

## **ВВЕДЕНИЕ**

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 процессы 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 наноэлектроники 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.

Целесообразность выбора углерода в качестве подложки и объекта модификации определяется множественностью его аллотропных форм, соединений и широкого диапазона электрофизических свойств: от диэлектрических до полупроводниковых. Поэтому формирование электронных объектов на основе модифицированных углеродных структур является задачей актуальной и современной, так как при использовании подобных структур появляется возможность получать как изолирующие, полупроводниковые так и электропроводящие объекты. Необходимо разработать технологические основы синтеза перспективных материалов, являющихся составной частью системы управления качеством.

Контроль параметров производимых материалов в значительной степени определяется метрологическим обеспечением, которое должно удовлетворять уровню технологии. Необходимо строить систему контроля качества на неразрушающих методах, что позволит приблизить эту систему к идеализированной схеме, подобной системе мониторинга.

Книга предназначена для инженерно-технических и научных работников, занимающихся информационными технологиями в проектировании объектов электронного машиностроения, в том числе управлением качества производства материалов на основе углеродных структур. Издание может быть рекомендовано аспирантам вузов и студентам, обучающимся по специальности 210107 – «Электронное машиностроение» и направлению 210100 – «Электроника и наноэлектроника».

# **ГЛАВА 1. МОДЕЛИРОВАНИЕ И АЛГОРИТМИЗАЦИЯ ПРОЦЕССА УПРАВЛЕНИЯ КАЧЕСТВОМ ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ УСТРОЙСТВ**

## **1.1. Основная задача и состояние работ в области управления качеством при производстве наноустройств**

### **Анализ современных мировых тенденций управления качеством**

В современной теории и практике управления качеством выделяют следующие пять основных этапов [61]:

1. Принятие решения «что производить?» и подготовка технических условий – маркетинговые исследования.
2. Проверка готовности производства и распределение организационной ответственности.
3. Процесс изготовления продукции или предоставления услуг.
4. Устранение дефектов и обеспечение информацией обратной связи в целях внесения в процесс производства и контроля изменений, позволяющих избегать выявленных дефектов в будущем.
5. Разработка долгосрочных планов по качеству.

Осуществление перечисленных этапов невозможно без взаимодействия всех отделов, органов управления фирмой. Такое взаимодействие называется *единой системой управления качеством*. Это обеспечивает системный подход к управлению качеством.

Рассмотрим более подробно содержание этапов управления качеством.

На первом этапе качество означает ту степень, в которой товары или услуги фирмы соответствуют её внутренним техническим условиям. Этот аспект качества называют **качеством соответствия техническим условиям**.

На втором этапе оценивается качество конструкции. Качество может отвечать техническим требованиям фирмы на конструкцию изделия, однако, сама конструкция может быть как высокого, так и низкого качества.

На третьем этапе качество означает ту степень, в которой работа или функционирование услуг (товаров) фирмы удовлетворяет реальные потребности потребителей. Товары фирмы могут соответствовать внутренним техническим условиям; сама конструкция изделия может быть выдающейся; услуга или изделие могут не подходить для удовлетворения конкретных нужд потребителя.

Рассмотрено содержание трёх основополагающих этапов, которые одинаково важны. Любая недоработка в любом из них может создать проблемы с качеством [63].

Система управления качеством продукции опирается на следующие взаимосвязанные категории управления: объект, цели, факторы, субъект, методы, функции, средства, принцип, вид, тип критериев и т.д. [64].

Под управлением качеством продукции понимают постоянный, планомерный, целеустремленный процесс воздействия на всех уровнях на факторы и условия, обеспечивающий создание продукции оптимального качества и полноценное её использование.

**Система управления качеством продукции** включает следующие функции:

1. Функции стратегического, тактического и оперативного управления.
2. Функции принятия решений, управляющих воздействий, анализа и учёта, информационно-контрольные.

3. Функции специализированные и общие для всех стадий жизненного цикла продукции.
4. Функции управления по научно-техническим, производственным, экономическим и социальным факторам и условиям.

**Стратегические функции включают:**

- прогнозирование и анализ базовых показателей качества;
- определение направлений проектных и конструкторских работ;
- анализ достигнутых результатов качества продукции;
- анализ информации о рекламациях;
- анализ информации о потребительском спросе.

**Тактические функции включают:**

- управление сферой производства;
- поддержание на уровне заданных показателей качества;
- взаимодействие с управляемыми объектами и внешней средой [65].

*Система управления качеством продукции* представляет собой совокупность управленческих органов и объектов управления, мероприятий, методов и средств, направленных на установление, обеспечение и поддержание высокого уровня качества продукции.

В 1987 г. Международной организацией по стандартизации (ИСО) при участии США, Канады, ФРГ были разработаны и утверждены пять международных стандартов серии 9000 (по системам качества), в которых были установлены требования к системам обеспечения качества продукции, в том числе к разработке продукции, изготовлению, к организации контроля и испытаний продукции, к ее эксплуатации, хранению и транспортированию [62]. Международные стандарты ИСО 9000 по системам качества включают четыре наименования:

1. ИСО 9000 «Общее руководство качеством и стандарты по обеспечению качества. Руководящие указания по выбору и применению».

2. ИСО 9001 «Система качества. Модель для обеспечения качества при проектировании и (или) разработке, производстве, монтаже и обслуживании».

3. ИСО 9002 «Система качества. Модель для обеспечения качества при производстве и монтаже».

4. ИСО 9003 «Система качества. Модель для обеспечения качества при окончательном контроле и испытаниях».

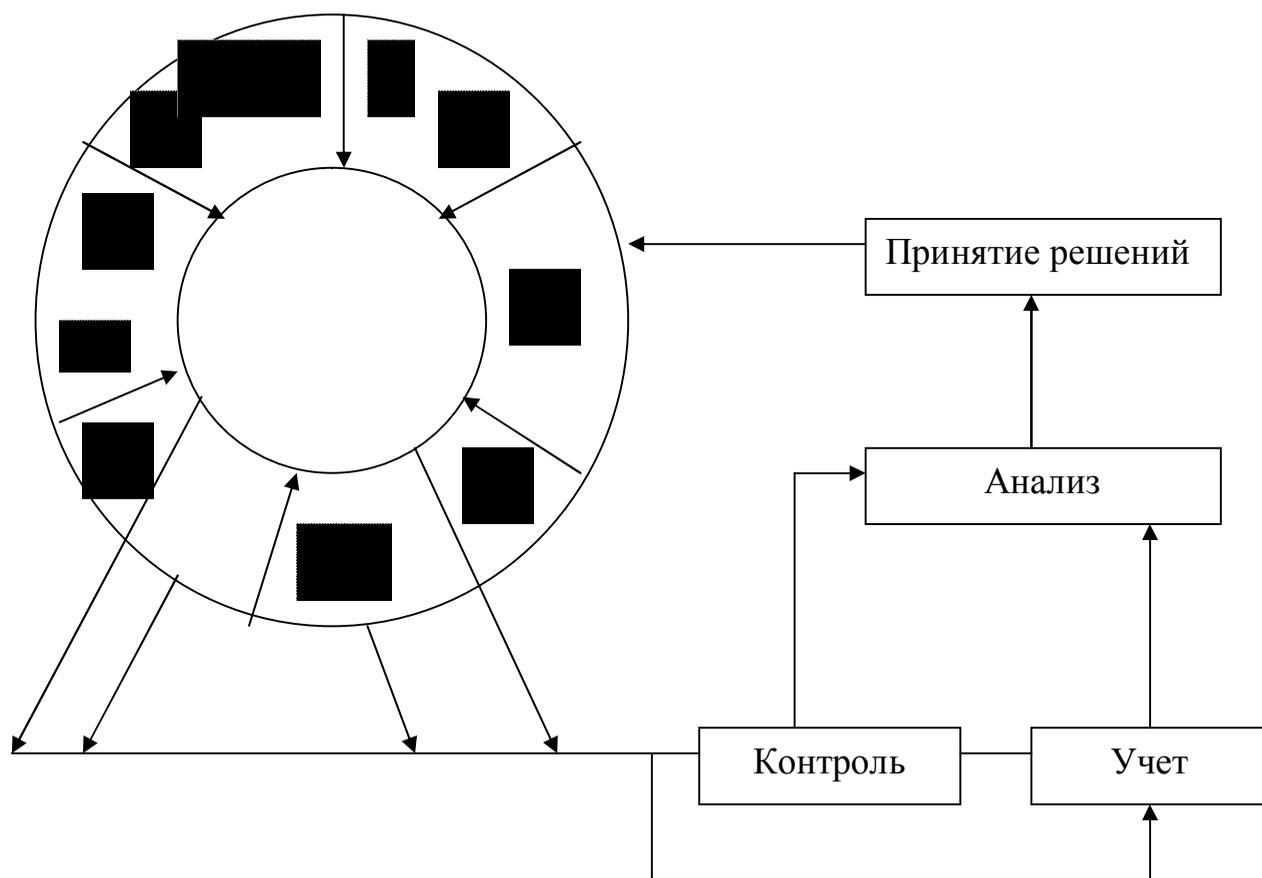
В соответствии со стандартом ИСО жизненный цикл продукции включает одиннадцать этапов:

1. Маркетинг, поиск и изучение рынка.
2. Проектирование и разработка технических требований, разработка продукции.
3. Материально-техническое снабжение.
4. Подготовка и разработка производственных процессов.
5. Производство.
6. Контроль, проведение испытаний и обследований.
7. Упаковка и хранение.
8. Реализация и распределение продукции.
9. Монтаж и эксплуатация.
10. Техническая помощь и обслуживание.
11. Утилизация после использования.

Перечисленные этапы представляются в литературе по менеджменту в виде «петли качества» (рис. 1.1.1).

Таким образом, обеспечение качества продукции – это совокупность планируемых и систематически проводимых мероприятий, создающих необходимые условия для выполнения каждого этапа петли качества, чтобы продукция удовлетворяла требованиям к качеству.

Управление качеством включает принятие решений, чему предшествует контроль, учёт, анализ.

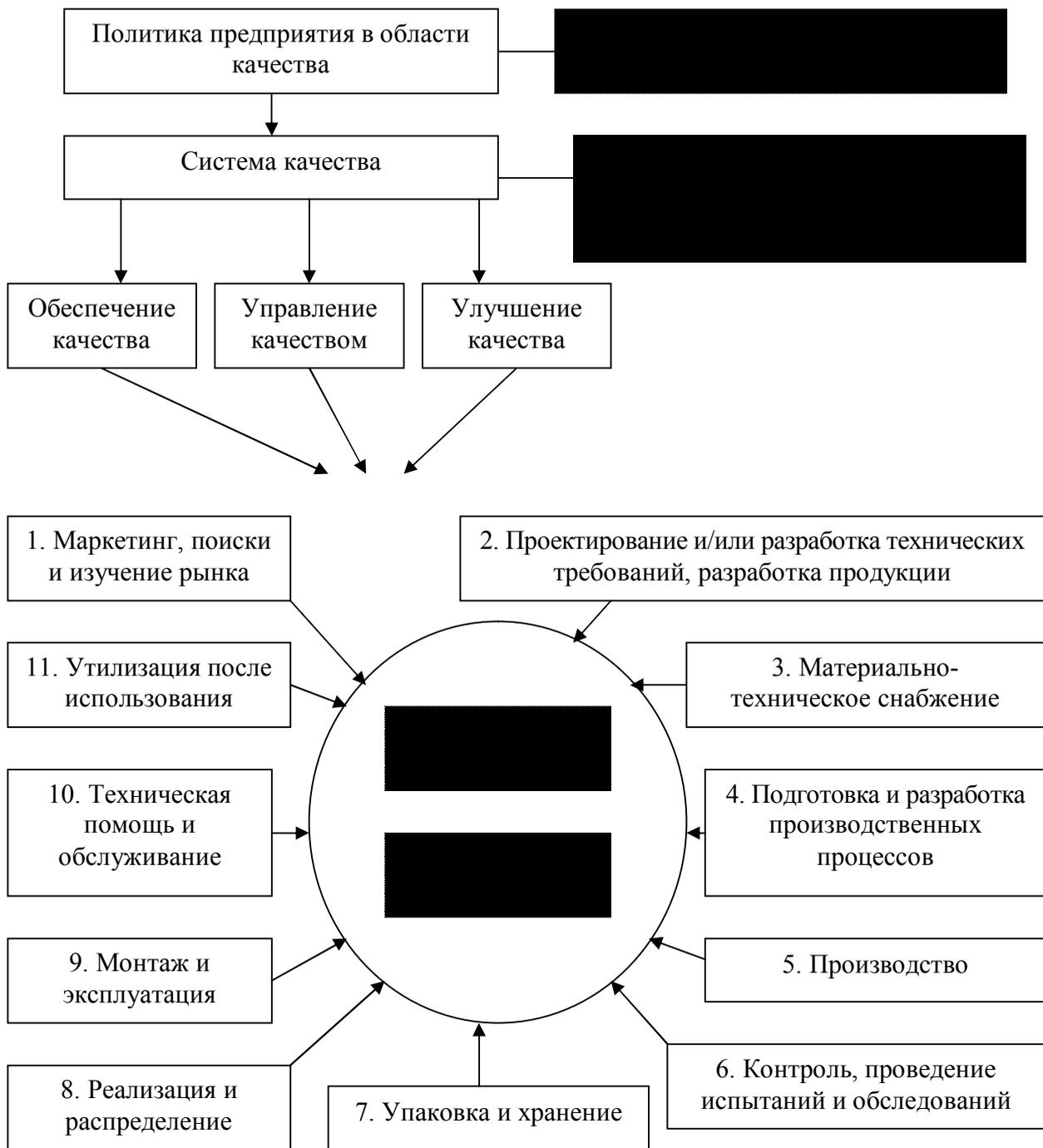


**Рис. 1.1.1. Схема обеспечения качества**

Улучшение качества – постоянная деятельность, направленная на повышение технического уровня продукции, качества её изготовления, совершенствование элементов производства и системы качества.

Механизм управления качеством показан на рис.1.1.2. На этом рисунке система управления качеством представлена в концентрированном виде. Здесь прежде всего выделена политика предприятия в области качества, собственно система качества, включающая обеспечение, управление и улучшение качества [67].

Наилучших результатов в создании и выпуске конкурентоспособной продукции добиваются предприятия, обладающие исчерпывающими сведениями о состоянии и возможностях производственных процессов, а также своевременно вырабатывающие управляющие воздействия по их совершенствованию.



**Рис. 1.1.2. Управление качеством продукции**

По мнению отечественных и зарубежных специалистов, качество продукции закладывается в конструкторской и технологической документации, и та, и другая должны соответствующим образом оцениваться.

Стремление России к интеграции в мировое сообщество, а также развитие рыночных отношений внутри страны предполагают всестороннее

и полное выявление свойств и оценку показателей, определяющих и характеризующих качество продукции и технический уровень производства.

Состав и взаимосвязь основных требований, предъявляемых к производству продукции в нормативной и технической документации, представлены на рис. 1.1.3.

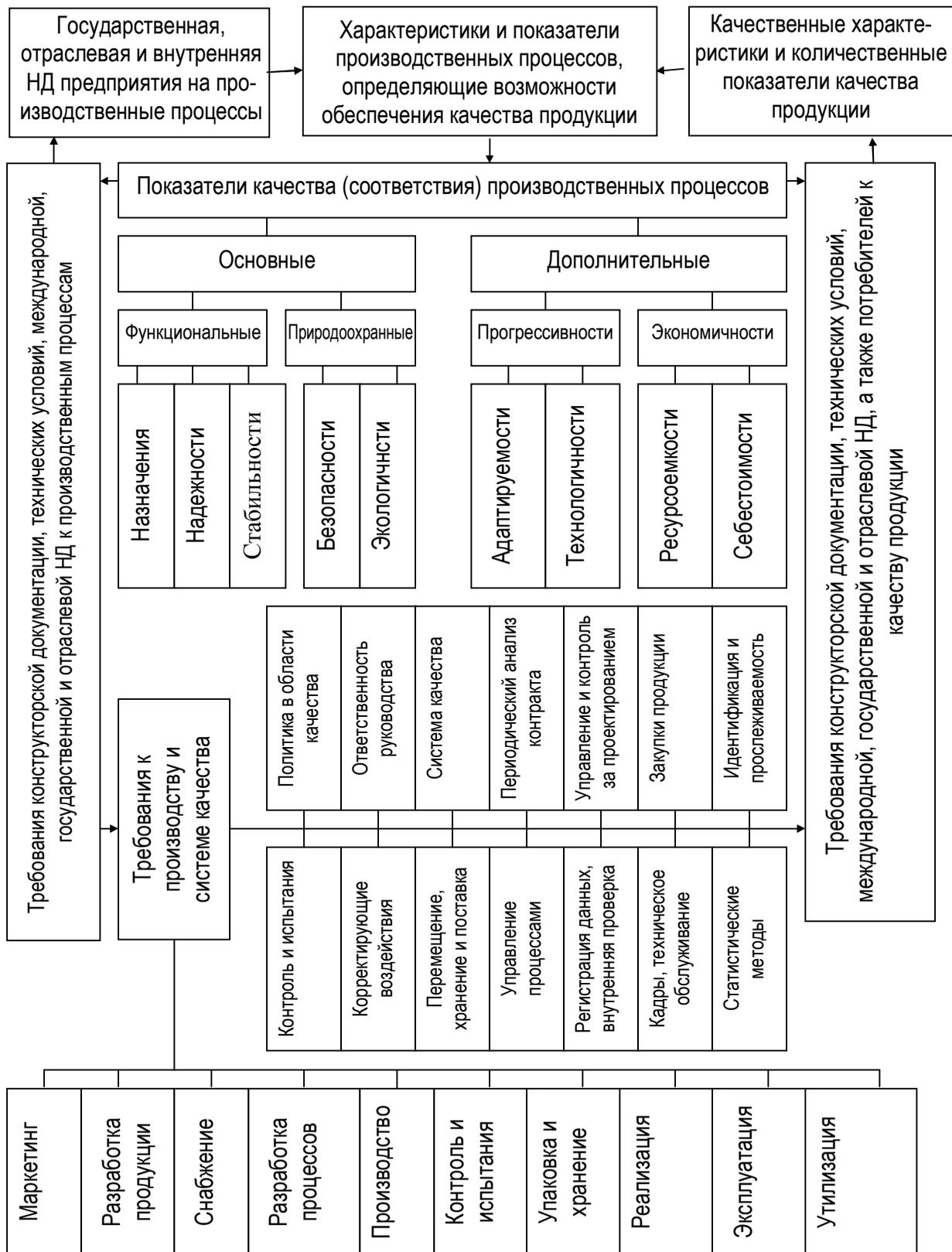
## **Научное обоснование системы управления качеством в процессе производства наноразмерных материалов**

Анализ российских и зарубежных систем управления качеством позволяет заключить:

1. Качество – одна из фундаментальных категорий, определяющих образ жизни, социальную и экономическую основу для развития человека и общества.
2. Проблема качества не может быть решена без участия ученых, инженеров, менеджеров.
3. Качество является важным инструментом в борьбе за рынки сбыта.
4. Качество определяется действием многих случайных, местных и субъективных факторов.

Для предупреждения влияния случайных факторов на уровень качества необходима *система управления качеством*.

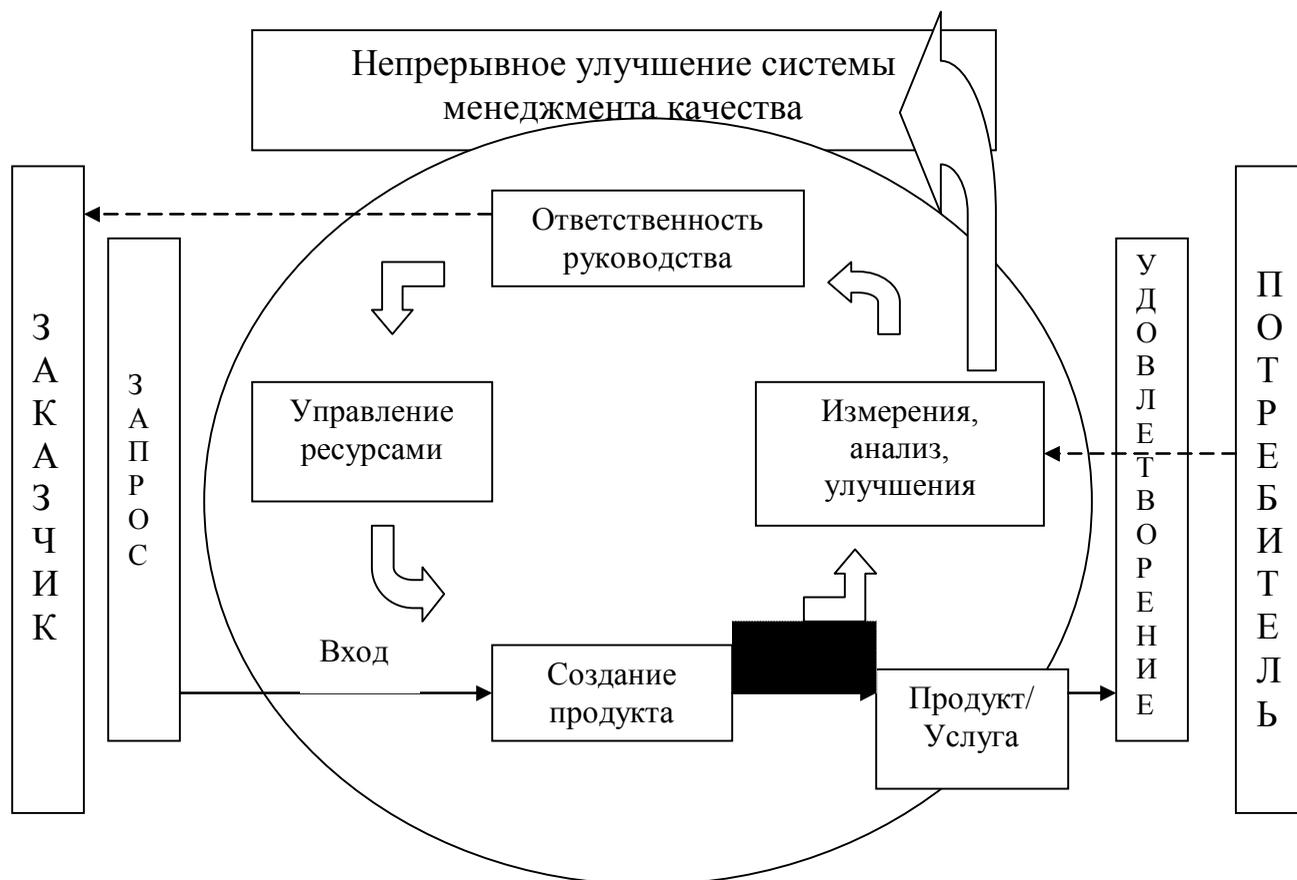
Для качества как объекта менеджмента свойственны все составные части менеджмента: планирование, анализ, контроль, совершенствование продукции. Формирование системы управления качеством в процессе производства углеродных наноматериалов также должно содержать в себе данные составляющие части менеджмента [79]. Таким образом, система



**Рис. 1.1.3.** Основные требования к процессу производства продукции в нормативной и технической документации

управления качеством в процессе производства углеродных наноматериалов, основанная на принципах Всеобщего Менеджмента Качества (рис. 1.1.4), включает в себя следующие этапы:

1. исследование состояния и перспектив развития, оценку конкурентоспособности наноматериалов;
2. исследование и разработка научных основ создания наноматериалов;
3. разработка технологических основ получения наноматериалов, а также технологических основ синтеза эпитаксиальных структур и многослойных токоведущих комплексов с учётом результатов научных исследований;



**Рис. 1.1.4.** Базовые принципы Всеобщего Менеджмента Качества

4. разработка методов контроля качества наноматериалов на этапах технологического производства, при исследовании их свойств, испытаниях, оценке экспериментальных характеристик и показателей надёжности;
5. анализ системы управления качеством в процессе производства, направленный на повышение экономической эффективности;
6. разработка стандартов предприятия электронной техники, регулирующих выпуск продукции.

Внедрение данных принципов на российских предприятиях позволит создавать высококачественные наукоёмкие электронные наноконпоненты, предназначенные для использования в различных отраслях промышленности.

### **Автоматизированная информационно-управляющая система управления качеством производства наноматериалов**

Любое современное российское предприятие стоит перед задачей обеспечения надежного управления и хранения всего объема разнородной информации, которая появляется и используется в процессе технической подготовки производства продукции, а также в течение всего жизненного цикла изделия. Эта информация может возникать в различных информационных системах, действующих на предприятии. С точки зрения любого потребителя информации, на каждом этапе жизненного цикла изделия требуемую для работы информацию необходимо получать в нужное время, в нужном виде и степени детализации и в конкретном месте [85]. Такую организацию информационного пространства предприятия принято называть CALS-технологиями.

На сегодня термин «CALS» интерпретируется как Computer Aided Life

Support, что предполагает компьютерное сопровождение всего жизненного цикла изделия и его модификаций – начиная от зарождения идеи его создания и заканчивая утилизацией.

Автоматизированная система управления качеством в процессе производства углеродных наноматериалов (АСУКПП), выступая в роли единого информационного пространства технической подготовки производства на предприятии, должна органично вписываться в модель CALS. Она обеспечивает как создание и накопление информации об изделии на первых этапах его жизненного цикла (разработка, проектирование, производство), так и дозацию информации по технологии промышленного производства, которая автоматически формируется для каждого специалиста, задействованного в процессе проектирования и производства, в нужном месте, в требуемой форме, с заданной степенью детализации и актуальностью на текущий момент времени. АСУКПП должна обеспечивать непрерывный поток информации от конструктора до производства, поддерживая при этом обратные связи, которые позволяют вести параллельную разработку модификаций, устранять конструкторские и технологические ошибки, выявленные на этапе производства, производить «виртуальный» контроль производства.

Важным идеологическим моментом в АСУКПП является возможность реализации в контуре обратной связи автоматизированной системы анализа информации, накопленного в процессе технологического производства. Для этого в структуре должны быть предусмотрены соответствующие базы данных (БД). Анализ содержания данных БД должен позволять проводить оперативную корректировку технологического процесса с целью снижения количества некондиционной продукции [94].

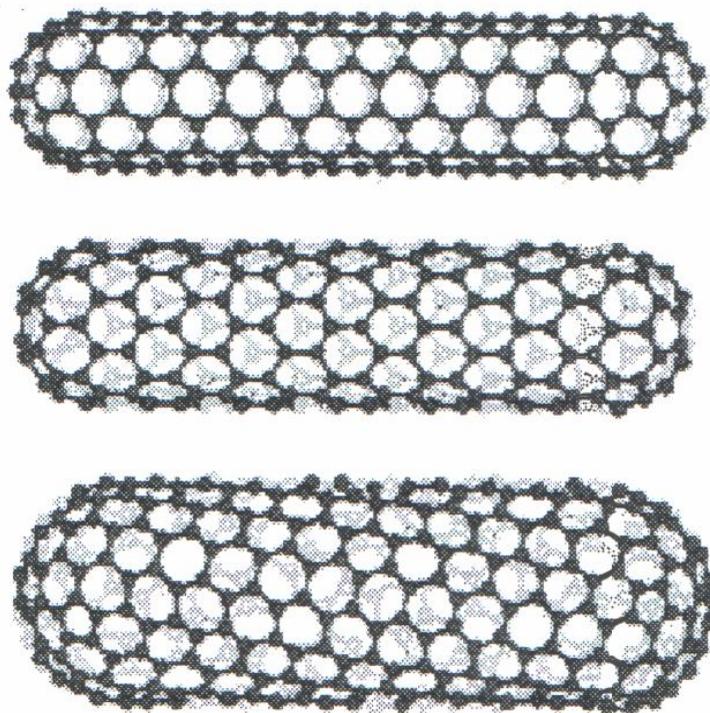
## ***1.2. Структура и свойства наноматериалов***

Многие из перспективных направлений в материаловедении, нано-

технологии, наноэлектронике, прикладной химии связываются в последнее время с фуллеренами, нанотрубками и другими похожими структурами, которые можно назвать общим термином – углеродные каркасные структуры [14]. Что же это такое?

Углеродные каркасные структуры – это большие (а иногда и гигантские!) молекулы, состоящие исключительно из атомов углерода. Можно даже говорить, что углеродные каркасные структуры – это новая аллотропная форма углерода [75]. Главная особенность этих молекул – это их каркасная форма: они выглядят как замкнутые, пустые внутри «оболочки» (рис. 1.2.1).

До 1985 г. об углероде было известно, что он может существовать в природе в двух аллотропных состояниях: 3D форме (структура алмаза) и слоистой 2D форме (структура графита). В вышеуказанном году была открыта новая 0D форма углерода: сферическая структура из 60 атомов



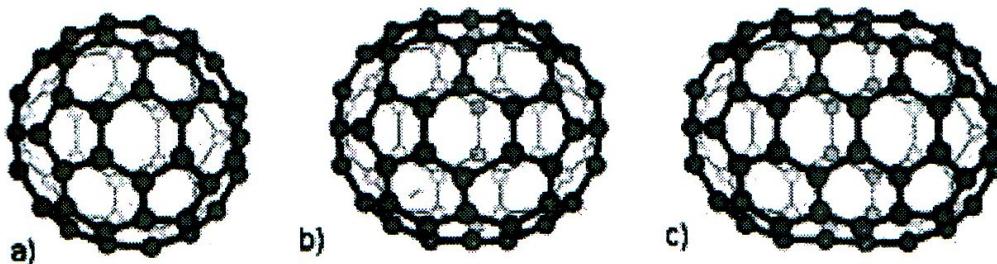
**Рис. 1.2.1.** Углеродные наноструктуры

углерода. Данный новый вид был открыт во время экспериментов по лазерному испарению углерода. Эти углеродные кластеры стали называть фуллеренами (рис. 1.2.2) [76].

Фуллерены – группа специфических молекул, состоящих только из атомов углерода, которые образуют каркас из 12 пятиугольников и нескольких шестиугольников. Своим названием эти соединения обязаны инженеру и дизайнеру Р. Бакминстеру Фуллеру, чьи геодезические конструкции построены по этому принципу. Впервые фуллерены были синтезированы в 1985 г., а в 1992 г. их обнаружили в породах докембрийского периода [18]. Сейчас их интенсивно изучают в лабораториях разных стран, пытаясь установить условия их образования, структуру, свойства и возможные сферы применения. Позже учёным в земных лабораториях удалось синтезировать и исследовать многочисленные производные этих шарообразных молекул. Возникла химия фуллеренов. Некоторые соединения включения в кристаллическую решетку фуллерена  $C_{60}$  оказались «горячими сверхпроводниками» с критической температурой до 117 К. Фуллерены обладают необычными химическими и физическими свойствами. Так, при высоком давлении  $C_{60}$  становится твёрдым, как алмаз [21]. Его молекулы образуют кристаллическую структуру, как бы состоящую из идеально гладких шаров, свободно вращающихся в гранцентрированной кубической решетке. Благодаря этому свойству  $C_{60}$  можно использовать в качестве твердой смазки. Фуллерены обладают также магнитными и сверхпроводящими свойствами.

Ведутся попытки создать на основе фуллеренов материалы для современной молекулярной электроники. Всё это интересно и важно. Но фуллерены, как выяснилось, есть и в земных породах. А последние открытия геохимиков заставляют вернуться к проблеме происхождения фуллеренов.

В 1991 г. японец Сумио Лиджима обнаружил другую новую 1D форму углерода: продолговатые трубчатые образования, названные «нано-



**Рис. 1.2.2.** Углеродные наноструктуры Некоторые представители семейства фуллеренов:

а) Фуллерен  $C_{60}$ ; б) Фуллерен  $C_{70}$ ; в) Фуллерен  $C_{80}$

трубками» [42]. Следует отметить, что примерно в это же время российские учёные объявили об открытии нанотрубок и их связок, имеющих, однако, намного меньший коэффициент отношения длины к диаметру и напоминавших скорее продолговатые фуллерены. Данные структуры состоят из сетки атомов углерода в форме гексагонов, и могут рассматриваться как цилиндр, скрученный из планарной графитовой плоскости.

Нанотрубки обладают уникальными свойствами: они имеют очень малую массу и в то же время рекордно высокий модуль упругости (до 1 ТПа). Нанотрубки на данный момент являются прочнейшими волокнами, которые когда-либо могли быть сделаны [69]. При этом их можно произвольно закручивать: они не ломаются, а только гнутся. Для лучшего понимания структуры и свойств нанотрубок следует начать с рассмотрения в качестве первого приближения структуры графита.

## Свойства углеродных нанотрубок

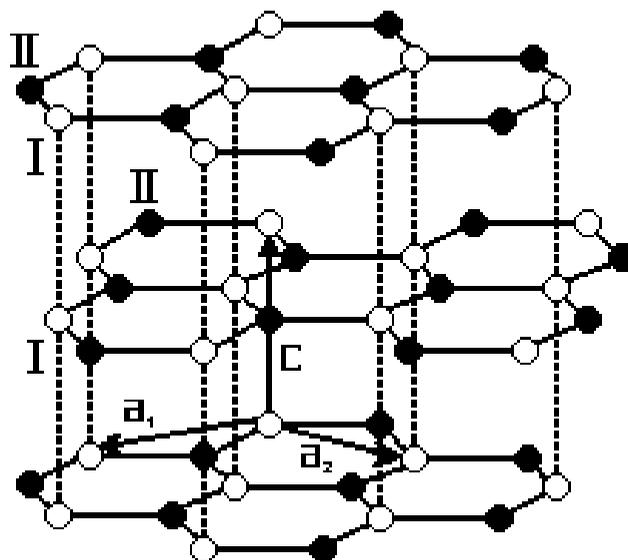
### Структура графита

Графит имеет слоистую структуру, в которой каждый слой форми-

рован их сетки гексагонов с расстоянием между ближайшими соседями  $d_{c-c}=0,142$  нм. Слои располагаются в АВВВ... последовательности (рис. 1.2.3), где атомы I – это атомы, лежащие непосредственно над атомами в смежных плоскостях, а атомы II – лежащие над центрами гексагонов в смежных областях. На этом рисунке также показана результирующая кристаллографическая структура, где  $a_1$  и  $a_2$  – единичные векторы в графитовой плоскости,  $c$  – единичный вектор, перпендикулярный гексагональной плоскости. Расстояние между плоскостями в решётке равно  $0,337$  нм [75].

### Угол хиральности и диаметр нанотрубок

Существует ограниченное число схем, с помощью которых из графитового слоя можно выстроить нанотрубку. Рассмотрим точки A и A' на рис. 1.2.4. Вектор, соединяющий A и A' определяется, как  $c_h=na_1+ma_2$ , где  $n, m$  – действительные числа,  $a_1, a_2$  – единичные векторы в графитовой



**Рис. 1.2.3.** Кристаллографическая структура графита. Решетка определяется единичными векторами  $a_1, a_2$  и  $c$

плоскости. Трубка образуется при сворачивании графитового слоя и соединении точек А и А'. Тогда она определяется единственным образом вектором  $c_h$ .

Диаметр трубки  $d$ , равный

$$d = \frac{|c_h|}{\pi} = \frac{a}{\pi} \sqrt{n^2 + nm + m^2},$$

где  $a = \sqrt{3} \cdot d_{c-c}$  – постоянная решетки и угол  $\theta = \arctg\left[-\sqrt{3} \frac{m}{2n+m}\right]$  количественно определяют хиральность нанотрубки, как показано на рис. 1.2.4. Можно выделить два предельных случая: нанотрубки типа «кресло» и «зигзаг» (рис. 1.2.5) [86].

«Зигзаг» определяется углом  $\theta = 0^\circ$ , что соответствует вектору  $(n, m) = (n, 0)$ . В ней связи С-С идут параллельно оси трубки.

Структура типа «кресло» характеризуется углом  $\theta = \pm 30^\circ$ , соответствующему вектору  $(n, m) = (2n, -n)$  или  $(n, n)$ . Эта группа трубок будет иметь С-С связи, перпендикулярные оси трубки. Остальные комбинации формируют трубки хирального типа, с углами  $0^\circ < \theta < 30^\circ$ . Как видно из рис. 1.2.4, шаг спирали  $P$  зависит от угла  $\theta$ .

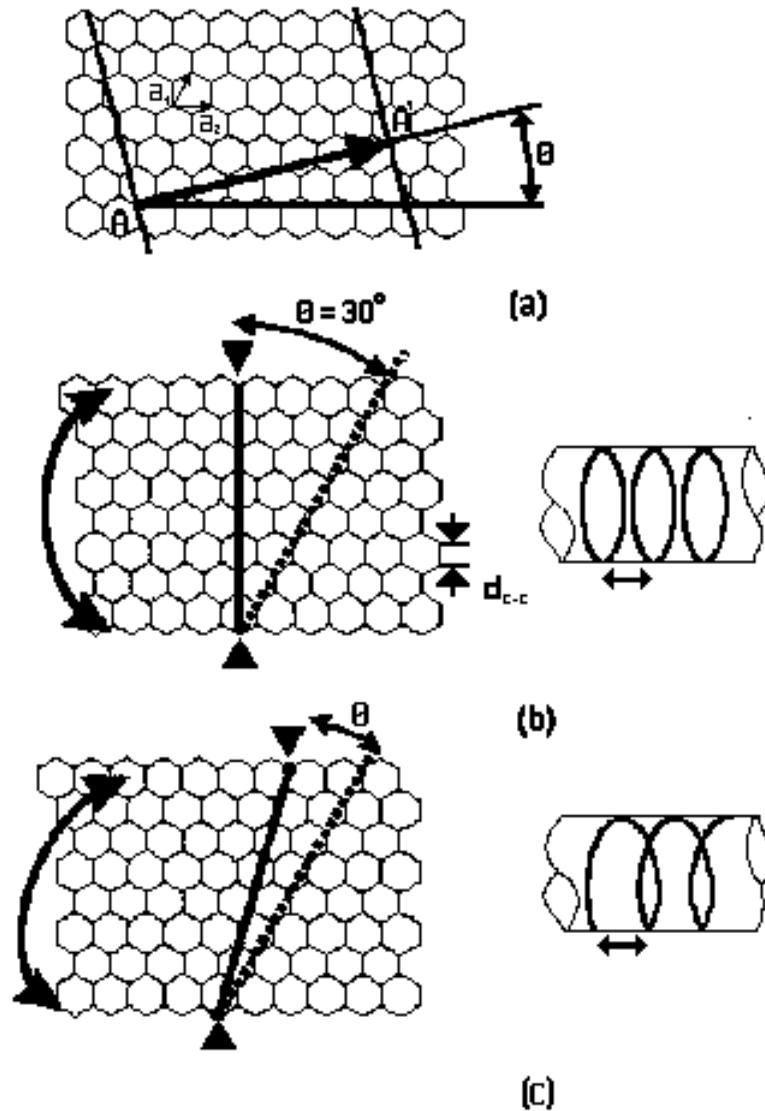
Одиночная трубка обычно называется однослойной нанотрубкой (ОСНТ). Известный наименьший диаметр нанотрубки – 0,7 нм, что является диаметром молекулы фуллерена  $C_{60}$ . Нанотрубки, состоящие из коаксиальных графитовых цилиндров, с расстоянием между слоями 0,34 нм, называются многослойными нанотрубками (МСНТ).

## Электронная структура нанотрубок

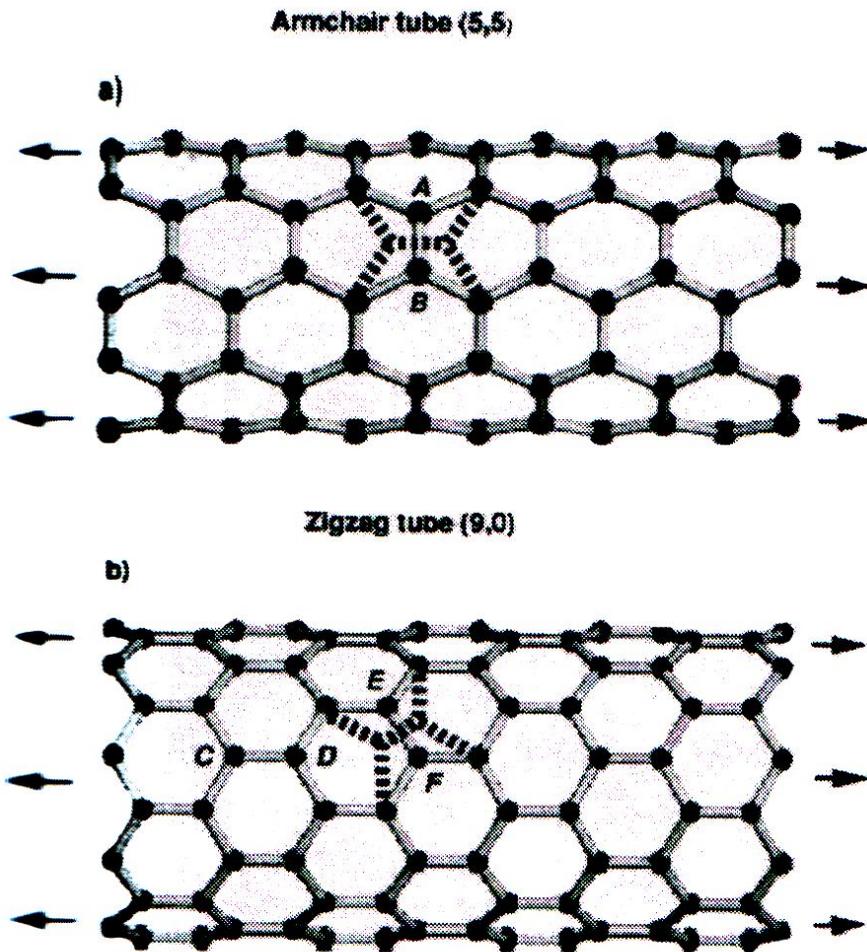
Нанотрубки являются квазиодномерными периодическими структурами. Вдоль окружности на нанотрубку налагается дополнительное гранич-

ное условие, которое можно выразить, как  $c_n \cdot k = 2pl$ , где  $l$  – действительное число [86, 87].

Расчёты зонной структуры предсказывают, что индексы  $(n, m)$ , т.е. диаметр и угол хиральности полностью определяют электрические свойства ОСНТ. Изолированная графитовая плоскость является полупроводником с



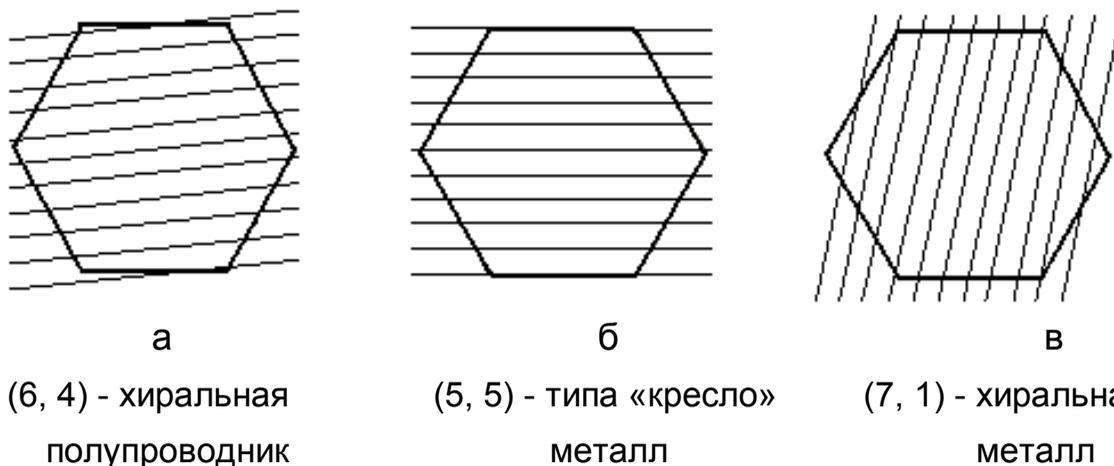
**Рис. 1.2.4.** Углеродные нанотрубки образуются при скручивании графитовых плоскостей в цилиндр, соединяя точку А с А'. Угол хиральности определяется как  $\theta$  (а); трубка типа «кресло» (b); шаг  $P$  зависит от угла  $\theta$  (c)



**Рис. 1.2.5.** Структура нанотрубки типа «кресло» а);  
Структура нанотрубки типа «зигзаг» б)

нулевой запрещенной зоной, чья электронная структура вблизи энергии Ферми задается заполненной  $p$  и пустой  $p^*$  зонами. Эти зоны имеют линейный закон дисперсии, и пересекаются в точке К зоны Бриллюэна на уровне Ферми [74].

Поверхность Ферми идеальной графитовой плоскости состоит из шести угловых К-точек. При образовании трубки, вследствие периодичности граничных условий вдоль окружности, могут быть разрешены только некоторые  $k$ -состояния планарной графитовой поверхности. Разрешенные  $k$ -состояния (рис. 1.2.6) зависят от диаметра и хиральности трубки.



**Рис. 1.2.6.** Периодические граничные условия в нанотрубке приводят к квантованию разрешенных  $k$ -значений.

Нанотрубка может быть: **(а)** полупроводниковой, **(б, в)** металлической в зависимости от того, совпадают ли точки Ферми с разрешенными  $k$ -значениями

Если разрешенные  $k$ -состояния включают точку  $K$ , то система является металлической с плотностью, отличной от плотности состояний на уровне Ферми, что соответствует одномерному металлическому проводнику. Если же точка  $K$  не включена, то система является полупроводниковой с различным размером энергетической щели. Важно отметить, что состояния вблизи энергии Ферми, как в металлических, так и в полупроводниковых трубках, являются состояниями вблизи точки  $K$ , и, следовательно, транспортные и другие свойства нанотрубок определяются состояниями на разрешенных уровнях. Так, например, зона проводимости и валентная зона полупроводниковой нанотрубки исходят из состояний вдоль уровня, ближайшего к точке  $K$ .

Общее правило для определения типа проводимости ОСНТ звучит следующим образом:

- $(n, n)$  трубки всегда металлические (рис. 1.2.6,б);
- $(n, m)$  трубки с  $n - m = 3j$ , где  $j$  не нулевое целое число, являются

полупроводниками с малой шириной запрещенной зоны (рис. 1.2.6,в);  
— все остальные являются полупроводниками с большой шириной запрещенной зоны (рис. 1.2.6,а)

В действительности зонная теория для  $n - m = 3j$  трубок дает металлический тип проводимости, но при искривлении плоскости открывается небольшая щель в случае ненулевого  $j$ . Нанотрубки типа «кресло»  $(n, n)$  в одноэлектронном представлении остаются металлическими вне зависимости от искривления поверхности, что обусловлено их симметрией. С увеличением радиуса трубки  $R$  ширина запрещенной зоны для полупроводников с большой и малой шириной уменьшается по закону  $1/R$  и  $1/R^2$  соответственно. Таким образом, для большинства экспериментально наблюдаемых нанотрубок, щель с малой шириной, которая определяется эффектом искривления, будет настолько мала, что в условиях практического применения все трубки с  $n - m = 3j$  считаются металлическими при комнатной температуре. Таким образом, трубка на рис. 1.2.6,в будет металлической.

Когда диаметр нанотрубок возрастает, все большее число волновых векторов может быть размещено по окружности. Это влечёт за собой превращение нанотрубки в структуру 2D типа. Для трубок малого диаметра было предсказано, то размер щели зависит от диаметра как  $E_{\text{gap}} = 2\gamma_0 d_{\text{c-c}}/d$ , где  $\gamma_0$  есть энергия связи С-С, а расстояние между атомами в графитовой решетке  $d_{\text{c-c}}=0,142$  нм.

Металлы имеют высокую проводимость благодаря тому, что огромное количество электронов может беспрепятственно проникнуть в зону проводимости. В полупроводниках электронам требуется придать дополнительную энергию, чтобы преодолеть энергетическую щель до первого свободного уровня проводимости. Графит является полуметаллом, т.е. имеет низкую проводимость без приложенного внешнего поля, т.к. только малая часть электронов может пройти сквозь узкий канал в зону проводимости.

## Основные характеристики нанотрубок

**Таблица 1.2.1.**

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°C в вакууме и 750°C на воздухе	Металлизация в схемах плавится при 600-1000°C
Цена	500 \$/г	Золото 10 \$/г

## 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 Атланты (Джорджия, США) [90]. В эксперименте использовали индивидуальные многослойные углеродные нанотрубки, полученные стандартным электроразрядным методом. К нанотрубке 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.

## 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. Действительно, не так-то просто изготовить качественные, не содержащие дефектов, нанотрубки, да еще и умудриться при этом посжимать их и порастягивать. Согласно теоретическим расчетам, критический уровень деформации однослойных нанотрубок составляет 5-6 %. В принципе, это согласуется с немногочисленными имеющимися на сегодня экспериментальными данными. 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.

### ***1.3. Методы производства наноматериалов Структура и свойства наноматериалов***

С момента обнаружения первых нанотрубок, основные усилия ученых сфокусировались на разработке метода, позволяющего производить нанотрубки с заданными свойствами и в больших количествах. Именно отсутствие такого метода объясняет высокую стоимость нанотрубок на сегодняшний день, связанную со сложной и дорогой технологией очищения и выделения нанотрубок, полученных каким-либо методом.

Можно выделить четыре основополагающих метода, использующихся в технологии получения углеродных наноструктур:

— метод термического разложения графита в дуговом разряде;

- метод химического осаждения из газовой фазы с использованием катализатора;
- метод лазерного испарения графита;
- метод холодной деструкции графита [12].

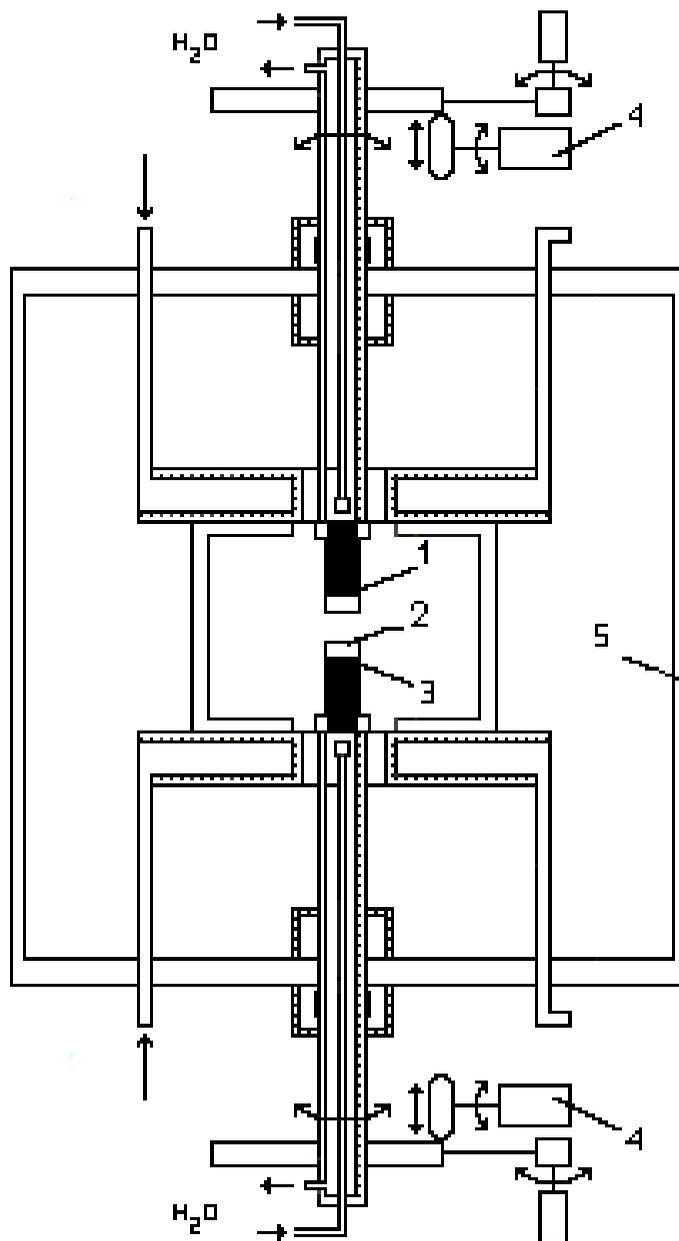
Отметим, что существует широкий набор их модификаций, возможны также сочетания нескольких методов в одном процессе.

### **Термическое разложение графита в дуговом разряде**

Этот метод основан на образовании углеродных нанотрубок при термическом распылении графитового электрода в плазме дугового разряда, горящего в атмосфере гелия. Этот метод, как и метод лазерного распыления, лежащий в основе эффективной технологии получения модифицированных углеродных структур, позволяет получать нанотрубки в количестве, достаточном для детального исследования их физико-химических свойств [84].

Трубка может быть получена из протяженных фрагментов графита, которые далее скручиваются в цилиндр. Для образования протяжённых фрагментов необходимы специальные условия нагрева графита. Оптимальные условия получения нанотрубок реализуются в дуговом разряде при использовании электролизного графита в качестве электродов.

Распыление графита осуществляется при пропускании через электроды тока с частотой 60 Гц, величина тока от 50 до 100 А, напряжение 10-20 В. Регулируя натяжение пружины, можно добиться, чтобы основная часть подводимой мощности выделялась в дуге, а не в графитовом стержне. Камера заполняется гелием с давлением 100-500 Па. Скорость испарения графитового анода в этой установке может достигать 10 г/В [13]. При этом поверхность медного кожуха, охлаждаемого водой, покрывается продуктом испарения графита, т.е. графитовой сажой. Если получаемый порошок соскоблить и выдержать в течение нескольких часов в кипящем толуоле, то



**Рис. 1.3.1.** Структура Схема электродуговой установки для производства наноразмерных материалов:

- 1 – графитовый анод;
- 2 – дуга;
- 3 – графитовый катод;
- 4 – устройство регулировки межэлектродного зазора;
- 5 – охлаждаемая стенка рабочей камеры

получается тёмно-бурая жидкость. При выпаривании ее во вращающемся испарителе получается мелкодисперсный порошок, вес его составляет не более 10% от веса исходной графитовой сажи, в нём содержится до 10% фуллеренов и нанотрубок.

В описанном способе получения нанотрубок гелий играет роль буферного газа. Атомы гелия наиболее эффективно по сравнению с другими атомами «тушат» колебательные движения возбужденных углеродных фрагментов, препятствующих их объединению в стабильные структуры. Кроме того, атомы гелия уносят энергию, выделяющуюся при объединении углеродных фрагментов. Опыт показывает, что оптимальное давление гелия находится в диапазоне 100 Па. При более высоких давлениях агрегация фрагментов углерода затруднена.

На выход нанотрубок влияет множество факторов. Среди различных продуктов термического распыления графита (фуллерены, наночастицы, частицы сажи) наибольшая часть (несколько процентов) приходится и на многослойные нанотрубки, которые частично прикрепляются к охлаждаемым поверхностям установки, частично осаждаются на поверхности.

Наиболее важным является давление гелия в реакционной камере, которое в оптимальных, с точки зрения производства нанотрубок, условиях составляет 500 Па, а не 100-150 Па, 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 во избежание растрескивания анода и его равномерного испарения, что влияет на содержание нанотрубок в катодном депозите.

Однослойные нанотрубки образуются при добавлении в анод небольшой примеси Fe, Co, Ni, Cd (т.е. добавлении катализаторов). Кроме того, ОСНТ получают при окислении МСНТ. С целью окисления, МСНТ обра-

батываются кислородом при умеренном нагреве, либо кипящей азотной кислотой, причем в последнем случае происходит удаление пятичленных графитовых колец, приводящее к открытию концов трубок. Окисление позволяет снять верхние слои с МСНТ и открыть её концы. Так как реакционная способность наночастиц выше, чем у нанотрубок, то при значительном разрушении углеродного продукта в результате окисления доля нанотрубок в оставшейся ее части увеличивается.

Изменение параметров процесса и конструкции установки ведёт к изменению эффективности процесса и состава продукта. Выход годных в различных модификациях колеблется от 20% до 80%, тем не менее нанотрубки растут достаточно короткими, а в случае ОСНТ наблюдаются множественные дефекты.

Использование автоматического устройства поддержания межэлектродного расстояния на фиксированном уровне способствует увеличению стабильности параметров дугового разряда и обогащению нанотрубками материала катодного депозита.

### **Химическое осаждение углерода из газовой фазы**

Метод плазмохимического осаждения углерода из газовой фазы (ПХО) основан на том, что газообразный источник углерода (чаще всего метан, ацетилен или монооксид углерода) подвергают воздействию какого-либо высокоэнергетического источника (плазмы или резистивно-нагреваемой катушки) для того, чтобы расщепить молекулу на реакционно-активный атомарный углерод. Далее происходит его распыление над разогретой подложкой, покрытой катализатором (обычно это переходные металлы первого периода Fe, Co, Ni и др.), на котором осаждается углерод [15]. Нанотрубки образуются только при строго соблюдаемых параметрах. Точное воспроизведение направления роста нанотрубок и их позиционирова-

ние на нанометровом уровне может быть достигнуто только используя метод каталитического ПХО. Возможен также точный контроль за диаметром нанотрубок и их скоростью роста. В зависимости от диаметра частиц катализатора могут расти исключительно ОСНТ или МСНТ. На практике данное свойство широко используется в технологии создания зондов для сканирующей зондовой микроскопии. Задавая положение катализатора на конце кремниевой иглы, выращивается нанотрубка, которая значительно улучшает воспроизводимость характеристик и разрешающую способность микроскопа, как при сканировании, так и при проведении литографических операций.

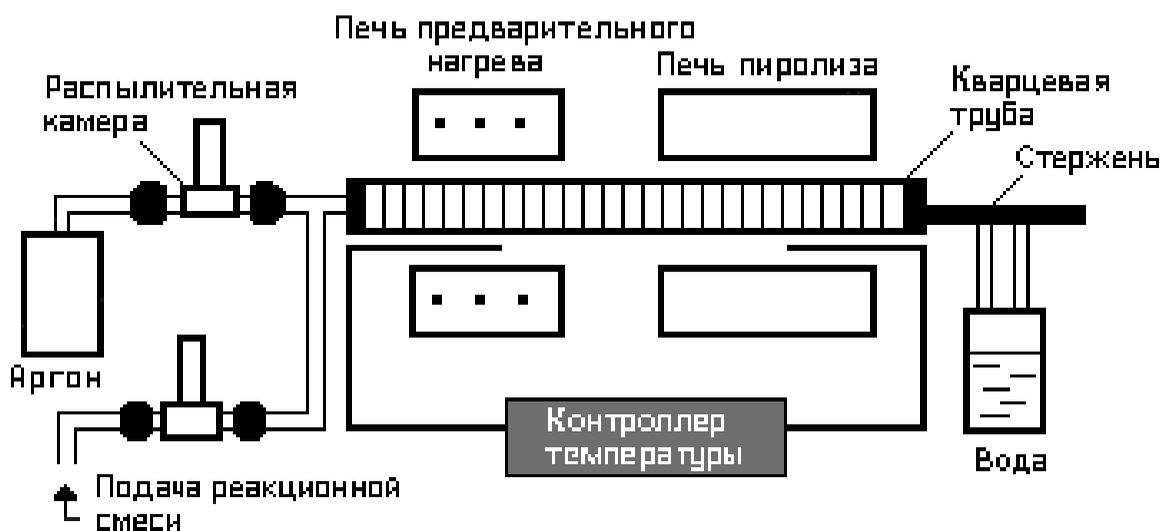
Обычно синтез нанотрубок по ПХО методу происходит в два этапа:

1. приготовление катализатора;
2. собственно рост нанотрубок.

Нанесение катализатора осуществляется распылением переходного металла на поверхность подложки, а затем, используя химическое травление или отжиг, инициализируют формирование частиц катализатора, на которых в дальнейшем происходит рост нанотрубок. Температура при синтезе нанотрубок варьируется от 600 до 900°C. Выход годных составляет примерно 30%. В последнее десятилетие были разработаны различные методы, основанные на ПХО: плазменно-индуцированное ПХО, термическое ПХО, спиртосодержащее каталитическое ПХО, рост в паровой фазе, лазерно-индуцированное ПХО.

Среди данных методов следует отметить метод каталитического пиролиза углеводородов, в котором возможно реализовать гибкое и раздельное управление условиями образования нанотрубок (рис. 1.3.2).

Широко используемый способ получения нанотрубок основан на использовании процесса разложения ацетилена ( $C_2H_2$ ) в присутствии катализаторов. В качестве катализаторов использовались частицы металлов Ni, Co, Cu и Fe размером несколько нанометров. Обычно используется железо,



**Рис. 1.3.2.** Схема установки каталитического пиролиза углеводородов

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  $250^{\circ}\text{C}$ , происходит испарение углеводорода и начинается процесс разложения металлсодержащей соли. Далее аэрозоль попадает в зону печи пиролиза, температура в которой составляет  $900^{\circ}\text{C}$ . 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 [55].

На эксперименте с никелевым катализатором наблюдались четыре типа структур:

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.

Выход и структура нанотрубок зависит от типа катализатора – замена Co на Fe даёт 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 образцах 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.

Было предложено использовать в качестве подложки для роста

нанотрубок пористый кремний, поры которого заполнены наночастицами железа. Подложка помещалась 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 формировались ориентированные многослойные нанотрубки.

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 нм с расстоянием между трубками 100 нм. Они имеют хорошо графитизированную структуру с расстоянием между графеновыми слоями 0,34 нм. Сообщается, что изменяя параметры и время обработки алюминиевой подложки можно менять как диаметр нанотрубок, так и расстояние между ними.

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 нанесённой 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 многослойные нанотрубки большого диаметра (больше 100 нм) прямые, и их оси направлены строго перпендикулярно подложке. Когда диаметр нанотрубки становится меньше 100 нм, преимущественная ориентация, перпендикулярная плоскости подложки, исчезает. Ориентированные массивы многослойных нанотрубок могут создаваться на площадях в несколько квадратных сантиметров.

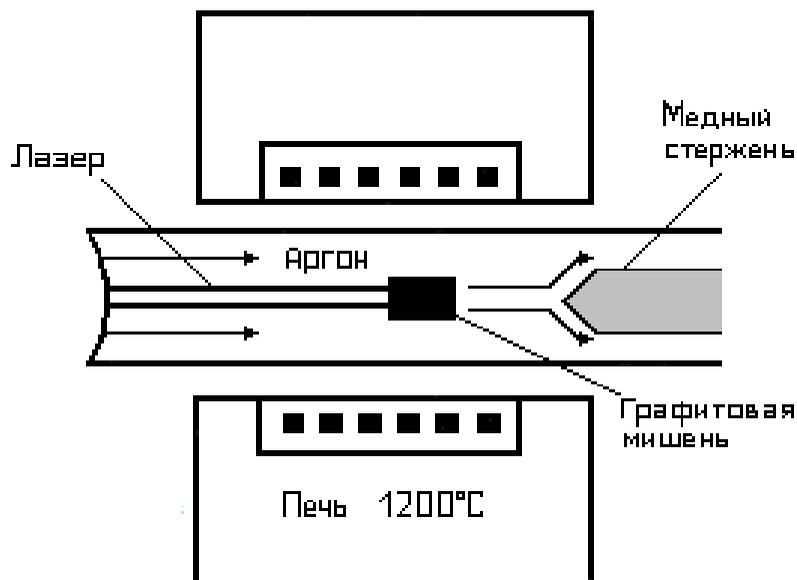
## Метод лазерного испарения графита

В 1995 г. появилось сообщение о синтезировании углеродных нанотрубок методом распыления графитовой мишени под воздействием импульсного лазерного излучения в атмосфере инертного газа.

Графитовая мишень находится в кварцевой трубке при температуре 1200°C, по которой течет буферный газ. В данном методе синтезируются в основном ОСНТ при испарении смеси углерода и переходных металлов лазерным лучом из мишени, состоящей из сплава металла с графитом. Фокусирующейся системой линз лазерный пучок 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 [58]. По сравнению с методом дугового разряда, прямое испарение позволяет обеспечить более детальный контроль условий роста, проводить длительные операции и производить нанотрубки с большим выходом годных и лучшего качества. Фундаментальные же принципы, лежащие в основе производства ОСНТ методом лазерного испарения такие же, как и в методе дугового разряда: атомы углерода начинают скапливаться и образовывать соединения в месте нахождения частиц металлического катализатора.

В установке, изображённой на рис. 1.3.3, сканирующий лазерный луч фокусируется в пятно диаметром 6-7 мм на металл-графит содержащую мишень. Мишень помещается в наполненную аргоном трубу при повышенном давлении и нагретую до 1200°C. Сажа, которая образуется при лазерном испарении, уносится потоком аргона из зоны высокой температуры и осаждается на охлаждаемый водой медный коллектор, находящийся на выходе из трубы.

Изготовление мишени требует нескольких сложных шагов при сме-



**Рис. 1.3.3.** Схема установки лазерного испарения графита

шивании металлической крошки и порошка природного графита с выдерживанием в течение нескольких часов под высоким давлением при температуре 1200°C.

Как и в методе термического разложения получается несколько видов конечного материала:

- в экспериментах, где в качестве мишени использовался чистый графит, получались многослойные 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. 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.
- при добавлении в мишень небольшого количества переходных металлов в продуктах конденсации наблюдались однослойные нанотрубки. Однако

в процессе испарения мишень обогащалась металлом, и выход ОСНТ снижался. 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 и их смеси с другими элементами. Получаемые однослойные 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 однослойных нанотрубок. Смесь Co/Cu даёт низкий выход ОСНТ, а использование чистой меди вообще не приводит к формированию ОСНТ. На окончаниях нанотрубок, свободных от частиц катализатора, наблюдались сферические шапочки [70].

В результате при лазерном испарении было обнаружено образование ОСНТ с большим процентом выхода годных. Оптимизация процесса позволила увеличить процент выхода ОСНТ до 70%. В качестве катализатора были взяты различные металлы и их смеси. В зависимости от выбранного катализатора и давления инертного газа получались разные структуры. При повышении давления появлялись и начинали доминировать однослойные нанотрубки диаметром 1-2 нм, они были объединены в пучки диаметром до 20 нм с поверхностью, свободной от аморфного углерода.

### **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 [78].

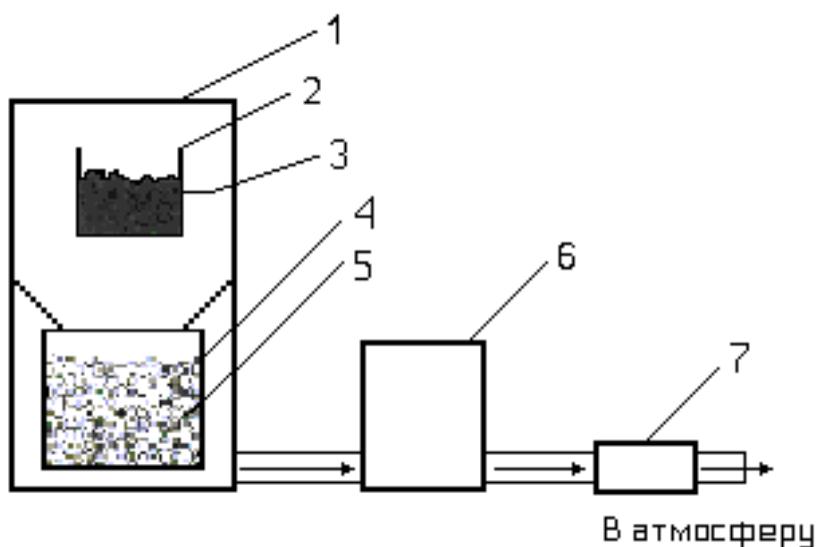
Для обеспечения оптимального режима изготовления смеси весовое соотношение исходного графитсодержащего сырья и галоген-кислородного соединения равно 2:1.

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  $n = 1 - 4$ , 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, иницируется 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 иници-

руется до взрыва. В результате энергий, высвобождаемых микровзрывом, происходят разрывы не только Ван-дер-ваальсовых, но и межтомарных связей с образованием не только свободных радикалов  $C$ ,  $C_2$ ,  $C_3$  и т.д., но и радикалов в виде гексагонов (одного или нескольких) с присоединенными к ним радикалами вида  $C$ ,  $C_2$ ,  $C_3$ ,  $C_4$  и  $C_5$ , обеспечивающими в совокупности высокую реакционную способность получаемой углеродной смеси [88].

После загрузки исходного сырья в установку (рис. 1.3.4), любым из вышеуказанных способов инициируется взрывообразный процесс. В результате преобразования графита и образования углеродной смеси наночастиц объем исходного вещества увеличивается в несколько сотен раз, и готовая смесь, поднимаясь, высыпается через край загрузочной ёмкости (2) и попадает в приёмный сетчатый бункер (4). Образующиеся внутри



**Рис. 1.3.4.** 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

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).

#### **1.4. Устройства на основе наноматериалов**

*Электронно-лучевая трубка на нанотрубках.* Способность нанотрубок к холодной эмиссии электронов под действием приложенного (к тому же умеренного) напряжения легла в основу электронно-лучевой трубки, разработанной недавно группой инженеров из нескольких японских университетов. В качестве электронных эмиттеров использовались нанотрубки двух типов – с открытыми и замкнутыми концами, полученные стандартными методами в электрической дуге. Нанотрубки формировали слой (по типу паутины) толщиной несколько сотых мм. Из этого слоя вырезались иглообразные фрагменты диаметром 0,1 мм и длиной 0,1-0,2 мм, которые укреплялись на конце вольфрамовой либо хромовой проволоки толщиной 0,2 мм с помощью углеродной пасты [16]. Из множества изготовленных таким образом нанотрубок составлялись жгуты, которые и служили электронными эмиттерами.

*Транзистор.* Путем прямых измерений, подкреплённых многочисленными расчетами, выяснилось, что углеродная нанотрубка по своим электронным свойствам представляет собой полупроводник с шириной запрещённой зоны, изменяющейся в зависимости от ее хиральности – угла ориентации графитовой поверхности относительно оси трубки. При этом,

в зависимости от типа хиральности и диаметра, нанотрубки обладают либо металлическими, либо полупроводниковыми свойствами [20].

На основе полупроводниковой или металлической нанотрубок возможно создание полевого транзистора, работающего при комнатной или сверхнизкой температуре соответственно.

В наноустройстве нанотрубка помещается на два тонких платиновых электрода, на которые подается основное напряжение для прохождения тока. В качестве третьего управляющего электрода (затвора) используется слой кремния. В полупроводниковой нанотрубке состояния валентной зоны отделены от зоны проводимости энергетической щелью, поэтому в отсутствие внешнего поля концентрация носителей в зоне проводимости мала и нанотрубка обладает высоким сопротивлением. При подаче на третий электрод электрического поля  $U$  в области нанотрубки возникает электрическое поле, что изменяет энергетическое распределение в зонах, и края зон смещаются относительно поверхности Ферми. При этом концентрация дырок в валентной зоне и электропроводность возрастают. Для потенциала затвора около 6В концентрация дырок достигает максимума, сопротивление – минимума и нанотрубка становится металлической.

При создании полевого транзистора на металлической нанотрубке используются эффекты туннельного переноса электронов через нанотрубку по отдельным молекулярным орбиталям.

*Микросхема из нанотрубок.* Изготовленная схема – многоуровневая. В нижнем уровне расположены кремниевые полевые транзисторы, а в верхнем – транзисторы на нанотрубках. Для выращивания нанотрубок методом химического осаждения из газовой фазы использовали специальный катализатор (неплохим катализатором служит окись алюминия). Нанотрубки имели диаметр 2-4 нм и микронную длину. О хороших характеристиках транзисторов говорить пока не приходится. Затвор находился довольно далеко от нанотрубки, поэтому для открывания канала пришлось

подавать на него напряжение 15В. Отношение тока в открытом состоянии к току в закрытом состоянии, когда напряжение на затворе равно нулю, редко превышало 100 [41].

*Эмиттер.* Электрофизические свойства углеродных нанотрубок дают основание рассматривать их как превосходный материал для получения низковольтных полевых эмиттеров с плотностью эмиссионного тока  $100 \text{ А/см}^2$ . В известных из публикаций конструкциях эмиттеров к нанотрубкам предъявляются жесткие требования – нанотрубки должны быть прямыми, расположенными перпендикулярно площадке, на которой они формируются, и иметь одинаковую длину и диаметр [50]. Выполнение этих условий – сложнейшая технологическая задача. Изогнутость и неперпендикулярность ответственны за нестабильность эмиссионного тока, заставляют увеличивать расстояние между управляющим электродом и эмиттером, что, в свою очередь, приводит к увеличению управляющего напряжения до 100 В. Кроме того, предлагаемые эмиттеры имеют вертикальную конструкцию и плохо подходят для планарных интегральных схем.

*Дисплей.* Углеродные нанотрубки способны испускать электроны, благодаря чему могут найти применение в сверхтонких дисплеях. Они обладают очень низким полевым напряжением, а это означает, что для излучения электронов и получения изображения не нужно подавать высокого напряжения. В настоящее время на мониторы подается очень высокое напряжение, а для этого нужны мощные источники питания. Если же высокого напряжения не требуется, то можно обойтись без мощного источника питания, следовательно, заметно уменьшаются и габариты всего устройства.

Созданы и опробованы прототипы тонких плоских дисплеев, работающих на матрице из нанотрубок [57]. Под действием напряжения, прикладываемого к одному из концов нанотрубки, с другого конца начинают испускаться электроны, которые попадают на фосфоресцирующий экран и вызывают свечение пикселя. Получающееся при этом зерно изображения

будет фантастически малым: порядка одного микрона!

**Диоды.** Еще одно применение многослойных нанотрубок – изготовление светодиодов на основе органических материалов. Как уже отмечалось ранее, УНТ образуются из углеродных шестичленных колец, как в графите. Если углеродный шестиугольник заменить на пятиугольник или семиугольник, то нанотрубка изогнётся. При этом изменяется ориентация шестичленных колец по отношению к оси нанотрубки, а значит меняется положение уровня Ферми, ширина запрещённой зоны и проводящие свойства. Слева относительно изгиба нанотрубка будет металлической, а справа – полупроводниковой. Таким образом, изогнутая нанотрубка представляет собой гетеропереход металл-полупроводник [73]. Теперь, если рассмотреть левый и правый отрезок нанотрубки изолированно с разных сторон относительно её изгиба, то электроны на уровне Ферми обладают разной энергией. При объединении отрезков выигрыш в энергии приведёт к перетеканию заряда и образованию потенциального барьера. Электрический ток будет течь только в том случае, если электроны переходят из области нанотрубки с большей энергией Ферми в область с меньшей энергией. В результате ток потечёт только в одном направлении. Эта односторонняя проводимость используется для создания выпрямляющего диода. Если нанотрубка находится в контакте с двумя проводниками из золота, нанесёнными на кварцевую подложку, то в зависимости от полярности напряжения электрический ток через устройство либо отсутствует, либо линейно нарастает при увеличении напряжения выше запорного.

**Электромеханический резонатор.** Преобразование электрических колебаний в механические требуется для создания различных устройств, например, электроакустических головок. Для возбуждения колебаний нанотрубки под действием электрического поля её закрепляют на одном из двух электродов, на этот раз под углом ко второму электроду. При подаче на электроды электрического напряжения трубка заряжается, и за счёт

электростатического притяжения отклоняется ко второму электроду. Если на электроды подать переменное напряжение, частота которого совпадает с частотой собственных колебаний нанотрубки, зависящих от её толщины и длины, возникнут механические колебания нанотрубки [90].

**Квантовые провода.** Теоретические и экспериментальные исследования электрических и магнитных свойств нанотрубок обнаружили ряд эффектов, которые указывают на квантовую природу переноса заряда в этих молекулярных проводах и могут быть использованы в электронных устройствах [89].

Проводимость обычного провода обратно пропорциональна его длине и прямо пропорциональна поперечному сечению, а в случае нанотрубки она не зависит ни от её длины, ни от её толщины, и равна кванту проводимости – предельному значению проводимости, которое отвечает свободному переносу делокализованных электронов по всей длине проводника. При обычной температуре наблюдаемое значение плотности тока на два порядка превосходит достигнутую сейчас плотность тока в сверхпроводниках.

**Химическая модификация.** Возможности использования нанотрубок в молекулярной электронике неизмеримо возрастают при переходе от чисто углеродных к химически модифицированным нанотрубкам. Например, благодаря наличию цилиндрической полости внутри углеродных нанотрубок удастся внедрить различные элементы, включая тяжёлые металлы. Возможно добавление аддендов (например, атомов фтора) на внешнюю поверхность трубки. Кроме углеродных, сейчас умеют получать и бор-азотные нанотрубки. Во всех этих случаях должны получаться материалы с новыми и пока ещё экспериментально не изученными свойствами [57].

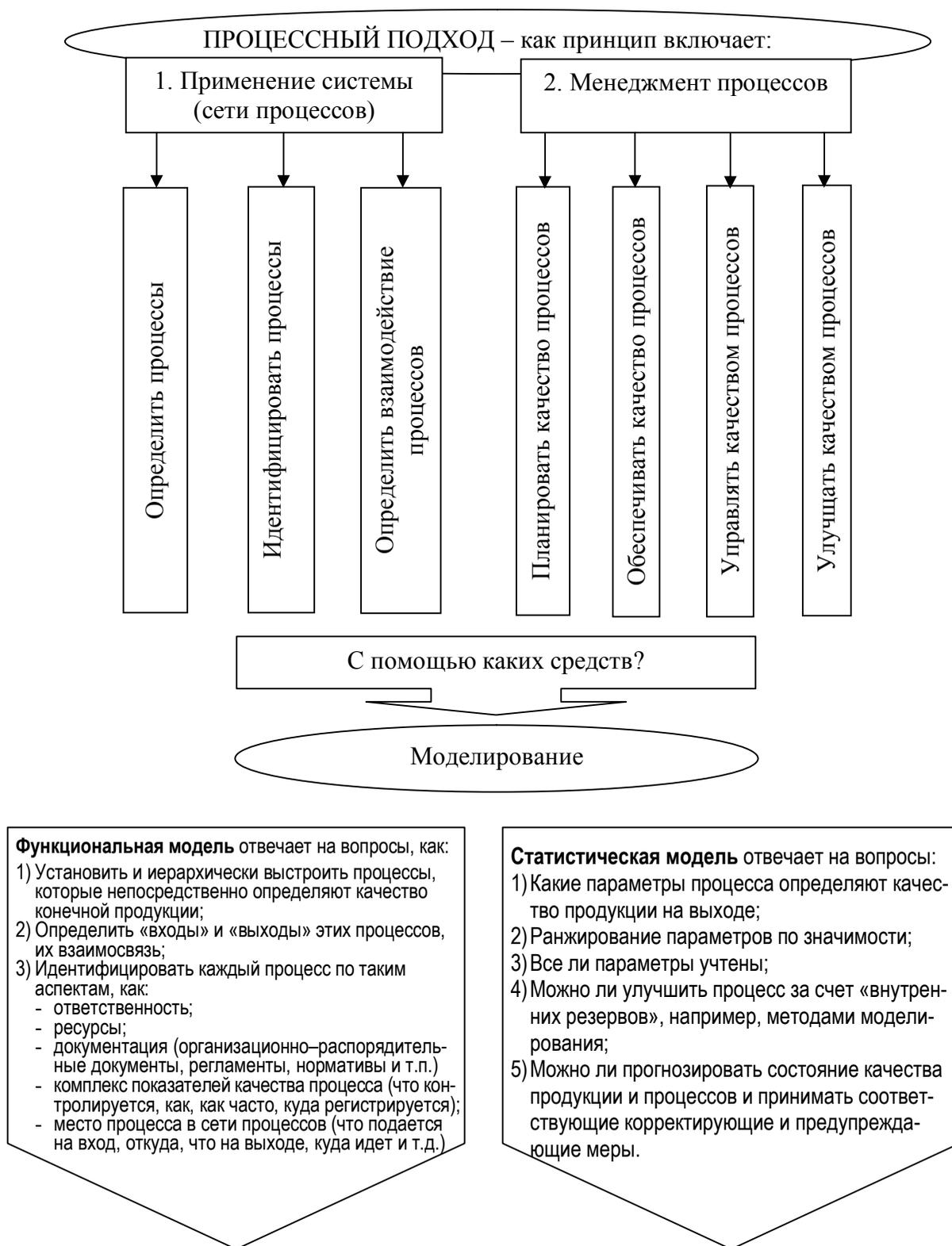
## **1.5. Функциональное моделирование системы менеджмента качества на основе IDEF0 в различных процессах производства наноматериалов и электронных устройств**

Для сложных систем очень сложно получить одно единственное описание, отвечающее на все вопросы, пригодное для достижения всех ключевых целей и задач. Являясь по своей природе многогранной по форме и содержанию представления, система как совокупность взаимосвязанных компонентов может быть описана в виде целого ряда самостоятельных, законченных моделей, количество которых определяется главным образом целями моделирования. Типичный пример сложной системы – предприятие, которое может быть представлено как:

- система процессов, с помощью которой предприятие осуществляет свою деятельность;
- совокупность источников и каналов связи потоков информации и типов данных;
- организационная структура;
- инфраструктура (территории, здания, сооружения, коммуникации).

Соответственно, различные представления предприятия или его системы менеджмента могут быть описаны соответствующими моделями, отвечающими на конкретные вопросы.

Функциональная модель представляет описание с требуемой степенью детализации сети процессов, например, с целью их планирования, обеспечения, управления и улучшения. Модели данных (информационные модели) представляют собой подробное описание объектов и типов информации данных, например, с целью оптимизации и последующей автоматизации и т.д., рис. 1.5.1.



**Рис. 1.5.1. Идеология процессного подхода**

## Методология функционального моделирования IDEF0

Методология IDEF0 является одной из основных методологий функционального моделирования сложных систем и процессов, которая регламентирована рекомендациями Р50.1.028-2001.

Общая методология IDEF состоит из частных методологий, основанных на графическом представлении систем:

- IDEF0 используется для создания *функциональной модели*, отображающей структуру и функции системы, а также потоки информации и материальных объектов, связывающие эти функции.
- IDEF1 применяется для построения *информационной модели*, отображающей структуру и содержание информационных потоков, необходимых для поддержки функций системы.
- IDEF3 является альтернативой широко распространенному подходу описания процессов с помощью блок-схем. Её важным достоинством является возможность описания в рамках единой модели параллельно исполняемых, а также конкурирующих (альтернативно исполняемых) процессов.

Функциональная модель является базовым понятием в методологии IDEF0, основанным на понятии «функция». Функции являются содержанием элементов (блоков), из которых строится модель изучаемой или проектируемой системы.

Понятие функции в методологии IDEF0 есть интерпретация понятия преобразования, где под преобразованием понимается выражение (рис. 1.5.2):

$$Y = WX,$$

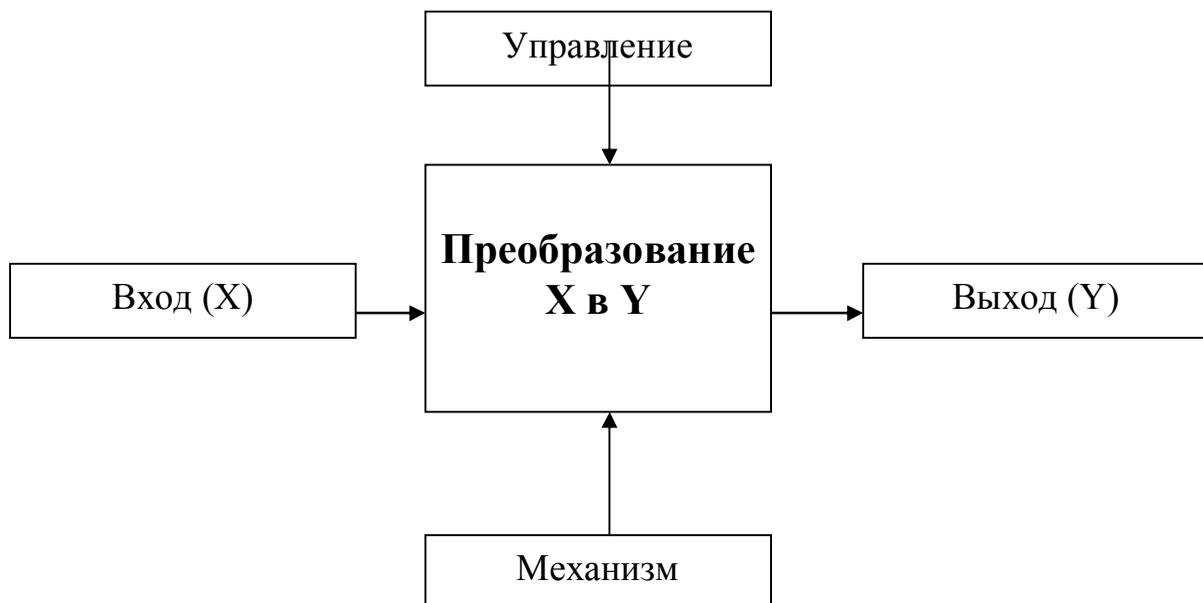
где  $Y$  – множество выходов;  $X$  – множество входов;  $W$  – оператор преобразования, т.е. правило, согласно которому множество  $X$  преобразуется в

множество  $Y$ .

В методологии IDEF0 функция интерпретируется как преобразование входа в выход под воздействием управления и при помощи механизма.

Модель есть образ реального объекта (оригинала), созданный при помощи определенного набора средств. Набором средств для создания образов-моделей реальных систем является графический язык. Модели, создаваемые в вычислительной среде (в компьютере), есть отображение объекта реального мира в мир информационный. В этом качестве модель есть информационный объект.

Модель и оригинал связаны между собой некоторыми условиями соответствия. Выполнение этих условий есть адекватность модели оригиналу, т.е. модель и оригинал находятся между собой в отношении гомоморфизма. Это значит, что всё, что имеется в модели, обязательно имеется и в оригинале. Обратное утверждение неверно: не все свойства и элементы оригинала находят отражение в модели, т.к. модель в каком-то смысле должна быть проще оригинала.



**Рис. 1.5.2.** Графическая интерпретация функции как преобразования в методологии IDEF0

Функциональная модель есть образ реальной системы (оригинала), описанный графическим языком и отображающий функции, выполняемые системой, и отношения (связи) между ними.

Метод IDEF0 предназначен для функционального моделирования, т.е. моделирования выполнения функций объекта, путем создания описательной графической модели, показывающей, что, как и кем делается в рамках функционирования предприятия. Функциональная модель представляет собой структурированное изображение функций производственной системы или среды, информации и объектов, связывающих эти функции.

Модель строится методом декомпозиции: от крупных составных структур к более мелким, простым. Элементы каждого уровня декомпозиции представляют собой процессы (операции, действия) по переработке информационных или материальных ресурсов при определённых условиях с использованием заданных механизмов. Каждая операция раскладывается на более мелкие составляющие – действия по переработке определённой части информационных или материальных ресурсов при определённых условиях с использованием части заданных механизмов. Аналогично раскладываются операции. Последний шаг декомпозиции должен приводить к получению модели, степень детализации которой удовлетворяет требованиям, заданным в начале процесса её создания.

## ГЛАВА 2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ УПРАВЛЕНИЯ КАЧЕСТВОМ ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ УСТРОЙСТВ НА ОСНОВЕ УГЛЕРОДНЫХ СТРУКТУР

### 2.1. Классификация задач оптимизации

В общем виде задача оптимизации может быть записана следующим образом:

$$F(X) \rightarrow \underset{X \in D \subseteq S}{opt}, \quad (2.1.1)$$

где:  $F = \{f_1, f_2, \dots, f_k\}$  – 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 [46].

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 заменя-

ют "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 [53]. 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:

$$F(X) = \{f_1(X), f_2(X), \dots, f_p(X)\} \rightarrow \min_{X \in D \subset R^n}, \quad (2.1.2)$$

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).

(2.1.3)

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

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

$$\Pi = f(D_{\Pi}) = \{(f_1(X), f_2(X), \dots, f_p(X)) \in R^p, X \in D_{\Pi}\} \quad (2.1.4)$$

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 [63]. 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 походы 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

$$\sum_{i=1}^p \lambda_i \cdot f_i(X) \rightarrow \min, \quad \lambda_i \geq 0, \quad \sum_{i=1}^p \lambda_i = 1, \quad (2.1.5)$$

причём, если все  $\lambda_i > 0$ , то данное решение является недоминируемым (точка Парето) для задачи (2.1.4).

Если существует несколько решений задачи (2.1.5) и существуют некоторые  $\lambda_i = 0$ , то решение принадлежит множеству Парето задачи (2.1.4), 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

$$\phi(X, \lambda) = \max_{i \in \{1, \dots, p\}} \lambda_i \cdot f_i(X) \rightarrow \min, \quad \lambda_i > 0, \quad \sum \lambda_i = 1, \quad (2.1.6)$$

то данное решение является недоминируемым (точка Парето) для задачи (2.1.4).

Если существует несколько решений задачи (2.1.6), то решение принадлежит множеству Парето задачи (2.1.4) тогда и только тогда, когда оно не доминируется по совокупности критериев точками из  $Arg \min \phi(X, \lambda)$ . [62].

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

$$\phi_z(X, \bar{\lambda}) = \sum_{i=1}^p \lambda_i \cdot [f_i(X)]^{z_i} \rightarrow \min_{X \in D}, \quad (2.1.7)$$

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.

$$D_{\Pi} = \bigcup_{\bar{\lambda} \in \Lambda} Arg \min \sum_{i=1}^p \lambda_i \cdot [f_i(X)]^{z_i}. \quad (2.1.8)$$

К сожалению, прямое использование результатов этой теоремы чрезмерно затруднено из-за сложности вычисления пороговых величин, для

которых неизвестны оценки снизу, а оценки сверху очень сложно получить.

Если допустимое множество задачи (2.1.4) конечно и все целевые функции принимают значения в интервале  $[c, d]$ ,  $c > 0$ , то множество Парето задачи (2.1.4) совпадает с множеством решений задач

$$\phi_\alpha(X, \bar{\lambda}) = \left[ \sum_{i=1}^p (\lambda_i \cdot f_i(X))^\alpha \right]^{1/\alpha} \rightarrow \min_{X \in D}, \quad (2.1.9)$$

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  $\bar{\alpha} \geq 1$ .

$$D_\Pi = \bigcup_{\bar{\lambda} \in \Lambda^\sigma} \text{Arg min}_{X \in D} \left[ \sum_{i=1}^p (\lambda_i \cdot f_i(X))^\alpha \right]^{1/\alpha}. \quad (2.1.10)$$

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 построению множества Парето, в частности основанные на методе после-

довательного 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

$$f_i(X) \rightarrow \min_{X \in D_i}; \quad (2.1.11)$$

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

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

Тогда, в случае единственности решения задачи (2.1.11)  $X^{*i}$  принадлежит множеству Парето задачи (2.1.4). Если же решением задачи (2.1.11) является конечное множество точек, то среди них найдётся хотя бы одна недоминируемая по совокупности критериев задачи (2.1.4).

**Теорема 2.**

Пусть  $X^0 \in D_{\Pi}$  – решение задачи (2.1.4). 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 (2.1.11).

В многокритериальных задачах дискретной оптимизации с конечным множеством допустимых альтернатив методом варьирования ограни-

чений 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:*

$$f(X) = w_1 f_1(X) + \dots + w_k f_k(X), \quad w_1 + \dots + w_k = 1, \quad w_i > 0. \quad (2.1.12)$$

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) [72].

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:

$$f_i(X) \rightarrow \min_{X \in D},$$
$$f_j(X) \leq \left(1 + \frac{\Delta_j}{100}\right) f_k^*, \quad j = 1, \dots, i-1. \quad (2.1.13)$$

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 [80].

### Дискриминационный подход

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

$$\begin{aligned} f_i(X) &\rightarrow \min_{X \in D}, \\ f_j(X) &\leq \Delta_j, \quad j = 1, \dots, K, \quad i \neq j. \end{aligned} \tag{2.1.14}$$

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 [81].

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  $y_i$ :

$$f(X) = |f_1(X) - y_1| + \dots + |f_k(X) - y_k|. \quad (2.1.15)$$

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  $y_i$ :

$$f(X) = (f_1(X) - y_1)^2 + \dots + (f_k(X) - y_k)^2. \quad (2.1.16)$$

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 [93].

## 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 скаляр-

ная функция, представляющая собой сумму относительных расстояний частных критериев от их минимальных значений:

$$f(X) = \frac{f_1(X) - f_1^*}{|f_1^*|} + \dots + \frac{f_k(X) - f_k^*}{|f_k^*|}, \quad (2.1.17)$$

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:

$$f(X) = \left( \frac{f_1(X) - f_1^*}{f_1^*} \right)^2 + \dots + \left( \frac{f_k(X) - f_k^*}{f_k^*} \right)^2. \quad (2.1.18)$$

Полностью аналогичен предыдущему методу.

Здесь неявно предполагается, что критерии равноценны, что редко бывает на практике [82].

### **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:

$$\max \{ |f_i(X)|, i = 1, \dots, k \} \rightarrow \min. \quad (2.1.19)$$

Пользователь выбирает алгоритм оптимизации для обобщённого критерия.

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

$$\max \{f_i(X), i = 1, \dots, k\} \rightarrow \min. \quad (2.1.20)$$

Пользователь выбирает алгоритм оптимизации для обобщённого критерия.

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

$$\max \{|f_i(X) - y_i|, i = 1, \dots, k\} \rightarrow \min. \quad (2.1.21)$$

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:

$$\max \left\{ \frac{f_i(X) - f_i^*}{|f_i^*|}, i = 1, \dots, k \right\} \rightarrow \min, \quad (2.1.22)$$

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 [91].

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

$$\max \left\{ w_i \left( \frac{f_i(X) - f_i^*}{|f_i^*|} \right), \quad i = 1, \dots, k \right\} \rightarrow \min, \quad (2.1.23)$$

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:

$$\max \{ w_i f_i(X), \quad i = 1, \dots, k \} \rightarrow \min. \quad (2.1.24)$$

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  $f_i^0$ :

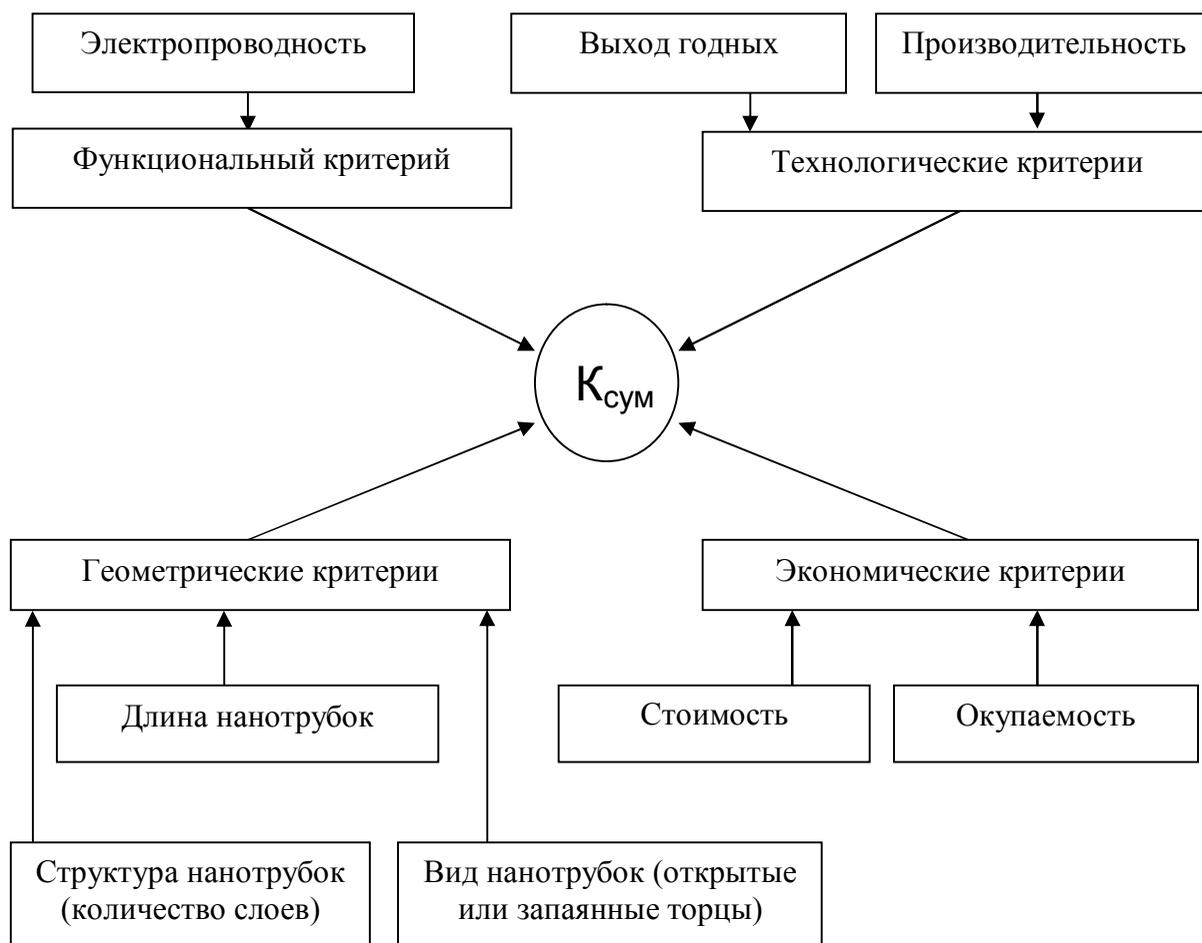
$$\begin{aligned} \min_{x \in X'} f_1(x), \\ X' = X \cap \left\{ x \mid f_i(x) \leq f_i^0, i = 2, \bar{k} \right\}. \end{aligned} \quad (2.1.25)$$

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.

## **2.2. Оптимизация обобщённого критерия качества дискриминационным методом**

Семейство дискриминационных методов предполагает наложение ограничений на критерии на основе некоторой системы требований, предъявляемой к оптимальному решению [1].

Дискриминационный метод применим к оптимизации сложной системы (в данном случае обобщённого критерия качества (рис. 2.2.1)), которая состоит из отдельных подсистем [2]. 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



**Рис. 2.2.1.** Обобщенный критерий качества

высокое качество изготавливаемых нанообъектов, с вероятностями не ниже заданных, при этом затраты не должны превосходить заданной границы [3].

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

$$P(v_i) = \max \left\{ \sum_{p=1}^q g_p(v_i) \right\} \quad (2.2.1)$$

при наличии ограничений:

$$g_p(v_i) = \sum_{j=1}^n g_p(u_j); \quad (p = 1, \dots, q); \quad (i = 1, \dots, k); \quad (j = 1, \dots, n); \quad (2.2.2)$$

$$g_p(u_i) \geq g_p^*, \quad (2.2.3)$$

$$v_i \in V,$$

$$u_i \in U$$

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;

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

$P(v_i)$  – целевая функция – обобщённый 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  $o_j$  типов элементов по возрастанию, согласно значениям рассматриваемого критерия [5]. 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 [51].

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  $l_j+1$  элементов ( $l_j$  резервных и 1 основной), то выход её из строя происходит при выходе из строя всех  $l_j+1$  элементов (так называемое «параллельное резервирование»). Максимальная кратность резервирования определяется по формуле

$$\lambda_{j(l_j)}^* = \min_{p=1, \dots, q} \left[ \frac{\Delta g_p^j}{g_p(u_{j(l_j)})} - 1 \right], \quad (2.2.4)$$

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

Определив по формуле (2.2.4) максимальную кратность резервирования оставшихся после отсева по типам элементов подсистем, можно их продублировать. Завершающий этап предлагаемого метода – поиск опти-

мального решения [61]. 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 [54]. Это ограничение задано не жестким образом, а находится на каждой итерации по формуле

$$f(v) \geq f_i^* = \frac{f_{\max} + f_{i-1}^*}{2}, \quad (2.2.7)$$

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

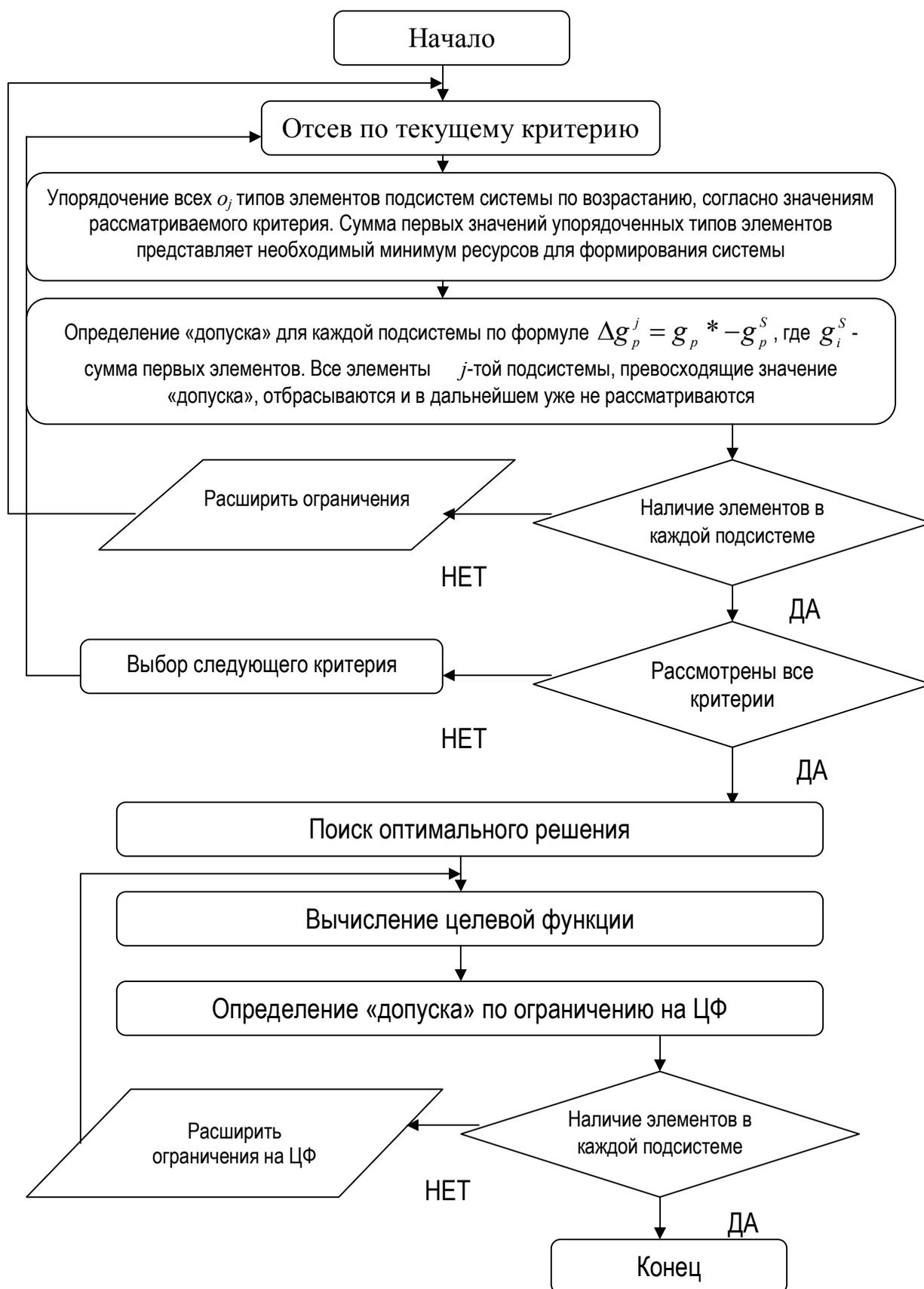
где  $f_{\min}$  – соответственно сумма максимальных и минимальных в каждой подсистеме значений целевой функции.

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 [60].

Алгоритм работы метода выделения главного критерия качества представлен на рис. 2.2.2.

### **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 [17]. Сравнивать 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 противоре-



**Рис. 2.2.2.** Алгоритм работы метода выделения главного критерия качества

чивы. 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 [35].

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 [36].

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,  $J = G$ ;

$$a_i^R = \frac{\sum_{j=1}^G \frac{K_i^R}{K_j^R}}{\sum_{i=1}^J \left( \sum_{j=1}^G \frac{K_i^R}{K_j^R} \right)} \quad - \quad (2.2.8)$$

– 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 [77].

### 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. 2.2.1).

Обобщённый критерий качества рассчитывается по следующей формуле

$$K_{\text{сум}} = \frac{1}{\sum_{i=1}^N \frac{a_i}{k_i}},$$

**Таблица 2.2.1.**

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  $K_{сум1}$  для первой альтернативы А1 (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:

$$K_{сум3} = \frac{1}{n_{3,1} + n_{3,2} + n_{3,3} + n_{3,4} + n_{3,5}} = \frac{1}{1,52 + 0,215 + 1,81 + 2,645 + 0,665} = 0,15 .$$

4. Рассчитаем обобщённый критерий  $K_{сум4}$  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:

$$K_{\text{сум4}} = \frac{1}{n_{4,1} + n_{4,2} + n_{4,3} + n_{4,4} + n_{4,5}} = \frac{1}{1,085 + 1,115 + 0,725 + 1,475 + 1,185} = 0,18$$

Сведем полученные данные в общую таблицу 2.2.2 :

**Таблица 2.2.2.**

Общая таблица суммарных критериев качества

Суммарные критерии качества	Сравнительные значения
К1	0,13
К2	0,21
К3	0,15
К4	0,18

Расчёты суммарных критериев качества с учётом себестоимости получаемых наноматериалов показали, что наиболее эффективными методами производства углеродных нанотрубок можно считать метод термического 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.

## 2.3. 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 применение для исследования функционирования систем, имитируются на ЭВМ изменения состояний системы («имитационное моделирование») [53].

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 [62].

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 [96]. 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, моделирова-

ние 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 процессов 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 [63].

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 [66]. 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 [67]. В этом случае не возникает необходимости в имитации физических параметров элементов. При оценке и прогнозировании

нии 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 материи, в частности при механическом движении [62].

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 [80]. 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:

$$I = \sum_{t=t_n+1}^{t_k-1} f^0(t, x(t), u(t)) + F(x(t_n), x(t_k), u(t_n), u(t_k)), \quad (2.3.1)$$

где  $f^0(\cdot)$ ,  $F(\cdot)$  – скалярные функции (функционалы). Они отражают представления о качестве процесса (точность, стоимость, затраченная энергия и т.п.) [93].

## 2.4. 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 [70]. 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

$$\begin{aligned} x^1(t+1) &= x^1(t) + uK + (1-u)(a + bx^2(t)), \\ x^2(t+1) &= (1-u)(x^2(t) + 1). \end{aligned} \tag{2.4.1}$$

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  $E$ .

$$L = G(x(t_h), x(t_k), u(t_h), u(t_k)) - \sum_{t_h}^{t_k-1} R(t, x(t), u(t)), \quad (2.4.2)$$

где

$$G = F(\cdot) + \phi(t_k, x(t_k)) - \phi(t_h, x(t_h)), \quad (2.4.3)$$

$$R = \phi(t+1, f(t, x(t), u(t))) - f^0(t, x(t), u(t)) - \phi(t, x(t)), \quad (2.4.4)$$

$\phi(t, x(t))$  – вспомогательная неопределённая пока функция (функционал),

$$f^0(t_h, \cdot) = 0.$$

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  $\varphi(t, x(t))$ , 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  $\varphi$ , которая и доопределит 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  $\varphi$  существует много, и это создаёт возможность приспособливаться к специфике задачи. В настоящее время разработаны по крайней мере три характерных способа задания  $\varphi$ , два из которых ведут к процедурам Понтрягина и Беллмана, а третий аналогов не имеет и хорошо приспособлен к решению широко распространённых на практике так называемых вырожденных задач [43]. 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. Заддим  $\varphi$  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  $\dot{x} = ax - u$  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.

# 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 [6]. Из-за воздействия неконтролируемых переменных технологические операции теряют такое свойство, как установившееся состояние, и процесс может дрейфовать во времени [7].

Предположим, что математическая модель процесса производства электронных нанообъектов 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 [8]. 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 [9]. 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 [30]. 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  $X_{опт}$  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  $X_{опт}$ . Так, в одномерном случае, при использовании линейной математической модели имеем [31]:

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

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

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

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

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

Для прогнозирования результатов реального процесса производства по найденному значению переменной управления подставим полученное значение  $X'_{опт}$  в математическую модель с истинными значениями коэффициентов [32]. 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 [33].

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 [34].

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 [37]. 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 [38].

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 [52]. В качестве экспериментальных данных для регрессионного анализа используются данные табл. 3.1.1.

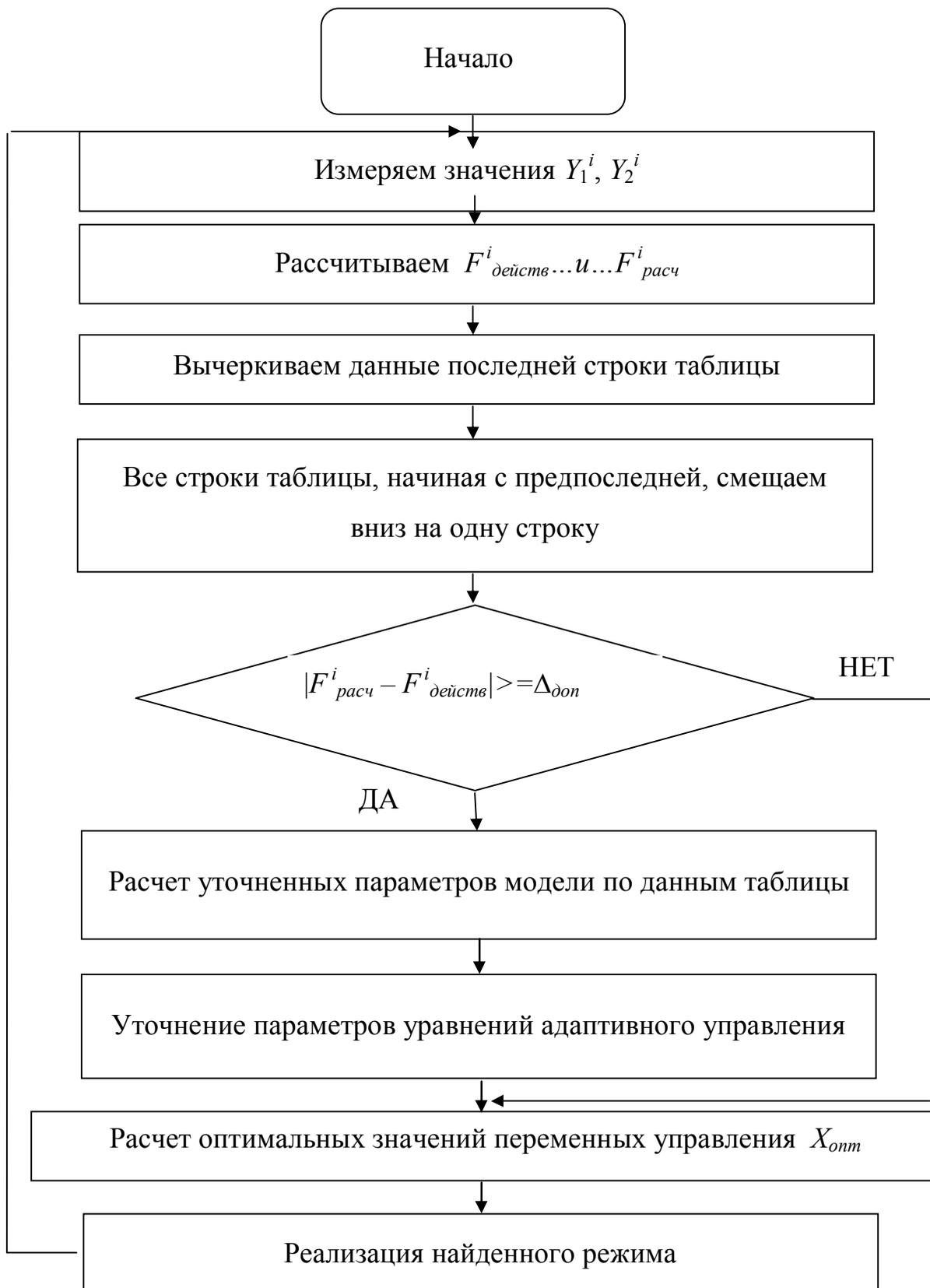
В целом алгоритм вышеописанного подхода, соответствующий моменту времени  $t_i$ , описывается следующей блок-схемой (рис. 3.3.1):

## **3.2. Регрессионный анализ**

Для построения экспериментально-статистических моделей используются статистические методы, такие, как корреляционный, регрессионный, конфлюэнтный 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 [95].

### 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 [97]. 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 искомых 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 искомым 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:

$$\begin{aligned} \det_0 &= \begin{bmatrix} \sum X_i^2 & \sum X_i \\ \sum X_i & N \end{bmatrix} = N \cdot \sum X_i^2 - (\sum X_i)^2; \\ \det_A &= \begin{bmatrix} \sum X_i Y_i & \sum X_i \\ \sum Y_i & N \end{bmatrix} = N \cdot \sum X_i Y_i - \sum X_i \cdot \sum Y_i; \\ \det_B &= \begin{bmatrix} \sum X_i^2 & \sum X_i Y_i \\ \sum X_i & \sum Y_i \end{bmatrix} = \sum X_i^2 \cdot \sum Y_i - \sum X_i \cdot \sum X_i Y_i. \end{aligned} \quad (3.2.9)$$

Таким образом, оценки коэффициентов регрессии вычисляются:

$$\begin{aligned} A &= \frac{\det_A}{\det_0} = \frac{N \cdot \sum X_i Y_i - \sum X_i \cdot \sum Y_i}{N \cdot \sum X_i^2 - (\sum X_i)^2}, \\ B &= \frac{\det_B}{\det_0} = \frac{\sum X_i^2 \cdot \sum Y_i - \sum X_i \cdot \sum X_i Y_i}{N \cdot \sum X_i^2 - (\sum X_i)^2}. \end{aligned} \quad (3.2.10)$$

### 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:

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

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

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

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

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

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

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 [30]. 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 [31].

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 [32]. Для учёта влияния неконтролируемых переменных и ошибок измерения на выходные переменные составляется

**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 $z$	Случайная составляющая погрешности $y$
1	$X_1$	$\bar{Y}_1$	$Y_1$	$z_1$	$y_1$
...	...	...	...	...	...
$i$	$X_i$	$\bar{Y}_i$	$Y_i$	$z_i$	$y_i$
...	...	...	...	...	...
$q$	$X_q$	$\bar{Y}_q$	$Y_q$	$z_q$	$y_q$

табл. 3.3.1, рассчитанная на  $q$  моментов времени (строк), предшествующих 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, соответственно, 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 [33].

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 [34]. 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 [2].

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 [4, 5].

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 [48].

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:

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

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

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 [51]. Другими словами,  $i$ -й критерий, не отличающийся по большинству альтернатив, имеет более

существенное значение для принятия решения. Определим «удовлетворяющий» 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) [54].

**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:

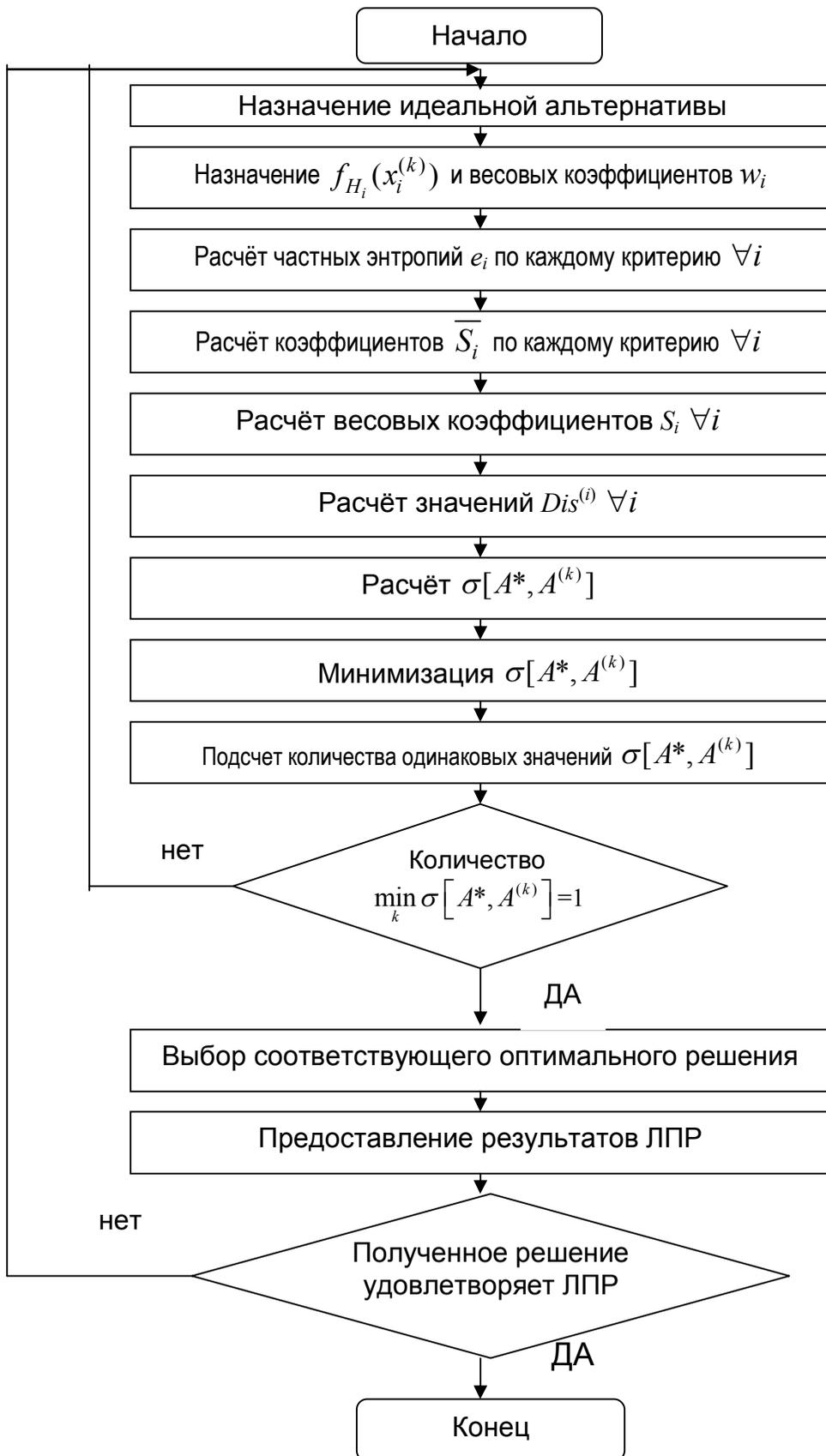
1. Расчёт частных энтропий  $e_i$  по каждому критерию  $\forall i$ .
2. Расчёт «удовлетворяющих» коэффициентов  $\overline{S}_i$  по каждому критерию  $\forall i$ .
3. Расчёт весовых коэффициентов  $S_i \forall i$ , комбинирующих нечёткость и субъективную оценку.
4. Расчёт значений  $Dis^{(i)} \forall i$ , 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**.

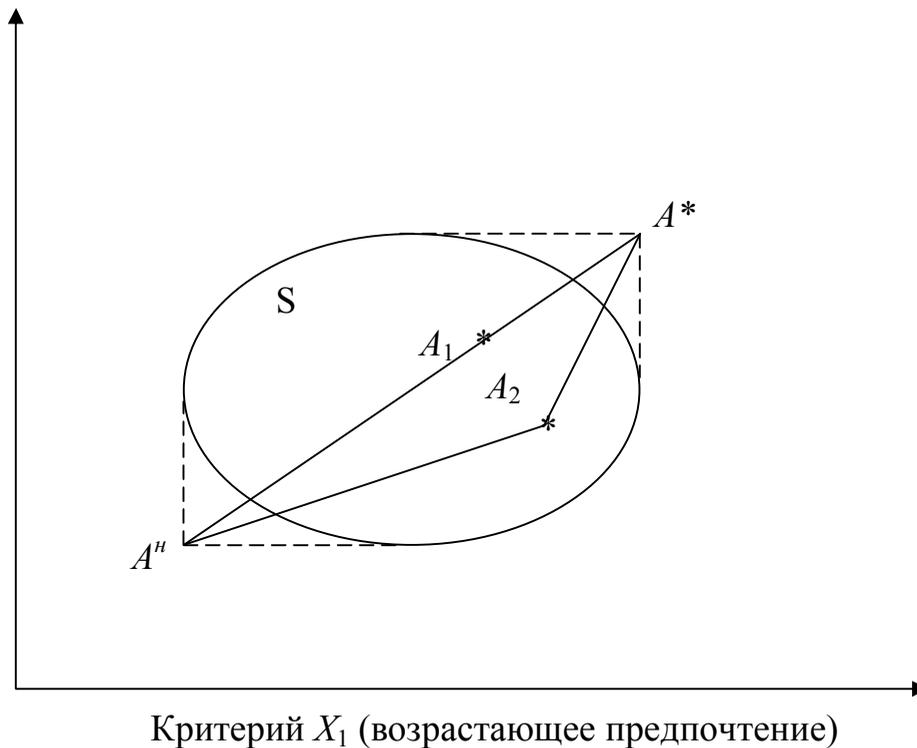
### Метод упорядоченного предпочтения через сходство с идеальным решением

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 [2].

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 гео-



**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

метрическом смысле. Однако этот подход противоречит утверждению о наибольшем расстоянии до идеального негативного решения [60]. Иногда выбранная альтернатива, имеющая наименьшее евклидово расстояние до 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 [72]. 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.

$$D = \begin{array}{c|cccccc} & X_1 & X_2 & \dots & X_j & \dots & X_n \\ \hline A_1 & x_{1;1} & x_{1;2} & \dots & x_{1j} & \dots & x_{1n} \\ \hline A_2 & x_{2;1} & x_{2;2} & \dots & x_{2j} & \dots & x_{2n} \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline A_i & x_{i;1} & x_{i;2} & \dots & x_{ij} & \dots & x_{in} \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \hline A_m & x_{m;1} & x_{m;2} & \dots & x_{mj} & \dots & x_{mn} \end{array}$$

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. Опре-

делим две искусственные альтернативы  $A^*$  и  $A^H$  как

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

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

где  $J = \{j = 1, 2, \dots, n \mid j \text{ — связанные с критерием выгоды}\}$ ,

$J' = \{j = 1, 2, \dots, n \mid j \text{ — связанные с критерием затраты}\}$ .

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 компромиссного» решения. Для решения задач многоцелевого принятия решений (multiple objectives decision making) было предложено большое число методологий [49].

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

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

где  $x \in X = \{x | g_h(x) \{>, =, <\} 0, \quad h = 1, 2, \dots, m\}$ , (3.5.1)

$f_j(x)$ : 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 [97].

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 двуцелевой 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 [98].

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  $d_1$  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:

$$\begin{aligned} \min d_p^{III} (x) \\ \max d_p^{III} (x) , \\ x \in X \end{aligned} \tag{3.5.6}$$

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. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ И АЛГОРИТМОВ УПРАВЛЕНИЯ КАЧЕСТВОМ ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ УСТРОЙСТВ**

### ***4.1. Экспериментальные исследования процессов формирования электронных устройств на основе модифицированных углеродных структур***

Модификация поверхности образцов наноматериалов была осуществлена в нанотехнологической установке «Луч-2» (производство Института Нанотехнологий Международного Фонда Конверсии), специально разработанной для проведения нанотехнологических операций в химически активных газовой-паровых средах [11].

Установка предназначена для разработки нанотехнологий, создания объектов нанометровых размеров, а также неразрушающего исследования поверхностей методами туннельной микроскопии и спектроскопии.

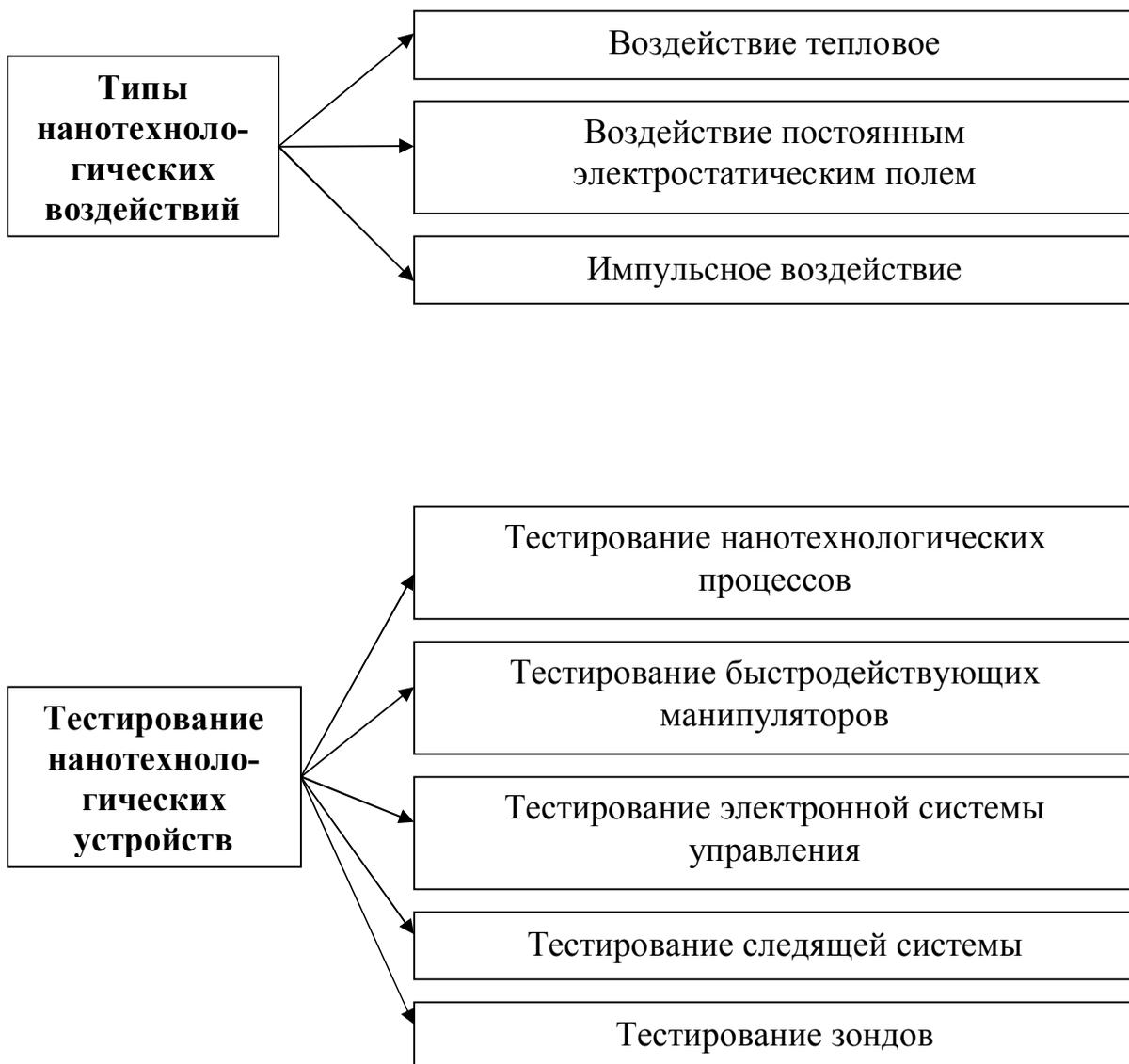
Нанотехнологические процессы в установке осуществляются за счёт инжекции технологических газов (жидкостей) и создания полей сверхвысокой напряженности между зондом и образцом.

При этом в зависимости от цели производится фиксация и перемещение объектов нанометрового диапазона, полимеризация, осаждение, травление. В результате в установке могут изготавливаться объекты с двумерной поверхностной, тонкопленочной или трехмерной структурой [10].

На рис. 4.1.1 представлена блок-схема режимов работы технологического модуля установки «Луч-2», содержащего герметизированную технологическую камеру с системой прецизионного позиционирования,



**Рис. 4.1.1.** Блок-схема режимов работы нанотехнологической установки



**Рис. 4.1.1.** Блок-схема режимов работы нанотехнологической установки (продолжение)

блок ввода-вывода технологических газов (жидкостей), систему защиты от внешних сейсмических и акустических воздействий. Технологический модуль имеет разъёмы для соединения с системой управления и резервуарами газов (жидкостей).

Модуль рекомендуется размещать в вытяжном шкафу, расположенном в месте минимальных производственных (сейсмических и акустических) шумов.

Созданы различные варианты технологических модулей: с внутренним хранением технологических газов, мини реакторы для работы с малыми технологическими объёмами [43].

Технологический модуль соединяется с форвакуумным насосом и системой напуска технологических газов вакуумными шлангами. Блок сопряжения и преобразования сигналов располагается вне вытяжного шкафа. Блок электроники располагается на компьютерной стойке или на столе рядом с компьютером. Управление нанотехнологической установкой производится по локальной сети или с удаленного компьютера [11].

Подготовка нанотехнологической установки к работе включает в себя несколько этапов:

1. Установка зонда.
2. Установка образца.
3. Герметизация технологической и защитной камер наноконтекста.
4. Запуск системы управления составом технологических сред.
5. Включение электронной системы управления.
6. Первичный подвод зонда к поверхности образца.

Дистанционно-управляемая работа на нанотехнологической установке производится в режимах измерения параметров нанообъектов и поверхности, а также в режиме проведения технологических воздействий (см. рис. 4.1.1).

Для начала работы необходимо мышью ЭВМ активировать значок с надписью «Nanotechnology», в результате на экране появится основная панель с изображением перечня выполняемых функций. Окно «System condition» содержит информацию о параметрах туннельного зазора, режима подвода, а также индикаторы текущих значений ЦАПов и АЦП. Раздел «Tunnel Function» позволяет изменять параметры туннельного зазора: напряжение между образцом и зондом «Voltage» и величину туннельного тока «Current», а также показывает текущее значение туннельного тока. Изменение параметров производится путём набора соответствующих новых величин вместо существующих. Для ввода измененных параметров необходимо нажать кнопку «Set» в этом же разделе.

Подвод и отвод зонда к образцу осуществляется в разделе «Approach». Существует несколько режимов подвода/отвода: «Fine» – «Точный» подвод образца (для сканирования быстродействующими манипуляторами) и «Rough» – «Грубый» подвод образца (для сканирования широкодиапазонными манипуляторами). Выбор режима подвода осуществляется выбором соответствующей кнопки. Существует также режим «Fast» для ускорения операций подвода/отвода, активируемый флажком. Кнопка «In» активирует процесс подвода, кнопка «Out» - процесс отвода. Кнопка «Options» вызывает окно, в котором содержатся параметры режима подвода/отвода. Подвод производится до появления скачка туннельного тока.

Подвод зонда к образцу, в соответствии с выбранным режимом, осуществляется автоматически после нажатия кнопки «In» в разделе «Approach» и проходит в два этапа: грубый подход «шагающим» двигателем до обнаружения туннельного тока и «плавный» за счёт широкодиапазонного Z-двигателя. При активации операции подвода/отвода открывается окно состояния «Wait please», в котором отображается стадия процесса в виде соответствующих пиктограмм [11].

При подводе зонда на расстояние в несколько ангстрем начинает появляться туннельный ток (скачкообразно или по другому закону).

Проведение нанотехнологических процессов является основным режимом работы нанотехнологической установки. В данном режиме осуществляется задание области проведения нанотехнологических процессов, типов и параметров воздействий.

Режим активируется после проведения сканирования нажатием правой кнопки мыши в окне результатов сканирования. В возникшем вспомогательном меню нужно выбрать пункт «Lithography». При этом открывается одноименное окно.

Тип воздействия устанавливается с помощью опции «Influence Type». В программе допускается активировать устройства, обеспечивающие следующие воздействия: электрическим полем, механическим воздействием, тепловым нагревом, ультразвуковым и колебаниями, СВЧ и лазерным излучением.

В базовой конфигурации «Луч-2» встроены устройства для выполнения первых двух типов воздействий. Однако программное обеспечение допускает сопряжение с программами управления дополнительных устройств, осуществляющих другие типы воздействий. Подключение дополнительных устройств позволяет расширить спектр выполняемых воздействий.

После выбора типа воздействия производится задание его параметров. Для электрического воздействия, применяемого в данной работе, это: «Pulse Voltage» – величина напряжения импульсов литографии, «Pulse Number» – количество импульсов литографии, «Pulse Duration» – длительность импульсов литографии, и «Pulse Period» – период импульсов литографии.

При проведении процесса производится сканирование в области предыдущего кадра. В процессе сканирования изображение областей, по которым уже произведено сканирование, покрывается сеткой белых точек.

В выделенных для проведения режима воздействий областях, в каждой точке включаются соответствующие устройства, осуществляющие воздействие с заданными параметрами.

По завершении процесса литографии автоматически запускается режим сканирования поверхности, результаты которого выводятся в самостоятельное окно результатов.

При сканировании поверхностей с большой шероховатостью или расположенных под большим наклоном, перепад высот рельефа может превысить ход быстродействующего Z-двигателя. При этом может произойти потеря туннельного тока или столкновение с поверхностью. Для продолжения сканирования в подобных ситуациях необходимо скорректировать положение зонда за счёт широкодиапазонного Z-двигателя. Для этого в программе предусмотрен режим коррекции положения широкодиапазонного Z-двигателя во время сканирования – «Hardware». Если значение высоты рельефа окажется больше верхнего заданного порога или меньше нижнего, то развёртка приостанавливается и подается команда на перемещение широкодиапазонного Z-двигателя до такого положения, при котором значение высоты рельефа попадет в заданный интервал. В режиме ожидания система также следит за тем, чтобы текущее значение Z находилось в заданном интервале. Режим «Hardware» активируется установкой соответствующего флажка в разделе «Z-guard».

Сканирование кадра ведётся с учётом заданных параметров начальной точки, размеров области, числа измерений в точке, задержек на перемещение и измерения, а также установленного режима. Во время сканирования получаемое изображение в реальном масштабе времени отображается построчно на экране в окне результатов, что дает возможность наблюдать за ходом выполнения операции. Процесс сканирования может быть прерван кнопкой «Stop» в окне состояния. После каждой отсканированной строки происходит нормировка полученного изображения поверхности. По завер-

шении процесса сканирования окно состояния закрывается.

На экране могут одновременно находиться несколько окон результатов. Так как результаты сканирования представлены в псевдоцвете, каждое окно результатов содержит шкалу, отображающую последовательность цветов в представлении высот рельефа, работы выхода и т.п.

Далее полученные результаты сканирования можно сохранить, обработать, использовать для литографии или для исследования локальных характеристик в выбранных точках поверхности и т.п. Для операций сохранения/считывания полученных изображений нужно использовать следующие команды главного меню: «File/Save» для сохранения и «File/Open» для считывания [11].

## ***4.2. Методы сканирующей туннельной микроскопии***

Зонды для СТМ, как правило, получают электрохимическим травлением проволочек из вольфрама в растворе едкого калия. По внешнему виду трудно судить об их качестве, так как в реальных технологических процессах «работает» самый последний кончик, который практически невозможно визуализировать с помощью даже достаточно мощного электронного микроскопа [19].

Для диагностики качества зонда производится измерение зависимости туннельного тока от расстояния зонд–образец. Этот режим является модификацией режима измерения ВАХ. Только в отличие от случая в вольтамперной характеристике здесь как параметр задается относительное смещение зонда от текущего положения в ангстремах. Во время измерения своего начального положения и на всём протяжении отрезка смещения производятся замеры величины туннельного тока. Снятые таким образом данные

отображаются в виде кривой, по форме которой можно судить о качестве зонда [10].

Объектами исследований служили углеродные наночастицы, сформированные на поверхности алмазоподобной углеродной плёнки типа б-С:Н с помощью СТМ-литографа. В отличие от «самоорганизующихся» углеродных наноструктур, которые представляют собой существенно нерегулируемые системы, такие нанообъекты могут быть контролируемо «нарисованы» на поверхности подложки с различными размерами и конфигурацией [47]. Немаловажным обстоятельством является также то, что это планарные наноструктуры и, следовательно, они удобны для зондовых исследований.

Тонкие пленки б-С:Н толщиной 4–7 нм были выращены на металлизированных поликорковых подложках методом электрически активированного газофазного химического осаждения в кольцевом ионном источнике ИИ4-015 типа «Радикал». Исходным реагентом для получения плёнки служил изопропиловый спирт.

Установка вакуумного напыления состоит из блока откачного и камеры с вращающимся столиком, в которую входит ионный источник ИИ4-015, который предназначен для получения трубчатого пучка ионов.

Откачка агрегата от атмосферного давления до давления 6,3 Па производится насосом НВР-16Д. Высокий вакуум в агрегате обеспечивается за счет откачки газов турбомолекулярным насосом 01 АБ-1500-004.

Типичные параметры процесса осаждения алмазоподобных тонких пленок таковы:

1) Очистка поверхности ионами Ar:

- давление Ar – 0,17 Па;
- ускоряющее напряжение – 3 кВ;
- ток разряда – 100 мА;
- время процесса – 15ч20 мин.

## 2) Осаждение пленки б-С:Н:

- давление паров изопропилового спирта – 0,1 Па;
- ускоряющее напряжение – 0,6ч1,0 кВ;
- ток разряда – 10ч20 мА;
- скорость осаждения – 0,5ч2 нм/мин [11].

Выбор сверхтонких алмазоподобных пленок б-С:Н в качестве основы для образования наноструктур обусловлен следующими соображениями. Этот рентгеноаморфный твердый углерод является либо диэлектриком, либо проводником (в зависимости от толщины), сохраняет характерную для алмаза и графита инертность по отношению ко многим средам и обладает прочностью, близкой к прочности алмаза. Однако в отличие от гомогенной поверхности пиролитического графита, поверхность пленки гетерогенна вследствие того, что сама пленка, как установлено многими исследованиями [23], представляет собой алмазную матрицу, разупорядоченную внедренными в нее кластерами графитовой структуры с поперечным размером 2-3 нм. Вследствие этого и поверхность пленки представляет собой совокупность островов алмазной и графитовой субструктур, обладающих разными химическими свойствами: в то время как на поверхности графита вообще нет болтающихся связей, на поверхности алмазной фазы эти связи имеются и они насыщены атомами водорода. Это кажущееся на первый взгляд неблагоприятное обстоятельство для пленки по сравнению с однородной поверхностью графита сыграло определяющую роль при выборе рабочей поверхности для образования устойчивых наноструктур с контролируемыми свойствами. Технологическая задача формировалась следующим образом: структуры наноразмерного диапазона должны быть связаны с поверхностью прочными химическими связями и должны содержать в себе возможность наращивания вещества посредством образования новых химических связей [24, 25]. С этой точки зрения именно алмазная фаза сверхтонкой плёнки б-С:Н представляется наиболее перспективной благодаря высокой вероят-

ности к образованию устойчивых С-С связей с молекулами адсорбата при определенных технологических условиях. Полученные вышеописанным способом очень прочные сверхтонкие плёнки б-С:Н содержат до 80% [26] алмазной фазы, обладает диэлектрическими свойствами.

Поликристаллические плёнки состоят из кристаллов, размер которых во внешнем слое плёнки по мере её роста увеличивается приблизительно пропорционально её толщине и достигает нескольких микрометров или даже десятков микрометров. Поверхности этих кристаллитов создают характерную гранную морфологию «ростовой» поверхности плёнки. Встречаются и другие типы морфологии поверхности алмазных плёнок.

Совершенство кристаллической структуры выращенных плёнок контролируют, снимая спектры комбинационного рассеяния (Романовские спектры).

Неалмазному углероду с преимущественной  $sp^2$ -гибридизацией орбиталей отвечают две размытые полосы при приблизительно 1350 и 1550  $cm^{-1}$ . Сравнивая высоты пика алмаза и максимума  $sp^2$ -углерода, следует иметь в виду, что чувствительность Романовского сигнала к неалмазному углероду приблизительно в 50 раз выше, чем кристаллическому алмазу [27]. В лучших образцах примесь неалмазного углерода не превышает 1%.

Содержание неалмазного углерода уменьшается в направлении от подложки к поверхности плёнки [28]; это связано, скорее всего, с укрупнением кристаллитов, вследствие чего вклад от межкристаллитных границ снижается.

Примесь  $sp^2$ -углерода оказывает существенное влияние на качество алмазных плёнок как электродов. И всё же современные физические и оптические методы исследования (Рамановская и оже-спектроскопия, атомно-силовой микроскоп) не дают исчерпывающей характеристики включений неалмазного углерода в поликристаллические алмазные плёнки; более тонкие эффекты выявляются с помощью электрохимических измерений.

Неалмазную фазу в поликристаллических алмазных плёнках часто называют графитом, но это утверждение не совсем точно. Неалмазная фаза – это прежде всего разупорядоченный углерод межкристаллитных границ. Его удаётся непосредственно наблюдать на поверхности алмаза с помощью электронного микроскопа с высоким разрешением или рамановского микросонда; толщина межкристаллитных границ в алмазных плёнках достигает нескольких нанометров. Помимо этого, в алмазных плёнках присутствуют различные дефекты кристаллической решётки. В ряде случаев поверхность алмазной плёнки покрыта тонким (порядка нескольких нанометров) слоем неалмазного углерода, который часто образуется на последних плохо контролируемых стадиях роста плёнок, когда активация газовой фазы уже прекращена и реактор начинает остывать, но в него продолжает поступать реакционная газовая смесь. Этот слой оказывает существенное влияние на поверхностную проводимость алмаза, повышая её у диэлектрических плёнок и снижая у сильно легированных (проводящих). Для его удаления образцы прогревают на воздухе при  $t=520^{\circ}\text{C}$  или кипятят в концентрированной  $\text{HClO}_4$ , хромовой смеси или растворах других сильных окислителей; предполагается, что при этом неалмазный углерод окисляется до  $\text{CO}_2$  [92].

Удельное сопротивление плёнок алмаза определяется не только концентрацией носителей заряда (свободных дырок), но и их подвижностью. Из-за рассеяния носителей заряда на межкристаллитных гранях и других структурных дефектах подвижность дырок в поликристаллических плёнках обычно невелика, гораздо ниже, чем в монокристаллах. Обычно эта подвижность – порядка  $1 \text{ см}^2/(\text{В}\cdot\text{с})$ , хотя в некоторых случаях она в десятки и сотни раз выше. Подвижность в монокристаллических (гомоэпитаксиальных) плёнках на 1-2 порядка величины выше, чем в поликристаллических, выращенных в тех же условиях [70]. Следует отметить, что некоторое нарушение кристаллической решётки алмаза при его легиро-

вании, обусловленное введением бора, может отрицательно сказаться на величине подвижности. В плёнках аморфного алмазоподобного углерода величина подвижности на несколько порядков ниже, чем в кристаллическом алмазе.

Сопротивление межкристаллитных границ может вносить нелинейность в полное сопротивление умеренно легированных поликристаллических алмазных плёнок. Был сделан вывод о том, что к обсуждавшейся выше электропроводности, обусловленной движением свободных дырок в зоне проводимости, при более высоких температурах добавляется вторая составляющая, связанная с перескоками носителей заряда между локализованными ловушками для электронов.

С помощью сильного электрического поля в межэлектродном зазоре СТМ возможна заметная поляризация молекул среды и их перестройка. За счёт дипольного взаимодействия молекул и допирующих примесей возможно образование проводящих молекулярных мостиков из адсорбата электродов или из жидкой и газообразной диэлектрической фазы.

В качестве адсорбата газообразной фазы необходимо было выбрать углеводород с легко «отрывающимся» атомом углерода, сильно поляризованной молекулой или образующей при диссоциации свободные связи предпочтительно на атомах углерода (или других атомах). Эти соображения привели к выбору трихлорэтилена  $C_2HCl_3$  в качестве газовой среды и, впоследствии, части адсорбата.

### **4.3. Расчётные данные оптимизации процессов производства с применением математических моделей**

Рассмотрим применение дискриминационного метода оптимизации обобщенного критерия качества на условном примере.

Обозначим отдельные параметры качества следующими переменными:

$u_1$  – проводимость, Ом<sup>-1</sup>;

$u_2$  – длина нанотрубок, нм;

$u_3$  – структура нанотрубок (количество слоев), шт.;

$u_4$  – стоимость, у.е.;

$u_5$  – окупаемость (время), г.

Введём следующие приблизительные ограничения значений отдельных параметров качества:

$$g_1(u_1) \geq 10^{-7}$$

$$g_2(u_2) \geq 50$$

$$g_3(u_2) \leq 2000$$

$$g_4(u_3) \geq 1$$

$$g_5(u_3) \leq 20$$

$$g_6(u_4) \leq 500$$

$$g_7(u_5) \leq 2$$

Вычислим весовые значения видовых критериев качества:

$$g_1(v_1) = g_1(u_1)$$

$$g_2(v_2) = g_2(u_2) + g_3(u_2)$$

$$g_3(v_3) = g_4(u_3) + g_5(u_3)$$

$$g_4(v_4) = g_6(u_4)$$

$$g_5(v_5) = g_7(u_5)$$

Нормируем весовые значения с помощью процедуры Саати (табл. 4.3.1).

**Таблица 4.3.1.**

Весовые значения, номинированные с помощью процедуры Саати

		Альтернативы				Предельные значения
		A1	A2	A3	A4	
Критерии	K1	Весовые коэффициенты				
	K2					
	K3					
	K4					
	K5					
	K6					
	K7					
Обобщённый критерий						Max (ai)

Определим множество обобщенных критериев качества для каждой альтернативы:

$$P(v_i) = \sum_{p=1}^q g_p(v_i). \quad (4.3.1)$$

Среди найденного множества определим Max обобщенный критерий:

$$P(v_i) = \max \left\{ \sum_{p=1}^q g_p(v_i) \right\}. \quad (4.3.2)$$

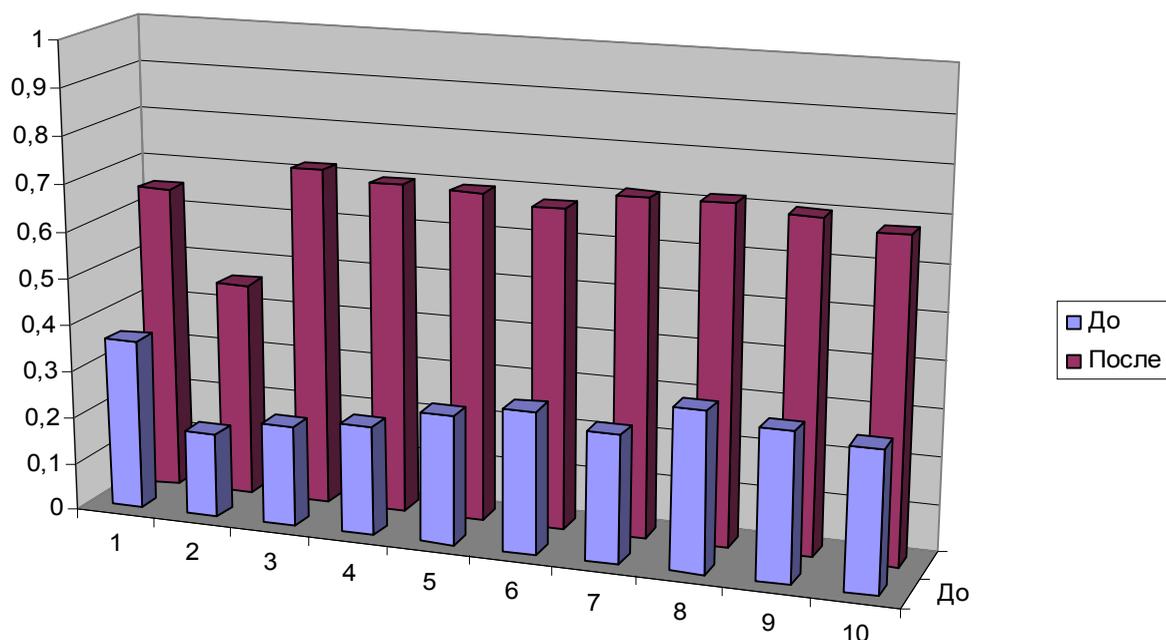
## Метод термического разложения графита в дуговом разряде

Результаты расчётов ОКК до и после оптимизации представлены в табл. 4.3.2 и на гистограмме результатов (рис. 4.3.1).

**Таблица 4.3.2.**

Результаты расчётов ОКК до и после оптимизации

До	После
0,365	0,653
0,181	0,459
0,215	0,72
0,233	0,701
0,276	0,698
0,302	0,678
0,273	0,713
0,34	0,716
0,315	0,702
0,298	0,68



**Рис. 4.3.1.** Гистограмма результатов

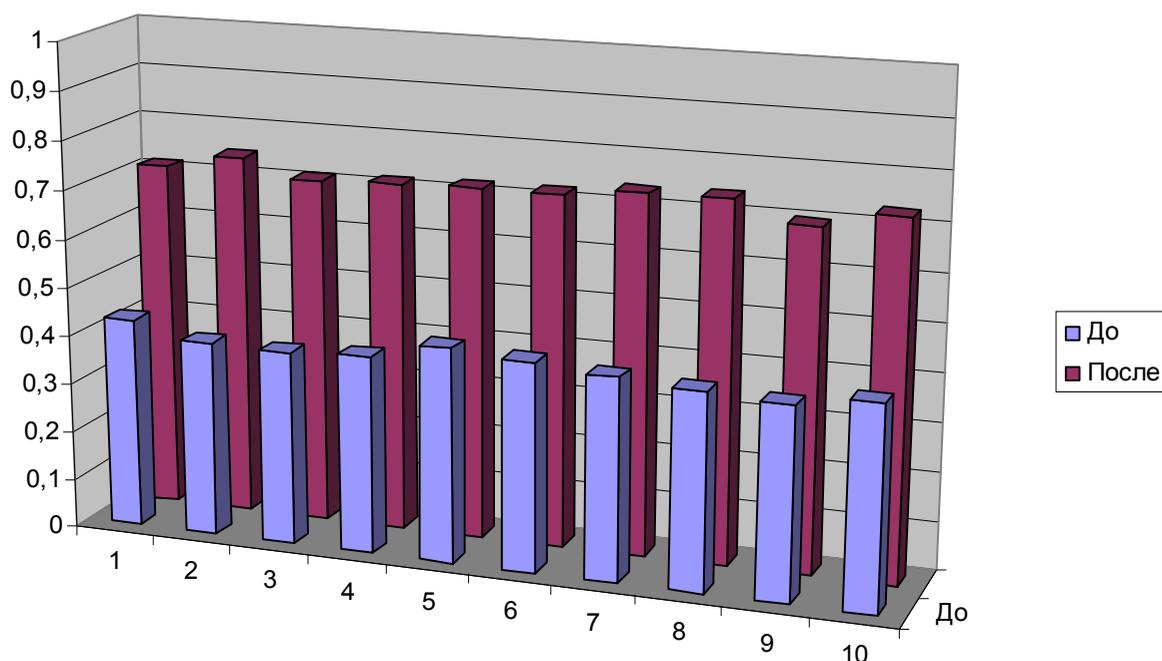
## Химическое осаждение углерода из газовой фазы

Результаты расчётов ОКК до и после оптимизации представлены в табл. 4.3.3 и на гистограмме результатов (рис. 4.3.2).

**Таблица 4.3.3.**

Результаты расчётов ОКК до и после оптимизации

До	После
0,433	0,715
0,402	0,745
0,398	0,71
0,405	0,716
0,44	0,72
0,427	0,721
0,415	0,736
0,403	0,739
0,395	0,698
0,416	0,73



**Рис. 4.3.2.** Гистограмма результатов

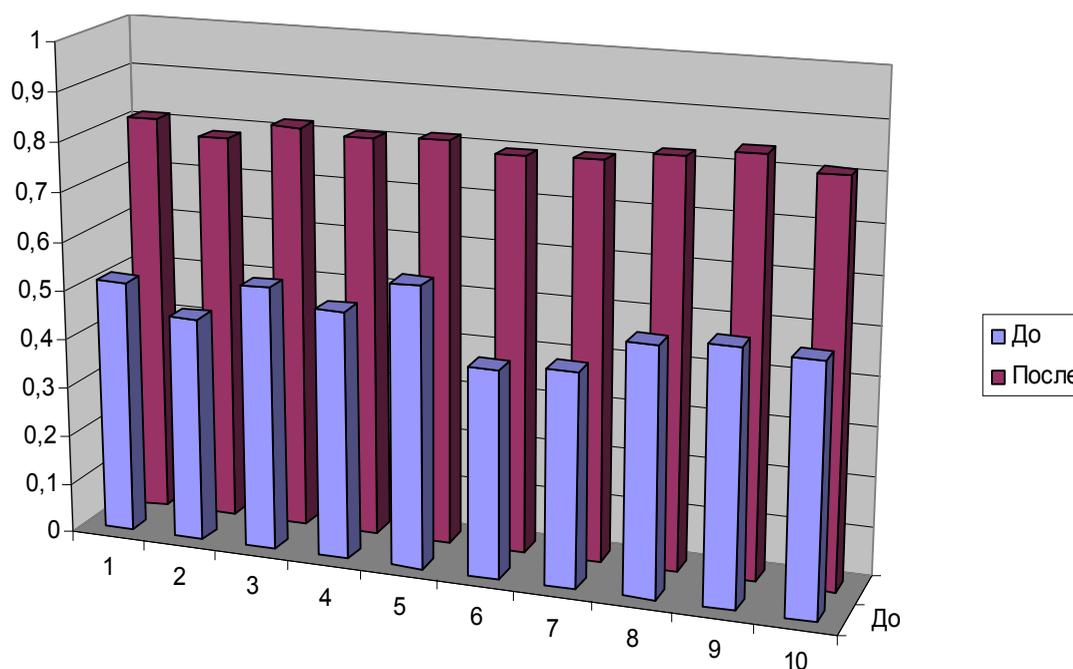
## Метод лазерного испарения графита

Результаты расчётов ОКК до и после оптимизации представлены в табл. 4.3.4 и на гистограмме результатов (рис. 4.3.3).

**Таблица 4.3.4.**

Результаты расчётов ОКК до и после оптимизации

До	После
0,516	0,815
0,497	0,787
0,536	0,82
0,523	0,81
0,53	0,817
0,48	0,799
0,498	0,803
0,502	0,823
0,513	0,84
0,504	0,811



**Рис. 4.3.3.** Гистограмма результатов

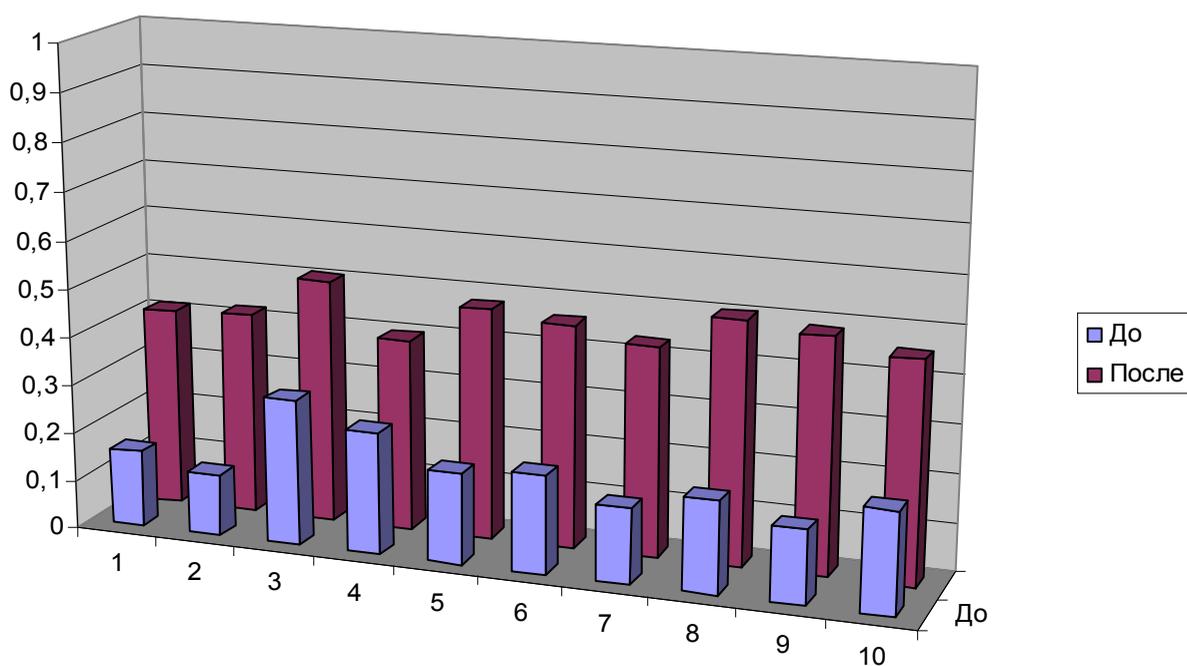
## Метод холодной деструкции графита (микровзрывы)

Результаты расчётов ОКК до и после оптимизации представлены в табл. 4.3.5 и на гистограмме результатов (рис. 4.3.4).

**Таблица 4.3.5.**

Результаты расчётов ОКК до и после оптимизации

До	После
0,16	0,412
0,127	0,42
0,302	0,506
0,254	0,398
0,189	0,477
0,203	0,46
0,156	0,432
0,19	0,503
0,153	0,486
0,21	0,457



**Рис. 4.3.4.** Гистограмма результатов

## Расчёт себестоимости материалов

Для анализа и сравнения технико-экономических показателей, а также определения экономической эффективности различных методов производства модифицированных углеродных наноструктур, необходимо рассчитывать полную себестоимость наноматериалов для каждого метода, которая включает следующие статьи затрат:

1. Основные и вспомогательные материалы.
2. Основная и дополнительная заработная плата разработчиков.
3. Единый социальный налог.
4. Затраты на использование оборудования.
5. Прочие денежные расходы.

Требования к стоимости научно-технической продукции устанавливают предельное значение себестоимости разработки образца, превышение которого приводит к выводу о нецелесообразности выполнения его разработки. Себестоимость научно-технической продукции представляет собой стоимостную оценку используемых в процессе научно-исследовательских и опытно-конструкторских работ (НИОКР) природных ресурсов, сырья, материалов, топлива, энергии, основных фондов, трудовых ресурсов, а также других затрат на выполнение соответствующих технологических работ.

Приведём расчёт себестоимости углеродного субмикронного материала, произведенного только методом термического разложения графита. Расчёты себестоимости продукта в других методах производства производятся аналогично.

Расчёт себестоимости наноматериала, произведённого  
методом термического разложения графита

Смета затрат на материалы и расчет основной и дополнительной заработной платы разработчиков приведены в табл. 4.3.6 и 4.3.7.

**Таблица 4.3.6.**

Смета затрат на материалы

№ п/п	Наименование материала	Единица измерения	Стоимость единицы измерения материала, у.е.	Количество	Общая стоимость, у.е.
<b>Основные материалы</b>					
1	Графит	кг	21	10	210
2	Гелий	л	80	10	800
3	Вода	л	0,1	1000	100
4	Толуол	л	25	10	250
5	Кислород	л	20	10	200
<b>Итого на основные материалы</b>					1560
Вспомогательные материалы (10% от основных)					156
<b>Итого на основные и вспомогательные материалы</b>					1716
Транспорно-заготовительные расходы (5% от суммы расходов на основные и вспомогательные материалы)					85,8
<b>Итого затрат на основные и вспомогательные материалы с учетом транспортных расходов</b>					<b>1801,8</b>

**Таблица 4.3.7.**

Расчёт основной и дополнительной  
заработной платы разработчиков

Должность	Оклад, у.е./мес.	Оплата, у.е./день	Продолжительность работ, дни	Итого, у.е.
Главный инженер	10000	454,5	2	909
Инженер-технолог	8000	363,6	2	727,2
Старший техник	6000	272,7	2	545,4
Техник	5000	227,3	2	454,6
<b>Итого</b>				<b>2636,2</b>
Дополнительная заработная плата (20% от основной)				527,24
<b>Основная и дополнительная заработная плата</b>				<b>3163,44</b>
Отчисление на социальные нужды (35,6% от основной и дополнительной зарплаты)				1126,18

Прочие денежные расходы составляют в среднем 50% от затрат на основную и дополнительную заработную плату:  $3163,44 \cdot 0,5 = 1581,72$  у.е.

Полученные данные сводим в табл. 4.3.8:

Расходы на содержание и эксплуатацию оборудования рассчитываем исходя из их стоимости (табл. 4.3.9).

Стоимость машинного времени рассчитывается по формуле:

$$S = \frac{S_{ам.} + S_{обсл.}}{T},$$

где  $S_{ам.}$  – годовые амортизационные отчисления, составляют 27,5% от стоимости оборудования;

**Таблица 4.3.8.**  
Затраты по элементам

№ п/п	Затраты по элементам	Сумма, у.е.
1	<b>Материалы</b>	1801,8
2	Основная заработная плата производственного отдела	2636,2
3	Дополнительная заработная плата	527,24
4	Отчисления на социальные нужды	1126,18
5	Прочие денежные расходы	1581,72
<b>Итого</b>		<b>7673,14</b>

**Таблица 4.3.9.**  
Стоимость оборудования

№ п/п	Наименование	Цена, у.е.
1	Электродуговая установка для производства углеродных наноструктур	300000
2	Установка для выпаривания толуола	50000
3	Вспомогательные инструменты для соскабливания углеродной сажи	10000
4	<b>Установка для окисления нанотрубок</b>	50000
5	Контейнеры для хранения и транспортировки углеродного материала	1000
<b>Итого:</b>		<b>411000</b>

$S_{обсл.}$  – затраты на обслуживание и электроэнергию, составляют 20% от стоимости оборудования;

$T$  – количество рабочих дней в году.

Эксплуатационные расходы составят:

$$S = \frac{411000 \cdot 0,275 + 411000 \cdot 0,2}{260} = 750,8 \text{ у.е./день.}$$

Смета затрат на производство углеродного наноматериала методом термического разложения графита приведена в табл. 4.3.10.

**Таблица 4.3.10.**

Смета затрат на производство углеродного наноматериала методом термического разложения графита

№ п/п	Статья расходов	Сумма, у.е.
1	Затраты на основные и вспомогательные материалы	1801,8
2	Основная заработная плата	2636,2
3	Дополнительная заработная плата	527,24
4	Отчисления на социальные нужды	1126,18
5	Расходы на содержание и эксплуатацию оборудования	1501,6
6	Прочие расходы	1581,72
Итого:		9174,74

Анализ полученных результатов показывает, что себестоимость углеродного наноматериала, полученного методом термического разложения графита составляет 9174,74 у.е.

## **ГЛАВА 5. ПРИПОВЕРХНОСТНЫЕ СВОЙСТВА НАНООБЪЕКТОВ, ГЕТЕРОСТРУКТУР И НАНООБЛАСТЕЙ С РАЗЛИЧНЫМ ХИМИЧЕСКИМ И ФАЗОВЫМ СОСТАВОМ**

Изучение наноструктур знаменует новый этап в развитии физики и биологии, поскольку по своему размеру они занимают промежуточное положение между молекулами и микроскопическими объектами. Они содержат поддающееся счёту число атомов и, следовательно, подходят для решения технологических задач на атомном уровне.

Последние 30 лет микроструктура поверхности твёрдых тел интенсивно изучалась методами дифракции и рассеяния электронных и ионных пучков, а также электронной спектроскопии. Однако большинство этих методов первоначально разрабатывалось для исследования объёмной структуры твёрдых тел, поэтому они не всегда годятся для получения информации о структуре поверхности, тем более на атомном уровне.

Изучение природы сил трения на атомарном уровне стало возможным после создания атомно-силовых микроскопов. За десятилетний отрезок времени интенсивное развитие этой принципиально новой техники физических исследований привело к становлению чрезвычайно богатой области физики — нанотрибологии, объединившей экспериментальное и теоретическое изучение адгезии, трения, износа и смазки, химической активности и трибоэлектромагнетизма на наноструктурном уровне.

Изучение трения и связанных с ним процессов на наноструктурном уровне представляет значительный интерес для широкого спектра технических применений, таких как технология производства и поверхностного покрытия жёстких магнитных дисков для компьютеров, изготовление

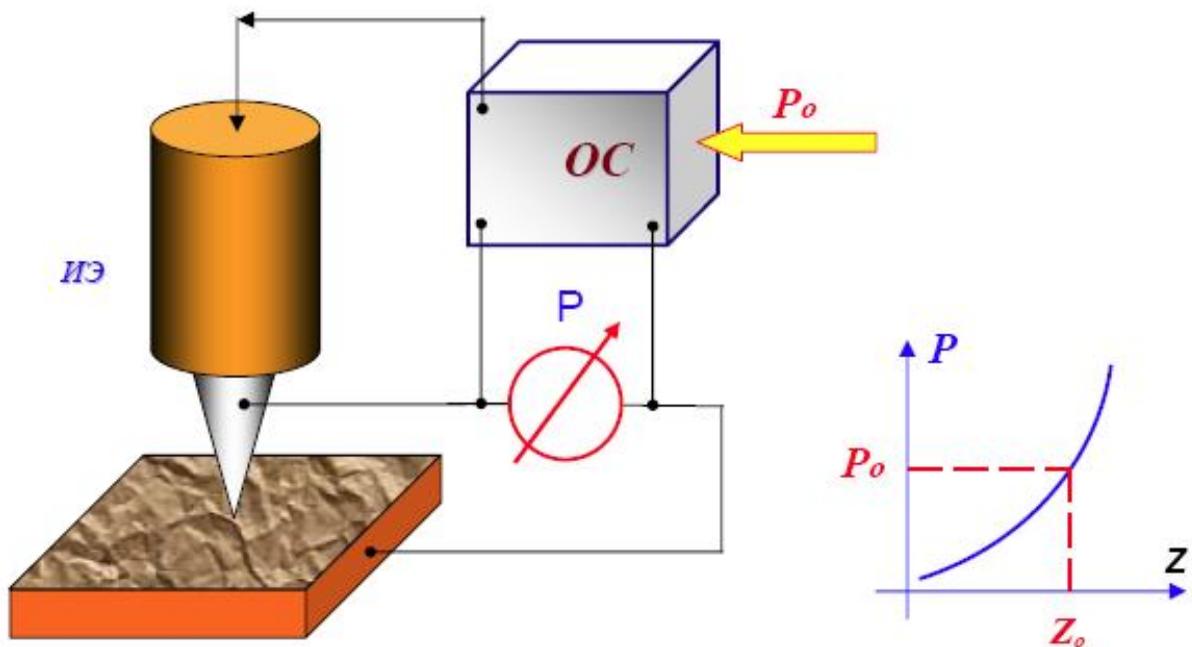
микросенсорных датчиков и т.д. Эти и традиционно важные инженерные применения трибологии в машиностроении требуют более глубокого понимания свойств материалов на атомарном уровне с целью оптимизации и прогнозирования трибологических характеристик трущихся поверхностей. Успешное решение этих задач требует сокращения существующего разрыва между макро- и наноструктурным уровнями знаний о свойствах материалов вообще, и о трении в частности.

Вместе с атомно-силовой и фрикционной микроскопией (АФМ) важное место в изучении нанотрибоэффектов занимают также исследования процессов наноиндентирования поверхностей.

### ***5.1. Основные принципы сканирующей зондовой микроскопии***

В сканирующих зондовых микроскопах исследование микрорельефа поверхности и её локальных свойств проводится с помощью специальным образом подготовленных зондов в виде игл. Рабочая часть таких зондов (острие) имеет размеры порядка десяти нанометров. Характерное расстояние между зондом и поверхностью образцов в зондовых микроскопах по порядку величины составляет 0,1-10 нм.

Контакт зонда и образца подразумевает их взаимодействие. Чтобы осуществлять исследование с помощью конкретного прибора, из широкого спектра выбирается какое-либо одно рабочее взаимодействие. Природа этого выбранного взаимодействия и определяет принадлежность прибора к тому или иному типу в рамках семейства зондовых микроскопов. Информация о поверхности извлекается путём фиксации (при помощи системы обратной связи (рис. 5.1.1)) или детектирования взаимодействия зонда и образца.



**Рис.5.1.1.** Схема организации системы обратной связи зондового микроскопа

В основе атомно-силовой микроскопии лежит взаимодействие зонда и образца с силами притяжения или отталкивания. Миниатюрные размеры зонда и высокая чувствительность детектирующей системы микроскопа позволяет достигать нано- и субнанометрового пространственного разрешения при детектировании поверхностных свойств (разрешающая способность прибора, как правило, тем выше, чем более короткодействующий характер имеет взаимодействие зонда и образца.)

Рассмотрим общие черты, присущие различным зондовым микроскопам. Пусть взаимодействие зонда с поверхностью характеризуется некоторым параметром  $P$ . Если существует достаточно резкая и взаимно однозначная зависимость параметра  $P$  от расстояния зонд-образец  $P=P(z)$ , то данный параметр может быть использован для организации системы обратной связи (ОС), контролирующей расстояние между зондом и образцом.

Система обратной связи поддерживает значение параметра  $P$  постоянным, равным величине  $P_0$ , задаваемой оператором. Если расстояние зонд-поверхность изменяется (например, увеличивается), то происходит изме-

нение (увеличение) параметра  $P$ . В системе ОС формируется разностный сигнал, пропорциональный величине  $\Delta P = P - P_0$ , который усиливается до нужной величины и подаётся на исполнительный элемент (ИЭ). Исполнительный элемент обрабатывает данный разностный сигнал, приближая зонд к поверхности или отодвигая его до тех пор, пока разностный сигнал не станет равным нулю. Таким образом можно поддерживать расстояние зонд–образец с высокой точностью. В существующих зондовых микроскопах точность удержания расстояния зонд–поверхность достигает величины  $\sim 0,01$  Е. При перемещении зонда вдоль поверхности образца происходит изменение параметра взаимодействия  $P$ , обусловленное рельефом поверхности. Система ОС обрабатывает эти изменения, так что при перемещении зонда в плоскости  $X, Y$  сигнал на исполнительном элементе оказывается пропорциональным рельефу поверхности.

Процесс сканирования осуществляется при помощи пьезокерамического манипулятора (или системы манипуляторов). Зонд движется последовательно, строка за строкой, вдоль поверхности (изменяются координаты  $X$  и  $Y$ ). Для оцифровки данных участок сканирования разбивается на  $N$  строк, а каждая строка на  $M$  точек. Таким образом, положение иглы в плоскости  $XY$  описывается двумя координатами  $X_i, Y_j$  из множества  $\{X_i, Y_j\}$   $N \times M$  точек (обычно выбирают  $N=M$ ). Результатом работы сканирующего зондового микроскопа является установление соответствия между каждой парой координат из множества  $\{X_i, Y_j\}$  и некоторым числовым значением (или рядом значений), характеризующим анализируемый параметр поверхности (или ряд параметров).

**По способу движения зонда над поверхностью можно провести следующую дифференциацию работы СЗМ.**

- Если зонд движется над поверхностью при постоянной координате  $Z$ , то говорят, что сканирование осуществляется по способу *постоянной высоты*. В этом случае в каждой точке из множества  $\{X_i, Y_j\}$  измеряется

*интенсивность* рабочего взаимодействия  $\Phi_{ij|Z=\text{const}}$ . Результатом исследования является массив (или несколько массивов, если измерялось несколько параметров)  $\{\Phi_{ij|Z=\text{const}}, X_i, Y_j\}$ , описывающий зависимость функции двух переменных  $\Phi|_{Z=\text{const}}(X, Y)$ .

- Если же *система обратной связи* фиксирует в процессе сканирования на заданном уровне величину рабочего взаимодействия  $A(X, Y, Z)$  вариацией вертикальной  $Z$  координаты зонда, то говорят, что сканирование осуществляется по способу *постоянного взаимодействия*. Результатом работы СЗМ в этом режиме будет массив  $\{Z_{ij|A=\text{const}}, X_i, Y_j\}$ , коррелирующий с топографией исследуемой поверхности. Помимо «топографического» массива можно, проводя в каждой точке измерения какого-либо дополнительного параметра (или нескольких), получать зависимости вида  $\Phi_{ij|A=\text{const}}(X_i, Y_j)$ .

Таким образом, результатом СЗМ-исследования является получение функциональных зависимостей двух типов: по способу постоянной высоты:  $\Phi|_{Z=\text{const}}(X, Y)$  и по способу постоянного взаимодействия:  $Z|_{A=\text{const}}(X, Y)$  («топография»), плюс какая-либо дополнительная зависимость  $\Phi|_{A=\text{const}}(X, Y)$ . С помощью компьютерного программного обеспечения можно проводить анализ полученных зависимостей (анализ характерных латеральных и вертикальных размеров поверхностных особенностей, построение сечений, фурье-анализ, оценка шероховатости и т.п.), отображать полученные зависимости на экране монитора и выводить их на принтер.

Следует учитывать отличие «топографического» массива, полученного в режиме постоянного взаимодействия:  $\{Z_{ij|A=\text{const}}, X_i, Y_j\}$  от реальной топографии поверхности. В случае неоднородного распределения поверхностных свойств, определяющих интенсивность взаимодействия зонда и образца, для извлечения точной информации о топографии объекта необходимо в каждой точке проводить дополнительный анализ взаимодействия зонда и образца.

## **5.2. Силовое взаимодействие зонда с поверхностью**

При сканировании зондом поверхности исследуемого образца на зонд действуют различные силы, возникающие в приповерхностных областях образца. Эти силы можно регистрировать в различных режимах работы зондового микроскопа. Но при исследовании каждого конкретного вида взаимодействия остальные исключить невозможно, поэтому при получении экспериментальных данных необходимо проводить их обработку, чтобы окончательные результаты исследования соответствовали действительности.

Для анализа получаемых результатов необходимо представление о видах силового взаимодействия зонда с исследуемой поверхностью.

### **5.2.1. Электростатическое взаимодействие**

Даже если к зонду и образцу не приложена разность потенциалов, различные участки поверхности могут содержать не скомпенсированные поверхностные заряды, находящиеся в материалах зонда и образца. Также друг с другом будут взаимодействовать диполи, находящиеся в материале зонда и образца. Эти взаимодействия носят силовой характер, и тем самым могут приводить к искажению других измеряемых величин в процессе сканирования зондом по поверхности образца.

#### **Отдельные заряды**

Сила взаимодействия двух электрических зарядов выражается известным законом Кулона:

$$F = \frac{kq_1q_2}{\epsilon r^2},$$

где  $q_1, q_2$  – электрические заряды;

$r$  – расстояние между ними;

$\epsilon$  – диэлектрическая проницаемость;

$k = 9 \cdot 10^9$  м/Ф.

Свободные заряды индуцируют поверхностные заряды на ближайших поверхностях, известным и наглядным образом описания которых является метод электрических изображений. Его удобство состоит в том, что наведенное распределение зарядов можно заменить одним или несколькими точечными зарядами–изображениями и затем для расчёта силы взаимодействия пользоваться законом Кулона. В случае заряда над плоскостью этот метод приводит к следующей формуле:

$$F = \frac{-kq^2}{4r^2} \left( \frac{\epsilon_s - \epsilon_m}{\epsilon_s + \epsilon_m} \right),$$

где  $\epsilon_m, \epsilon_s$  – диэлектрическая проницаемость среды и образца соответственно;

$r$  – расстояние от заряда до плоскости.

В случае металлов диэлектрическая проницаемость  $\epsilon_s$  бесконечна, и поэтому выражение в скобках равно единице, и взаимодействие сводится к кулоновскому притяжению заряда к своему заряду–изображению другого знака, находящегося на расстоянии  $2r$ . Другой практический случай – погружение системы в жидкую среду с некоторым  $\epsilon_m$ . Электростатическое взаимодействие может значительно уменьшиться и даже превратиться в отталкивание при  $\epsilon_m > \epsilon_s$ .

## Диполи

Помимо нескомпенсированных зарядов в толще кантилевера и образца для анализа электростатического взаимодействия необходимо принимать во внимание диполи.

Энергия диполя  $d$ , помещенного в поле  $E$  (рис. 5.2.1),

$$W_D = -dE.$$

Электрическое поле, создаваемое диполем (рис. 5.2.2),

$$E = \frac{3(nd)n - d}{r^3}, \quad |E| \sim \frac{1}{r^3}$$

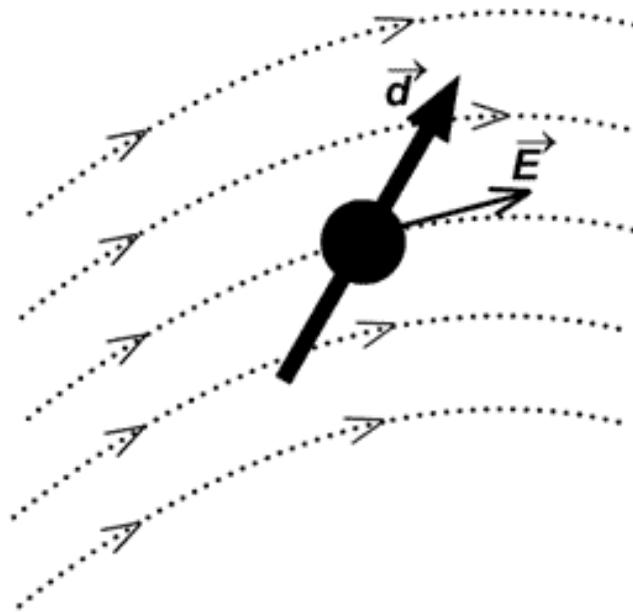


Рис. 5.2.1. Диполь, помещённый в электрическое поле

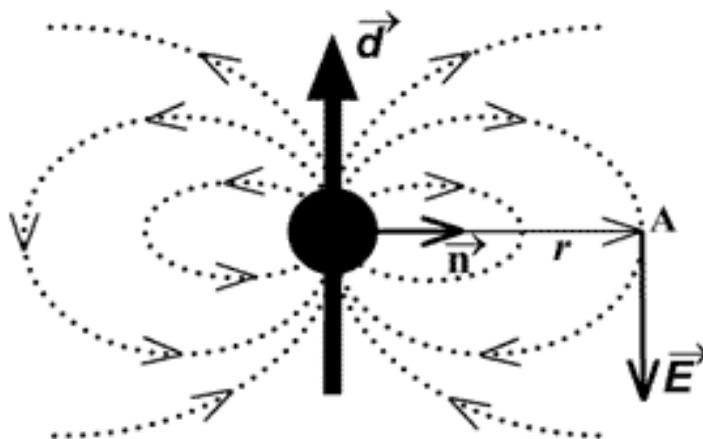


Рис. 5.2.2. Электрическое поле, созданное диполем

где  $n$  – единичный вектор в направлении на диполь из точки, где ищется поле.

Сила взаимодействия двух диполей убывает с расстоянием гораздо быстрее кулоновской:

$$F = \frac{3d_1d_2 - (nd_1)(nd_2)}{r^4},$$

где  $d_1, d_2$  – векторы дипольных моментов;

$r$  – расстояние между ними;

$n$  – вектор нормали вдоль линии, их соединяющей.

Малость силы дипольного взаимодействия по сравнению с Кулоновской компенсируется большим количеством диполей в веществе, в роли которых выступают молекулы. Полярные молекулы, один атом или группа атомов которых более электроотрицательны, чем другие, обладают дипольным моментом.

## Контактные явления.

### Работа выхода полупроводников

Работой выхода  $A$  (термодинамической работой выхода) электрона из полупроводника называют наименьшую энергию, которую нужно затратить для удаления электрона из полупроводника в вакуум, если начальная энергия электрона в полупроводнике равна электрохимическому потенциалу. Уровень энергии, равной электрохимическому потенциалу, обычно называют уровнем химического потенциала или уровнем Ферми полупроводника. Его положение по отношению к дну зоны проводимости определяется как и для металлов, значением химического потенциала  $m$  (обычно  $m < 0$ , т.е. уровень Ферми в полупроводниках, в отличие от металлов, располагается ниже дна зоны проводимости).

Внешней работой выхода  $A_{внеш}$  электрона из полупроводника называют наименьшую работу, которую нужно затратить, для удаления в вакуум электрона, начальная энергия которого в полупроводнике равна энергии, соответствующей дну зоны проводимости. Связь между  $A$  и  $A_{внеш}$  имеет вид  $A = A_{внеш} - m$ . Внешняя работа выхода определяется свойствами кристаллической решетки полупроводника и для различных полупроводников составляет от 1 до 6 эВ.

Для чистого (беспримесного) полупроводника

$$\mu = -\frac{\Delta W_0}{2} + \frac{kT}{2} \ln \frac{v_h}{v_e},$$

где  $\Delta W_0$  – энергия активации собственной проводимости;

$k$  – постоянная Больцмана.

$$v_e = \frac{2}{h^3} (2\pi m_e^* kT)^{\frac{3}{2}};$$

$$v_h = \frac{2}{h^3} (2\pi m_h^* kT)^{\frac{3}{2}},$$

$h$  – постоянная Планка;

$m_e^*$ ,  $m_h^*$  – эффективные массы электрона и дырки. При  $T = 0$  К

$\mu = -\frac{\Delta W_0}{2}$ , т.е. уровень Ферми располагается посередине запрещенной

зоны.

### Контакт металл–полупроводник

Если зонд выполнен из полупроводникового материала, а поверхность исследуемого образца металлическая, то при подводе зонда микроскопа к исследуемой поверхности электроны начинают переходить из тела с мень-

шей работой выхода в тело с большей работой выхода. При этом между металлом и полупроводником устанавливается контактная разность потенциалов, обусловленная двойным электрическим слоем, образующимся в области контакта и называемым контактным слоем. Вследствие малой концентрации носителей тока в полупроводнике (порядка  $(10^{14} \text{ ч } 10^{16}) \text{ см}^{-3}$  вместо  $10^{22} \text{ см}^{-3}$  в металлах) толщина контактного слоя в полупроводнике достигает  $(10^{-5} \text{ ч } 10^{-4}) \text{ см}$ , т.е. на несколько порядков больше толщины контактного слоя в металлах. В пределах этого слоя в полупроводнике распределён объёмный заряд и существует контактное электрическое поле. Знак объёмного заряда зависит от соотношения между работами выхода электрона из металла ( $A_1$ ) и из полупроводника ( $A_2$ ): если  $A_1 > A_2$ , то объёмный заряд положительный, а если  $A_1 < A_2$ , – отрицательный.

Вследствие существования электрического поля контактного слоя потенциал в точках этого слоя полупроводника при  $A_1 > A_2$  ниже, чем в остальном объёме полупроводника, а при  $A_1 < A_2$  – выше. Поэтому, если  $A_1 > A_2$ , то при прочих равных условиях энергия электронов в контактном слое полупроводника больше, чем в остальном объёме. Так как электрохимический потенциал во всех частях полупроводника одинаков, то в контактном слое дно зоны проводимости поднимается вверх, удаляясь от уровня Ферми, а верхний край валентной зоны также поднимается вверх, приближаясь к уровню Ферми. При этом ширина запрещённой зоны  $DW_0$  между верхним краем валентной зоны и дном зоны проводимости остается такой же, как и в остальном объёме полупроводника.

Если  $A_1 < A_2$ , то направление вектора напряженности контактного электрического поля изменяется на противоположное, так что потенциал в точках контактного слоя полупроводника выше, чем в остальном объёме. Поэтому в контактном слое дно зоны проводимости опускается вниз, приближаясь к уровню Ферми, а верхний край валентной зоны также опускается вниз, удаляясь от уровня Ферми.

Возможны следующие четыре случая контакта металла с примесными полупроводниками (рис. 5.2.3).

$A_1 > A_2$ , полупроводник *n*-типа. Объёмный положительный заряд обусловлен избытком в контактном слое полупроводника положительных ионов донорных примесей. Контактный слой полупроводника обеднён основными носителями тока – электронами в зоне проводимости. Поэтому

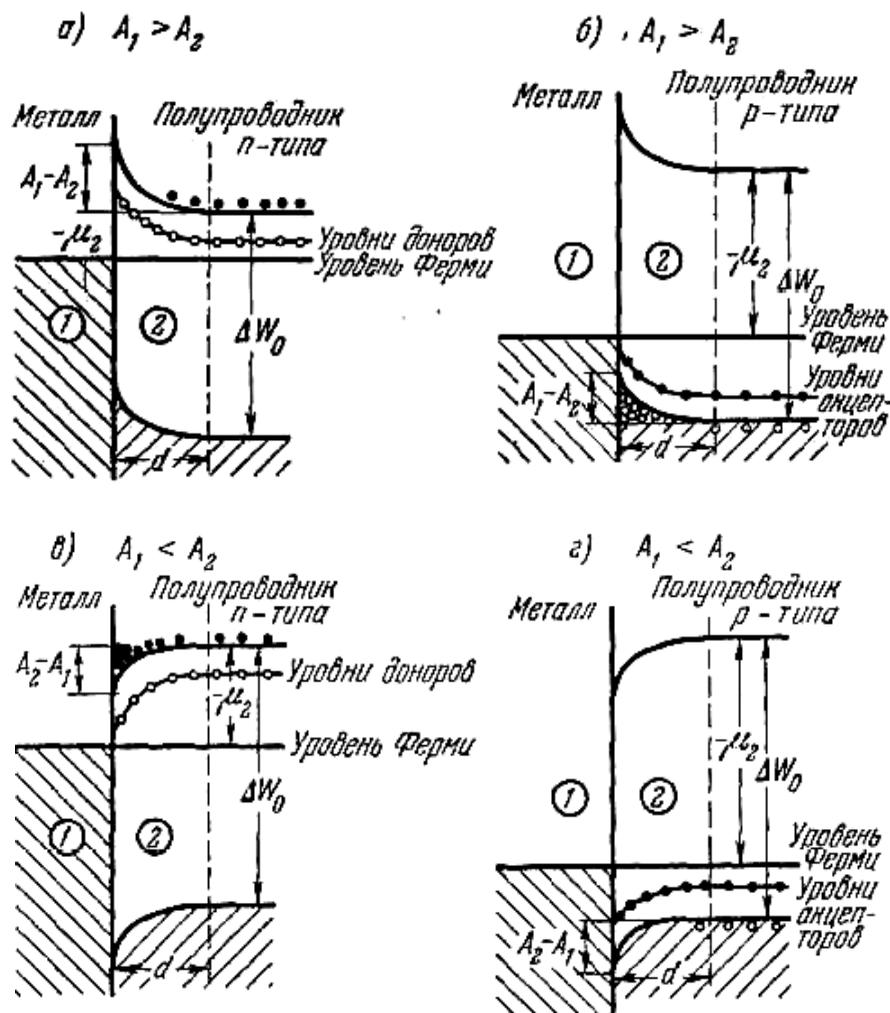


Рис. 5.2.3. Контакт металла с примесными полупроводниками:

- – электроны в зоне проводимости и отрицательные ионы акцепторных примесей;
- – дырки в валентной зоне и положительные ионы донорных примесей;
- $d$  – толщина контактного слоя

удельное сопротивление контактного слоя во много раз больше удельного сопротивления остальной части полупроводника. Такой контактный слой называют запирающим.

$A_1 > A_2$ , полупроводник *p*-типа. В контактном слое полупроводника имеется избыток основных носителей тока – дырок в валентной зоне. Поэтому контактный слой обладает повышенной проводимостью. Такой контактный слой называют антизапирающим.

$A_1 < A_2$ , полупроводник *n*-типа. Отрицательный объёмный заряд контактного слоя полупроводника обусловлен избытком основных носителей тока – электронов в зоне проводимости. Поэтому контактный слой полупроводника обладает повышенной проводимостью, т.е. представляет собой антизапирающий слой.

$A_1 < A_2$ , полупроводник *p*-типа. В контактном слое полупроводника имеется избыток отрицательных ионов акцепторных примесей и недостаток основных носителей тока – дырок в валентной зоне. Поэтому контактный слой запирающий.

### Контакт полупроводник–полупроводник

Граница соприкосновения двух полупроводников с различными, *n*- и *p*-типами проводимости называется электронно-дырочным переходом. Двойной электрический слой *p-n*-перехода образуется в результате перемещения электронов из *n*- в *p*-полупроводник (так как  $A_p > A_n$ , где  $A_p$  и  $A_n$  – работы выхода электрона из полупроводников *p*- и *n*-типа), а дырок – в противоположном направлении. При этом в контактном поле *n*-полупроводника оказывается избыток положительных ионов донорных примесей, а в контактном слое *p*-полупроводника – избыток отрицательных ионов акцепторных примесей. Таким образом, контактный слой в обоих полупроводниках обеднён основными носителями тока и обладает пониженной

проводимостью, т.е. является запирающим. Толщина  $d$   $p$ - $n$ -перехода в практически важных случаях составляет  $(10^{-4} \text{ч} 10^{-5})$  см.

## Трибоэлектрический эффект

Разность потенциалов между проводниками также может появиться вследствие возникновения электрических зарядов при трении – трибоэлектрический эффект. Этот эффект наблюдается при взаимном трении двух диэлектриков, полупроводников или металлов различного химического состава или одинакового состава, но разной плотности, при трении металлов о диэлектрики, при трении двух одинаковых диэлектриков, при трении жидких диэлектриков друг о друга или о поверхность твёрдых тел и др. При этом электризуются оба тела, их заряды одинаковы по величине и противоположны по знаку.

При трении двух химически одинаковых тел положительные заряды получает более плотное из них. При трении двух диэлектриков положительно заряжается диэлектрик с большей диэлектрической проницаемостью  $\epsilon$ . Тела можно расположить в трибоэлектрические ряды, в которых предыдущее тело электризуется положительно, а последующее — отрицательно (ряд Фарадея: (+) мех, фланель, слоновая кость, перья, горный хрусталь, флинтглас, бумажная ткань, шёлк, дерево, металлы, сера (–)). Для диэлектриков, расположенных в трибоэлектрический ряд, наблюдается убывание твёрдости (ряд Гезехуса: (+) алмаз (твёрдость 10), топаз (8), горный хрусталь (7), гладкое стекло (5), слюда (3), кальцит (3), сера (2), воск (1) (–)). Для металлов характерно возрастание твёрдости. У жидких диэлектриков положительный заряд приобретает вещество с большей  $\epsilon$  или поверхностным натяжением.

У твёрдых тел трибоэлектричество объясняется переходом носителей тока при трении от одного тела к другому. В случае двух металлов, двух

полупроводников или металла и полупроводника оно обусловлено переходом электронов от вещества с меньшей работой выхода к веществу с большей. При контакте металла с диэлектриком трибоэлектричество возникает за счёт перехода электронов из металла в диэлектрик и перехода ионов того или иного знака из диэлектрика на поверхность металла. При трении двух диэлектриков - диффузия электронов и ионов. Существенную роль может играть также разное нагревание тел при трении, что вызывает переход носителей тока с локальных неоднородностей более нагретой поверхности. При трении двух диэлектриков, отличающихся только плотностью, из более плотного вещества будет диффундировать больше электронов и оно зарядится положительно. При контакте двух разных диэлектриков из вещества с большей  $e$  будет переходить больше электронов.

### 5.2.2. Сила Ван-дер-Ваальса

При дальнейшем сближении зонда АСМ и исследуемой поверхности, на зонд начинают действовать межмолекулярные силы Ван-дер-Ваальса.

Эта сила имеет три составляющие  $W_{ориент}$ ,  $W_{инд}$ ,  $W_{дисп}$  :

$$\bullet \quad W_{ориент} = \frac{\int W_D \exp\left(-\frac{W_D}{kT}\right) dv}{\int \exp\left(-\frac{W_D}{kT}\right) dv},$$

где  $W_D = \frac{d_1 d_2 - 3(d_1 n)(d_2 n)}{r^3} \sim \frac{1}{r^3}$ ;

$$\bullet \quad W_{инд} = d_{инд} E = \chi E^2 = \chi \frac{-3(nd_1)^2 + d_1^2}{r^6} = -\chi \frac{d_1^2}{2r^6} \sim \frac{1}{r^6},$$

где  $d_{инд} = \chi E$  – индукционный момент;

$\chi$  – поляризуемость молекулы;

$$\bullet \quad W_{\text{дисп}} = -\frac{A_3}{r^6}, \text{ где } A_3 = \frac{3I_1I_2}{2(I_1 + I_2)} \chi_1\chi_2,$$

где  $I_1, I_2$  – потенциалы ионизации;

$\chi_1, \chi_2$  – поляризуемости молекул.

Эти составляющие обладают несколько отличной физической природой, но их потенциал зависит от расстояния между молекулами одинаково — как  $l/r^6$ . Это обстоятельство позволяет непосредственно сравнивать константы взаимодействия, соответствующие трём составляющим силы Ван-дер-Ваальса, причём по причине их одинаковой зависимости от расстояния пропорция между компонентами будет сохраняться при различных  $r$ . Сами же константы при множителе  $l/r^6$  будут отличаться для разных веществ:

$$W = W_{\text{ориент}} + W_{\text{инд}} + W_{\text{дисп}} \sim \frac{l}{r^6}.$$

В основе всех трёх составляющих силы Ван-дер-Ваальса лежит взаимодействие диполей. Проводя оценки для типичных условий АСМ-эксперимента в режиме контакта сила ван-дер-ваальсовского притяжения:

$$F_{\text{ВдВ}} \sim 10^{-8} \text{ ч } 10^{-9} \text{ Н.}$$

Сила, соответствующая потенциалу парного взаимодействия равна производной по расстоянию  $r$ :

$$f = \frac{6A_3}{r^7},$$

где  $A_3$  – константа Гамакера.

Исходя из этого микроскопического описания, можно определить силу притяжения, возникающую между зондом и образцом в целом. Оно равно сумме всех парных взаимодействий молекул кантилевера и исследуемой поверхности:

$$F = \sum_{\substack{\text{по молекулам} \\ \text{зонда}}} \left( \sum_{\substack{\text{по молекулам} \\ \text{образца}}} f \right).$$

Понятно, что при этом результат будет существенно зависеть от пространственной геометрии задачи.

Пренебрегая дискретным распределением взаимодействующих центров (молекул), несложно от парного суммирования перейти к двойному интегралу:

$$F = \int n_1 dV_1 \left( \int n_2 dV_2 f \right),$$

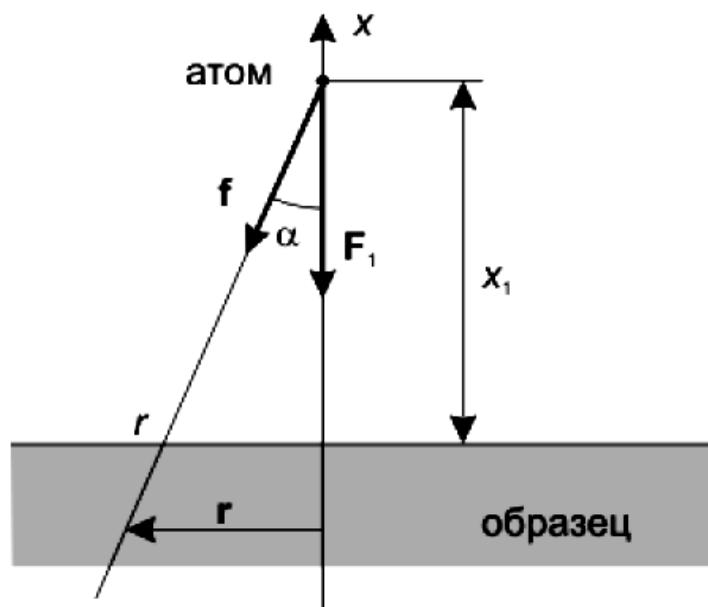
где  $n_1$  и  $n_2$  – концентрации молекул зонда и образца (плотности).

Физический смысл внутреннего интеграла – сила взаимодействия одной молекулы с плоскостью. Сила притяжения очень быстро спадает с увеличением расстояния ( $r^{-7}$ ), поэтому отдаленные части системы не вносят в интеграл существенного вклада. Благодаря этому, интегрирование можно распространить на все полупространство, как если бы оно было однородным образцом.

Для вычислений внутреннего интеграла (обозначив его через  $F_1$ ) вводится цилиндрическая система координат (рис. 5.2.4). Начало отсчёта совмещается с молекулой. Из соображений симметрии понятно, что результирующая сила будет направлена вертикально вниз. При этом горизонтальные составляющие силы притяжения к двум молекулам, расположенным симметрично относительно оси  $x$ , компенсируются. Поэтому проще изначально учитывать лишь вертикальную компоненту силы

$$f_1 = \cos \alpha \frac{6A_3}{r^7}.$$

Эта сила будет одинакова для всех точек кольца с радиусом  $s$ , поэтому интегрирование по углу вокруг оси  $x$  сводится просто к умножению на  $2\pi$ . Дальнейшие вычисления достаточно просты:



**Рис. 5.2.4.** Система атом – плоский образец

$$F_1 = n_2 \int_0^{\infty} 2\pi\rho d\rho \int_{x_0}^{\infty} dx \frac{z}{\sqrt{\rho^2 + x^2}} \frac{6A_3}{\left(\sqrt{\rho^2 + x^2}\right)^7} = \frac{\pi A_3 n_2}{2x_0^4} .$$

Чтобы взять внешний интеграл нужно проинтегрировать по объёму зонда:

$$F = \int F_1 n_1 dV_1 .$$

Дальнейшие вычисления необходимо проводить для конкретной модели острия кантилевера.

### 5.2.3. Капиллярные силы

В случае, если исследование поверхности проводится на воздухе, то на поверхности исследуемого образца обязательно присутствует жидкая плёнка, состоящая в основном из воды.

В тот момент, когда кантилевер касается жидкой плёнки на плоской поверхности, происходит смачивание поверхности кантилевера (рис. 5.2.5),

т.к. энергетически более выгоден контакт жидкости с материалом кантилевера (если он гидрофильный), чем с воздухом.

Искривлённая поверхность шейки будет стремиться выпрямиться, что возможно только за счёт опускания кантилевера. Это означает возникновение притяжения кантилевера к образцу.

Расчёт этой силы довольно прост. Пусть радиус кривизны зонда намного превосходит все остальные характерные размеры задачи (рис. 5.2.6).

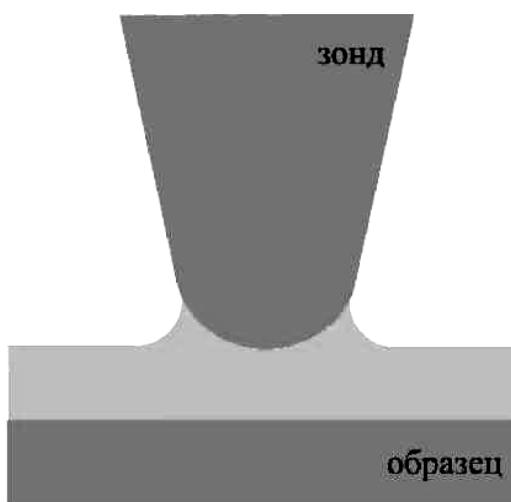


Рис. 5.2.5. Образование «шейки»

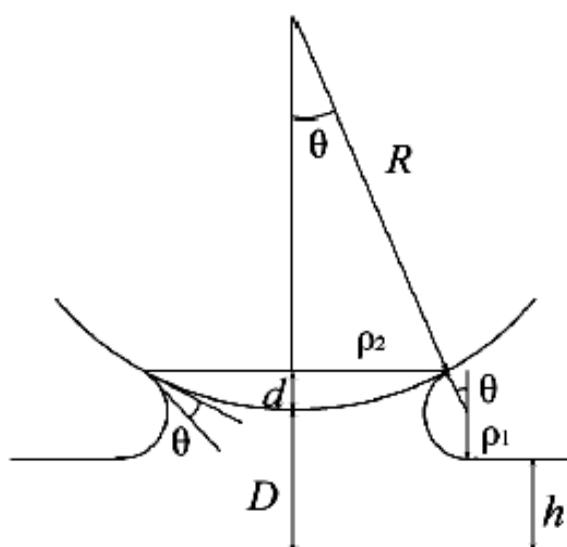
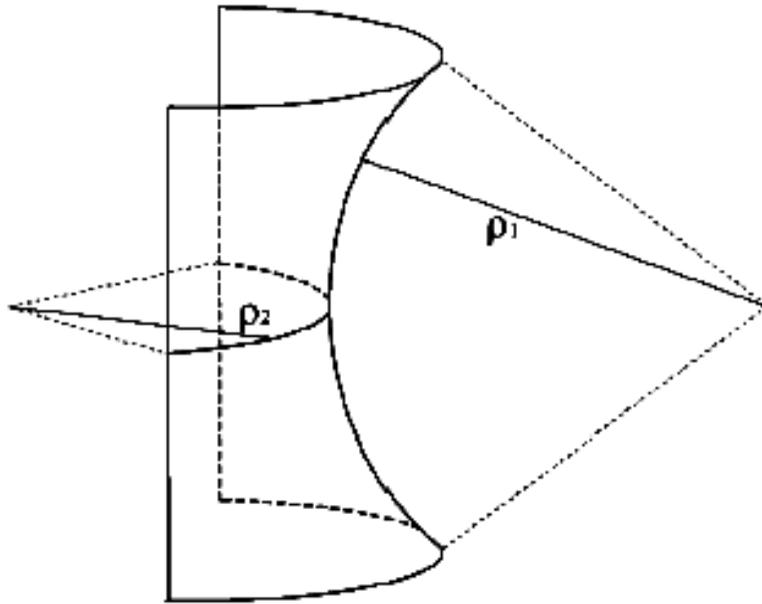


Рис. 5.2.6. К расчёту капиллярной силы



**Рис. 5.2.7.** Пояснение к формуле Лапласа

На рис. 5.2.6 введены следующие обозначения:  $D$  – расстояние зонд-образец,  $d$  – «глубина погружения»,  $h$  – толщина плёнки,  $c_1$  – меньший радиус кривизны жидкой поверхности,  $c_2$  – радиус площадки соприкосновения кантилевера с жидкостью.

Согласно формуле Лапласа, давление в жидкости будет меньше атмосферного на величину:

$$\Delta p = \sigma \left( \frac{1}{\rho_1} + \frac{1}{\rho_2} \right) \approx \frac{\sigma}{\rho_1}, \quad \rho_1 \ll \rho_2 .$$

Это давление приложено к поверхности соприкосновения зонда с жидкостью площадью  $\pi \rho_2^2 \approx 2\pi R d$  (рис. 5.2.7). Сила притяжения кантилевера к образцу, обусловленная капиллярным эффектом, составляет:

$$F = P \cdot 2\pi R d = 2\pi R d \frac{\sigma}{\rho_1} .$$

Для простоты можно считать краевые углы  $u$  для образца и зонда равными. Тогда  $c_1 \cos u + c_1 = D + d - h$ . Таким образом,

$$F_{\text{кап}} = \frac{2\pi R\sigma(1 + \cos\theta)}{1 + \frac{D-h}{d}} . \quad (5.2.1)$$

Если можно пренебречь толщиной жидкой плёнки ( $h=0$ ), то справедливо равенство  $2c_1 \cos u \approx D + d$ , и формула (5.2.1) упрощается до формулы

$$F_{\text{кап}} = \frac{4\pi R\sigma \cos\theta}{1 + \frac{D}{d}} .$$

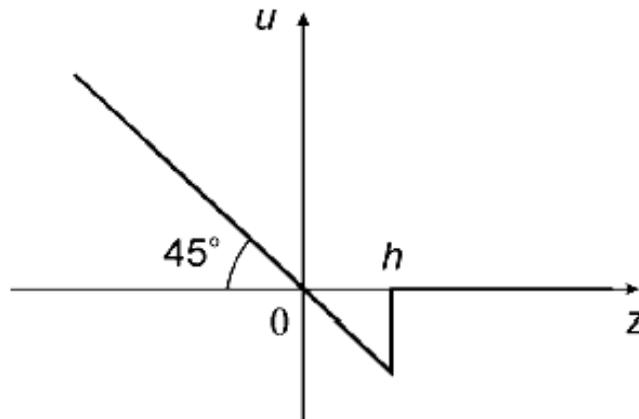
Для оценок достаточно воспользоваться максимальным значением капиллярной притягивающей силы, которое достигается при  $D = 0$ . При этом из формулы исчезает неизвестный параметр  $d$ :

$$F_{\text{кап}} = F_{\text{макс}} = 4\pi R\sigma \cos\theta .$$

Из данной формулы можно установить, что для кантилевера, радиус  $R$  которого равен 10 нм, капиллярная сила  $F_{\text{кап}} \sim 10^{-8}$  ч  $10^{-9}$ . Таким образом, по порядку величины капиллярная сила совпадает с взаимодействием Ван-дер-Ваальса и электростатическими силами.

При подводе и отводе кантилевера от поверхности образца возникает гистерезис. При движении вверх перетяжка остается дольше. Поверхность кантилевера уже смочена, и жидкая шейка тянется за зондом. Когда связи прерываются, капиллярная сила притяжения перестаёт действовать и кантилевер резко выпрямляется.

На рис. 5.2.8 по вертикальной оси отложена деформация кантилевера  $u$ , а по горизонтальной – расстояние  $z$ , на которое он подводится к образцу. На больших расстояниях деформация отсутствует, график горизонтален. В левой части зависимость линейна и имеет наклон  $45^\circ$ . Это означает контакт с поверхностью, так как острие уперлось в образец и возникающая деформация равна величине прижатия зонда к поверхности.

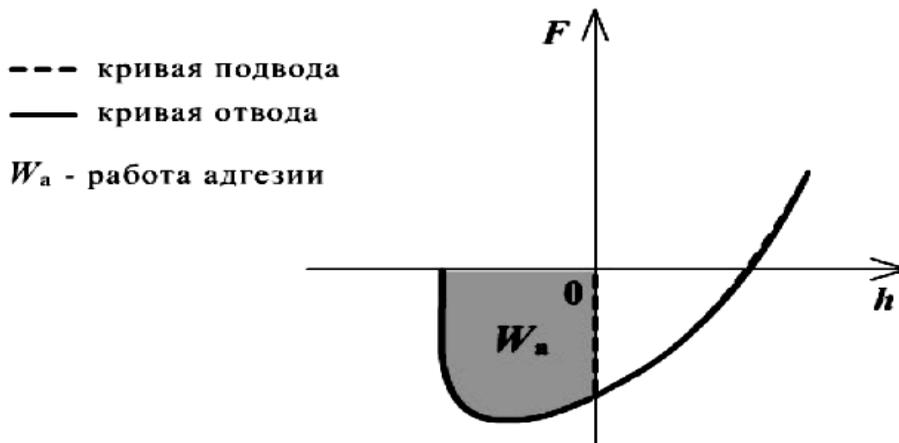


**Рис. 5.2.8** График зависимости деформации кантилевера от расстояния зонд-образец

При отведении кантилевера острие залипает: капиллярная сила удерживает его, и линейная зависимость продолжается уже ниже горизонтальной оси до тех пор, пока перетяжка не исчезнет. После чего кантилевер скачком выпрямляется до недеформированного состояния.

#### 5.2.4. Адгезионные силы

В промежуточной области, когда между одними парами молекул зонда и образца действуют силы притяжения (потенциал  $-1/r^6$ ), а между другими — отталкивания  $1/r^{12}$ , точное нахождение силы взаимодействия всего зонда с образцом невозможно. Более того, в переходной области возникает качественно новое явление — адгезия. Своим происхождением оно обязано наличию короткодействующих молекулярных сил. Следует различать две разновидности адгезии: зонд–жидкая плёнка на поверхности и зонд–твёрдый образец. Если первый случай сводится к капиллярному взаимодействию, которому посвящён раздел 5.2.3, то причина адгезионных сил между зондом и твёрдым образцом — молекулярное электростатическое взаимодействие.



**Рис. 5.2.9.** Работа адгезии. Показаны зависимости адгезионных сил при подводе и отводе зонда от поверхности образца

Адгезия – неконсервативный процесс. Силы при подводе кантилевера к образцу отличаются от сил при отрыве зонда от поверхности. На подобную операцию будет потрачена работа, называемая работой адгезии (рис. 5.2.9).

Эта работа состоит из следующих компонентов:

$$W_a = W_a^d + W_a^p + W_a^i + W_a^h + W_a^\pi + W_a^{da} + W_a^e ,$$

где  $d$  – дисперсионное взаимодействие Лондона;

$p$  – диполь–дипольное (ориентационное) взаимодействие;

$i$  – индукционное взаимодействие;

$h$  – водородная связь;

$p$  –  $p$ -связь;

$da$  – донорно-акцепторная связь;

$e$  – электростатическое взаимодействие.

Причём первые три слагаемых – работа сил Ван-дер-Ваальса.

Причина адгезии – электростатические силы на границе раздела двух тел, возникающие благодаря образуемому в контакте двойному заряженному слою. Его происхождение различно для материалов разных типов. Для металлов она определяется контактной разностью потенциалов, состо-

яниями внешних электронов атомов поверхностного слоя и дефектами кристаллической решётки; для полупроводников – поверхностными состояниями и наличием примесных атомов; для диэлектриков – дипольным моментом групп молекул на границе фаз.

Адгезия – необратимый процесс. При наличии контактной разности потенциалов, например, возникает дрейф электронов, который ведёт к увеличению энтропии. Именно поэтому силы при сведении и разведении кантилевера и образца различаются (см. ниже), и процесс, тем самым, оказывается неконсервативным.

Для количественного описания адгезии используются приближённые модели.

### Модель DMT адгезии твердых тел

Модель DMT (*Derjagin, Muller, Toropov* — 1975 г.) применима для зондов с малым радиусом закругления и большой жёсткостью. При этом считается, что геометрия деформированных поверхностей мало отличается от той, что даёт решение задачи Герца. Учет сил Ван-дер-Ваальса по периметру контактной площадки приводит к дополнительному притяжению между зондом и образцом, которое уменьшает силы упругого отталкивания.

Выражения для давления и глубины проникновения получают следующий вид:

$$F = \frac{Ka^3}{R} - 2\pi R\sigma; \quad h = \frac{a^2}{R},$$

где  $R$  – радиус закругления зонда;

$a$  – радиус контактной площадки;

$$K \left( \frac{1}{K} = \frac{3}{4} \left( \frac{1 - \mu'^2}{E'} + \frac{1 - \mu^2}{E} \right) \right) - \text{эффективный модуль Юнга;}$$

$y$  – работа адгезии (рис. 5.2.10).

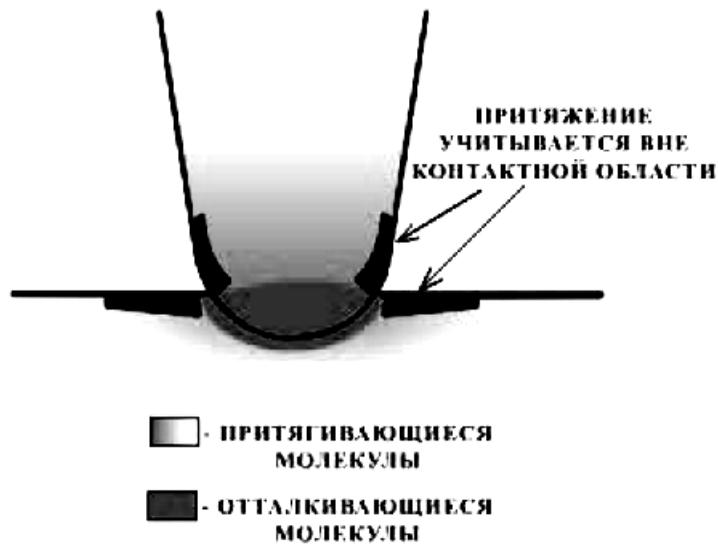


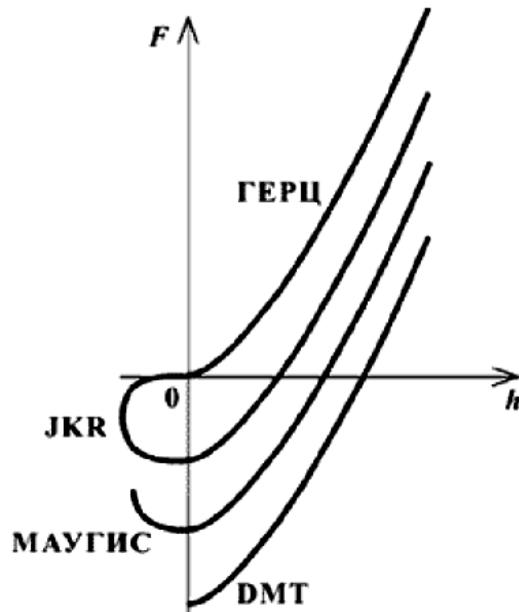
Рис. 5.2.10,а. Применимость моделей JKR и DMT



Рис. 5.2.10,б. Применимость модели Маугиса

### Модель JKR адгезии твердых тел

Модель JKR (*Johnson, Kendall, Roberts* — 1964-1971 г.г.) пригодна для зондов с большим радиусом кривизны и малым коэффициентом жёсткости. Такие системы называются сильноадгезивными. Модель учитывает вклад ван-дер-ваальсовых сил в пределах контактной площадки.



**Рис. 5.2.10,в.** Зависимость силы от глубины проникновения для задачи Герца, моделей JKR, DMT и Маугиса

За счёт этого возникает притяжение, которое не только снижает силу упругого отталкивания (правая часть графика, представленного на рис. 5.2.10,в), но и приводит к образованию шейки ( $h < 0$ , левая часть графика) и отрицательной силе. График соответствует формулам:

$$F = \frac{Ka^3}{R} - \sqrt{6\pi\sigma Ka^3};$$

$$h = \frac{a^2}{R} - \frac{2}{3} \sqrt{\frac{6\pi\sigma a}{K}}.$$

### Модель Маугиса адгезии твёрдых тел

Механика Маугиса (1992 г.) является наиболее сложным и точным подходом. Его можно использовать для любых систем: как с большой, так и с малой адгезией. Степень адгезии определяется параметром  $l$ :

$$\lambda = \frac{2,06}{\xi_0} \left( \frac{R\sigma^2}{\pi K^2} \right)^{\frac{1}{3}},$$

где  $\sigma_0$  – межатомное расстояние.

Модели DMT и JKR являются предельными случаями механики Маугиса, отвечающие разным параметрам  $l$ . Для жёстких материалов (DMT)  $l \rightarrow 0$ , для податливых (JKR)  $l \rightarrow \infty$ .

В модели Маугиса считается, что молекулярная сила притяжения действует в пределах кольцевой области на границе контактной площадки. Поправка Маугиса к решению задачи Герца выражается неявно, через параметр  $m$ :

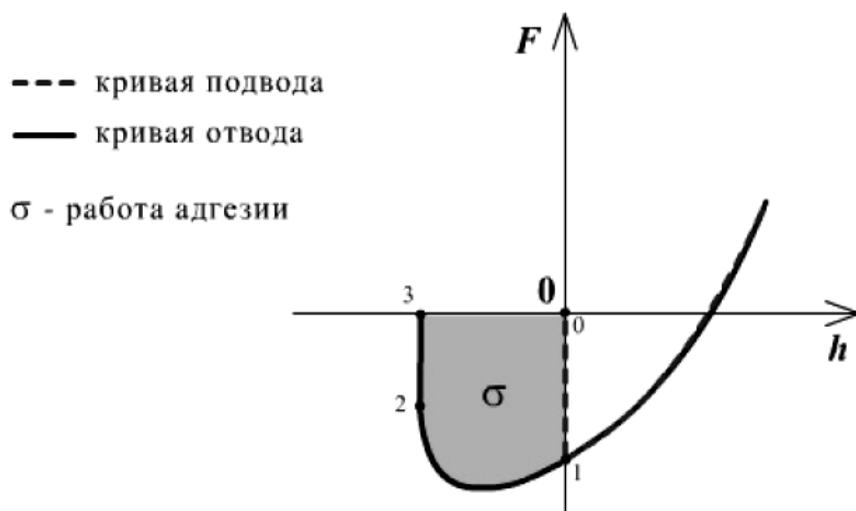
$$1 = \frac{\lambda a^2}{2} \left( \frac{K}{\pi R^2 \sigma} \right)^{\frac{2}{3}} \left[ \sqrt{m^2 - 1} + (m^2 - 2) \arctan \sqrt{m^2 - 1} \right] + \\ + \frac{4\lambda a^2}{3} \left( \frac{K}{\pi R^2 \sigma} \right)^{\frac{1}{3}} \left[ 1 - m + \sqrt{m^2 - 1} \arctan \sqrt{m^2 - 1} \right];$$

$$F = \frac{Ka^3}{R} - \lambda a^2 \left( \frac{\pi \sigma K^2}{R} \right)^{\frac{2}{3}} \left[ \sqrt{m^2 - 1} + m^2 \arctan \sqrt{m^2 - 1} \right];$$

$$h = \frac{a^2}{R} - \frac{4\lambda a}{3} \left( \frac{\pi \sigma}{RK} \right)^{\frac{1}{3}} \sqrt{m^2 - 1}.$$

В модель JKR и механику Маугиса изначально заложено существование гистерезиса при подведении и отведении. Считается, что при подведении кантилевера сила притяжения возникает резко в момент касания, и система переходит от точки 0 в точку 1 (рис. 5.2.11). При отведении кантилевера, система «движется» по другой траектории 1–2 до момента отрыва, который происходит скачкообразно 2–3.

Наличие на графике петли 0–1–2–3 означает затрату работы на отрыв зонда от образца, которая равна площади петли. Это и есть работа адгезии  $u$ .



**Рис. 5.2.11.** Зависимость силы от глубины проникновения для моделей Герца, DMT, JKR и Маугиса при одном  $l$

### 5.2.5. Упругие взаимодействия зонда с образцом

При исследовании любых приповерхностных свойств исследуемых объектов в зоне контакта зонда с образцом возникают механические напряжения, и для верной интерпретации получаемых результатов при сканировании необходимо учитывать деформации материала зонда и образца под воздействием таких механических напряжений. Для наноиндентирования задача определения механических нагрузок в зоне контакта зонда с исследуемой поверхностью и геометрических характеристик контактной зоны является сутью данного метода исследований.

#### Задача Герца

Несмотря на то что инденторы номинируются как остроконечные пирамиды, острие реального индентора в первом приближении можно считать сферическим с эффективным радиусом  $R$ . Вследствие этого можно свести задачу определения нагрузок в контактной области к задаче Герца.

Решение задачи Герца в общем виде имеет неявный вид. Для получения численного решения контактной задачи Герца можно упростить условия данной задачи, сделав следующие предположения:

- Материалы, как образца так и кантилевера, изотропны, а значит, все их упругие свойства описываются всего двумя парами параметров – модулями Юнга  $E, E'$  и коэффициентами Пуассона  $\nu, \nu'$ .
- Вблизи точки соприкосновения недеформированному участку поверхности каждого из тел можно приписать два радиуса кривизны  $r_1, r_2$  (для зонда) и  $r'_1, r'_2$  (для исследуемого участка образца) во взаимно перпендикулярных плоскостях, ортогональных самой поверхности в данной точке (рис. 5.2.12).
- Деформации малы по сравнению с радиусами кривизны поверхностей.

### Решение задачи Герца

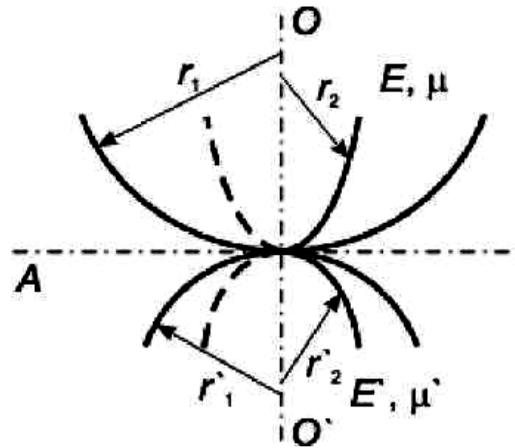
Чтобы получить общее представление о деформациях при упругом контакте и найти характерные числовые значения, достаточно ограничиться анализом взаимодействия сферических поверхностей – острия зонда и небольшого участка образца. Это означает, что  $r_1=r_2=r, r'_1=r'_2=r'$ .

При воздействии нагрузки соприкасающиеся тела деформируются так, что вместо точки касания образуется некоторая контактная площадка. Из осевой симметрии задачи следует, что эта площадка будет иметь форму круга с радиусом  $a$ .

Для расчётов будут удобны следующие величины:

$$\frac{1}{R} = \frac{1}{r} + \frac{1}{r'}$$

а также эффективный модуль Юнга для заданной пары материалов:



**Рис. 5.2.12.** Задача Герца ставится в предположении, что в точке касания до деформации поверхности описываются радиусами кривизны в двух взаимно перпендикулярных плоскостях, ортогональных касательной плоскости  $A$

$$\frac{1}{K} = \frac{3}{4} \left( \frac{1 - \mu'^2}{E'} + \frac{1 - \mu^2}{E} \right).$$

При малых деформациях (допущение 3 задачи Герца) выполняется следующее геометрическое соотношение между глубиной проникновения  $h$  и радиусом контактного круга  $a$ :

$$h = \frac{a^2}{2R},$$

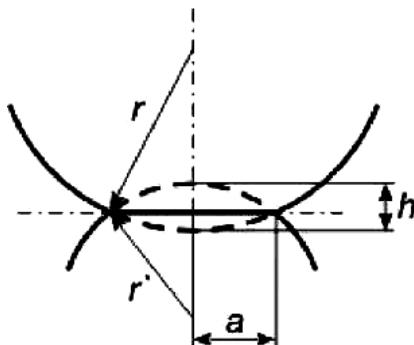
которое можно понять из рис. 5.2.13.

Решение задачи Герца дает соотношение между придавливающей силой  $F$  и глубиной проникновения  $h$ :

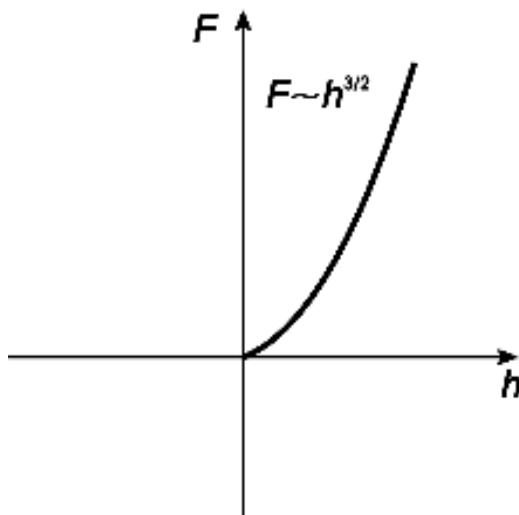
$$F = \frac{Ka^3}{R} = Kh^{\frac{3}{2}} R^{\frac{1}{2}}.$$

Соответственно, давление связано с силой следующим образом:

$$P = \frac{F}{\pi a^2} = \frac{1}{\pi} \sqrt[3]{\frac{FK^2}{R^2}}. \quad (5.2.2)$$



**Рис. 5.2.13.** Связь между радиусом контактной площадки  $a$  и глубиной проникновения  $h$  в деформированном состоянии



**Рис. 5.2.14.** Зависимость силы  $F$  от глубины проникновения  $h$

Приведенное решение для двух соприкасающихся сферических форм содержит важный частный случай: контакт плоского образца и зонда с радиусом кривизны  $R$  ( $r = R$ ,  $r' = \infty$ ). На рис. 5.2.14 изображено графическое решение задачи Герца. График определен при положительных  $h$ .

Решая задачу Герца в общем виде, можно получить решение для любых поверхностей, однако решение такой задачи будет иметь неявный вид. Тем не менее, результат по порядку величины будет такой же, что и в данном, крайне упрощенном случае. Поэтому приближенную оценку характерной величины давления в контакте можно провести на основе формулы (5.2.2). Типичные значения в АСМ: радиус зоны контакта до 10 нм, глубина проникновения до 20 нм, контактное давление до 10 ГПа.

Другое ограничение (условие 1 задачи Герца) состоит в рассмотрении соприкасающихся тел в рамках модели сплошной среды с изотропными характеристиками. Понятно, что на микроуровне молекулярная структура уже играет существенную роль, поэтому подобное приближение довольно условно. В связи с этим решение задачи Герца с более точными геометрическими характеристиками соприкасающихся поверхностей (в отличие от рассмотренного случая) лишено смысла, так как условие 1 раздела задачи Герца само по себе является весьма грубым приближением.

### **Разрушение материала при сканировании**

Оценив давление в месте контакта по формуле (5.2.2), нетрудно установить, разрушение каких материалов может происходить при сканировании. Достаточно сравнить предел прочности материалов (который измеряется в Па) и возникающее напряжение (давление  $P$ ). Впрочем, если в статическом режиме или при медленном сканировании нагрузка превышает критическую, то при большой скорости перемещения кантилевера разрушение может не происходить, даже если превышен порог прочности. Дело в том, что давление сверх критического должно воздействовать дольше, чем продолжается процесс разрушения (время релаксации упругих деформаций составляет примерно  $10^{-6}$  с.). При достаточно быстром сканировании больших площадей это условие может и не выполняться. Но эффект может неожиданно проявиться при изменении размеров кадра, когда его уменьшение приводит к разрушению образца.

Кроме того, при одинаковых условиях сканирования одним и тем же зондом может происходить разрушение твёрдого образца, в то время как более мягкий не разрушится. Это, как отмечалось, является следствием того, что для очень твёрдых веществ размеры контактной

площадки малы, и возникает гораздо большее давление по сравнению с более мягкими образцами.

### Поправки к решению задачи Герца.

#### Неупругие консервативные силы при контакте

Помимо упругости при контакте зонда с образцом возникают и другие силы. Например, взаимодействие Ван-дер-Ваальса (которое действует не только в соприкосновении, но и на некотором расстоянии) приводит к отрицательной добавке к давлению, так как силы Ван-дер-Ваальса, в отличие от сил упругости, являются притягивающими, а не отталкивающими.

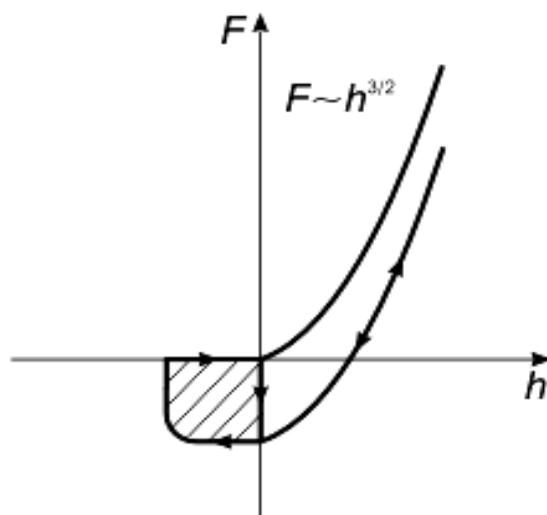
Это, наряду с другими притягивающими микроскопическими взаимодействиями, приводит к смещению вниз графика (см. рис. 5.2.14) – решения задачи Герца. Видно, что при  $h = 0$  сила отрицательна. Это означает, что при незначительном касании зондом образца действует притягивающая сила.

#### Неконсервативные эффекты

Кроме сил упругости и Ван-дер-Ваальса, существуют ещё неконсервативные силы: от трения и капиллярной силы до рассеивания энергии возникающими упругими волнами – фононами. Их наличие приводит к ещё большему искажению решения задачи Герца.

Благодаря неконсервативным силам, в частности, возникает адгезия (или прилипание) зонда к поверхности. При этом прикосновение и отрыв происходят по-разному, т.е. в системе появляется гистерезис.

Зонд, прилипший к поверхности, при движении вверх увлекает за собой небольшой «приклеившийся» участок образца, который в течение некоторого времени, пока не произойдет отрыв, поднимается, образуя «шейку» (см. рис. 5.2.5).



**Рис. 5.2.15.** Зависимость силы  $F$  от глубины проникновения  $h$   
(график решения задачи Герца, а также решение с петлёй гистерезиса, учитывающее неконсервативные силы)

Таким деформациям можно сопоставить отрицательные значения глубины проникновения  $h$ . Это означает, что при обратном движении кантилевера вверх рассмотренный график может уйти левее вертикальной оси, пока не произойдет скачкообразный отрыв. На рис. 5.2.15 стрелками показана траектория в координатах  $F - h$  при перемещении зонда вниз и вверх.

Наличие петли гистерезиса на графике говорит о том, что чтобы прижать зонд к поверхности образца и затем его оторвать, вернув в первоначальное положение, необходимо затратить работу.

### **5.3. Эффекты, возникающие при исследовании приповерхностных свойств**

При исследовании приповерхностных свойств нанобъектов, гетероструктур и нанобластей с различным химическим и фазовым составом взаимодействие зонда микроскопа с исследуемым образцом приводит к появлению различных эффектов, которые могут также оказывать силовое воздействие на зонд, искажая тем самым получаемые результаты при ис-

следовании какого-либо свойства приповерхностной области исследуемого образца.

### **5.3.1. Эффект прилипания–скольжения**

Эффект прилипания–скольжения (stick–slip effect) является ключевым для контактной моды АСМ.

Упрощенная, но наглядная картина эффекта ПС выглядит следующим образом. Первоначально зонд находится в точке минимума потенциальной энергии системы зонд–поверхность. Это взаимодействие характеризуется периодическим потенциалом, отражающим трансляционную симметрию атомарной структуры поверхности. Существенной особенностью модели является предположение о применимости адиабатических условий на каждом шаге движения зонда. При латеральном нагружении контакта, вызванном сканированием кантилевера, энергия запасается в форме упругой энергии контакта, кантилевера и образца. Относительное скольжение зонда и поверхности начинается в тот момент, когда запасенная энергия достаточно велика для того, чтобы зонд мог "выскочить" из потенциальной ямы и зафиксироваться в другой точке поверхности. Затем система релаксирует, при этом "лишняя" энергия быстро диссипирует из области контакта через электрон–фононную подсистему. Время диссипации очень мало, поскольку характерные скорости электронов и фононов на много порядков превосходят типичные скорости сканирования зонда АСМ ( $10^{-7}$  ч  $10^{-4}$   $\text{мс}^{-1}$ ).

Были проведены сравнительные исследования эффекта ПС для различных сочетаний материалов зонда и образца. В экспериментах наблюдавшаяся периодичность скольжения зонда по поверхности соответствовала топографии атомного рельефа поверхности, полученной в нормальной моде АСМ, но положения максимумов латеральных и нормальных сил были слегка смещены относительно друг друга.

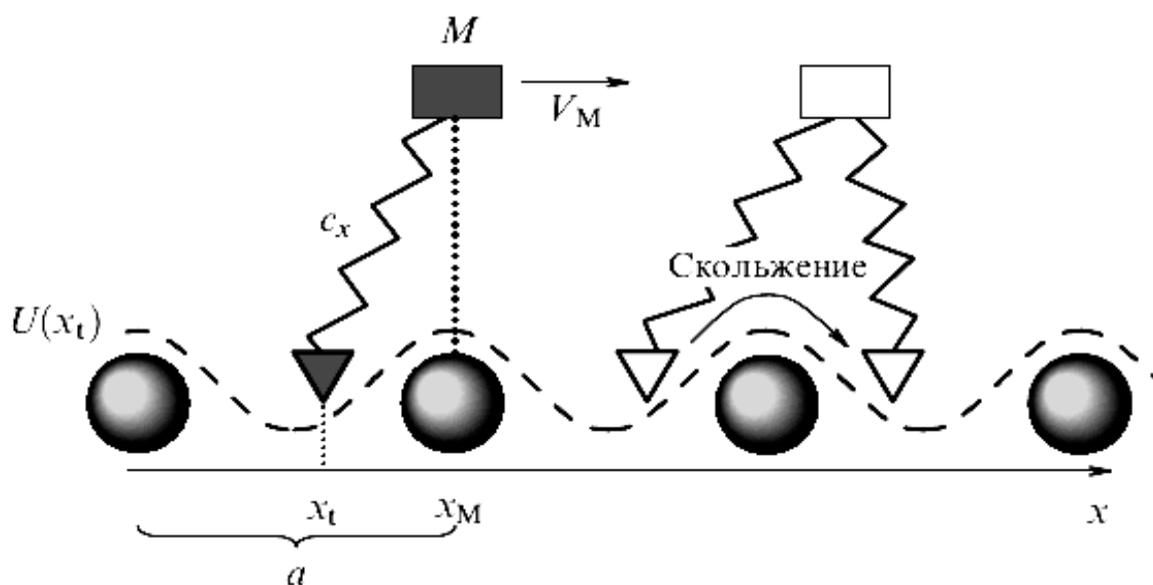
## Механическая модель и результаты исследований

Механическая модель латерального движения зонда по поверхности основывается на ранней работе Томлинсона и приближении независимого осциллятора. Схема, иллюстрирующая её, показана на рис. 5.3.1.

Зонд упруго прикрепляется к телу массы  $M$ , связанному с кантилевером пружиной жесткостью  $c_x$ , и взаимодействует с образцом через периодический потенциал  $U(x_t)$ , где  $x$  – латеральная координата зонда. Движение кантилевера осуществляется с постоянной скоростью  $V_M$ . Диссипация энергии в единицу времени, независимо от микроскопических механизмов, определяется коэффициентом затухания и пропорциональна скорости, а сила трения от скорости не зависит.

Эффективные массы и жёсткости пружин удовлетворяют условию критического демпфирования:

$$\gamma = 2\sqrt{cm}.$$



**Рис. 5.3.1.** Модель эффекта прилипания–скольжения

(Слева: простая модель скольжения зонда по атомно-гладкой поверхности. Справа: схема движения зонда в потенциале поверхности)

Без этого предположения регулярность эффекта ПС с периодом решетки нарушается, и наблюдается нерегулярное скольжение с кратными "прыжками".

В целом, наиболее слабым пунктом данной теории является несоответствие модели точечного осциллятора действительной ситуации с зондом АСМ, когда зона контакта имеет размеры порядка десятков нанометров и, следовательно, включает большое число атомов. В этом случае, как показывают расчёты, при различных положениях аперкса алмазного зонда на поверхности графита изменение потенциальной энергии системы после проведения релаксации структуры не превышает 0,2 эВ, поэтому нельзя говорить о том, что скачки зонда идут от одного положения минимума энергии до другого. Экспериментальный факт заключается в том, что период, связанный с эффектом ПС, совпадает с периодом поверхностной атомной структуры. Таким образом, в режиме контактной латеральной моды АСМ следует говорить не об "истинном атомном разрешении" поверхности, как это имеет место в модуляционном режиме, а об "атомарном контрасте". Об этом свидетельствует и факт отсутствия разрешения точечных атомных дефектов поверхности (в контактной моде).

Используя синусоидальное приближение для периодической силы зонд-поверхность, уравнение движения зонда в одномерном случае можно записать в виде:

$$m \frac{d^2 x}{dt^2} = c(x_M - x) - U_0 \frac{2\pi}{a} \sin\left(\frac{2\pi}{a} x\right) - \gamma \frac{dx}{dt}, \quad (5.3.1)$$

где  $m$  – эффективная масса осциллятора;

$x_M = V_M t$  – равновесное положение недеформированной пружины в момент времени  $t$ ;

$a$  – период решётки;

$\gamma$  – постоянная затухания.

Решая уравнение (5.3.1) относительно  $x(t)$ , величину латеральной силы, действующей на зонд, можно найти из уравнения  $F = c(x_M - x)$ .

Диссипативная сила трения определяется путем усреднения по времени величины латеральной силы  $F_d = \langle F \rangle$ . Если  $V_M = 0$  и зонд находится в положении устойчивого равновесия, то критическим условием начала скачкообразного движения является малость коэффициента жёсткости кантилевера по сравнению с латеральной жёсткостью контакта:

$$c < -\frac{d^2U}{dx^2} = \frac{4\pi^2U_0}{a^2}.$$

В этой модели трение появляется только в результате движения и не зависит от скорости. Для больших скоростей доминирует "вязкое" затухание, и сила трения пропорциональна скорости скольжения,  $F_d \approx \zeta V_M$ .

В случае двумерного движения зонда уравнение (5.3.1) обобщается путем введения двух эффективных масс, коэффициентов жёсткости и постоянных затухания, соответствующих направлениям  $x, y$ .

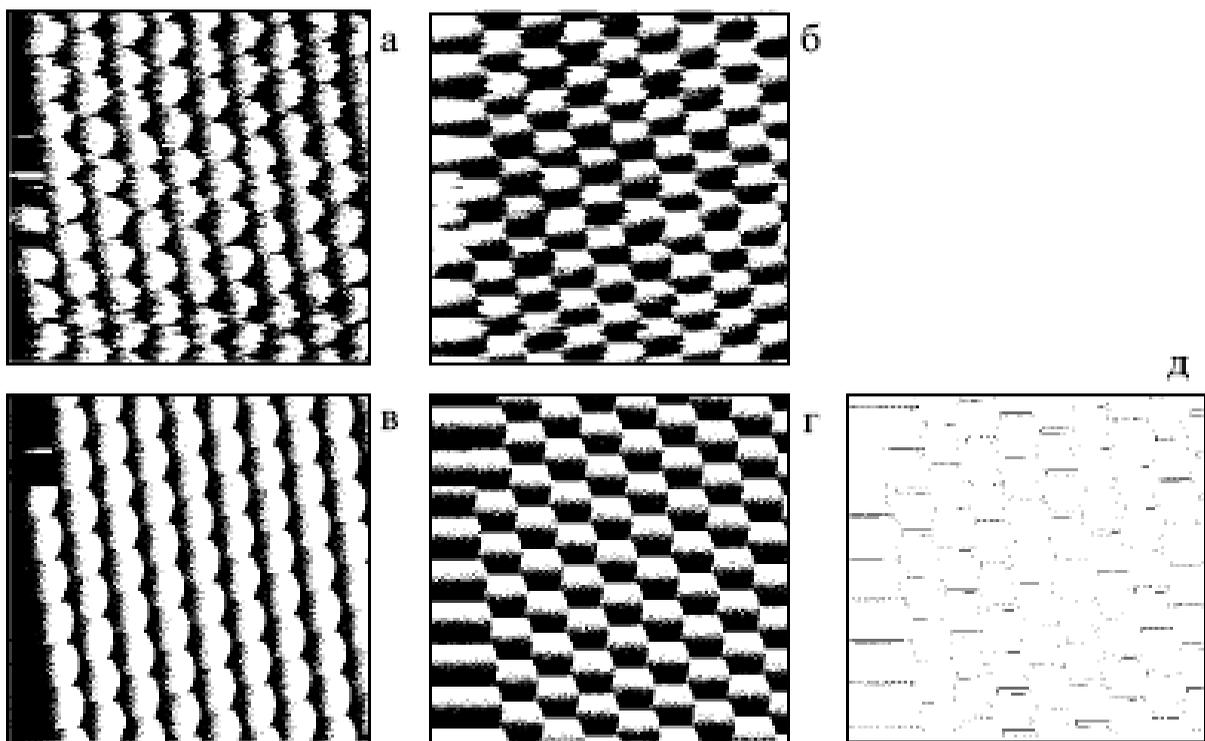
Для высокоориентированного пиролиитического графита (ВОПГ), являющегося тестовым материалом АСМ, потенциал взаимодействия зонда с поверхностью аппроксимируется модельным выражением:

$$U_{НОРГ}(x, y) = -U_0 \left[ 2 \cos\left(\frac{2\pi}{a}x\right) \cos\left(\frac{2\pi}{a\sqrt{3}}y\right) + \cos\left(\frac{4\pi}{a\sqrt{3}}y\right) \right],$$

где  $a = 0,246$  нм;  $U_0 = 0,5$  эВ.

На рис. 5.3.2 показаны экспериментальные и полученные в результате моделирования распределения латеральных сил, действующих на зонд АСМ на поверхности ВОПГ.

На рис. 5.3.2 кадры (а) и (б) соответствуют экспериментальным данным латеральных распределений сил в направлении сканирования  $F_x$  (а) и в перпендикулярном ему  $F_y$  (б), полученным с АФМ на поверхности свежего



**Рис 5.3.2.** Сравнение экспериментальных и полученных при моделировании силовых распределений

скола ВОПГ (размер кадра 242 нм). Теоретические силовые распределения, соответствующие рассмотренной модели, приведены для  $F_x$  (в) и  $F_y$  (г). Сравнение рисунков показывает, что моделирование воспроизводит даже нерегулярности структуры в момент начала скольжения (левая часть изображений). Расстояние между отдельными максимумами (0,246 нм) совпадает как на экспериментальных, так и на теоретических изображениях. Максимумы теоретических силовых распределений показаны точками для проекции  $F_x$  и сплошными линиями для  $F_y$ . Виден сдвиг между ними.

Из расчетов следует, что диссипативные силы трения не зависят от скорости скольжения при  $V < 10^{-2} \text{ см} \cdot \text{с}^{-1}$ . При больших скоростях сканирования начинает преобладать вязкое трение. Это согласуется с экспериментальными измерениями сил трения в алмазе, графите и на плёнках

аморфного углерода. Но экспериментальный интервал исследованных скоростей составлял  $2 \cdot 10^{-4} < V < 2,5 \cdot 10^{-3} \text{ см} \cdot \text{с}^{-1}$ , который в дальнейших исследованиях целесообразно увеличить.

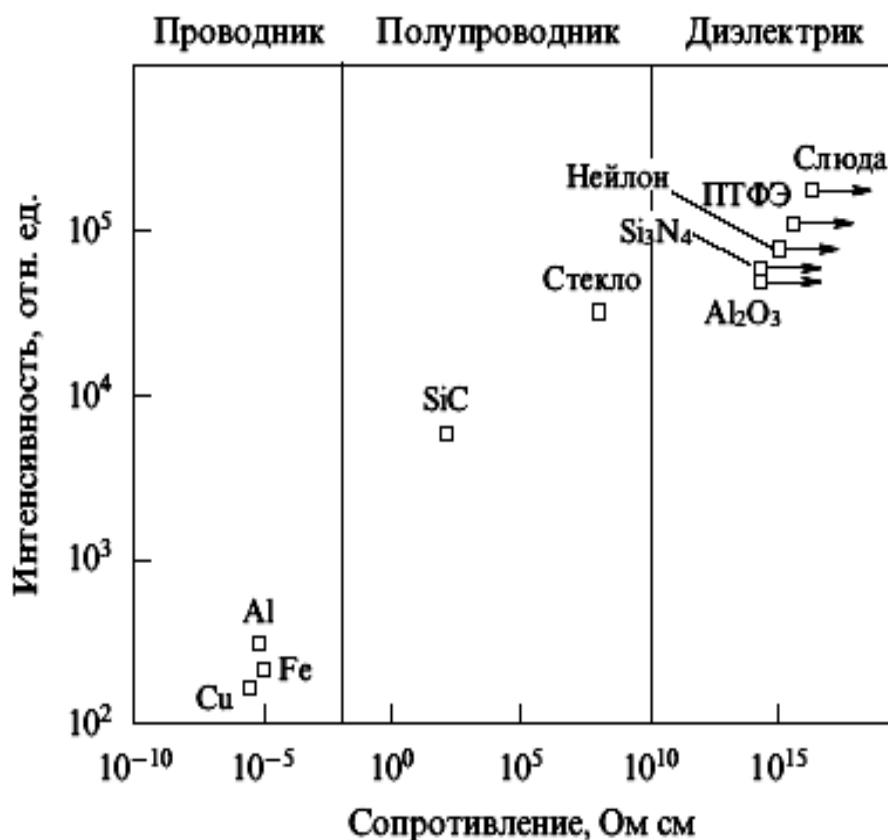
Из экспериментов и модельных расчётов следует, что для наблюдения нестабильности, связанной с эффектом ПС, необходимо сочетание "мягкого" кантилевера с "жёсткой" поверхностью при условии сильного взаимодействия с последней, причём больше энергии диссипирует в случае более мягких контактов. В моделях, однако, не учитываются возможные механизмы диссипации энергии, в которых сила трения пропорциональна скорости.

### **5.3.2. Трибоэмиссия частиц, электромагнитных и акустических волн**

При трении в зоне контакта зонда с исследуемым образцом возникают трибоэлектромагнитные явления, которые могут сильно искажать результаты исследований различных параметров в АСМ.

Для различных контактов скольжения, используемых в устройствах магнитной записи, проводились исследования, в которых измерялся выход заряженных частиц и фотонов в процессе скольжения алмазного зонда с радиусом кривизны острия 10 мкм, при силе нормальной нагрузки 0,5 нН и скорости скольжения  $2 \cdot 7 \text{ см} \cdot \text{с}^{-1}$  (в атмосферных условиях). Результаты этих экспериментов выявили корреляцию, характеризующую увеличение выхода частиц с увеличением удельного сопротивления материала поверхности.

Можно предположить, что образующиеся при трении зонда заряженные частицы оказываются способны покинуть поверхность и могут быть зарегистрированы лишь в том случае, если время разряда возникающих электрических полей достаточно велико. Для металлов это время примерно равно  $10^{-15} \text{ с}$ , поэтому даже быстрые электроны, имеющие скорости порядка скорости Ферми ( $2 \cdot 10^8 \text{ см} \cdot \text{с}^{-1}$ ), не имеют времени для того, чтобы покинуть



**Рис. 5.3.3.** Корреляция между интенсивностью заряженных частиц, эмиттированных из зоны контакта зонд–поверхность, и проводимостью твёрдых тел. (Стрелки показывают, что значения сопротивления выходят за пределы чувствительности прибора)

поверхность. Следовательно, наблюдаемый в случае металлов выход положительных и отрицательных частиц, по всей видимости, связан с химической реактивностью поврежденной зондом поверхности, взаимодействующей с окружающими молекулами воздуха.

Для полупроводников и диэлектриков ситуация противоположная, и значительная часть частиц может уйти с поверхности в процессе полевой десорбции или испарения. Эти соображения дают качественное объяснение наблюдаемым зависимостям выхода частиц в случае проводников и диэлектриков (см. рис. 5.3.3).

Очевидно, необходимо проведение более детальных экспериментов такого рода для определения специфики соответствующих механизмов эмиссии в вакуумных условиях.

Эффекты нанотрибоэмиссии электронов, очевидно, имеют много общего с экзоэлектронной эмиссией, наблюдаемой в различных условиях, измерение которой является одним из перспективных методов изучения поверхностей.

Так при скольжении магнитных записывающих головок с промежуточным типом контакта наблюдалась акустическая эмиссия. Такая эмиссия служит сигналом о контакте зонда с поверхностью при образовании царапин.

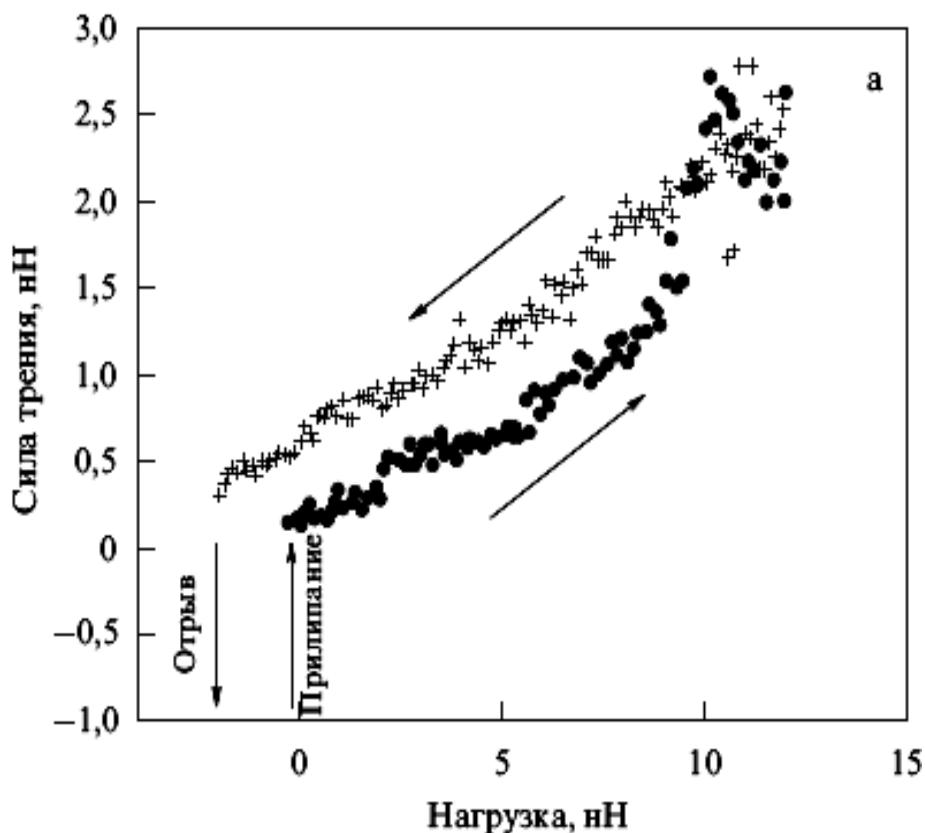
В целом, быстропротекающие процессы, происходящие в зоне контакта поверхностей, могут генерировать излучение разных типов, поэтому соответствующие эффекты целесообразно использовать для изучения нанотрибологических воздействий на поверхность.

### **5.3.3. Трибохимические эффекты**

Трение может стимулировать и подавлять химические процессы, а последние, в свою очередь, влияют на трение, чувствительное к химическому составу зоны контакта. Зависимость нормальных и латеральных сил может быть, в частности, использована для получения соответствующих изображений и исследования химической реактивности поверхности.

Измеренные зависимости сил трения от силы нагрузки для поверхности КС1 приведены на рис. 5.3.4, и качественно резко отличаются от аналогичных зависимостей в случае упругих контактов.

Из экспериментальных данных видно, что сила нагрузки изменяется в относительно широких пределах. Сила трения плавно возрастает до тех пор, пока сила нагрузки не становится равной 5,5 нН, после чего наблюда-



**Рис. 5.3.4.** Зависимость сила трения – сила нагрузки для KCl

ется более резкий подъём. При силах нагрузки, превышающих 9 нН, наблюдаются нерегулярные флуктуации силы трения, причём при уменьшении нагрузки величина силы трения обнаруживает гистерезис и не возвращается к первоначальным значениям.

Для интерпретации данных кривых можно сделать предположение о структурной модификации контактной зоны. Аналогичные зависимости наблюдаются также на плёнках золота, нитрида кремния и органических материалов. Возможно, именно такие (линейные) зависимости специфичны для трибохимического износа.

#### **5.4. Анализ химического состава поверхности**

При проведении исследований приповерхностных свойств химический состав поверхности исследуемых образцов и распределение поверхностных

потенциалов учитывается косвенно или не учитывается совсем. А ведь именно эти свойства поверхности могут существенно влиять на получаемые экспериментальные данные и приводить к расхождениям между значениями параметров, предсказанных теоретически и результатами эксперимента.

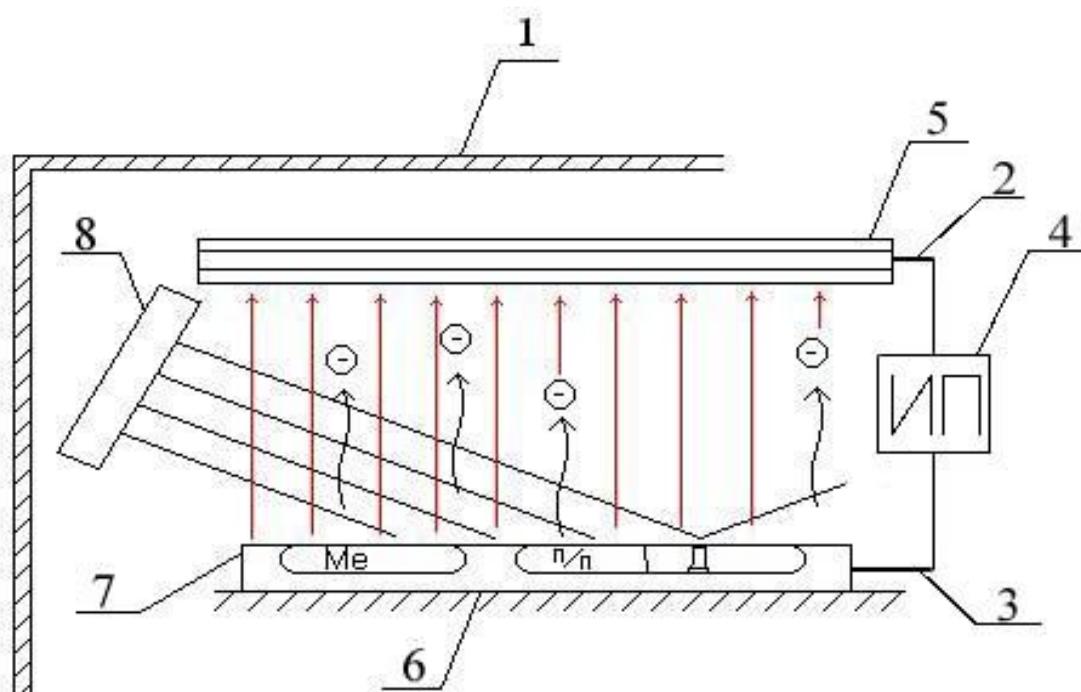
Для получения информации о химическом составе поверхности исследуемого образца, а также о распределении поверхностных электрических потенциалов предлагается использовать устройство определения химического состава поверхности подложки.

В основу разработки положена задача обеспечить возможность определения химического состава на поверхности материала туннельно-зондовым (неконтактным) методом.

Устройство для регистрации химического состава поверхности подложек (рис. 5.4.1) содержит вакуумную камеру 1, внутри которой расположены анод 2 и катод 3, связанные с источником питания 4. Анод 2 выполнен в виде люминесцентного экрана 5, а катод 3 в виде подложкодержателя 6, на котором установлена подложка 7, внутри вакуумной камеры 1 под углом к подложке 7 расположен источник электромагнитного излучения 8, с возможностью подачи квантов излучения на поверхность подложки 7 и отражения от неё без взаимодействия с люминесцентным экраном 5, который расположен параллельно подложке 7.

Устройство работает следующим образом.

При подаче электромагнитного излучения на подложку 7 кванты излучения начинают выбивать электроны с её поверхности, которые под действием электрического поля, создаваемого между анодом 2 и катодом 3, начинают «бомбить» люминесцентный экран 5. В результате того, что подложка 7 легирована, электроны, вылетая из разных областей и осаждаваясь на люминесцентном экране 5, образуют на нём «картину», по которой можно определить химический состав и области легирования материала подложки 7.



**Рис. 5.4.1.** Схема устройства регистрации химического состава подложки

Применение устройства для регистрации химического состава позволяет определять химический состав и области легирования материала подложки туннельно-зондовым (неконтактным) методом, что в свою очередь означает возможность применения данного устройства для исследования образцов, обладающих малой механической жесткостью (органические структуры и биологические объекты).

Рассмотренные выше свойства приповерхностных областей твердых тел являются лишь частью всех свойств, которыми обладают данные области. Также ясно, что для более полного описания этих свойств необходимо проводить дальнейшие исследования в данной области с использованием различных методов исследования, таких как сканирующая туннельная микроскопия, ближнепольная оптическая микроскопия, электронная спектроскопия и т.д.

Задача исследования приповерхностных свойств твёрдых тел является первостепенной при моделировании любых устройств, принцип действия которых основан на данных свойствах. Поэтому для составления полной картины о свойствах приповерхностных областей необходимо объединить исследования в различных направлениях, что исключит возможность повторений, ускорит процесс накопления знаний о данных свойствах, и в итоге приведёт к ускорению темпов развития современных технологий.

## ЗАКЛЮЧЕНИЕ

В заключение необходимо отметить следующее.

Построенная адаптивная модель управления качеством производства материалов создаёт предпосылки обеспечения необходимого уровня качества материалов на стадии проектирования процесса производства и исключает проведение непосредственных текущих доработок.

Построенная дискриминационная модель оптимизации параметров качества производимых объектов обеспечивает выбор наиболее рационального уровня качества на стадиях проектирования производства и проведения инженерного анализа.

Разработанный алгоритм оценки параметров качества позволяет эффективно оценивать уровень качества каждой производимой партии материалов.

Разработанный алгоритм выбора оптимального управленческого решения на основе многокритериальной задачи принятия решений обеспечивает наиболее рациональный выбор технологических решений из множества альтернатив, рассматриваемых в процессе производства, исключая возможность пропуска некоторых мероприятий, порождающих дефекты конечного продукта и материальные затраты, связанные с устранением этих дефектов на заключительных этапах изготовления.

Предложенная практическая реализация разработанных математических моделей и алгоритмов управления качеством, обеспечивает возможность контроля качества в процессе производства электронных устройств, предельно минимизируя возможность появления дефектов.

Проведённый сравнительный анализ результатов машинного эксперимента с известными данными показал сходимость этих результатов, что свидетельствует о корректности выбора математических моделей и алго-

ритмов управления качеством.

Предложенная статистическая оценка качества измеряемых параметров объектов выявила повышение уровня качества по сравнению с показателями, существовавшими ранее и известными в научной и технической литературе.

Построенные математические модели и алгоритмы обеспечивают возможность принимать научно обоснованные, технически целесообразные и экономически выгодные технологические решения для повышения качества электронных устройств на всех стадиях производственного процесса.

## ЛИТЕРАТУРА

1. 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.
2. 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.
3. 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.
4. 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.
5. 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.
6. Кравчук И.С., Тихоглаз Ю.С., Ву Тхи Тует Ланг. Эволюционная стратегия управления в задачах распознавания образов // Системы

- управления и информационные технологии, 2008, 2.3(32). – С. 358-360.
7. 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.
  8. 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.
  9. 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.
  10. Кузькин В.И. Формирование электронных нанообъектов на основе модифицированных углеродных структур. – М.: МИЭМ, 2005.
  11. Нанотехнологическая установка «Луч-2». Руководство пользователя. – М. – 1999.
  12. Ивашов Е.Н., Реутова М.В. Модель влияния магнитного поля на формирование углеродных нанотрубок и фуллеренов. Деп. рукопись ВИНТИ № 2327 В. 2003: 31.12.2003. – 6 с.
  13. 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.
  14. Ивашов Е.Н., Реутова М.В. Углеродные нанотрубки в системах измерения и контроля качества // Сборник докладов научно-технической конференции «Датчики и преобразователи систем измерения, контроля и управления» – М.: МИЭМ, 2003.

15. Ивашов Е.Н., Реутова М.В. Моделирование процессов, протекающих при образовании углеродных наноструктур. Деп. рукопись ВИНТИ № 2324 В. 2003: 31.12.2003. – 6 с.
16. 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.
17. 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.
18. Ивашов Е.Н., Пак М.М. и др. Механическое и полевое тестирование модифицированных наноструктур. Деп. рук. ВИНТИ № 1202 В. 2004: 13.07.04. – 11 с.
19. Ивашов Е.Н., Пак М.М. Туннельный метод измерения нанорельефа поверхности. Деп. рук. ВИНТИ № 1201 В. 2004: 13.07.04. – 10 с.
20. 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.
21. Пак М.М. Методы измерения наноструктур материалов на основе устройств получения нанодорожек, определения химического состава и нанорельефа. // Сборник докладов научно-технической конференции «Вакуумная наука и техника». – М.: МИЭМ, 2003.
22. Пак М.М. и др. Формирование наноструктур на атомарном уровне // Материалы конференции «Образование через науку». – М.: МИЭМ, 2005.
23. J.Robertson, *Advanced Physics*, v.35, P.317 (1986).
24. E.F. Sheka, V.D. Khavryutchenko and V.A. Zayetz, *Physical Low-Dimension Structures*. 2/3, 59 (1995).

25. Шека Е.Ф., Берзигияров П.К., Заец В.А., Разумов В.Ф. Параллельный программный комплекс MP-ZAVA для проведения квантовохимических расчетов в spd-базисе // Труды научного семинара «Математическое моделирование нанотехнологических процессов и наноструктур». Вып. 1. – М., 2001.
26. 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.
27. Bou P., Vandembulcke L. // J. Electrochem. Soc. 1991, v.138. – P. 2991-2994.
28. Hashimoto K., Muraki Y., Baba R., Fujishima A. // J. Electro anal. Chemistry. 1995, v.397. – P. 339-341.
29. Резников Б.А. Методы и алгоритмы оптимизации на дискретных моделях сложных систем. – Л.: ВИКИ им. А.Ф.Можайского, 1983. – 215 с.
30. Автоматизированные системы управления оборудованием. Адаптивное управление: Методические указания по проведению лабораторных работ // Московский государственный институт электроники и математики. Сост. В.П. Чулков. – М., 2006. – 36 с.
31. Ли Т.Т., Адамс Т.Э., Гейнз У.М. Управление процессами с помощью ЭВМ. Моделирование и оптимизация. – М.: Советское радио, 1972.
32. Фаронов В.В. ТУРБО Паскаль 7.0. Начальный курс: Учебное пособие. – М.: Норидж, 1997. – 616 с.
33. Чулков В.П. САПР оборудования и технологии производства СБИС: Учебное пособие // Московский государственный институт электроники и математики. – М., 2003. – 180 с.
34. В.П. Чулков. Моделирование технологических операций производства СБИС: Учебное пособие // Московский государственный институт электроники и математики. – М., 1994. – 73 с.

35. Саати Т.Л. Элементы теории массового обслуживания и её приложения. – М.: Советское радио, 1977.
36. Беленов В.К., Блинникова-Вяземская Е.В., Иоаннисянц Т.А. Количественный анализ технологического процесса // Микроэлектроника, 1975, вып. 8.
37. Вальков В.М., Вершин В.Е. Автоматизированные системы управления технологическими процессами. – М.: Машиностроение, 1977.
38. Норенков И.П. Введение в автоматизированное проектирование технических устройств и систем. – М.: Высшая школа, 1980. – 311 с.
39. Золотухин И.В., Калинин Ю.Е., Стогней О.В. Новые направления физического материаловедения: Учебное пособие. – Воронеж: Издательство Воронежского государственного университета, 2000. – 360 с.
40. Нанотехнологии в электронике / Под ред. Ю.А. Чаплыгина. – М.: Техносфера, 2005. – 448 с.
41. 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.
42. Суздальев И.П. Нанотехнология: физико-химия нанокластеров, наноструктур и наноматериалов. – М.: КомКнига, 2006. – 592 с. (Синергетика: от прошлого к будущему.)
43. 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.
44. Головин Ю.И. Введение в нанотехнологию. М.: «Издательство Машиностроение-1», 2003. – 112 с.
45. Кобаяси Н. Введение в нанотехнологию. – Пер. с японск. – М.: БИНОМ. Лаборатория знаний, 2005. – 134 с.
46. Чуличков А.И. Математические модели нелинейной динамики. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 296 с.

47. Коротков В.П., Тайц Б.А. Основы метрологии и теории точности измерительных устройств. – М.: Издательство стандартов, 1978. – 352 с.
48. Кравченко В.А., Цидилин С.М., Федосеева Т.Л. Алгоритмы решения задач многокритериальной оптимизации: Учебное пособие. – М.: МИЭМ, 1988. – 74 с.
49. Выявление экспертных знаний (процедуры и реализации) / О.И. Ларичев, А.И. Мечитов, Е.М. Мошкович, Е.М. Фуремс. – М.: Наука, 1989. – 128 с.
50. 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.
51. Рапопорт Э.Я. Анализ и синтез систем автоматического управления с распределенными параметрами: Учеб. пособие. – М.: Высш. шк., 2005. – 292 с.
52. Чулков В.П. Системы машинного проектирования технологических процессов производства БИС: Учеб. пособие. – М.: МИЭМ, 1990. – 77 с.
53. Советов Б.Я., Яковлев С.А. Моделирование систем. Практикум: Учеб. пособие для вузов. – 3-е изд., стер. – М.: Высш. шк., 2005. – 295 с.
54. Оптимальное управление. Сборник. – М.: «Знание», 1978. – 144 с. (Нар. ун-т. Естественнонаучный фак. Издается с 1961 г.).
55. Быков В.П. Методическое обеспечение САПР в машиностроении. – Л.: Машиностроение. Ленингр. отд-ние, 1989. – 255 с.
56. Элементная база микро- и нанoeлектроники: физика и технология // Сборник научных трудов – М.: МГИЭТ (ТУ), 1994. – 180 с.
57. Виброзащита радиоэлектронной аппаратуры полимерными компаундами // Ю.В. Зеленев, А.А. Кирилин, Э.Б. Слободник, Е.Н. Талицкий; Под ред. Ю.В. Зеленева. – М.: Радио и связь, 1984. – 120 с.
58. Норенков И.П., Маничев В.Б. Основы теории и проектирования САПР: Учеб. для вузов по спец. «Вычислительные машины, комплексы,

- системы и сети». – М.: Высш. шк., 1990. – 335 с.
59. Олфрей Г.Ф. Физическая электроника // Под ред. М.Б. Великовского. – Пер. с англ. В.И. Гайдука. – М.: Издательство «Мир», 1966. – 316 с.
  60. Семенкин Е.С., Семенкина О.Э., Коробейников С.П. Оптимизация технических систем: Учебное пособие. – Красноярск: СИБУП, 1996. – 284 с.
  61. Дружинин Г.В. Методы оценки и прогнозирования качества. – М.: Радио и связь, 1982. – 160 с. – (Межиздательская серия: Надежность и качество).
  62. Информационные технологии управления: Учебное пособие // Под ред. Ю.М. Черкасова. – М.: ИНФРА-М, 2001. – 216 с. – (Серия «Высшее образование»).
  63. Варакута С.А. Управление качеством продукции: Учебное пособие. – М.: ИНФРА-М, 2001. – 207 с. – (Серия «Вопрос – ответ»).
  64. Аристов О.В. Управление качеством: Учебник. – М.: ИНФРА-М, 2008. – 240 с. – (Серия «Высшее образование»).
  65. Шестопап Ю.Т., Дорофеев В.Д., Шестопап Н.Ю., Андреева Э.А. Управление качеством: Учебное пособие. – М.: ИНФРА-М, 2008. – 331 с. – (Серия «Высшее образование»).
  66. Клыков Ю.И., Горьков Л.Н. Банки данных для принятия решений. – М.: Советское радио, 1980. – 208 с.
  67. Компьютерные технологии обработки информации: Учеб. пособие // С.В. Назаров, В.И. Першиков, В.А. Тафинцев и др.; Под ред. С.В. Назарова. – М.: Финансы и статистика, 1995. – 248 с.
  68. Борисенко В.Е. Нанoeлектроника – основа информационных систем XXI века. – М.: ИНФРА-М, 2001.
  69. Рыбалко В.В. Наноразмерные углеродсодержащие материалы. – М.: МИЭМ, 2003. – 50 с.
  70. Лускинович П.Н. и др. Нанотехнологические процессы и установки // Сб. докладов «Математическое моделирование нанотехнологических

- процессов и наноструктур»: Труды научных семинаров. – М.: МИФИ, 2001.
71. Чумаченко Б.О., Лавров К.Н. Нанотехнологии – ключевой приоритет обозримого будущего. – М.: Олма-пресс, 2004.
  72. Еленин Г.Г. Нанотехнологии и вычислительная математика // Сб. докладов «Математическое моделирование нанотехнологических процессов и наноструктур»: Труды научных семинаров. – М.: МИФИ, 2001.
  73. Неволин В.К. Электронные устройства с элементами нанометровых размеров // Электронная промышленность. – М., 2002.
  74. Золотухин И.В. Углеродные нанотрубки. – М.: ФИЗИКА, 1999.
  75. Соколов В.И., Станкевич И.В. Фуллерены – новые аллотропные формы углерода: структура, электронное строение и химические свойства // Успехи химии. Т. 62 (5), 1993. – С. 455.
  76. Шпилевский М.Э., Стельмах В.Ф. Фуллерены и фуллереноподобные структуры – основа перспективных материалов. – М.: ЮНИТИ, 2004.
  77. Саати Т. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993.
  78. Норенков И.П. Введение в автоматизированное проектирование технических устройств и систем: Учеб. пособие для вузов. – 2-е изд., перераб. и доп. – М.: Высшая. школа, 1986. – 304 с.
  79. Алексеев О.В., Головков А.А., Пивоваров И.Ю. и др. Автоматизация проектирования радиоэлектронных средств. – М.: Высш. шк., 2000. – 479 с.
  80. Зародов А.Ф. Представление иерархии моделей при проектировании систем // Аэрокосмические технологии. – М.: МГТУ им Н.Э. Баумана, 2003.
  81. Солодовников И.В., Зародов А.Ф. Использование методов имитационного моделирования в исследовании экологических систем // Труды

- семинара «Новые информационные технологии». – М.: МИЭМ, 2001.
82. Киндлер Е. Языки моделирования. – М.: Энергоатомиздат, 1985. – 288 с.
  83. Антамошкин А.Н. и др. Системный анализ: Проектирование, оптимизация и приложения: Учебное пособие. В 2-х томах. Том 2. – Красноярск: Сибирская аэрокосмическая академия, 1996. – 290 с.
  84. Фоминых В.П. Оборудование и технология дуговой сварки: Учеб. пособие для проф.-тех. училищ. – М.: Машиностроение, 1996. – 312 с.
  85. Элементарный учебник физики. Т. 2 / Под ред. Г.С. Ландсберга. – М.: Электричество и магнетизм, 1973. – 528 с.
  86. Трофимова Т.И. Физика в таблицах и формулах: Учеб. пособие для студентов вузов. – 2-е изд., испр. – М.: Дрофа, 2004. – 432 с.
  87. Гершензон Е.М., Малов Н.Н., Мансуров А.Н. Молекулярная физика: Учеб. пособие для студентов высших пед. уч. заведений. – М.: Издательский центр «Академия», 2000. – 272 с.
  88. Савостьянов В.П., Филатова Г.А., Филатов В.В. Расчёт и конструирование деталей аппаратуры САУ: Учеб. пособие. – М.: Машиностроение, 1982. – 328 с.
  89. Слепов Н.Н. Современные технологии цифровых оптоволоконных сетей связи. – М.: Радио и связь, 2000. – 468 с.
  90. Иванов И.П. Углеродные нанотрубки: их свойства и применение. – М.: ИНФРА-М, 2005. – 345 с.
  91. Смирнов С.А. Оценка интеллектуальной собственности. – М.: Финансы и статистика, 2002. – 352с.
  92. Неволин В.К. Физические основы туннельно-зондовой нанотехнологии: Учеб. пособие. – М.: МГИЭТ, 2000.
  93. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных. Докл. АН СССР, 1956. Т. 108. – С. 179-182.

94. Герасименко Н.Н., Джаманбалин К.К., Медетов Н.А. Самоорганизованные структуры на поверхности и в объёме полупроводников. – Алматы: LEM, 2002. – 192 с.
95. 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.
96. Фаронов В.В. Турбо Паскаль 7.0. Начальный курс: Учеб. пособие. – М.: КНОРУС, 2006. – 576 с.
97. Клименко Ю.И. Высшая математика для экономистов: теория, примеры, задачи: Учебник для вузов. – М.: Издательство «Экзамен», 2005. – 736 с. (Серия «Учебник для вузов»).
98. Пискунов Н.С. Дифференциальное и интегральное исчисления. Т.2. – М.: Наука, 1976. – 576 с.
99. Binning G, Quate C F, Gerber C Phys. Rev. Lett. 56 930 (1986).
100. Mate C M et al. Phys. Rev. Lett. 69 1942 (1987).
101. Israelachvili J.N. Intermolecular and Surface Forces. – Academic Press, 1998. – 450 p.
102. Зимон А. Д. Адгезия жидкости и смачивание. – М.: Химия, 1974. – 416 с.
103. Derjaguin B. V., Muller V.M., Toporov Yu.P., J.Colloid. Interface Sci. 53, 314 (1975).
104. Дерягин Б.В., Чураев Н.В., Муллер В.М. Поверхностные силы. – М.: Наука, 1985.
105. Джонсон К.Л. Механика контактного взаимодействия. – М.: Мир, 1987.
106. Maugis D.J., J.Colloid. Interface Sci. 150, 243 (1992).
107. Tomlinson I Philos. Mag. 7 905 (1929).
108. Дедков Г. В., Дышеков М. Б. ЖТФ 70 (7) 96 (2000).

109. Ландау Л. Д., Лившиц Е. М. Теория упругости. – М.: Наука, 1987. – 246 с.
110. Головин Ю.И. Тюрин А.И., Иволгин В.И., Коренков В.В. ЖТФ 70 (5)82 (2000).
111. Миронов В.Л. Основы сканирующей зондовой микроскопии: Учебное пособие для студентов старших курсов высших учебных заведений. – Н.Новгород: РАН, Институт физики микроструктур, 2004. – 114 с.
112. Ивашов Е.Н., Степанчиков С.В., Мисюк Р.В. Приповерхностные свойства нанобъектов, гетероструктур и нанобластей с различным химическим и фазовым составом. – М.: МИЭМ, 2006. – 35 с.
113. Васин В.А., Ивашов Е.Н., Степанчиков С.В. Нанотехнологические процессы и оборудование электронной техники. – М.: МИЭМ, 2009. – 264 с.
114. Информационные технологии в проектировании объектов электронного машиностроения: Монография. В 5-ти кн. Кн. 1. В.А. Васин, А.С. Вишневский, Е.Н. Ивашов, С.В. Степанчиков. Информационная технология в проектировании пластин плотной записи для накопителей на жёстких магнитных дисках. – М.: Издательство НИИ ПМТ, 2010. – 227 с.
115. Информационные технологии в проектировании объектов электронного машиностроения: Монография. В 5-ти кн. Кн. 2. В.А. Васин, К.И. Домась, Е.Н. Ивашов, С.В. Степанчиков. Информационная технология в проектировании объектов лазерной обработки. – М.: Издательство НИИ ПМТ, 2011. – 211 с.

# ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
<b>ГЛАВА 1. Моделирование и алгоритмизация процесса управления качеством производства электронных устройств.....</b>	<b>5</b>
1.1. Основная задача и состояние работ в области управления качеством при производстве наноустройств.....	5
1.2. Структура и свойства наноматериалов.....	15
1.3. Методы производства наноматериалов. Структура и свойства наноматериалов.....	28
1.4. Устройства на основе наноматериалов.....	42
1.5. Функциональное моделирование системы менеджмента качества на основе IDEF0 в различных процессах производства наноматериалов и электронных устройств.....	47
<b>ГЛАВА 2. Математические модели управления качеством производства электронных устройств на основе углеродных структур.....</b>	<b>52</b>
2.1. Классификация задач оптимизации.....	52
2.2. Оптимизация обобщённого критерия качества дискриминационным методом.....	67
2.3. Вероятностное моделирование.....	79
2.4. Задача о замене оборудования.....	85
<b>ГЛАВА 3. Алгоритмы поиска оптимальных управленческих решений качества производства материалов на основе углеродных структур.....</b>	<b>88</b>

3.1. Адаптивное управление качеством производства наноматериалов.....	88
3.2. Регрессионный анализ.....	94
3.3. Моделирование процесса адаптивного управления с подстройкой параметров модели.....	102
3.4. Методы выбора недоминирующего решения из ряда недоминируемых альтернатив.....	106
3.5. Метод упорядоченного предпочтения через сходство с идеальным решением для задач многоцелевого принятия решений.....	118
<b>ГЛАВА 4. Практическая реализация математических моделей и алгоритмов управления качеством производства электронных устройств.....</b>	<b>123</b>
4.1. Экспериментальные исследования процессов формирования электронных устройств на основе модифицированных углеродных структур.....	123
4.2. Методы сканирующей туннельной микроскопии.....	130
4.3. Расчётные данные оптимизации процессов производства с применением математических моделей.....	135
<b>ГЛАВА 5. Приповерхностные свойства нанообъектов, гетероструктур и нанобластей с различным химическим и фазовым составом.....</b>	<b>147</b>
5.1. Основные принципы сканирующей зондовой микроскопии.....	148
5.2. Силовое взаимодействие зонда с поверхностью.....	152
5.2.1. Электростатическое взаимодействие.....	152
5.2.2. Сила Ван-дер-Ваальса.....	161

5.2.3. Капиллярные силы.....	164
5.2.4. Адгезионные силы.....	168
5.2.5. Упругие взаимодействия зонда с образцом.....	174
5.3. Эффекты, возникающие при исследовании приповерхностных свойств.....	180
5.3.1. Эффект прилипания-скольжения.....	181
5.3.2. Трибоэмиссия частиц, электромагнитных и акустических волн.....	186
5.3.3. Трибохимические эффекты.....	188
5.4. Анализ химического состава поверхности.....	189
ЗАКЛЮЧЕНИЕ.....	193
ЛИТЕРАТУРА.....	195



