.1) для всех строк и тем самым, для всех эталонов.

Скалярные произведения параллельно передаются в блок управления и анализа, в котором проводится нормировка (6.1.2) и выбор максимального значения  $b_j$ . Это максимальное значение определяет меру соответствия эталону, наиболее похожему на изображение  $X$ .

Матрицы формальных нейронов (см. рис. 6.1.9) образуют первый слой нейронной сети (рис. 6.1.10). Число матриц (модулей) соответствует числу групп изображений-эталонов  $M_j$ .

Значения нормированных скалярных произведений  $b_j$  с выходов модулей  $M_j$  передаются на входы выходного модуля  $N$ , который проводит окончательное распознавание изображения.



**Рис. 6.1.10.** Схема двухслойной нейронной сети

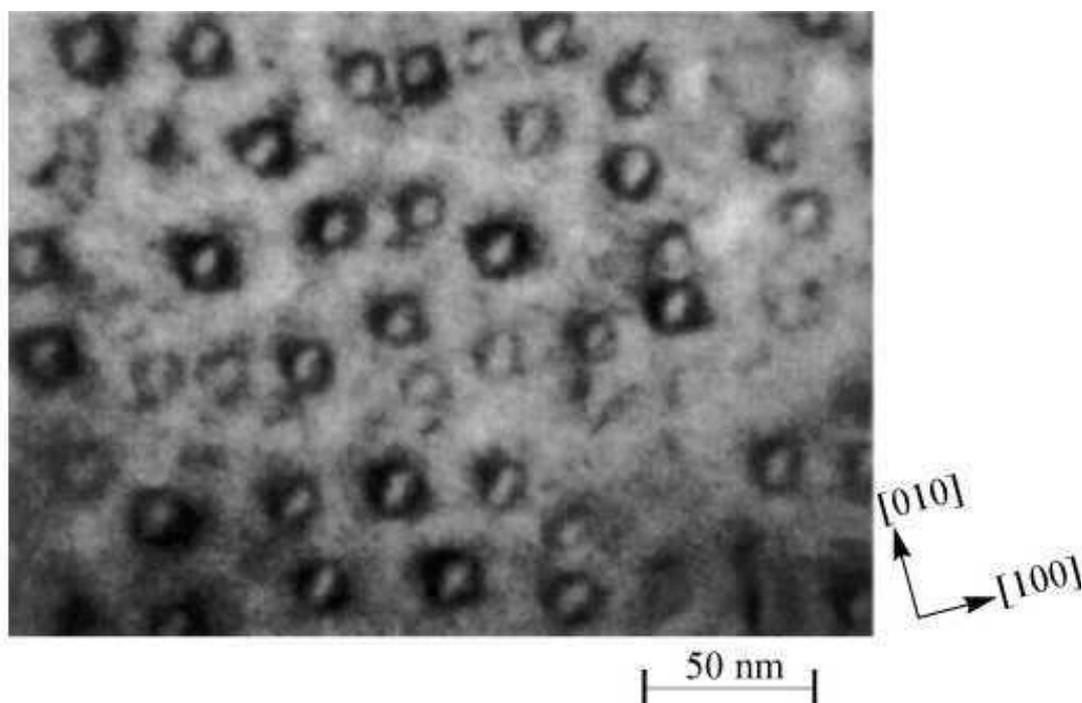
Компьютерное моделирование алгоритма распознавания изображения в нейронной сети показало его эффективность. В качестве эталонов  $Y_j$  принимались однотипные графические объекты разного масштаба (кресты, треугольники, квадраты и т.д.). В качестве входного вектора  $X$  выбирался или один из эталонов, или графическое изображение, близкое к одному из эталонов.

Предложенная структура нейронной сети хорошо соответствует современной планарной полупроводниковой технологии. Использование в качестве базового элемента транзисторов на основе нанокристаллов Si, Ge в пленке  $\text{SiO}_2$  позволяет разместить на одном кристалле плоскую нейронную сеть гигабайтного объема. Параллельный алгоритм вычислений, реализованный в плоской нейронной сети, позволяет проводить распознавание изображений в реальном масштабе времени.

## 6.2. Твёрдые объекты

Наиболее часто, говоря о квантовых точках, имеют в виду спонтанно сформировавшиеся в процессе роста массивы островков-включений одного полупроводникового материала (с меньшей шириной запрещённой зоны) в матрице другого (с большей шириной запрещённой зоны) – рис.6.2.1,а. Из-за различия ширины запрещённых зон электроны и дырки оказываются локализованы во всех трёх направлениях, как бы "заперты" в квантовой точке, следствием чего и является квазиатомный спектр.

На первом рисунке представлен "вид сверху", однако для более детального изучения структуры квантовых точек требуются изображения с большим разрешением и, конечно, исследователи имеют возможность получать такие изображения. В работе [1] с помощью сканирующей тун-



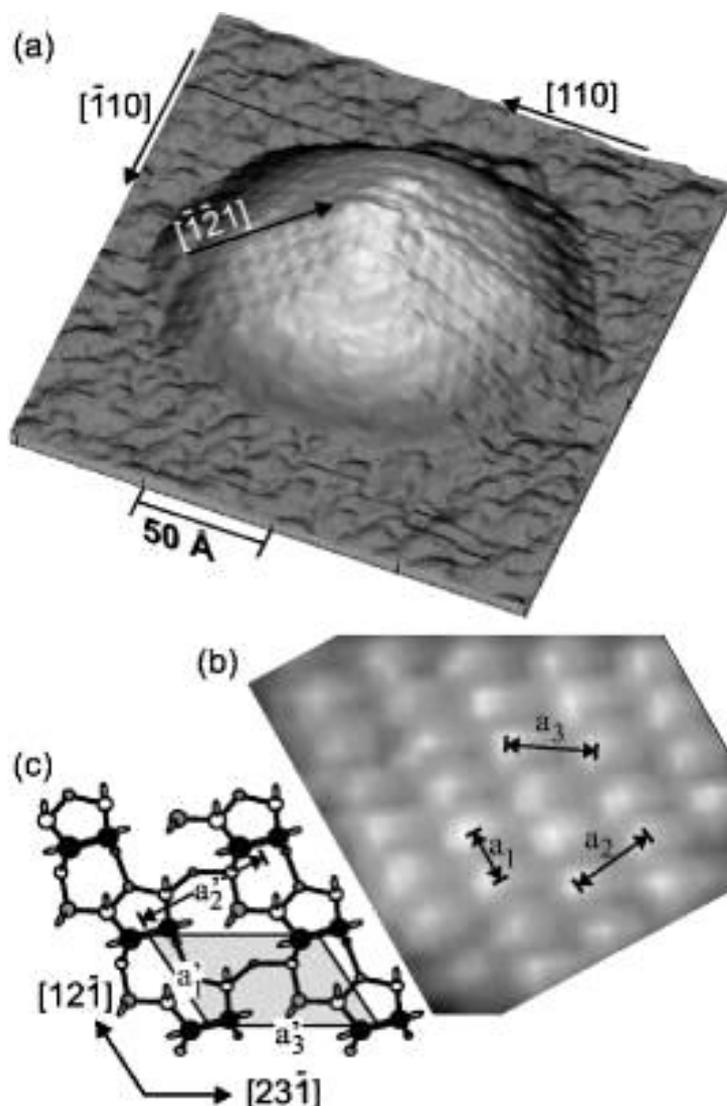
**Рис. 6.2.1,а.** Изображение квантовых точек InAs в матрице GaAs (вид сверху), полученное с помощью просвечивающей электронной микроскопии

нельной микроскопии (СТМ) были получены изображения квантовых точек с атомным разрешением. Непосредственно после формирования массива квантовых точек в результате выращивания ультратонкого слоя InAs на поверхности GaAs (массив квантовых точек, естественно, не защищался сверху слоем GaAs) образец был в условиях высокого вакуума перемещён в специальную аналитическую камеру. На рисунках 6.2.1,б и 6.2.1,в показаны полученные с помощью СТМ трёхмерные изображения одиночной квантовой точки [41÷43].

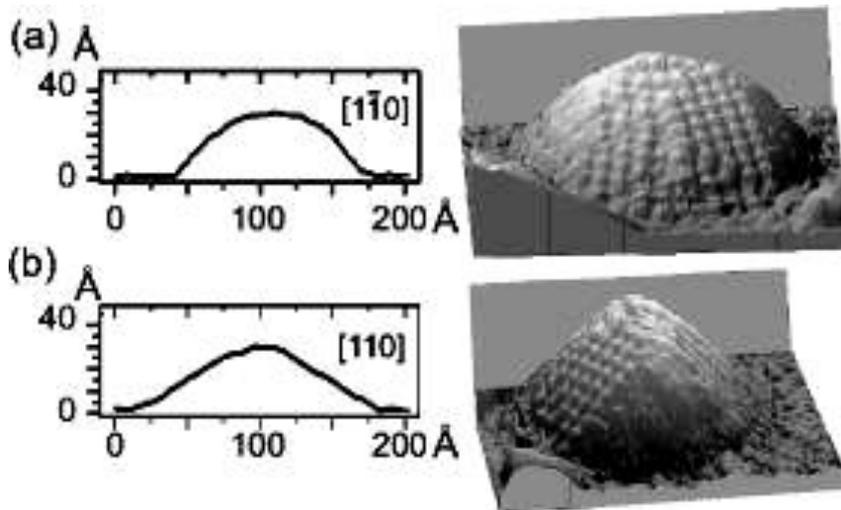
Конечно, надо иметь в виду, что даже в случае тех же соединений InAs и GaAs размеры и форма квантовой точки зависят от ростовых условий. В случае других соединений квантовые точки могут формироваться и в форме плоских блинчиков, и в форме микросфер (в диэлектрической матрице).

Квантовые точки, их иногда ещё называют искусственными атомами, представляют собой специальным образом выращенные наноразмерные островки-включения одного полупроводникового материала (с меньшей шириной запрещённой зоны) в матрице другого (с большей шириной запрещённой зоны). Из-за различия ширины запрещённых зон носители заряда оказываются локализованы в пределах островка, следствием чего является квазиатомный (представляющий собой набор отдельных уровней) энергетический спектр [42].

Отдельная квантовая точка представляет собой специальным образом полученный наноразмерный объект, обладающий дискретным энергетическим спектром. Способы получения полупроводниковых квантовых точек весьма различны: они могут создаваться из планарных полупроводниковых гетероструктур с помощью литографии, могут получаться химическими методами. Наиболее широко распространённым способом получения квантовых точек является спонтанное формирование наноразмерных островков-включений одного полупроводникового материала (с меньшей



**Рис. 6.2.1,б.** Трёхмерное изображение квантовой точки. Видно, что она имеет пирамидальную форму с достаточно острой вершиной: Стрелки и цифры в квадратных скобках обозначают различные кристаллографические направления; b – более детальное изображение одной из граней; c – модель реконструированной поверхности грани; чёрные и серые шарики – атомы мышьяка (As); светлые шарики – атомы индия (In)



**Рис. 6.2.1,в.** Профили (по высоте) квантовой точки и соответствующие трехмерные СТМ-изображения, вид с разных направлений

шириной запрещённой зоны) в матрице другого (с большей шириной запрещённой зоны) [44÷47]. Из-за различия ширины запрещённых зон носители заряда оказываются, локализованы в пределах островка, следствием чего и является квазиатомный (представляющий собой набор отдельных уровней) энергетический спектр.

### **6.3. Схема образования двумерных электронов в гетероструктуре**

Прежде чем перейти к методам формирования полупроводниковых элементов нейрочипов, рассмотрим схему образования двумерных структур. Двумерные электроны образуются на плоской границе контакта двух полупроводников с разной шириной запрещённой зоны – так называемой гетероструктуре. Рассмотрим образование гетероперехода на примере двух полупроводников – GaAs и Ga<sub>1-x</sub>Al<sub>x</sub>As (ширина запрещённой зоны  $E_{g2}$  увеличивается при увеличении  $x$ ). На рис. 6.3.1,а представлены зонные диаграммы двух разделённых в пространстве полупроводников разного состава

ва, причём энергия электрона в вакууме выбрана в качестве точки отсчёта. Таким образом, внутри полупроводника энергия электрона понижается, то есть для того, чтобы электрон удалить из полупроводника, необходимо затратить определённую энергию [41÷43, 49].

Когда два различных полупроводника соединяются, у границы их раздела происходит перераспределение электрического заряда и образуется так называемый гетеропереход (рис. 6.3.1,б). Электрическое поле, создаваемое электронами в арсениде галлия и ионизированными примесями в твёрдом растворе арсенида галлия с алюминием (показаны на рис. 6.3.1,б светлыми кружками), приводит к изгибу зон, и в возникающей квантовой яме образуются несколько уровней энергии.

Характерный размер потенциальной ямы в GaAs в направлении, перпендикулярном гетерогранице, порядка или меньше длины волны де Бройля для электронов в данном полупроводнике, поэтому движение электронов в этом направлении квантовано. При этом электроны могут свободно двигаться вдоль границы раздела материалов, то есть ведут себя как двумерные. Типичной является гетероструктура GaAs/Ga<sub>1-x</sub>Al<sub>x</sub>As. Ширина запрещённой зоны  $E_g$  в GaAs составляет 1,52 эВ. При добавлении Al величина  $E_g$  растёт. Для стандартной гетероструктуры при концентрации алюминия  $x = 0,3$  разность запрещённых зон составляет – 0,4 эВ. На границе возникает скачок потенциала, – 60% которого приходится на зону проводимости и – 40% — на валентную зону [46].

В настоящее время гетероструктуры созданы в самых различных полупроводниках и полупроводниковых соединениях Ge/Si, InAs/GaAs и т.д.



**Рис. 6.3.1.** Зонная диаграмма двух различных полупроводниковых материалов и профиль дна зоны проводимости гетеро-перехода. Индексы 1 и 2 относятся к GaAs и  $\text{Ga}_{1-x}\text{Al}_x\text{As}$  соответственно. Все энергии отсчитываются от уровня энергии электрона в вакууме. Двумерные электроны в гетеропереходе заштрихованы; светлые кружки – ионизированные, тёмные – неионизированные примеси:

а – зонная диаграмма двух различных полупроводниковых материалов (GaAs и  $\text{Ga}_{1-x}\text{Al}_x\text{As}$ )  $E_0$  – дно зоны проводимости,  $E_v$  – потолок валентной зоны,  $E_g$  – ширина запрещённой зоны; б – профиль дна зоны проводимости  $E_0$  – гетероперехода  $\Delta E_0$  – разрыв зоны проводимости,  $E_0$  и  $E_1$  – уровни размерного квантования

## **6.4. Теоретический подход к росту твёрдых объектов как элементов нейронной сети**

Рассмотрим результаты исследований свойств самоорганизованных квантовых точек  $\text{Si}_{1-x}\text{Ge}_x$  ( $x = 0;3$ ), сформулированных методом ионного синтеза.

В пластины кристаллического кремния ориентации (111) имплантировали ионы германия  $^{74}\text{Ge}^+$  на сильноточном ускорителе SCI-218 «BALZERS». Дозы имплантации составили  $D = 5 \cdot 10^{16}, 1 \cdot 10^{17} \text{ см}^{-2}$ , энергия ионов 50 кэВ. Для предотвращения эффектов каналирования падающий на кремниевую подложку ионный поток направляли с отклонением  $7^\circ$  от нормального падения. После имплантации образцы подвергались фотонному импульсному отжигу при температуре  $900^\circ\text{C}$  в атмосфере азота в течении 3 с. В результате подобного воздействия в тонком слое твёрдого раствора SiGe удалось сформировать области с повышенной концентрацией атомов Ge, протяжённость которых составляла несколько десятков нм и высота до 10 нм (наноразмерные структуры) [38÷42].

На электронном оже-спектрометре (ЭОС) PHI-680 фирмы Physical Electronics (США) проводились исследования локального элементного состава структур, а так же оценивались геометрические размеры структур и пространственное расположение квантовых точек в приповерхностной области. Ускоряющее напряжение первичных электронов составляло 10 кэВ, ток – 10 нА, диаметр первичного пучка 15÷20 нм, а глубина анализа не более 5 нм. Топографию поверхности изучали на атомно-силовом микроскопе (АСМ) Solver-47 фирмы НТ-МДТ. Для исследования формы наноразмерных структур и элементного состава поверхность образцов обрабатывали раствором КОН (33%) в течении 25 с при  $100^\circ\text{C}$ , что позволило выявить области с максимальным содержанием германия [49].

Имплантация проводилась ионами  $\text{Ge}^+$  с дозой  $D = 10^{16} \text{ см}^{-2}$  в двух режимах:

- 1) с энергией ионов 50 кэВ, проецированный пробег при этом был равен  $R_p = 35,5 \text{ нм}$ , а толщина скрытого слоя  $\Delta R_p = 13 \text{ нм}$ ;
- 2) с энергией ионов 150 кэВ,  $R_p = 89 \text{ нм}$ ,  $\Delta R_p = 30,6 \text{ нм}$ .

Отжиг проводился при  $\approx 1000^\circ \text{ C}$  в течении 15 минут [45÷47,49].

Ионы Si с энергией 140 кэВ имплантировали в слой  $\text{SiO}_2$  толщиной 0,6 мкм, выращенные термически на кремниевых подложках. Плотность ионного тока не превышала  $5 \text{ мкА/см}^2$ . Ионный синтез проводился в трёх вариантах так, чтобы во всех случаях сохранить одну и ту же дозу и сопоставимые термические бюджеты отжига. Таким образом, имелись образцы трёх типов, полученные в следующих режимах:

- 1) доза  $10^{17} \text{ см}^{-2}$  с последующим однократным отжигом при  $1100^\circ \text{ C}$  в течение 2 ч;
- 2) доза  $5 \cdot 10^{16} \text{ см}^{-2}$  с последующим отжигом при  $1100^\circ \text{ C}$  в течение 1 ч, и затем эта процедура повторялась ещё раз;
- 3) доза  $3,3 \cdot 10^{16} \text{ см}^{-2}$  с последующим отжигом при  $1100^\circ \text{ C}$  в течение 40 мин, и затем эта процедура повторялась дважды.

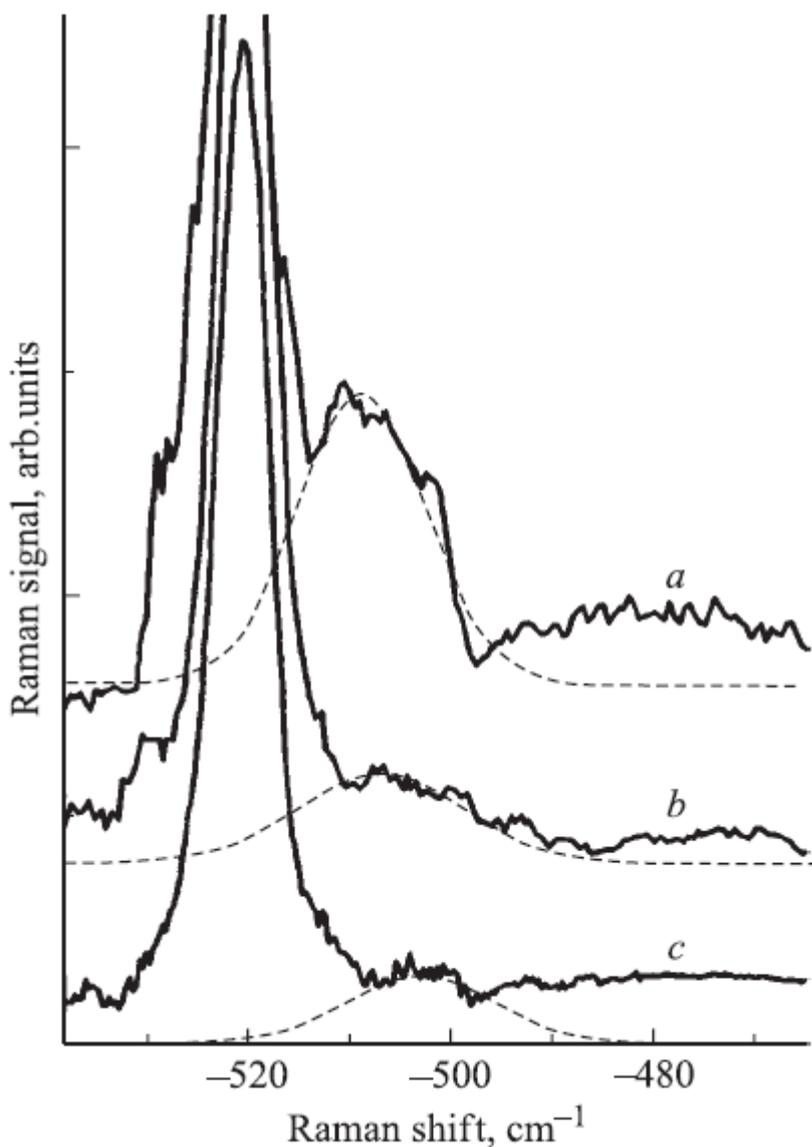
Все отжиги проводились в атмосфере азота. Согласно расчетам пробегов ионов по программе TRIM-95 для дозы  $10^{17} \text{ см}^{-2}$  в максимуме распределения концентрация избыточных атомов Si составляла 10 ат% .

Образцы исследовались методами фотолюминесценции (ФЛ), рамановского рассеяния и высокоразрешающей электронной микроскопии на поперечных срезах. Для возбуждения ФЛ использовался азотный лазер с длиной волны излучения  $\lambda = 337 \text{ нм}$ , а регистрация проводилась с помощью фотоумножителя ФЭУ-79. Все спектры нормировались на спектральную чувствительность аппаратуры. Рамановское рассеяние возбуждалось излучением аргонового лазера с  $\lambda = 514 \text{ нм}$ . Для снижения сигнала от кремниевой подложки была выбрана квазиобратная геометрия рассеяния

$Z(XX)Z$ , где  $Z$  – направление (001),  $X$  – направление (100). Спектры как рамановского рассеяния, так и ФЛ снимались при комнатной температуре. Поперечные срезы готовили по стандартной методике, а электронно-микроскопические исследования были проведены на микроскопе JEM-4000EX фирмы JEOL [49].

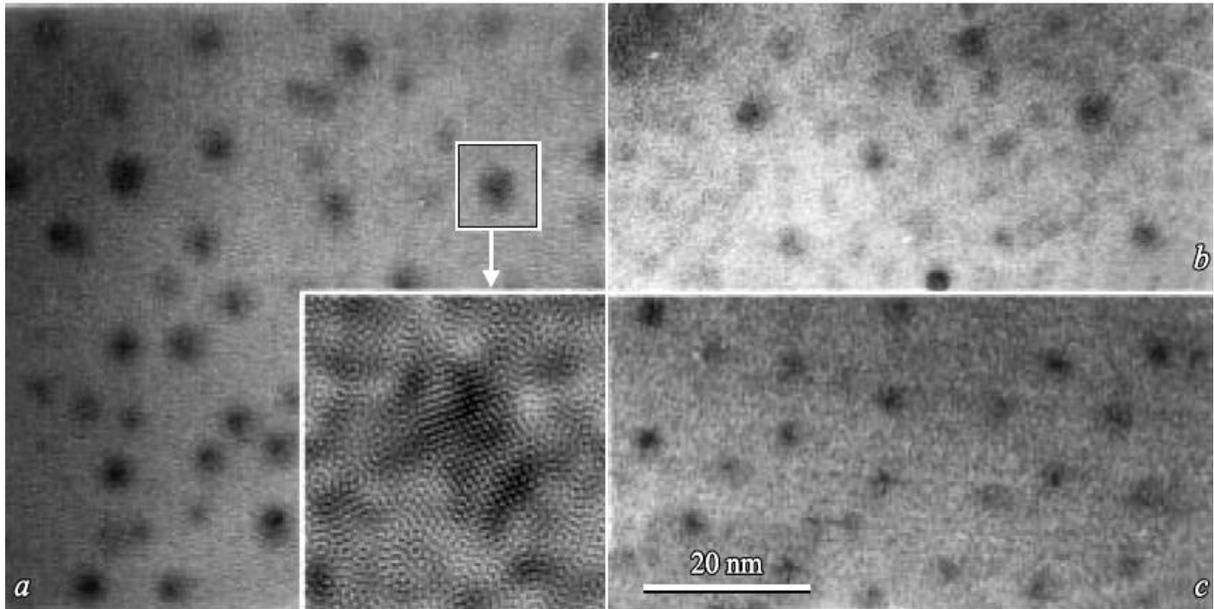
На рис. 6.4.1 показаны спектры рамановского рассеяния от образцов, полученных при трёх режимах ионно-лучевого синтеза нанопреципитатов. После имплантации полной дозы ионов Si и отжига вблизи полосы  $520 \text{ см}^{-1}$ , обусловленной рассеянием от кристаллической кремниевой подложки, появлялся чётко выраженный дополнительный пик с максимумом около  $510 \text{ см}^{-1}$ . Он свидетельствует об образовании нанокристаллов Si. Кроме того, просматривается слабая широкая полоса рассеяния в области с центром вблизи  $480 \text{ см}^{-1}$ , где рассеивают связи Si–Si аморфного кремния. Переход на режим имплантации с одним промежуточным отжигом существенно меняет спектр. Интенсивность дополнительного пика сильно понижается, а его максимум смещается в длинноволновую область  $k \sim 507 \text{ см}^{-1}$ . Отмеченные тенденции в ещё большей степени проявились после ионного синтеза с двумя промежуточными отжигами. Как видно из рис. 6.4.1, дополнительное рассеяние, присущее кремниевым квантово-размерным кристаллам, практически полностью исчезает. Существует лишь некоторый намёк на дополнительное рассеяние около  $504 \text{ см}^{-1}$ , но его интенсивность сопоставима с шумами [38÷40,49].

По данным высокоразрешающей электронной микроскопии, на поперечном срезе однократное введение дозы  $10^{17} \text{ см}^{-2}$  приводит после отжига к образованию кремниевых нанопреципитатов, у которых выявляется кристаллическая структура (рис. 6.4.2,а). Размеры нанокристаллов составляют  $4\div 5 \text{ нм}$ , а плотность  $10^{11}\div 10^{12} \text{ см}^{-2}$ . Если доза набиралась с промежуточными отжигами, в  $\text{SiO}_2$  были видны нанопреципитаты в виде тёмных пятен на изображении скола. Выявить в них признаки кристаллической



**Рис. 6.4.1.** Спектры рамановского рассеяния образцов, полученные в режимах 1 (*a*), 2 (*b*) и 3 (*c*)

структуры не удается. Подобные пятна наблюдались ранее неоднократно разными исследователями, когда условия синтеза оказывались недостаточными для формирования различных нанокристаллов. Промежуточные отжиги приводят к уменьшению средних размеров преципитатов до 3÷4 нм и к некоторому снижению их концентрации (рис. 6.4.2, *b*, *c*). Делать здесь какие-либо количественные сравнения затруднительно из-за малости площади обзора.

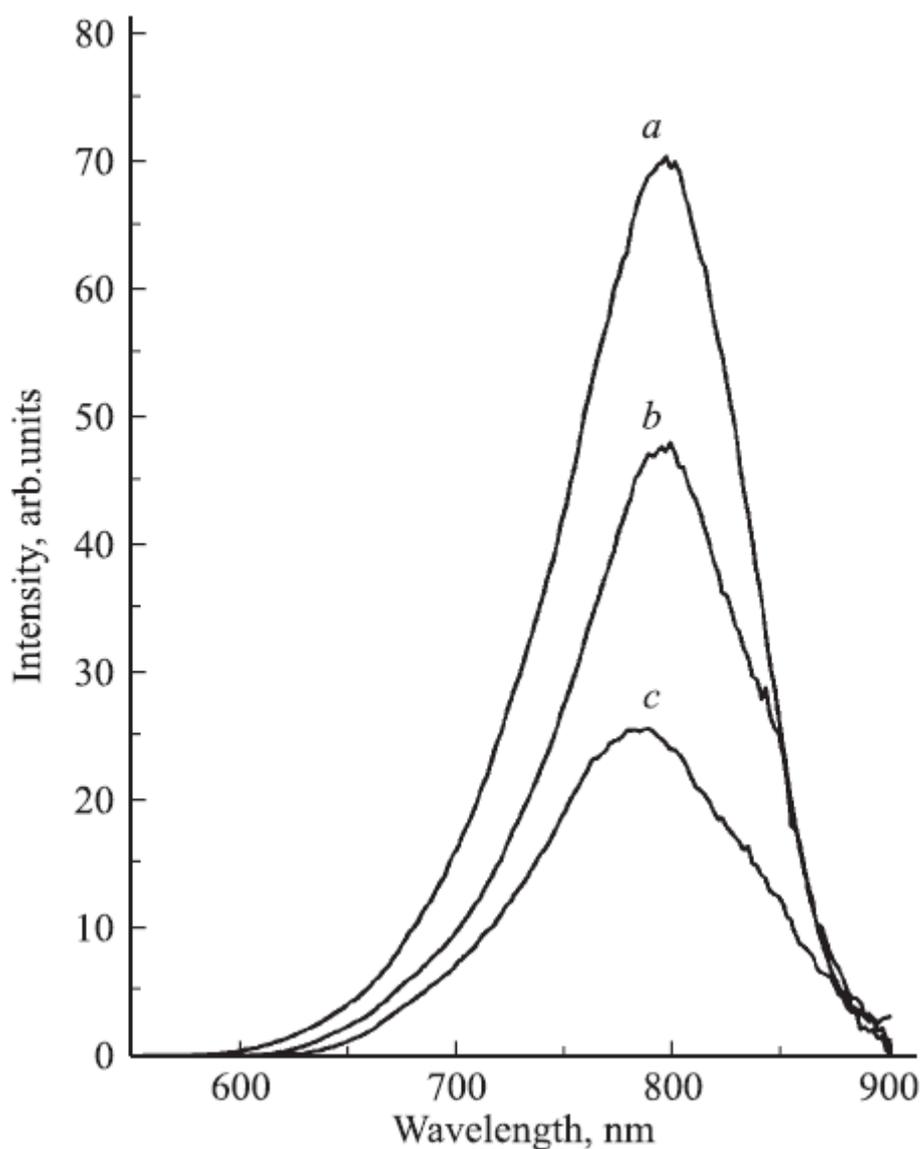


**Рис. 6.4.2.** Электронная микроскопия высокого разрешения на поперечных срезах образцов, полученных в режимах 1 (*a*), 2 (*b*) и 3 (*c*). Режим *a* – после Фурье-фильтрации на выделенном участке выявляется кристалличность включений

Спектры ФЛ после ионно-лучевого синтеза в каждом из трёх режимов представлены на рис. 6.4.3.

В отличие от данных по рамановскому рассеянию и электронной микроскопии, где при использовании промежуточных отжигов существенно ослабли признаки присутствия кремниевых нанокристаллов, их люминесценция оказалась затронута в меньшей степени. В случае имплантации полной дозы с последующим отжигом в спектре возникала интенсивная полоса с максимумом вблизи 795 нм. В настоящее время практически все исследователи связывают её с излучательной рекомбинацией в образующихся квантово-размерных кристаллах кремния. Имплантация с одним промежуточным отжигом приводила к некоторому снижению интенсивности свечения, причём заметного смещения максимума не проис-

ходило. Набор дозы  $10^{17} \text{ см}^{-2}$  в три приёма ведет к дальнейшему понижению интенсивности ФЛ, и здесь уже становится заметным коротковолновое смещение максимума полосы к  $\sim 785 \text{ нм}$ . Таким образом, использование двух промежуточных отжигов снижало интенсивность ФЛ всего в 2 с небольшим раза, в то время как возможности обнаружения признаков нанокристаллов по рамановскому рассеянию или с помощью высококоразрешающей электронной микроскопии практически исчерпывались.

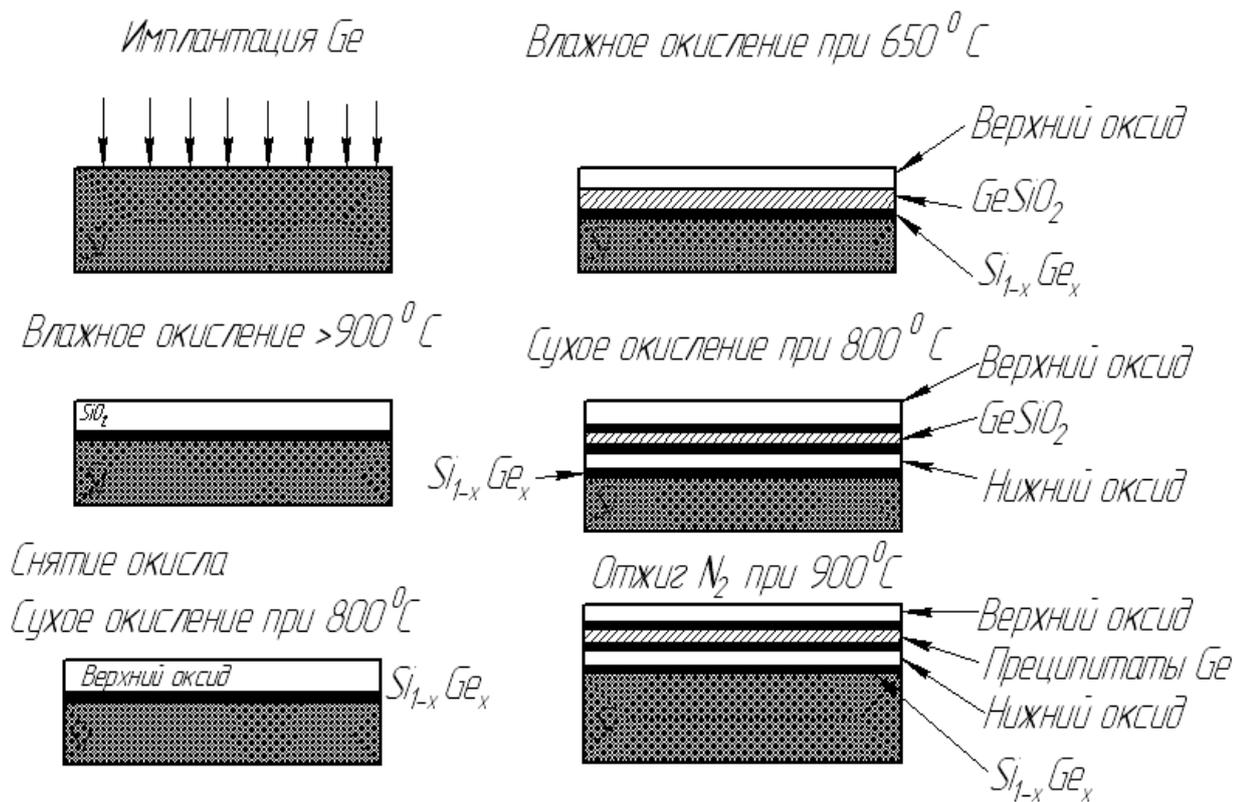


**Рис. 6.4.3.** Спектры фотолюминесценции образцов, полученных в режимах 1 (*a*), 2 (*b*) и 3 (*c*)

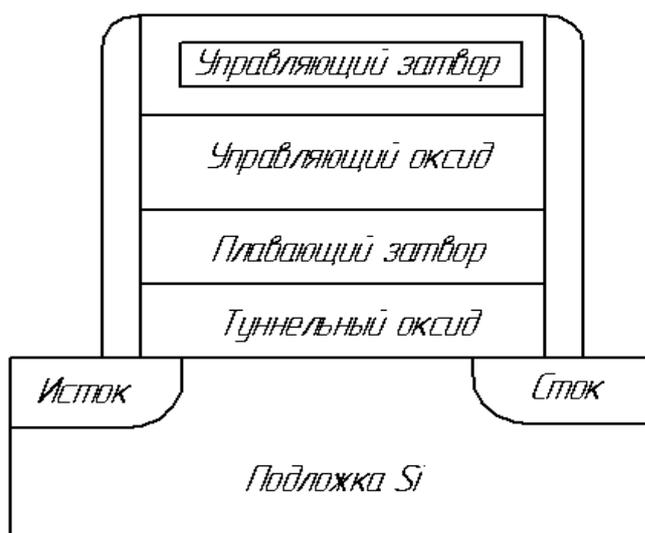
Данный механизм создания области затвор-диэлектрик, состоящий из управляющего оксида (верхний слой), слоя оксида с захороненными нанокристаллами германия (средний слой) и туннельного оксида (нижний слой) позволяет очень точно контролировать толщину оксидов окружающих захороненный слой нанокристаллов. В первую очередь на поверхности кремниевой подложки создается тонкий эпитаксиальный слой  $\text{Si}_{1-x}\text{Ge}_x$  с последующим высокотемпературным влажным окислением, которое заставляет имплантированный Ge скапливаться на границе раздела  $\text{Si}/\text{SiO}_2$  (вследствие отделения растущего оксида в данных условиях окисления). После удаления слоя высокотемпературного оксида верхний слой оксида выращивается при  $800^\circ\text{C}$  в атмосфере сухого  $\text{O}_2$ . В таких условиях Ge не взаимодействует с растущим оксидом. Затем слой  $\text{Si}_{1-x}\text{Ge}_x\text{O}_2$  выращивается при  $800^\circ\text{C}$  во влажной среде таким образом, что Ge проникает в область затвор-диэлектрик. На заключительном этапе выращивается туннельный оксид при  $800^\circ\text{C}$  в атмосфере сухого  $\text{O}_2$ , а затем структура отжигается при температуре  $900^\circ\text{C}$ , что приводит к образованию преципитатов Ge и формированию захороненного слоя нанокристаллов.

Схема эксперимента и основные результаты представлены на рис. 6.4.4 [48,49].

Элементарные ячейки нейронной сети могут хранить один бит информации и состоят из одного полевого нанотранзистора с электрически изолированной областью (плавающим затвором – floating gate), способного хранить заряд многие годы. Ячейка нейронной сети представляет собой одиночный многоходовой МОП транзистор (рис. 6.4.5). Проводящий поликремниевый слой находится между внешне доступным затвором (обозначенный как управляющий затвор) и плавающим затвором. Диэлектрики между плавающим затвором и подложкой (термический оксид кремния), а так же между плавающим затвором и контрольным входом являются туннельными и управляющими диэлектриками соответственно.



**Рис. 6.4.4.** Формирование слоя нанокристаллов Ge в подзатворном диэлектрике для хранения заряда в элементах нейронной сети



**Рис. 6.4.5.** Схематическое сечение МОП транзистора с плавающим затвором

Наличие или отсутствие заряда на плавающем затворе кодирует один бит информации. При записи заряд помещается на плавающий затвор одним из двух методов (зависит от типа ячейки): методом инжекции электронов или методом туннелирования электронов. Стирание содержимого ячейки (снятие заряда с плавающего затвора) производится методом туннелирования. Как правило, наличие заряда на транзисторе понимается как логический «ноль», а его отсутствие – как логическая «единица».

Изменение порогового напряжения  $\Delta F_{TH}$ , вызванное хранением заряда  $Q_{FG}$  определяется как

$$\Delta F_{TH} = \frac{-Q_{FG}}{C_{CG}},$$

где  $C_{CG}$  – емкость между контрольным и изолированным входом и задается формулой

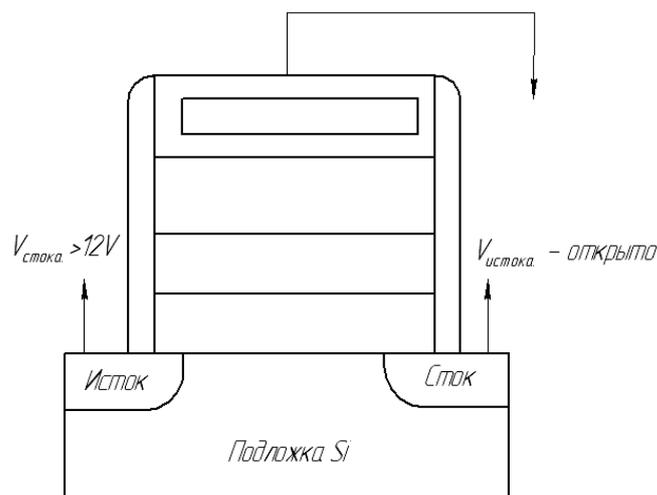
$$C_{CG} = \varepsilon \cdot \frac{A}{t},$$

где  $A$  – площадь конденсатора, а  $\varepsilon$  и  $t$  – диэлектрическая константа и толщина управляющего диэлектрика соответственно (рис. 6.4.6).

На рис.6.4.7 представлена схематическая диаграмма энергетических зон для кремниевой нанокристаллической памяти в процессе: (а) записи (инжекция электронов), и (б) стирания (экстракция электронов). В процессе записи (стирания) электроны инжектируются (экстрагируются) в/из нанокристаллы за счёт приложенного положительного (отрицательного) напряжения смещения на затвор по отношению к истоку и стоку. Толщина управляющего оксида должна быть относительно мала для выбора приемлемого низкого напряжения, но не такой маленькой, чтобы привести к утечке заряда в управляющий затвор. Эти ограничения приводят к выбору оптимальной толщины управляющего оксида в 5÷15 нм. В случае тонкого туннельного оксида толщиной менее 3 нм передача заряда осуществляется

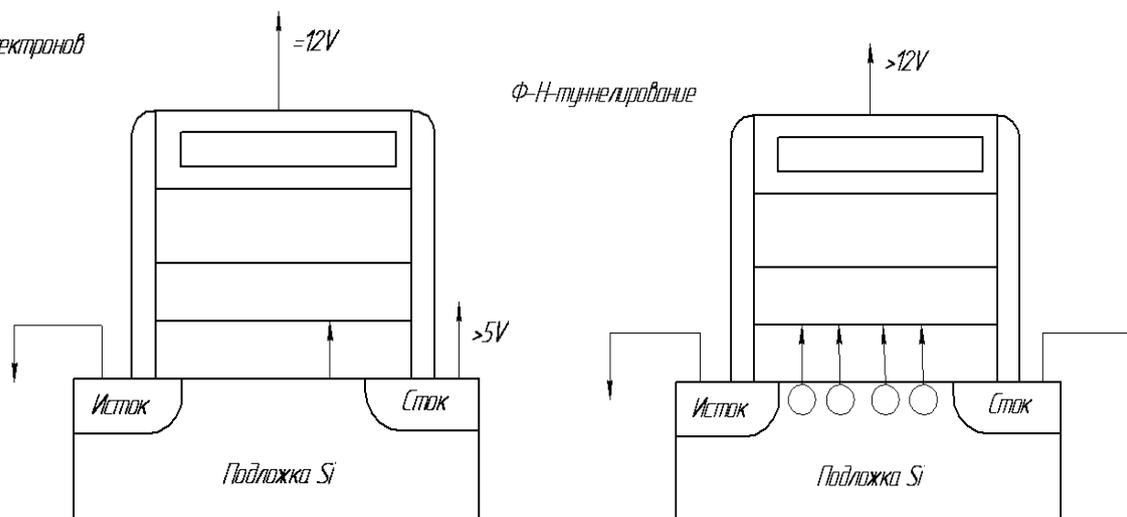
## СТИРАНИЕ

*Ф-Н-туннелирование*



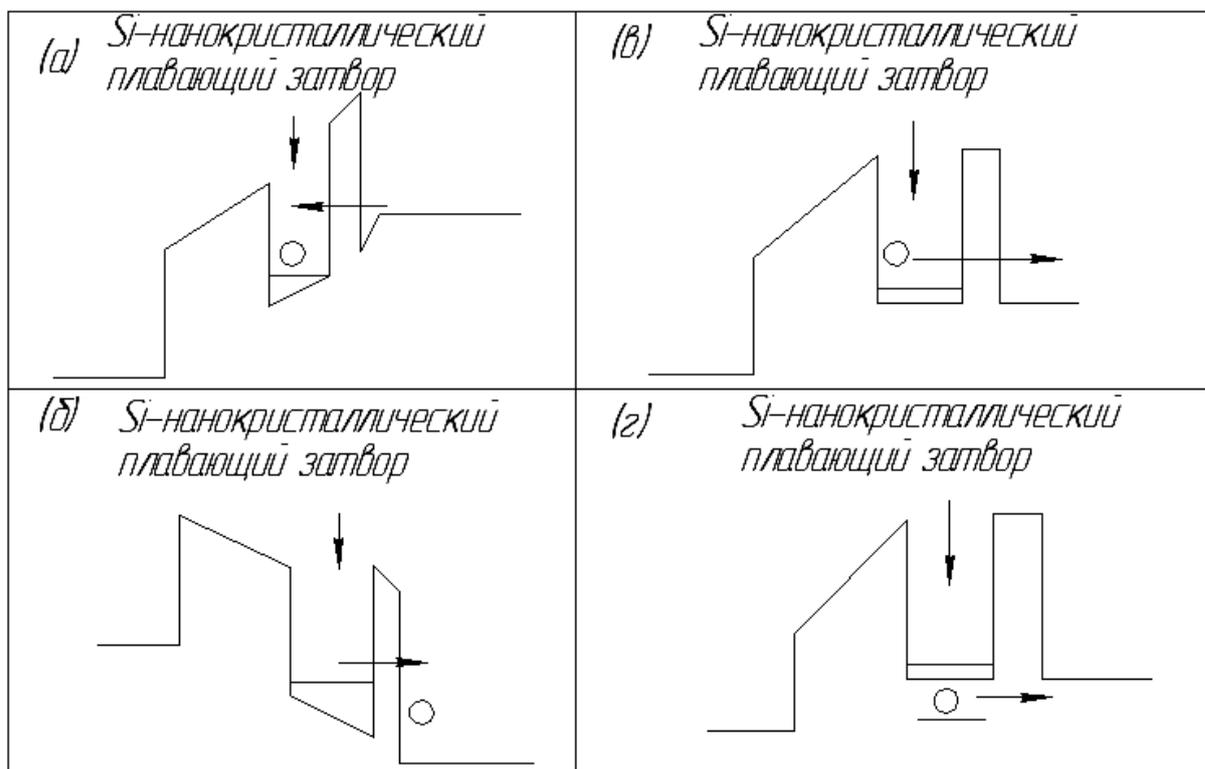
## ЗАПИСЬ

*инжекция электронов*



**Рис. 6.4.6.** Схематическое описание операций стирания/записи для ячейки элемента нейронной сети

через прямое туннелирование (потока электронов через весь оксид) вместо туннелирования по механизму Фаулера-Нордгейма. Генерация носителей низкими энергиями (с энергиями много меньше чем 3 эВ, что является порогом для некоторых главных механизмов деградации оксидов от горя-



**Рис. 6.4.7.** Диаграммы потенциала зоны проводимости для элементов нейронной сети на основе кремниевых нанокристаллов во время:

- (а) записи (инъекции электронов в нанокристалл);
- (б) стирания (экстракция электронов из нанокристалла);
- (в) хранение электронов на квантовых уровнях энергии нанокристалла;
- (г) хранение электронов на низких энергетических уровнях в нанокристалле и/или на границе раздела SiO<sub>2</sub>/нанокристалл

чих носителей заряда) во время операции программирования, понижает деградацию оксида во время Ф-Н-инъекции, в результате улучшая износостойчивость и характеристику заряд/пробой [49], что может быть реализовано, например, с помощью баллистического транзистора (рис. 6.4.8).



Одно из главных преимуществ ячеек нейронной сети на нанокристаллах, по сравнению с обычными устройствами на основе плавающих затворов, состоит в использовании взаимно изолированных узлов хранения заряда, вместо непрерывного поликремниевого слоя. Такой неоднородный плавающий затвор уменьшает потерю зарядов через дефекты в окисле, где происходит туннелирование, позволяя все более уменьшать толщину туннельного окисла.

## ЛИТЕРАТУРА К ГЛАВАМ 5 и 6

1. Анил К. Джейн, Жанчанг Мао, Моиуддин К.М. Введение в искусственные нейронные сети // *Artificial Neural Networks: A Tutorial, Computer*, Vol.29, No.3, 1996. – P. 31-44.
2. McCulloch W.S. and Pitts W. A logical Calculus of Ideas Immanent in Nervous Activity // *Bull. Mathematical Biophysics*, Vol. 5, 1943. – P. 115-133.
3. Hertz J., Krogh A. and Palmer R.G. Introduction to the Theory of Neural Computation, Addison-Wesley, Reading, Mass., 1991.
4. Haykin S, *Neural Networks: A Comprehensive Foundation*, MacMillan College Publishing Co., New York, 1994.
5. Rosenblatt R. *Principles of Neurodynamics*, Spartan Books, New York, 1962.
6. Mimnsky M. and Papert S. *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, Mass., 1969.
7. Hopfield J.J. *Neural Networks and Physical Systems with Emergent Collective Computational Abilities* // *In Proc. National academy of Sciences, USA* 79, 1982. – P. 2554-2558.
8. Werbos P. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences* // PhD Thesis, Dept. of Applied Mathematic, 1974.
9. Дорогов А.Ю., Алексеев А.А. Категории ядерных нейронных сетей // *Труды Всерос. науч.-техн. конф. “Нейроинформатика-99”* // Сб. науч. тр. Часть 1. – М., 1999. – С. 55-64.
10. Дорогов А.Ю., Алексеев А.А. Быстрые нейронные сети // *Труды международной научно-технической конференции “Пятьдесят лет развития кибернетики”*. – Санкт-Петербург, 1999. – С. 120-121.
11. Пак М.М. Моделирование нейронных сетей на основе твёрдотельных объектов. Автореферат дисс. на соиск. уч. ст. к.т.н. – М., 2007. – 23 с.
12. Zurada J. M. *Introduction to artificial neural systems* // PWS Publishing Company, 1992. – 785 p.

13. Haykin S. Neural networks. A comprehensive foundations. McMillan College Publ. Co. N.Y., 1994. – 696 p.
14. Ezhov A.A., Ventura D. Quantum neural networks. In: Future directions for Intelligent Information Systems and Information Sciences. Kasabov N, Physica-Verlag, Heidelberg, 2000. – P. 213-235 (Studies in Fuzziness and Soft Computing, vol.45).
15. Deutsch D. The fabric of reality. Alen Lane: The Penguin Press, 1997.
16. Narayanan A. Quantum algorithms // Technical Report 374; Department of Computer Science, University of Exeter, 1998.
17. Perus M. Common mathematical foundations of neural and quantum informatics // Zeitschrift fur Angewandte Mathematik und Mechanik, 1998; 78: – P. 23-26.
18. Behrman E.C., Niemel J., Steck J.E. et al. A quantum dot neural network. In: Proceedings of the 4th Workshop on Physics of Computation, Boston, 1996. – P. 22-24.
19. Menneer T., Narayanan A. Quantum inspired neural networks // Department of Computer Sciences, University of Exeter, UK, 1995.  
<http://www.dcs.ex.ac.uk/reports/reports.html>
20. Menneer T. Quantum artificial neural networks // PhD Thesis, Faculty of Science, University of Exeter, UK, 1998.
21. Menneer T., Narayanan A. Quantum artificial neural networks vs. Classical artificial neural networks: Experiments in Simulation // Proceedings of the Fifth Joint Conference on Information Sciences, Atlantic City, 2000; 1. – P. 757-759.
22. Ventura D. and Martinez T. Quantum associative memory with exponential capacity // Proceedings of the International Joint Conference on Neural Networks, 1998. – P. 509-513.
23. Chrisley R.L. Learning in Non-superpositional Quantum Neurocomputers, In Pylkkanen, P. and Pylkko, P. (Eds.) Brain, Mind and Physics. IOS Press,

Amsterdam, 1997. – P. 126-139.

24. Крисли Р.Л. Бомовские квантовые нейронные сети (в сборнике [23]).
25. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука (Сиб. отделение), 1996. – 276 с.
26. Le Cun Y., Denker J.S., Solla S.A. Optimal Brain Damage // Advances in Neural Information Processing Systems II (Denver 1989). – San Mateo, Morgan Kaufman, 1990. –P. 598-605.
27. Горбань А.Н. Обучение нейронных сетей. – М.: Изд-во СССР- США СП "ParaGraph", 1990. – 160 с (English Translation: Training Neural Networks // AMSETransaction, Scientific Siberian, A, 1993, Vol. 6. Neurocomputing. – P. 1-134).
28. Prechelt L. Comparing Adaptive and Non-Adaptive Connection Pruning With PureEarly Stopping // Progress in Neural Information Processing, Springer, 1996, Vol. 1. – P. 46-52.
29. Cybenko G. Approximation by superposition of a sigmoidal function // Mathematics of Control, Signals, and Systems, 1989, Vol. 2. – P. 303 - 314.
30. Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators // Neural Networks, 1989, Vol. 2. – P. 359 - 366.
31. Kochenov D.A., Rossiev D.A. Approximations of functions of  $C[A,B]$  class by neural-net predictors (architectures and results) // AMSE Transaction, Scientific Siberian, A. 1993, Vol. 6. Neurocomputing. – P. 189-203. Tassin, France.
32. Gilev S.E., Gorban A.N. On completeness of the class of functions computable by neural networks // Proc. of the World Congress on Neural Networks (WCNN '96), San Diego, CA, Lawrence Erlbaum Associates, 1996. – P. 984-991.
33. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных // Докл. АН СССР, 1956. Т. 108, №. 2. – С. 179-182.

34. Арнольд В.И. О функциях трёх переменных // Докл. АН СССР, 1957. Т. 114, №4. – С. 679-681.
35. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного // Докл. АН СССР, 1957. Т. 114, № 5. – С. 953-956.
36. Витушкин А.Г. О многомерных вариациях. – М.: Физматгиз, 1955.
37. Stone M.N. The generalized Weierstrass approximation theorem // Math. Mag., 1948.V.21. – PP. 167-183, 237-254.
38. Шефер Х. Топологические векторные пространства. – М.: Мир, 1971. – 360 с.
39. Шик А.Я. Квантовые нити // Соросовский Образовательный Журнал, 1997. № 5. – С. 87-92.
40. Демиховский В.Я. Квантовые ямы, нити, точки: Что это такое // Соросовский Образовательный Журнал, 1997. № 5. – С. 81-86.
41. Молекулярно-лучевая эпитаксия и гетероструктуры / Под ред. Л. Ченга, Л. Плога. – М.: Мир, 1989. – 582 с.
42. Беляховский В.И. Физические основы полупроводниковой технологии // Соросовский Образовательный Журнал. 1998. № 10. – С. 92-98.
43. Крисилов В.А., Олешко Д.Н., Трутнев А.В. Применение нейронных сетей в задачах интеллектуального анализа информации // Труды Одесского политехнического университета, 1999. Вып.2 (8). – С. 134.
44. Авеньян Э.Д. Алгоритмы настройки многослойных нейронных сетей // Автоматика и телемеханика, 1995. № 5. – С.106-118.
45. Минский М., Пейперт С. Перцептроны. – М.: Мир, 1971. – С. 261.
46. Tiwary S., Rana F., Chan K., Hanafi H., Chan Wei, Buchanan D. Volatile and non-volatile memories in silicon with nano-crystalstorage // IEEE International Electron Device Meeting. – Washington, DC, USA, 1995. – P. 521-524.
47. Tiwary S., Rana F., Hanafi H., Harstein A., Crabbe E.F. Chan K. A silicon

nanocrystals based memory // Appl. Lett., 1996. 68. – P. 1377-1379.

48. Bonafos C., Carrada M., Cherkashin N., Coffin H., Chassaing D., Assayag G.B. Claverie A., Muller T., Heinig K.H., Perego M., Fanciulli M., Dimitrakis P., Normand P. Manipulation of two-dimensional arrays of Si nanocrystals embedded in thin SiO<sub>2</sub> layers by low energy ion implantation // J. appl. Phys., 2004. 48. – P. 1175-1179.
49. Чаплыгин Ю.А. Нанотехнология в электронике. – М.: МИЭТ, 2005. – С. 153- 170.

рк

# СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ .....	3
ВВЕДЕНИЕ .....	7
<b>Глава 1. ФОРМИРОВАНИЕ УГЛЕРОДНЫХ НАНООБЪЕКТОВ В ТОНКОПЛЕНОЧНЫХ СТРУКТУРАХ .....</b>	<b>13</b>
1.1. Основная задача формирования углеродных нанобъектов .....	13
1.2. Физические аспекты формирования наноструктур на поверхности алмазоподобных тонких пленок .....	19
1.3. Технологические особенности получения тонких алмазоподобных пленок .....	28
1.4. Формирование нанобъектов фрактального типа .....	35
Литература к главе 1 .....	45
<b>Глава 2. ПРОИЗВОДСТВО И ПРИМЕНЕНИЕ ФУЛЛЕРЕНОВ И УГЛЕРОДНЫХ НАНОТРУБОК .....</b>	<b>50</b>
2.1. Автоматизированное проектирование процесса производства углеродных нанотрубок и фуллеренов .....	50
2.2. Применение углеродных наноматериалов и технологий на их основе .....	52
2.3. Определение основных признаков и свойств углеродных нанотрубок и фуллеренов .....	69
2.4. Методы получения углеродных нанотрубок и фуллеренов .....	80
2.5. Аналитическая процедура Саати в автоматизированном проектировании .....	98
Литература к главе 2 .....	106
<b>Глава 3. АЛГОРИТМЫ ПОИСКА ОПТИМАЛЬНЫХ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В ОБЛАСТИ КАЧЕСТВА ПРОИЗВОДСТВА МАТЕРИАЛОВ НА ОСНОВЕ УГЛЕРОДНЫХ СИСТЕМ .....</b>	<b>108</b>
3.1. Адаптивное управление качеством производства наноматериалов .....	108
3.2. Регрессионный анализ .....	114

3.3	Моделирование процесса адаптивного управления с подстройкой параметров модели .....	122
3.4.	Методы выбора недоминируемого решения из ряда недоминируемых альтернатив .....	126
3.5.	Метод упорядоченного предпочтения через сходство с идеальным решением для задач многоцелевого принятия решения .....	137
<b>Глава 4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ УПРАВЛЕНИЯ КАЧЕСТВА ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ УСТРОЙСТВ НА ОСНОВЕ УГЛЕРОДНЫХ СТРУКТУР .....</b>		<b>142</b>
4.1.	Классификация задач оптимизации .....	142
4.2.	Оптимизация обобщенного критерия качества дискриминационным методом .....	157
4.3.	Вероятностное моделирование .....	168
4.4.	Задача о замене оборудования .....	174
Литература к главам 3 и 4 .....		178
<b>Глава 5. Проектирование элементов нейронных сетей .....</b>		<b>184</b>
5.1.	Особенности автоматизированного проектирования искусственных нейронных сетей .....	184
5.2.	Модель искусственного нейрона .....	190
5.3.	Многослойные сети прямого распространения .....	196
5.4.	Твердотельные объекты в нейронных сетях .....	202
5.5.	Схема работы нейрона .....	204
<b>Глава 6. Создание элементов автоматизированной системы проектирования нейронных сетей .....</b>		<b>210</b>
6.1.	Элементы нейронных сетей .....	210
6.2.	Твердотельные объекты .....	221
6.3.	Схемы образования двумерных электронов в гетероструктуре .....	224
6.4.	Теоретический подход к росту твердотельных объектов как элементов нейронных сетей .....	227
Литература к главам 5 и 6 .....		240