

**С.И. ДВОРЕЦКИЙ, С.В. МАТВЕЕВ,  
С.Б. ПУТИН, Е.Н. ТУГОЛУКОВ**

**ОСНОВЫ МАТЕМАТИЧЕСКОГО  
МОДЕЛИРОВАНИЯ И  
ОПТИМИЗАЦИИ ПРОЦЕССОВ И  
СИСТЕМ ОЧИСТКИ И  
РЕГЕНЕРАЦИИ ВОЗДУХА**

**ИЗДАТЕЛЬСТВО ТГТУ**

УДК 519.711.3:62-784.43(075)  
ББК Л111-1с116я73  
О-753

Рецензенты:

Доктор технических наук, профессор, зав. кафедрой  
компьютерно-интегрированных систем РХТУ им. Д.И. Менделеева  
*А.Ф. Егоров,*

Доктор технических наук, начальник отдела ФГУ  
«1 Центральный научно-исследовательский институт  
Минобороны России»  
*Г.А. Родин*

**Дворецкий, С.И.**

О-753 Основы математического моделирования и оптимизации процессов и систем очистки и регенерации воздуха : учебное пособие / С.И. Дворецкий, С.В. Матвеев, С.Б. Путин, Е.Н. Туголуков. – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. – 324 с. – 100 экз. – ISBN 978-5-8265-0775-9.

Содержит систематическое изложение фундаментальных основ теории моделирования и оптимизации процессов и систем очистки и регенерации воздуха в условиях неопределенности. Обсуждаются механизмы процесса адсорбции, термодинамические основы поверхностных явлений, кинетика и динамика процесса адсорбции, а также тепловые эффекты, имеющие место при очистке и регенерации воздуха в изолирующих средствах защиты органов дыхания людей. Формулируются математические постановки задач исследования взаимосвязанных процессов адсорбции, тепло- и массопереноса при очистке и регенерации воздуха.

Предназначено для студентов высших учебных заведений, обучающихся по специальностям 240801, 251600, направлению 150400 и повышения квалификации специалистов в научно-образовательном центре «ТГТУ – ОАО «Корпорация «Росхимзащита».

УДК 519.711.3:62-784.43(075)  
ББК Л111-1с116я73

ISBN 978-5-8265-0775-9 © ГОУ ВПО «Тамбовский государственный  
технический университет» (ТГТУ), 2008

Министерство образования и науки Российской Федерации  
ГОУ ВПО «Тамбовский государственный технический университет»

**С.И. ДВОРЕЦКИЙ, С.В. МАТВЕЕВ,  
С.Б. ПУТИН, Е.Н. ТУГОЛУКОВ**

**ОСНОВЫ МАТЕМАТИЧЕСКОГО  
МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ  
ПРОЦЕССОВ И СИСТЕМ ОЧИСТКИ И  
РЕГЕНЕРАЦИИ ВОЗДУХА**

Утверждено Учёным советом университета  
в качестве учебного пособия для студентов специальностей  
240801, 251600 и направления 150400



---

Тамбов  
◆ Издательство ТГТУ ◆  
2008

ебное издание

ДВОРЕЦКИЙ Станислав Иванович,  
МАТВЕЕВ Сергей Витальевич,  
ПУТИН Сергей Борисович,  
ТУГОЛУКОВ Евгений Николаевич

ОСНОВЫ МАТЕМАТИЧЕСКОГО  
МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ  
ПРОЦЕССОВ И СИСТЕМ ОЧИСТКИ И  
РЕГЕНЕРАЦИИ ВОЗДУХА

Учебное пособие

Редактор Е.С. Мордасова  
Компьютерное макетирование М.А. Филатовой

Подписано к печати 25.11.2008  
Формат 60 × 84/16. Объем: 18,8 усл. печ. л.; Тираж 100 экз. Заказ № 593.

Издательско-полиграфический центр ТГТУ  
392000, Тамбов, Советская, 106, к. 14

Моделирование можно рассматривать как один из самых мощных методов и инструментов познания, анализа и синтеза, которым располагают специалисты, ответственные за разработку и функционирование сложных технических устройств и технологических объектов (систем).

Идея моделирования состоит в замене реальной системы некоторым «образом» – математической моделью – и в дальнейшем изучении модели на компьютере с целью получения новых знаний об этой системе. При этом у исследователя появляется возможность экспериментировать с моделью системы в тех случаях, когда делать это на реальном объекте практически невозможно или нецелесообразно. Работа не с самим объектом (явлением, процессом), а с его математической моделью дает возможность относительно быстро и без существенных затрат исследовать его свойства и поведение в любых мыслимых ситуациях (преимущества теории). В то же время вычислительные (имитационные) эксперименты с моделями систем позволяют подробно и глубоко изучать системы в достаточной полноте, недоступной чисто теоретическим подходам (преимущества эксперимента).

Собственно моделирование представляет собой процесс конструирования математической модели реальной системы (технической и/или технологической) и постановки вычислительных экспериментов на этой модели с целью либо понять (исследовать) поведение этой системы, либо оценить эффективность различных стратегий (алгоритмов) ее функционирования с помощью реализуемых на компьютерах вычислительно-логических алгоритмов. Таким образом, процесс моделирования включает и конструирование модели, и ее применение для решения поставленной задачи: анализа, исследования, оптимизации или синтеза (проектирования) системы.

## 1. ОСНОВЫ ТЕОРИИ МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ И СИСТЕМ

### 1.1. ПРИНЦИПЫ СИСТЕМНОГО ПОДХОДА

Основу современного подхода к решению задач моделирования и оптимизации технологических процессов и систем составляет *системный анализ* – метод научного познания, состоящий в том, что любой объект по отношению к субъекту рассматривается как система (сложное образование), состоящая из большого числа элементов, связанных между собой вещественными, энергетическими, информационными и другими связями сильнее, чем с окружающей средой [1].

Сущность системного анализа определяется тем, что мы вкладываем в понятие *система*. Будем называть системой совокупность элементов, обладающую следующими свойствами: 1) целостности и членимости; 2) связности; 3) интегративности (эмерджентности); 4) организованности; 5) наличием цикла существования – «жизненного цикла» системы.

Рассмотрим подробнее каждое из этих свойств.

*Свойство целостности и членимости.* Система – это прежде всего целостная совокупность взаимосвязанных элементов. Она допускает детализацию (расчленение) на подсистемы, т.е. пространственно-временные агрегаты, состоящие из взаимосвязанных элементов. Элементы системы при определенных условиях могут рассматриваться как самостоятельные системы, а сама система – как элемент более высокого уровня иерархического расчленения. Расчленение системы на подсистемы и элементы часто является условным и зависит от целей проводимого исследования.

*Свойство связности системы.* Система существует как целостное образование тогда и только тогда, когда сила межэлементных связей внутри системы выше, чем сила связи этих элементов с окружающей средой. Связи между элементами различают по физическому наполнению (вещественные, энергетические, информационные) и направлению (прямые, обратные).

Упорядоченная в пространстве и времени совокупность элементов и связей образует *структуру системы*. Наглядно структура системы может быть представлена в виде графа, узлы (вершины) которого соответствуют элементам, а дуги – связям.

*Интегративное качество (эмерджентность) системы* состоит в том, что она обладает свойствами, присущими системе в целом, но не присущими ни одному ее элементу в отдельности. Наличие интегративных качеств показывает, что свойства системы хотя и зависят от свойств элементов, но не определяются ими полностью. Отсюда следуют два важных вывода: 1) система не сводится к простой совокупности элементов; 2) расчленив систему на отдельные элементы, изучая каждый из них в отдельности, нельзя познать все свойства системы в целом.

*Свойство организованности системы* проявляется в снижении информационной энтропии (степени неопределенности) системы. Переход к более высокому уровню упорядоченности и организованности системы означает уменьшение ее текущей неопределенности (энтропии) за счет накопления информации.

*Свойство жизненного цикла системы* характерно для систем любого типа – естественных и искусственных, технических, биологических, социальных и др. Жизненный цикл системы – это промежуток времени от момента принятия решения о создании (проектировании) до момента окончания эксплуатации (утилизации) системы.

Жизненный цикл системы, в частности, определяет характер изменения во времени главных характеристик системы (работоспособности, силы, мощности, энергии, производительности и т.д.).

Системный анализ определяется его стратегией, в основе которой лежат общие принципы, применимые к решению любой системной задачи. К ним можно отнести: 1) четкую формулировку цели исследования, постановку задачи по достижению заданной цели и определение критерия эффективности решения задачи; 2) разработку развернутой стратегии исследования с указанием основных этапов и направлений в решении задачи: последовательно-параллельное продвижение по всему комплексу взаимосвязанных этапов и возможных направлений; организацию последовательных приближений и повторных

циклов исследований на отдельных этапах; принцип нисходящей иерархии анализа и восходящей иерархии синтеза при решении составных частных задач.

Можно выделить три основные задачи, связанные с изучением и созданием систем: *анализ, синтез и принятие решений*.

*Анализ* состоит в изучении свойств и поведения систем в различных условиях функционирования. В ходе анализа устанавливаются численные значения показателей эффективности систем.

*Синтез* заключается в построении возможных (допустимых) вариантов систем. Различают структурный и параметрический синтез. *Структурный синтез* – это построение структуры системы (элементов и связей). *Параметрический синтез* – определение параметров элементов системы при заданной структуре. Задача синтеза решается при аппаратурно-технологическом оформлении процессов.

*Оценка и принятие решений* (оптимизация) заключаются в выборе наилучшего варианта системы из нескольких альтернативных и оценке эффективности его функционирования. Функционирование системы характеризуется некоторым количественным или качественным *функционалом*, который называют *показателем эффективности*.

В общем случае систему можно охарактеризовать кортежем символов, например

$$ТС = \langle \Phi Н, \Phi, СТ, К, О, \Xi \rangle ,$$

где ТС – техническая (технологическая) система; ФН – функциональное назначение; Ф – функция; СТ – структура; К – компоновка; О – организация,  $\Xi$  – вектор показателей эффективности (качества).

*Функциональное назначение* (ФН) ТС характеризуется, например, ассортиментом выпускаемой продукции и объемом производства, а также другими директивными показателями.

*Функция* (Ф) ТС вводится для представления согласованной последовательности операций, посредством которых совершают физико-химические и другие преобразования продукта согласно комплексу заданных свойств. Функцию Ф удобно описывать ориентированным графом  $\Phi = (ТО; U)$ , в котором в качестве множества вершин ТО выступают технологические операции, а в качестве множества ребер  $U$  – материальные и информационные связи, характеризующие в процессе обработки потоки продукции, передающейся от одной операции к другой. Вариантность Ф определяется качественными изменениями на множестве ТО, а именно, заключается в возможности получения одного и того же результата различными способами (методами) обработки.

Под *структурой* (СТ) будем понимать некоторую организацию системы посредством синтеза из отдельных элементов, обладающих определенными свойствами и характеризующих цель и назначение ТС. Структура отражает качественный и количественный состав, множество связей между элементами и определяет основные свойства ТС. В качестве элемента СТ можно считать ТО в совокупности с определенным технологическим оборудованием и вспомогательными техническими средствами.

*Компоновка* (К) сводится к задаче геометрического размещения системных элементов в заданном (например, производственном) объеме или на заданной площади.

*Организация* (О) – это актуализация и упорядочение связей и самих элементов. Она заключается в разработке схем взаимодействия материальных, энергетических и трудовых ресурсов со средствами производства во времени и пространстве. Важнейшим компонентом организации является система управления, призванная обеспечить целенаправленное поведение ТС в условиях изменения ее внешних и внутренних параметров.

Вектор показателей эффективности  $\Xi$  задается совокупностью качественных и количественных показателей эффективности функционирования ТС, характеризующих полезный эффект от использования ТС.

Достаточно полно методологию системного подхода отражают три основных принципа: физичности, моделируемости и целенаправленности.

*Принцип физичности*: всякой системе присущи физические законы, определяющие внутренние причинно-следственные связи, существование и функционирование.

*Принцип моделируемости*: система представима конечным множеством моделей, каждая из которых отражает определенную грань ее сущности. Выявление новых свойств и сущностей не обязательно должно сопровождаться построением обобщающих моделей, а может ограничиваться наращиванием числа упрощенных моделей, взаимодействие которых обеспечивает отражение сложной системы в целом.

*Принцип целенаправленности*: сложной системе присуща функциональная тенденция, направленная на достижение некоторого состояния или на усиление (сохранение) некоторого процесса, при этом система способна противостоять внешнему воздействию.

Сложные системы имеют возможность выбора и обладают способностью выбирать поведение, т.е. реагировать на внешние воздействия в зависимости от внутренних критериев целенаправленности. Постулат выбора позволяет сложной системе, в соответствии с ее целенаправленностью, использовать редкие благоприятные события, возникающие во взаимодействии со средой, блокируя остальные (неблагоприятные) события и процессы.

## 1.2. ОБЩАЯ ХАРАКТЕРИСТИКА ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ СИСТЕМ

Замещение одной системы другой с целью получения информации о важнейших свойствах системы-оригинала с помощью системы-модели называется *моделированием*. Моделирование может быть определено как представление системы моделью для получения информации об этой системе путем проведения экспериментов (натурных, вычислительных и т.п.) с моделью [2, 3].

Обобщенно *моделирование* можно определить как метод опосредованного познания, при котором изучаемая система-оригинал находится в некотором соответствии с другой системой-моделью, причем модель способна в том или ином отношении замещать оригинал на некоторых стадиях познавательного процесса. Стадии познания, на которых происходит такая замена, а также формы соответствия модели и оригинала могут быть различными.

Моделирование может рассматриваться как познавательный процесс, содержащий переработку информации, поступающей из внешней среды, о происходящих в ней явлениях, в результате чего в сознании появляются образы, соответствующие системам.

Моделирование может заключаться в построении некоторой системы-модели (второй системы), связанной определенными соотношениями подобия с системой-оригиналом (первой системой), причем в этом случае отображение одной системы в другую является средством выявления зависимостей между двумя системами, отраженными в соотношениях подобия, а не результатом непосредственного изучения поступающей информации.

Отметим, что с точки зрения философии *моделирование* – эффективное средство познания природы. Процесс моделирования предполагает наличие *объекта (системы) исследования, исследователя*, перед которым поставлена конкретная задача; *модели*, создаваемой для получения информации о системе и необходимой для решения поставленной задачи. Причем по отношению к модели исследователь является, по сути дела, экспериментатором, только в данном случае эксперимент проводится не с реальной системой, а с ее моделью. Такой эксперимент для инженера есть инструмент непосредственного решения организационно-технических задач.

Следует заметить, что любой эксперимент может иметь существенное значение в конкретной области науки только при специальной обработке и обобщении его результатов. Единичный эксперимент никогда не может быть решающим для подтверждения гипотезы, проверки теории. Поэтому инженеры (исследователи и практики) должны быть знакомы с элементами современной методологии теории познания и, в частности, не должны забывать основного положения материалистической философии, что именно экспериментальное исследование, опыт, практика являются критерием истины.

Одна из проблем современной науки и техники – разработка и внедрение в практику проектирования новейших методов исследования характеристик сложных ТС. При проектировании таких систем возникают многочисленные задачи, требующие проведения структурного и параметрического синтеза, оценки количественных и качественных закономерностей процессов функционирования сложных систем и их подсистем.

Рассматриваемые в учебном пособии технологические процессы адсорбции, тепло- и массообменные процессы регенерации воздуха в ИДА относятся к классу сложных систем, исследование и оптимизация которых в настоящее время невозможны без использования различных видов и математических схем моделирования. При моделировании необходимо учитывать следующие особенности: сложность структуры и стохастичность связей между элементами, неоднозначность алгоритмов поведения при различных условиях, большое количество параметров и переменных, неполноту и недетерминированность (неопределенность) исходной информации, разнообразие и вероятностный характер воздействий внешней среды и т.д. Ограниченность возможностей экспериментального исследования сложных систем делает актуальной разработку методики их моделирования, которая позволила бы в соответствующей форме представить процессы функционирования исследуемых или проектируемых систем, оценить их характеристики.

Выбор метода моделирования и необходимая детализация моделей существенно зависят от этапа разработки сложной системы. На этапах разработки технического задания и рабочего проекта модели отдельных подсистем детализируются, и моделирование служит для решения конкретных задач проектирования, т.е. выбора оптимального варианта из множества допустимых по определенному критерию при заданных ограничениях.

Целевое назначение моделирования на этапе эксплуатации сложных систем – это рассмотрение возможных ситуаций для принятия обоснованных и перспективных решений по управлению объектом. Моделирование (имитацию) также широко применяют при обучении и тренировке персонала автоматизированных систем управления, вычислительных комплексов и сетей, информационных систем в различных сферах. В этом случае моделирование носит характер деловых игр. Модель, реализуемая обычно на ЭВМ, воспроизводит поведение управляемого объекта и внешней среды, а люди в определенные моменты времени принимают решения по управлению объектом.

С развитием системных исследований, с расширением экспериментальных методов изучения реальных явлений все большее значение приобретают абстрактные методы, появляются новые научные дисциплины, автоматизируются элементы умственного труда. Особое значение при создании реальных систем имеют математические методы анализа и синтеза, целый ряд открытий базируется на чисто теоретических изысканиях. Однако было бы неправильно забывать о том, что основным критерием любой теории является практика, и даже сугубо математические, отвлеченные науки базируются в своей основе на фундаменте практических знаний.

Одновременно с развитием теоретических методов анализа и синтеза совершенствуются и методы экспериментального изучения реальных объектов, появляются новые средства исследования. *Эксперимент* был и остается одним из основных и существенных инструментов познания. Подобие и моделирование позволяют по-новому описать реальный процесс и упростить его экспериментальное изучение. Совершенствуется и само понятие моделирования. Если раньше моделирование означало реальный физический эксперимент либо построение макета, имитирующего реальный процесс, то в настоящее время появились новые виды моделирования, в основе которых лежит постановка не только физических, но также и математических (вычислительных) экспериментов.

Разработка моделей выполняется для решения каких-либо проблем или задач, стоящих перед исследователем. В зависимости от характера задачи к модели предъявляются определенные требования, прежде всего, это пригодность модели для решения соответствующей задачи по форме, точности, быстродействию, диапазону изменения переменных, уровень формализации и т.д. Во многих исследовательских задачах модель служит источником информации при выполнении вычислений для получения конечного результата, а созданные математические модели являются своего рода хранилищем и носителем информации о связях между переменными, характером их изменения, свойствах системы (процесса, явления) и т.д.

Наиболее часто исследователю приходится иметь дело со следующими классами задач, которые решаются с использованием моделей:

- проектирование различного рода сложных систем;
- оптимизация и оптимальное управление сложными системами;
- исследование свойств систем, процессов, материалов;
- прогнозирование поведения систем и процессов;
- реинжиниринг организации;
- планирование деятельности и ресурсов;
- принятие различного рода управленческих решений;
- проблемы изучения и использования искусственного интеллекта.

Необходимо отметить, что многие задачи решаются в условиях неопределенности, отсутствия достаточного объема исходных данных. Для такого рода задач наблюдается некоторое объединение модели и алгоритма решения, т.е. алгоритм, процедура решения задачи в значительной степени берет на себя функции модели, модель и алгоритм рассматриваются как единое целое. Это наблюдается, например, при использовании Байесовского подхода или нечеткого логического вывода.

Понятие математической модели (ММ), как и ряд других понятий, используемых в математическом моделировании, не имеет строгого формального определения. Тем не менее, в это понятие вкладывают вполне конкретное содержание, с которым, в частности, тесно связано применение математики в инженерной практике.

Этапы развития многих естественно-научных направлений в познании законов природы и в совершенствовании техники и технологий – это построение последовательности все более точных и более полных ММ изучаемых процессов и явлений. Соответствующая реальности (адекватная) ММ является, как правило, большим научным достижением. Она позволяет провести детальное исследование изучаемой системы и дать надежный прогноз ее поведения в различных условиях. Но за адекватность ММ нередко приходится расплачиваться ее усложнением, что вызывает трудности при ее использовании. В этом случае на помощь математике и приходит современная вычислительная техника, существенно расширившая класс ММ, допускающих исчерпывающий количественный анализ.

*Определение.* Совокупность понятий и отношений, выраженных при помощи системы математических символов и обозначений, которые отражают наиболее существенные (характерные) свойства изучаемой системы, называют *математической моделью* этой системы.

В данном случае математика выступает, по существу, в роли универсального языка науки. Его универсальность французский математик Анри Пуанкаре определил всего одной фразой «Математика – это искусство называть разные вещи одним и тем же именем».

Применение математических методов при изучении реально существующей или мыслимой системы будет эффективным, если свойства ММ удовлетворяют определенным требованиям. Рассмотрим основные из этих свойств.

*Полнота* ММ позволяет отразить в достаточной мере именно те характеристики и особенности системы, которые интересуют нас с точки зрения поставленной цели проведения моделирования. Например, модель может достаточно полно описывать протекающие в системе процессы, но не отражать ее габаритные, массовые или стоимостные показатели.

*Точность* ММ дает возможность обеспечить приемлемое совпадение реальных и найденных при помощи ММ значений выходных переменных системы, составляющих вектор  $y = (y_1, y_2, \dots, y_n)^T \in R^n$ . Пусть  $y_i^M$  и  $y_i^P$  – найденное при помощи ММ и реальное значение  $i$ -й выходной переменной. Тогда относительная погрешность ММ по отношению к этой переменной при одних и тех же значениях входных переменных будет равна

$$\varepsilon_i = |y_i^M - y_i^P| / y_i^P, \quad i = \overline{1, n}.$$

В качестве скалярной оценки вектора погрешности модели  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T \in R^n$  можно принять какую-либо его норму, например

$$\varepsilon = \sqrt{\sum_i^n \varepsilon_i^2} \quad \text{или} \quad \varepsilon = \max_{i=1, n} |\varepsilon_i|.$$

*Адекватность* ММ – это способность ММ отображать выходные переменные системы с погрешностью не более некоторого заранее заданного значения  $\delta$ .

В общем смысле под адекватностью ММ понимают правильное качественное и достаточно точное количественное описание именно тех характеристик системы, которые наиболее важны в данном конкретном случае. В ряде прикладных областей, еще недостаточно подготовленных к применению количественных математических методов, ММ имеют, главным образом, качественный характер. Эта ситуация типична, например, для биологической и социальной сфер, в которых количественные закономерности не всегда поддаются строгой математической формализации. В таких случаях под адекватностью ММ естественно понимать лишь правильное качественное описание поведения изучаемых объектов.

*Экономичность* ММ оценивают затратами на вычислительные ресурсы (машинное время и память), необходимые для проведения вычислительного эксперимента с ММ на ЭВМ. Эти затраты зависят от числа арифметических операций при использовании модели, от размерности и пространства *фазовых переменных*, характеризующих состояние ТС и других факторов.

Очевидно, что требования экономичности, высокой точности и адекватности ММ противоречивы и на практике могут быть удовлетворены лишь на основе разумного компромисса.

*Робастность* ММ характеризует ее устойчивость по отношению к погрешностям исходных данных, способность нивелировать эти погрешности и не допускать их чрезмерного влияния на результат вычислительного эксперимента.

*Продуктивность* ММ связана с возможностью располагать достаточно достоверными исходными данными. Если они являются результатом измерений, то точность их измерения должна быть не ниже, чем для тех переменных, которые получаются при использовании ММ. В противном случае ММ будет непродуктивной и ее применение для анализа конкретной ТС теряет смысл.

В зависимости от масштаба ТС и наших предположений о ее свойствах ММ принимают конкретный вид. Можно говорить о ММ технологической машины или аппарата, технологического процесса, производства, предприятия и даже целой отрасли. Эти ММ отличаются одна от другой полнотой учета и глубиной описания различных процессов в системе. Если, например, ММ аппарата содержит чаще всего не более 10 – 15 уравнений, то в модель производства, предприятия и тем более отрасли может входить несколько десятков или сотен уравнений.

Достаточно общей формой представления модели исследуемой системы является модель динамической системы (ДС), которую будем обозначать символом  $\Sigma$ .

В простейшем случае ДС представляет собой систему, функционирование которой задается совокупностью обыкновенных дифференциальных уравнений в форме Коши обычно с достаточно гладкими правыми частями, обеспечивающими существование и единственность решения.

В более сложном случае система обыкновенных дифференциальных уравнений в форме Коши дополняется нелинейными алгебраическими уравнениями и набором вспомогательных формул.

В широком смысле под ДС понимается непрерывно наблюдаемая и изменяющая свое состояние под воздействием внешних и внутренних причин система, которая функционирует в непрерывном времени.

Основными системными объектами данной модели являются векторы входных переменных (входа)  $x$ , фазовых координат (переменных состояния)  $z$ , выходных переменных (выхода)  $y$  такие, что

$$\begin{aligned} x &= (x_1, x_2, \dots, x_m)^T \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_m; \\ z &= (z_1, z_2, \dots, z_n)^T \in \mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2 \times \dots \times \mathcal{Z}_n; \\ y &= (y_1, y_2, \dots, y_p)^T \in \mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2 \times \dots \times \mathcal{Y}_p, \end{aligned}$$

где  $\mathcal{X}, \mathcal{Z}, \mathcal{Y}$  – множества, в которых изменяются векторы  $x, z, y$  соответственно;  $\mathcal{X}_i (i = \overline{1, m}), \mathcal{Z}_j (j = \overline{1, n}), \mathcal{Y}_k (k = \overline{1, p})$  – множества значений компонент  $x_i, z_j, y_k$  векторов  $x, z, y$ ;  $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_m$  – декартово (прямое) произведение, т.е. множество  $\mathcal{X}$ , состоящее из всех упорядоченных совокупностей вида  $x = (x_1, x_2, \dots, x_m)$ , причем  $x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, \dots, x_m \in \mathcal{X}_m$ ,  $T$  – символ транспонирования; аналогичные определения имеют место и для  $\mathcal{Y}$  и  $\mathcal{Z}$ .

Вектор  $z(t_0)$  фазовых координат целиком определяет состояние системы  $\Sigma$  в фиксированный момент времени  $t_0$ . Положение системы в любой момент времени  $t$  в будущем, т.е. для  $t > t_0$  единственным образом определяется вектором  $z(t_0) = z_0$  и изменением входных воздействий (траекторий)  $x(\cdot) = (x(s), s \in [t_0, t])$  и не зависит от того, каким образом система пришла в состояние  $z_0$  (рассматриваются системы без последействия). Для таких систем имеет место следующее отображение

$$\Phi: \mathcal{F} \times \mathcal{F} \times \mathcal{Z} \times \mathcal{X}(\cdot) \rightarrow \mathcal{Z},$$

т.е. закон, по которому каждому элементу  $(t, t_0, z, x(\cdot))$  множества  $\mathcal{F} \times \mathcal{F} \times \mathcal{Z} \times \mathcal{X}(\cdot)$ , называемого областью определения отображения, ставится в соответствие некоторый элемент  $z$  множества  $\mathcal{Z}$ , называемый областью значений отображения. Здесь  $\mathcal{F}$  – множество значений моментов времени  $t(t_0)$  и  $\mathcal{X}(\cdot)$  – множество траекторий изменения входного воздействия  $x(\cdot)$ .

Можно использовать также более привычную форму записи  $\Phi$  в виде оператора, называемого переходной функцией, т.е.

$$z(t) = \Phi(t, t_0, z_0, x(\cdot)).$$

Связь между вектором переменных состояния  $z$  и контролируемым вектором выхода  $y$  задается некоторым выходным отображением

$$\Psi: \mathcal{F} \times \mathcal{Z} \rightarrow \mathcal{Y},$$

ставящим в соответствие каждой паре  $t, z$ , называемой событием или фазой, из множества  $\mathcal{F} \times \mathcal{Z}$  – пространства событий (фазового пространства) конкретный элемент из множества  $\mathcal{Y}$ . Эта зависимость между  $y$  и  $z$  может быть отражена также с помощью оператора

$$y(t) = \Psi(t, z).$$

Таким образом, динамическая система  $\Sigma$  задается четверкой множеств  $\mathcal{F}, \mathcal{X}, \mathcal{Z}, \mathcal{Y}$  и двумя операторами  $\Phi, \Psi$ , т.е.

$$\Sigma = (\mathcal{F}, \mathcal{X}, \mathcal{Z}, \mathcal{Y}; \Phi, \Psi),$$

ее общая схема представлена на рис. 1.1.

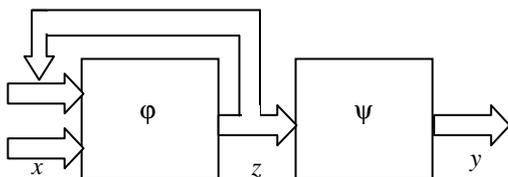


Рис. 1.1. Общая схема динамической системы

Приведем простейший пример ДС. В качестве  $\Sigma$  возьмем изотермический реактор идеального смешения (рис. 1.2) объема  $V$ , в который с постоянной объемной скоростью  $Q$  поступает поток вещества  $A$  с начальной концентрацией  $q_{A_0}$ . В аппарате протекает химическая реакция типа  $A \rightarrow B$ , причем скорость реакции равна  $r = kq_A$ , здесь  $k$  константа скорости реакции. Переменными в такой системе являются концентрации  $q_A$  и  $q_B$  веществ  $A$  и  $B$  соответственно, которые можно выбрать в качестве фазовых координат  $z_1$  и  $z_2$ , выходная координата  $y = z_2$ . Входом  $x$  является изменение начальной кон-

центрации вещества  $\mathcal{A}$  при подаче его в реактор (сырье переменного состава), т.е.  $q_{A_0}(t)$ . Таким образом,  $x = q_{A_0}(t)$ ,  $z_1 = q_A$ ,  $z_2 = q_B$ ,  $y = q_B$ .

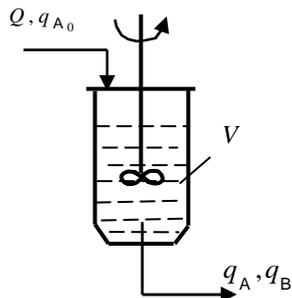


Рис. 1.2. Схема реактора идеального смешения

Уравнения динамики реактора имеют вид

$$\begin{cases} \frac{dq_A(t)}{dt} = \frac{Q}{V}(q_{A_0}(t) - q_A(t)) - kq_A(t); \\ \frac{dq_B(t)}{dt} = -\frac{Q}{V}q_B(t) + kq_A(t) \end{cases}$$

или

$$\begin{cases} \begin{pmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{pmatrix} = \begin{pmatrix} -(a+b) & 0 \\ b & -a \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} a \\ 0 \end{pmatrix} x(t); \\ y(t) = z_2(t), \end{cases}$$

где

$$a = \frac{Q}{V}, b = k.$$

Оператору или отображению соответствует решение системы уравнений в матричной форме, представленное в конечном виде

$$z(t) = e^{A(t-t_0)} z(t_0) + \int_{t_0}^t e^{A(t-s)} Bx(s) ds.$$

Известны различные классы динамических систем – детерминированные и стохастические, стационарные и нестационарные, линейные и нелинейные и т.д. Наиболее полно изучены линейные динамические системы  $\Sigma^l$ , уравнения которых имеют вид

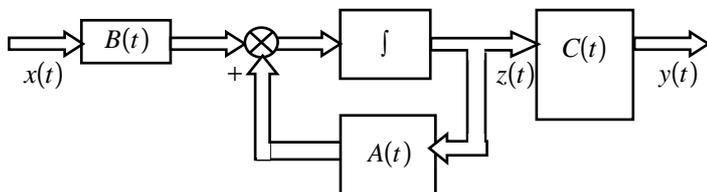
$$\begin{cases} \dot{z}(t) = A(t)z(t) + B(t)x(t); \\ y(t) = C(t)z(t), \end{cases}$$

где

$$A(t) = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & \dots & a_{2n}(t) \\ \dots & & \dots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix}; \quad B(t) = \begin{bmatrix} b_{11}(t) & \dots & b_{1n}(t) \\ b_{21}(t) & \dots & b_{2n}(t) \\ \dots & & \dots \\ b_{n1}(t) & \dots & b_{nm}(t) \end{bmatrix}; \quad C(t) = \begin{bmatrix} c_{11}(t) & \dots & c_{1n}(t) \\ c_{21}(t) & \dots & c_{2n}(t) \\ \dots & & \dots \\ c_{p1}(t) & \dots & c_{pn}(t) \end{bmatrix}.$$

Структурная схема  $\Sigma^l$  приведена на рис. 1.3. Для нее оператор  $\Phi$  определяется матрицами  $A(t)$  и  $B(t)$ , а оператор  $\Psi$  – матрицей  $C(t)$ . Если эти матрицы не зависят от времени, то систему называют линейной стационарной, будем ее обозначать  $\Sigma^{lc}$ . Модели систем  $\Sigma^{lc}$  широко используются на практике при исследовании сложных технических объектов.

Рис. 1.3. Схемы линейной динамической системы



Различие ММ обуславливается их предназначением, например, для исследования эффективности режимов функционирования систем, оптимизации установившихся (статических) и переходных (динамических) режимов их работы, оптимального проектирования систем и управления ими и т.д. требуются различные математические модели.

Для ММ типа «вход – выход» будем использовать следующие обозначения:  $x=(x_1, \dots, x_m)$  – вектор входных переменных системы;  $y=(y_1, y_2, \dots, y_n)$  – вектор выходных переменных системы,  $a=(a_1, a_2, \dots, a_l)$  – вектор параметров системы. Тогда поведение системы с сосредоточенными координатами  $y, x$  в статике и неизменными во времени  $t$  свойствами (стационарный объект) описывается уравнениями ММ вида

$$F[y, x, a] = 0 \text{ или } y = f(x, a).$$

ММ статики нестационарной системы с сосредоточенными координатами (квазистатическая модель) представляет собой систему уравнений вида:

$$F[y, x, a(t)] \approx 0, \quad \frac{da}{dt} = f_1(y, a).$$

Поведение системы с сосредоточенными координатами  $y, x$  в динамике и неизменными во времени  $t$  свойствами описывается уравнениями ММ вида

$$F\left[\frac{dy}{dt}, y(t), x(t), a\right] = 0 \text{ или } \frac{dy}{dt} = f(y(t), x(t), a).$$

ММ динамики нестационарной системы с сосредоточенными координатами представляет собой систему уравнений вида

$$F\left[\frac{dy}{dt}, x(t), a(t)\right] \approx 0, \quad \frac{da}{dt} = f_1(y(t), a).$$

Если координаты системы  $x, y$  распределены по пространственной переменной  $l$  (длина, радиус, высота) и его свойства неизменны во времени  $t$ , то мы имеем дело со стационарными ММ статики или динамики системы с распределенными координатами, которые имеют вид, соответственно:

$$F\left[\frac{dy}{dt}, y(l), x(l), a\right] = 0, \quad F\left[\frac{\partial y}{\partial t}, \frac{\partial y}{\partial l}, y(t, l), x(l), a\right] = 0.$$

По структуре  $F$  ММ систем разделяются на линейные и нелинейные. Решение  $y(x, a)$  системы уравнений ММ, линейной по  $y$ , удовлетворяет следующим условиям (принципу суперпозиции):

- 1) аддитивности  $y(x_1 + x_2, a) = y(x_1, a) + y(x_2, a)$ ;
- 2) однородности  $y(c \times x, a) = c \times y(x, a)$ ;

где  $x_1$  и  $x_2$  – произвольные функции аргументов  $t, l$  или некоторые числа;  $c$  – любое вещественное число.

Решение  $y(x, a)$  называется линейным по  $a$ , если

$$y(x, a_1 + a_2) = y(x, a_1) + y(x, a_2) \text{ и } y(x, c \times a) = c \times y(x, a),$$

где  $a_1, a_2$  – произвольные параметры ММ.

Если для некоторой ММ не выполняется хотя бы одно из условий принципа суперпозиции, то она относится к классу нелинейных.

Математические модели систем чаще всего описываются *нелинейными* уравнениями.

Один из существенных признаков классификации связан с отражением в ММ тех или иных особенностей систем. Если ММ отображает устройство системы и связи между составляющими ее элементами, то ее называют *структурной математической моделью*. Если же ММ отражает происходящие в системе физико-химические, механические или информационные процессы, то ее относят к *функциональным математическим моделям*.

*Структурные* ММ делят на топологические и геометрические, составляющие два уровня *иерархии* ММ этого типа. Первые отображают состав системы и связи между его элементами. *Топологическую* ММ целесообразно применять на начальной стадии исследования сложной по структуре системы, состоящей из большого числа элементов, прежде всего для уяснения и уточнения их взаимосвязи. Такие ММ имеют форму графов, таблиц, матриц, списков и т.п., ее построению обычно предшествует разработка структурной схемы системы.

*Геометрические* ММ дополнительно к информации, представленной в *топологической* ММ, содержат сведения о форме и размерах системы и ее элементах, об их взаимном расположении. *Геометрические* модели находят применение при проектировании систем, разработке технической документации и технологических процессов.

В соответствии с принципом моделируемости описать поведение достаточно сложной системы одной моделью, как правило, не удастся, а если такая ММ и была построена, то она оказалась бы слишком громоздкой для количественного анализа. Поэтому в таких случаях обычно применяют *постулат декомпозиции*. Он состоит в условном разбиении системы на отдельные более простые блоки (подсистемы) и элементы, допускающие их независимое исследование с последующим учетом взаимного влияния блоков (подсистем) и элементов друг на друга.

Сформулируем теперь конкретные критерии, которым должна удовлетворять «хорошая» модель. Такая модель должна быть: 1) целенаправленной, т.е. модель должна позволять решать определенный класс задач, для которых она предназначена, например, задач прогнозирования, оптимизации режимов работы, оптимального управления, проектирования систем и т.п.; 2) простой и понятной пользователю; 3) надежной в смысле гарантии от абсурдных ответов; 4) удобной в управлении и об-

рашении; 5) полной с точки зрения возможностей решения поставленных задач; 6) адаптивной, т.е. позволяющей легко переходить к другим модификациям или обновлять данные; 7) допускающей постепенные изменения в том смысле, что, будучи вначале простой, она может во взаимодействии с пользователем становиться все более сложной и точной.

### 1.3. ОСНОВНЫЕ ПОНЯТИЯ МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ ХТС В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Математические модели могут быть разделены на *детерминированные* и *стохастические, статические* и *динамические, дискретные, непрерывные* и *дискретно-непрерывные*. В зависимости от формы представления ХТС можно выделить *виртуальное* (мысленное) и *реальное* моделирование. Виртуальное моделирование часто является единственным способом моделирования объектов, которые либо практически нереализуемы в заданном интервале времени, либо существуют вне условий, возможных для их физического создания. Например, на базе виртуального моделирования могут быть проанализированы многие ситуации микромира, которые не поддаются физическому эксперименту. Виртуальное моделирование может быть реализовано в *наглядном, символическом* и *математическом* видах.

При *наглядном моделировании* на базе представлений человека о реальных объектах создаются различные визуальные модели, отображающие явления и процессы, протекающие в объекте. К ним относятся *гипотетическое* и *аналоговое моделирование*, а также *макетирование*. В основу *гипотетического моделирования* исследователем закладывается некоторая гипотеза о закономерностях протекания процесса в реальном объекте, базирующаяся на причинно-следственных связях между выходными и входными переменными изучаемого объекта. Гипотетическое моделирование используется, когда знаний об объекте недостаточно для построения формальных или аналитических моделей.

*Аналоговое моделирование* основывается на применении аналогий различных уровней. Наивысшим уровнем является полная аналогия, имеющая место только для достаточно простых объектов. С усложнением объекта используют аналогии последующих уровней, когда аналоговая модель отображает либо несколько, либо одну сторону функционирования объекта.

Существенное место при виртуальном моделировании занимает *макетирование*. Мысленный макет может применяться либо в случаях, когда протекающие в реальном объекте процессы не поддаются физическому моделированию, либо предшествовать проведению других видов моделирования. В основе построения мысленных макетов также лежат аналогии, обычно базирующиеся на причинно-следственных связях между явлениями и процессами в объекте.

*Символическое моделирование* представляет собой искусственный процесс создания логического объекта, который замещает реальный и выражает основные свойства его отношений с помощью определенной системы знаков или символов.

*Математическое моделирование* можно разделить на аналитическое, имитационное и комбинированное. Для аналитического моделирования характерно то, что процессы функционирования элементов системы записываются в виде некоторых функциональных соотношений (алгебраических, дифференциальных, интегро-дифференциальных и т.п.) или логических условий. Аналитическая модель может быть исследована следующими методами: а) аналитическим, когда возможно получить в явном виде зависимости для искомых характеристик; б) численным, когда невозможно получить аналитическое решение в виде конечных формул; в) качественным, когда, не имея решения в явном виде, можно исследовать некоторые его свойства (например, оценить устойчивость решения).

При *имитационном моделировании* воспроизводится процесс функционирования системы во времени, причем имитируются элементарные явления, составляющие процесс, с сохранением их логической структуры и последовательности протекания во времени, что позволяет по исходным данным получить сведения о состояниях процесса в определенные моменты времени, дающие возможность оценить характеристики системы [2]. Результаты, полученные при воспроизведении на имитационной модели процесса функционирования системы, как правило, являются реализациями случайных величин и функций. В этом случае для нахождения характеристик процесса требуется его многократное воспроизведение с последующей статистической обработкой информации, и целесообразно в качестве метода машинной реализации имитационной модели использовать метод *статистического моделирования (метод Монте-Карло)*.

Метод имитационного моделирования позволяет решать задачи анализа больших систем, включая задачи оценки: вариантов структуры системы, эффективности различных алгоритмов управления, влияния изменения различных параметров системы. Имитационное моделирование может быть положено также в основу структурного, алгоритмического и параметрического синтеза больших систем, когда требуется создать систему с заданными характеристиками при определенных ограничениях.

*Комбинированное* (аналитико-имитационное) *моделирование* при анализе и синтезе систем позволяет объединить достоинства аналитического и имитационного моделирования. При построении комбинированных моделей проводится предварительная декомпозиция процесса функционирования объекта на составляющие подпроцессы и для тех из них, где это возможно, используются аналитические модели а для остальных подпроцессов – имитационные модели. Такой комбинированный подход позволяет охватить качественно новые классы систем, которые не могут быть исследованы с использованием только аналитического и имитационного моделирования в отдельности.

Особое место в моделировании занимает *кибернетическое моделирование*, в котором отсутствует непосредственное подобие физических процессов, отображаемых в моделях, реальным процессам. В этом случае стремятся отобразить лишь некоторую функцию и рассматривают реальный объект как «черный ящик», имеющий входы и выходы, и моделируют связи между ними. Чаще всего при использовании кибернетических моделей проводят анализ поведенческой стороны объекта при различных воздействиях внешней среды. В основе кибернетических моделей лежит отражение некоторых информационных процессов управления, что позволяет оценить поведение реального объекта. Для построения имитационной модели в этом случае необходимо выделить исследуемую функцию реального объекта, попытаться формализовать эту функцию в виде некоторых операторов связи между выходными и входными переменными и воспроизвести с помощью модели данную функцию, причем на базе совершенно иных математических соотношений и, естественно, иной физической реализации процесса.

Обеспечение требуемых показателей качества функционирования сложных систем связано с необходимостью изучения протекающих в них значительного числа, в том числе стохастических, процессов на основе проведения комплекса теоретических и экспериментальных исследований, взаимно дополняющих друг друга. Эффективность экспериментальных исследований сложных систем оказывается крайне низкой, поскольку проведение натуральных экспериментов с реальной системой либо требует больших материальных затрат и значительного времени, либо вообще практически невозможно (например, на

этапе проектирования, когда реальная система отсутствует). Эффективность теоретических исследований с практической точки зрения в полной мере проявляется лишь тогда, когда их результаты с требуемой степенью точности и достоверности могут быть представлены в виде триады «модель–алгоритм–программа ЭВМ», пригодной для получения соответствующих характеристик процесса функционирования исследуемых систем.

Вычислительный эксперимент с имитационной моделью сложной системы требует также серьезной подготовки. Во многих случаях для решения задач имитационного моделирования создается информационная система, состав которой образует следующие виды обеспечения.

*Математическое и алгоритмическое обеспечение*, в основном это математические соотношения, описывающие поведение реального объекта, алгоритмы, обеспечивающие подготовку и работу с моделью, а также лингвистическое обеспечение, т.е. совокупность используемых языков программирования.

*Программное обеспечение*, состоящее из системного или общего и прикладного программного обеспечения. Последнее включает в себя программы планирования и проведения эксперимента, обработки и интерпретации результатов и др.

*Информационное обеспечение*, т.е. файлы с данными об исследуемой системе, базы данных, формы документов и т.п.

*Техническое (аппаратное) обеспечение*, т.е. средства вычислительной, коммуникационной и организационной техники, ввода, вывода и обмена информацией между участниками эксперимента.

*Организационное, правовое и методическое обеспечения*, включающие инструктивные и нормативные материалы по организации, проведению и обработке результатов вычислительного эксперимента при использовании конкретной информационной системы.

Немаловажное значение для успешного функционирования информационной системы имеет человеческий компонент, т.е. разработчики модели, программисты, пользователи и т.д.

Информационная система позволяет заменить эксперимент с реальным процессом функционирования сложного объекта экспериментом с математической моделью.

Основными достоинствами исследований сложных систем методом имитационного моделирования являются следующие факторы: 1) вычислительный эксперимент дает возможность исследовать процесс функционирования системы в любых условиях, в том числе ситуациях, недопустимых для нормальной работы системы, например предаварийных и аварийных; 2) применение ЭВМ существенно сокращает продолжительность испытаний по сравнению с натурным экспериментом; имитационная модель позволяет без больших затрат изменять структуру и параметры моделируемой системы, что важно при определении оптимального варианта системы на этапе проектирования.

Эффективность имитационного моделирования оценивается рядом показателей, к которым относятся точностные характеристики, затраты на разработку модели и информационной системы, время проведения эксперимента и требуемые машинные ресурсы. Для повышения эффективности организации и проведения вычислительных экспериментов применяются современные информационные технологии (например, CASE-технологии), языки программирования (C+, Delphi), пакеты прикладных программ (Matlab, Femlab, Cadfem), математический аппарат планирования эксперимента и др.

CASE-технология в широком смысле представляет собой совокупность методологий анализа, автоматического проектирования, разработки и сопровождения программного обеспечения сложных систем [4]. Аббревиатура CASE используется для двух направлений проектирования систем. Первое – Computer Aided System Engineering – направлено на решение задач концептуального проектирования сложных слабоструктурированных систем. CASE-технологии этого направления называют *системами CASE для концептуального проектирования*. Второе направление – Computer Aided Software Engineering решает задачи автоматизированного проектирования программного обеспечения. Эти CASE-системы называют инструментальными средами разработки программного обеспечения (*инструментальными CASE-системами*).

Среди систем CASE первого направления выделяют системы функционального и информационного (поведенческого) проектирования. Наиболее распространенной методикой *функционального проектирования* сложных систем является методика SADT (Structured Analysis and Design Technique). Эта методика стала основой стандарта IDEF0 (Integrated Definition 0). Программные средства информационного проектирования реализуют методики инфологического проектирования баз данных. Широкое распространение получила методика создания информационных моделей IDEF1X. Применение инструментальных CASE-систем позволяет сократить затраты на разработку программного обеспечения за счет уменьшения числа итераций и числа ошибок, а также улучшить качество программного обеспечения вследствие лучшего взаимопонимания разработчика и заказчика. При этом облегчается сопровождение готового программного продукта.

CASE-средства представляют собой программные продукты, используемые как для автоматизированной разработки определенных видов моделей, например, функциональных, информационных, так и для автоматизированного создания программного обеспечения информационных и других систем. В последнем случае CASE-средства поддерживают все этапы проектирования прикладного программного обеспечения и баз данных от формулировки требований до генерации кода, тестирования документирования и сопровождения в процессе эксплуатации.

С позиции решения задач моделирования CASE-технологии широко применяются для моделирования систем большинства предметных областей. Исключительно большое значение CASE-технологии имеют для разработки моделей деятельности предприятий, в частности функциональных и информационных, которые необходимы для решения задач системного анализа, проектирования, реинжиниринга и др.

При решении многих задач анализа и синтеза ХТС могут использоваться различные программные средства, предназначенные для математических вычислений в технических приложениях – Matlab, Matcad, Excel, Maple и др. Система MATLAB (MATrix LABoratory – матричная лаборатория, фирма MathWorks, Inc) создана «как язык программирования высокого уровня для технических вычислений» [5]. Система имеет открытую архитектуру, современные версии поставляются вместе с пакетом расширения Simulink. Наиболее полно функциональные возможности системы проявляются в рамках комплекса «MATLAB + Simulink + пакеты расширения». Число пакетов расширения насчитывает несколько десятков.

В системе реализован принцип визуально-ориентированного программирования; уравнения состояний, описывающие динамические системы, формируются автоматически; имеются виртуальные средства регистрации и визуализации результатов моделирования. Функции системы MATLAB позволяют в интерактивном режиме выполнять сложные математические

вычисления: разрабатывать алгоритмы; выполнять вычислительный эксперимент и имитационное моделирование; анализировать данные и визуализировать результаты.

Наиболее известны области применения системы MATLAB:

- математика и вычисления;
- разработка алгоритмов;
- вычислительный эксперимент, имитационное моделирование, макетирование;
- анализ данных, исследование и визуализация результатов;
- научная и инженерная графика;
- разработка приложений, включая графический интерфейс пользователя.

MATLAB – это интерактивная система, основным ее объектом является массив, для которого не требуется указывать размерность явно. Это позволяет решать многие вычислительные задачи, связанные с векторно-матричными формулировками, существенно сокращая время, которое понадобилось бы для программирования на скалярных языках типа C или FORTRAN. Система MATLAB – это одновременно и операционная среда и язык программирования. Одна из наиболее сильных сторон системы состоит в том, что на языке MATLAB могут быть написаны программы для многократного использования. Пользователь может сам написать специализированные функции и программы, которые оформляются в виде М-файлов. По мере увеличения количества созданных программ возникают проблемы их классификации и тогда можно попытаться собрать родственные функции в специальные папки. Это приводит к концепции пакетов прикладных программ, которые представляют собой коллекции М-файлов для решения определенной задачи или проблемы.

В табл. 1.1 представлены основные схемы моделирования, соответствующие этим подходам, области их применения и функции в пакете MATLAB.

### 1.1. Типовые математические схемы

Условное обозначение, название	Определение ММ (схемы)	Область применения, примеры [примечание, функции в MATLAB]
<i>S</i> -схемы	Непрерывно-детерминированные модели, отражающие статику изучаемого объекта	Модели процесса смешения потоков, теплообмена в теплообменнике, растворения, ректификации и др. Эти модели могут быть представлены системами алгебраических уравнений для объектов с сосредоточенными параметрами или дифференцированными уравнениями в частных производных для объектов с распределенными параметрами [polyfit (x, y, k)]
<i>D</i> -схемы	Непрерывно-детерминированные модели, отражающие динамику изучаемого объекта	Модели химического процесса конверсии, периодического процесса растворения, сушки. Математическая модель биосинтеза. Эти модели относятся к классу динамических нелинейных моделей с сосредоточенными координатами [ode45; ode23; ode113; ss; canstart]
<i>F</i> -схемы	Дискретно-детерминированные модели, используемые для создания вычислительных алгоритмов, позволяющих производить всестороннее численное исследование широкого класса моделей	Конечно-разностные схемы, конечные автоматы; <i>F</i> -автомат первого рода (автомат Мили); <i>F</i> -автомат второго рода (автомат Мура) [dss; oe]

Продолжение табл. 1.1

Условное обозначение, название	Определение ММ (схемы)	Область применения, примеры [примечание, функции в MATLAB]
-схемы	Дискретно-стохастические модели	<i>P</i> -автоматы (вероятностные автоматы) применяются для разработки методов проектирования дискретных систем, проявляющих статистически закономерное случайное поведение, для выяснения алгоритмических возможностей таких систем, обоснования границ целесообразности их использования, и решения задач синтеза по выбранному критерию дискретных стохастических систем; <i>P</i> -автомат (вероятностный автомат Мили); <i>P</i> -автомат (вероятностный автомат Мура) [arg, пакеты Stateflow, DSP Blockset]

<i>Q</i> -схемы	Непрерывно-стохастические модели – системы массового обслуживания	Процессы функционирования экономических, производственных, технических и других систем (потоки поставок продукции предприятиям, потоки деталей и комплектующих изделий, заявки на обработку информации ЭВМ от удаленных терминалов и т.д.); многоканальная <i>Q</i> -схема (приборы обслуживания соединены параллельно); многофазная <i>Q</i> -схема (приборы и их параллельные композиции соединены последовательно) [ivar, пакеты Signal Processing, Toolbox]
-----------------	---	---

Продолжение табл. 1.1

Условное обозначение, название	Определение ММ (схемы)	Область применения, примеры [примечание, функции в MATLAB]
<i>N</i> -схемы	Сетевые модели, решающие задачи, связанные с формализованным описанием и анализом причинно-следственных связей в сложных системах	Сети Петри, описывающие структуру и взаимодействие параллельных систем и процессов, необходимы для описания в моделируемой системе событий произвольной длительности, отражают порядок наступления событий в исследуемой системе [пакет Communications Block set]
<i>A</i> -схемы	Комбинированные модели	В качестве <i>A</i> -схемы выступает агрегат, а связь между агрегатами (внутри системы и с внешней средой) осуществляется с помощью оператора сопряжения <i>R</i> [пакет Power System Block set]
<i>H</i> -схемы	Модели на МСФ	Эти модели позволяют описывать процессы в сложных системах при длительной эксплуатации, когда отказ (нарушение) одного элемента или части элементов не приводит к отказу всей системы. Модели на МСФ могут быть представлены дифференциальными уравнениями с разрывной частью, уравнениями, параметры которых представляют собой случайные величины и др. Примерами систем, описываемых моделями на МСФ являются совокупность локальных систем контроля и регулирования технологической линии, парк станочного оборудования цеха и т.п.
<i>I</i> -схемы	Модели систем искусственного интеллекта	Примерами моделей СИИ являются модели представления знаний. С помощью логико-лингвистических моделей строятся нечеткие системы управления [среда fuzzy TECH]

Рассмотренные типовые математические схемы (табл. 1.1) (*S*-, *D*-, *F*-, *P*-, *Q*-, *N*-, *A*-, *H*-, *I*-схемы) позволяют формализовать задачи моделирования достаточно широкого класса сложных систем, с которыми приходится иметь дело в практике исследования и проектирования.

В 1984 г. группа предприимчивых профессоров начала разработку программы ChemCAD, и этот продукт стал собственностью издательского дома McGraw Hill. В 1988 г. фирма Chemstation овладела и полностью изменила программу, выпустив ChemCAD версии II. Цель – создать интерактивный продукт моделирования для ПК, используя графические ресурсы компьютера. Первыми покупателями программы ChemCAD стала химическая промышленность. На сегодняшний день реализован интегрированный программный продукт ChemCAD 5.

Основные модули универсальной моделирующей программы (симулятора процессов) были разработаны много лет назад. Например, база данных компонентов, основные термодинамические методы, оборудование для получения промышленных газов и компонентов коммунального обслуживания (вода, пар, топливо, воздух, электричество); а также система отчетов. Использование моделирующей программы, такой как ChemCAD, позволяет надежно выполнять все расчеты, необходимые в повседневной инженерной деятельности химика-технолога.

1. Оперативное получение сведений о физических свойствах чистых веществ и их смесей, основанное на надежной базе данных и современных термодинамических методах. К примеру, вычисление давления пара в химической смеси потребует: нахождения в таблицах или диаграммах давления пара всех компонентов системы; расчета парциальных давлений пара при заданных условиях технологического процесса, и после этого – вычисления суммарного давления пара. А если система неидеальная? В таком случае нужно будет воспользоваться равновесными диаграммами и вычислить активности всех компонентов смеси. Давление пара – это лишь одно из многих важных свойств смеси. А теперь сравните эти трудозатраты с нажатием нескольких клавиш, которое потребуется для решения всех этих задач компьютерной моделирующей программой (симулятором процессов).

2. Можно рассчитывать конструктивные характеристики единиц оборудования, например, трубопровода, и оценить перепад давления. Современный симулятор процессов включает модели всех единиц оборудования и расчет происходит достаточно быстро. В случае длинного трубопровода симулятор позволяет разделить трубопровод на короткие сегменты и точно рассчитать перепад давления. Более того, такие процессы, как ректификация, абсорбция и процессы в химических реакторах, слишком сложны для вычислений вручную. Нужно иметь в виду, что эмпирические методы, такие как Фенске-Андервуда-Джиллиленда для расчета ректификации, были разработаны в 1930 – 1940-х гг. Эти методы не удовлетворяют современным требованиям к точности расчетов. Только с помощью компьютерного расчета можно получить точные и надежные результаты.

3. Технологическая схема или модель технологического процесса обычно слишком сложна для вычислений вручную. Для моделирования простой технологической схемы из пяти аппаратов и нескольких потоков нужно будет последовательно рассчитывать все единицы оборудования. Для каждого потока следует обязательно рассчитывать его свойства – температуру, давление, расход и состав. Для всех аппаратов расчеты проводятся по упрощенному алгоритму. В случае циклического

процесса необходимо реализовать процедуру итерационных расчетов методом последовательных приближений до тех пор, пока не будет достигнута сходимость на двух последующих итерациях. Это очень трудоемкий и неточный способ. С помощью программы-симулятора процессов все вычисления выполняются автоматически.

4. Для нового технологического процесса не бывает простых решений. Возможно множество альтернатив. Только универсальная моделирующая программа-симулятор процессов поможет отбросить худшие варианты и даст возможность за короткое время найти наилучшие решения.

5. В условиях действующих производств всегда могут быть выбраны более эффективные решения и ежедневно приходится решать текущие производственные задачи. Программа-симулятор процессов поможет проанализировать реальную ситуацию и даст возможность определить сущность проблемы и пути ее наиболее эффективного и экономного решения.

При проектировании новых производств или реконструкции действующих сначала нужно решить, оправдано ли экономически получение производимых продуктов и правильно ли выбрана технологическая схема производства. Чтобы ответить на этот вопрос, следует оценить потребление сырьевых ресурсов и выход продуктов. Зная цены сырьевых материалов и конечных продуктов можно примерно оценить, выгодно ли вообще данное производство. Однако описанная процедура вовсе не отвечает на вопрос «Будет ли прибыльным производство?». Чтобы рассчитать прибыль, придется оценить будущие капиталовложения и эксплуатационные затраты (включая расходы на коммунальное обслуживание). В целом ряде случаев нужно рассчитать балансы технологических процессов, оценить затраты, и лишь после этого приступить к лабораторным экспериментам.

Применение универсальной моделирующей программы ChemCAD позволяет осуществлять:

- анализ возможных альтернативных решений с экономическими расчетами для поиска оптимальной технологической схемы и ее параметров;
- анализ параметрической чувствительности процессов с определением их критических значений;
- возможность синтеза экономически наиболее оправданной технологической схемы производства.

Расчет материальных и тепловых балансов – это первое, что необходимо сделать инжиниринговой компании при проектировании нового процесса или при реконструкции действующего. Универсальная моделирующая программа ChemCAD позволяет осуществить быстрый и точный расчет всех материальных и тепловых балансов технологических процессов, оптимизацию конфигурации технологической схемы производства, расчет всех промежуточных стадий производства, расчет конструктивных размеров технологического оборудования.

В случае периодических процессов условия материальных и тепловых балансов выполняются для каждого данного момента времени. Периодические процессы рассчитываются при моделировании нестационарных процессов. Тем не менее, для определения начального, конечного и критического состояния периодических процессов должны быть рассчитаны материальные и тепловые балансы стационарного состояния.

Компьютерное моделирование в системе ChemCAD позволяет определять кинетику химических или биохимических реакций на стадии разработки производства; осуществлять масштабные переходы при расширении производства: от лабораторных исследований через полупромышленную установку к промышленному внедрению; внедрять новые технологии для получения новых изделий на существующем оборудовании, что сокращает время, затраченное на внедрение новых технологий; разрабатывать новые типы оборудования для создания новых процессов; разрабатывать циклограммы периодических процессов; определять основные параметры процесса и находить наилучший вариант организации производства из многих альтернатив; оптимизировать производство путем определения лучших технологических режимов и настроек контроллера; рассчитывать параметры системы управления; определять степень безопасности процесса и, наконец, разрабатывать тренажеры для оперативного персонала.

При построении математических моделей процессов функционирования систем можно выделить ряд подходов: непрерывно-детерминированный (использует алгебраические и дифференциальные уравнения); дискретно-детерминированный (конечные-разностные схемы, конечные автоматы); дискретно-стохастический (вероятностные автоматы); непрерывно-стохастический (системы массового обслуживания); обобщенный (агрегативные системы) и др.

Экспертные системы (ЭС) с позиции задач моделирования систем следует рассматривать, во-первых, как средства или инструментариум для получения и использования моделей, и, во-вторых, как сложные объекты моделирования, в том числе моделирования представления знаний. ЭС относятся к классу систем искусственного интеллекта, они способны строить логические выводы, осуществлять обобщения и формировать заключения на основе использования знаний и данных подобно тому, как это делают специалисты в своей области при выработке умозаключений. Экспертная система представляет собой интеллектуальную программу, способную делать логические выводы на основании знаний в конкретной предметной области, и обеспечивающую решение специфических задач, в том числе в области моделирования. Экспертные системы обладают возможностями, позволяющими пользователю решать задачи, которые в отсутствие эксперта (специалиста в предметной области) правильно решить невозможно. Это обеспечивается тем, что они содержат в памяти необходимый багаж знаний, касающихся конкретной предметной области, используют результаты большого опыта работы в этой области, а также позволяют точно сформулировать и правильно решить исследуемую задачу.

Основными требованиями, предъявляемыми к экспертной системе, являются следующие: использование знаний, связанных с конкретной предметной областью; способность приобретения знаний от экспертов; решение реальных и достаточно сложных задач; наличие способностей, присущих экспертам. В качестве экспертов могут выступать опытные проектировщики, консультанты, экономисты, врачи, программисты, преподаватели, переводчики и другие квалифицированные специалисты в соответствующих областях.

Так как создание экспертных систем стало возможным лишь с развитием ЭВМ, то их можно рассматривать как компьютерные системы, использующие логику экспертов.

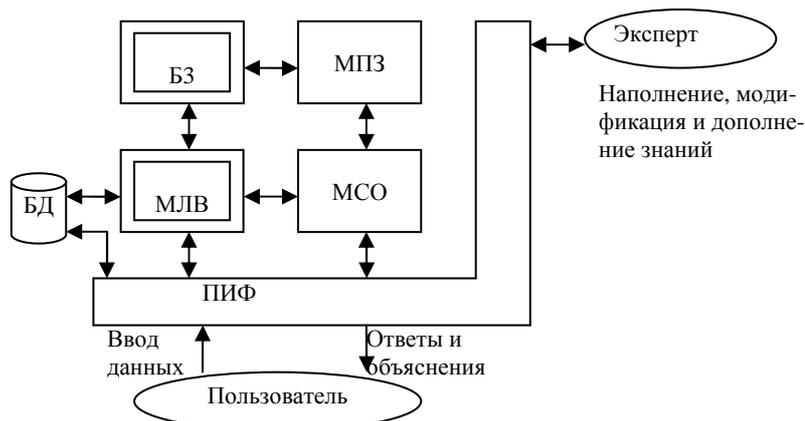
#### **ПРИМЕР СТРУКТУРНОЙ СХЕМЫ ЭКСПЕРТНОЙ СИСТЕМЫ ПРИВЕДЕН НА РИС. 1.4.**

Ядром экспертной системы является база знаний (БЗ), которая содержит знания из конкретной предметной области. В БЗ содержатся как общие знания, так и информация о частных случаях.

Механизм логического вывода или создатель заключения применяет знания и сведения из базы данных при решении реальных задач. Модуль приобретения знаний позволяет пополнять и модифицировать знания в процессе эксплуатации сис-

темы. Модуль советов и объяснений выдает заключения, позволяет программе пояснить пользователю свое «рассуждение», дать ответы на вопросы «Как» и «Почему?».

Экспертную систему можно рассматривать как прикладную диалоговую систему искусственного интеллекта (ИИ), спо-



собную получать, накапливать, корректировать знания из некоторой предметной

**Рис. 1.4. Структурная схема экспертной системы:**

БД – база данных; БЗ – база знаний; МЛВ – механизм логических выводов;  
 МПЗ – модуль приобретения знаний; МСО – модуль советов и объяснений;  
 ПИФ – пользовательский интерфейс

области (обычно предъявляемые специалистами-экспертами), выводить новые знания, находить на основе этих знаний решения практических задач, близкие по качеству к решениям экспертов, и по запросу пользователя объяснять ход решения в понятной для него форме.

Экспертные системы отличаются от обычных прикладных программ, решающих строго определенные математические задачи по точным детерминированным алгоритмам, тем, что они решают неформализованные (слабо формализованные), слабо структурированные задачи. Алгоритмическое решение таких задач или не существует в силу неполноты, неопределенности, неточности, расплывчатости рассматриваемых ситуаций и знаний о них, или решения неприемлемы на практике в силу сложности строгих и точных алгоритмов. Поэтому ЭС используют логический вывод и эвристический поиск решения.

От систем поддержки принятия решений, которые не используют экспертных оценок и опираются на математические методы и модели, ЭС отличаются тем, что базируются на эвристических, эмпирических знаниях, оценках, методах, которые получены от экспертов, и, кроме того, способны анализировать и объяснять пользователю свои действия и знания.

Для построения ЭС применяются специальные языки программирования, ориентированные на решения задач искусственного интеллекта, например LISP и PROLOG.

Язык LISP (Lisp Processing – обработка списков) поддерживает большое число встроенных функций со списками как связными структурами указателей.

Язык PROLOG (Programming in Logic – программирование в логике) представляет собой инструментальное средство логического программирования, основанное на теории предикатов первого порядка. Развитием этого языка является Lambda-PROLOG, а также создание сред декларативного логического программирования (языки Mercury, Goedel) и параллельные языки логического программирования.

К программным средствам, используемым при разработке экспертных систем относятся также EMYCIN, EXPERT, OPS5, ROSIE и многие другие.

Идея построения ЭС сформировалась в ходе исследований в области ИИ. С точки зрения задач, которые решают ЭС, они делятся на два класса. Системы первого класса предназначаются для повышения культуры работы и уровня знаний специалистов в различных предметных областях. Системы второго класса можно назвать консультирующими или диагностирующими. Для оказания помощи человеку в решении указанных задач разрабатываются комплексы программ для ПК, называемые интеллектуальными системами, основанными на знаниях.

Выделяют ЭС, основанные на правилах (ЭСП), ЭС на основе модели (ЭСМ), ЭС, использующие рассуждения на основе опыта (ЭСО), и гибридные экспертные системы (ГЭС).

Экспертные системы, основанные на правилах, используют знания экспертов в простой форме, например, в виде системы продукционных правил. Для этих ЭС характерна относительная простота разработки ЭС, так как правила легко накапливать, реализовывать и тестировать. ЭСП имеют хорошие результаты применения в узких предметных областях.

К недостаткам ЭСП относятся следующие: обычно правила носят эвристический характер и не охватывают знаний, основанных на моделях; в них отсутствует теоретическое обоснование; эвристические правила не проявляют робастности при «неожиданных» исходных данных, они оказываются неприменимыми на границе предметной области, при столкновении с новыми проблемами.

Экспертные системы, основанные на моделях, при решении задач используют функциональные и структурные знания о предметной области, что позволяет решать часть задач, не предусмотренных при создании ЭС. В ЭСМ имеется возможность возвращаться к исходным данным при столкновении с новой проблемой, кроме того, базирование на теоретических научных знаниях позволяет использовать знания в разных задачах.

В качестве недостатков ЭСМ укажем следующие. Они не используют известные экспериментальные данные и знания, относящиеся к предметной области. Кроме того, здесь необходимы точные модели описания предметной области, однако во многих реальных случаях хорошо определенная научная теория для построения точных моделей отсутствует. Разработка комплекса точных моделей, необходимых для ЭСМ, обычно связана с проведением сложных и трудоемких исследований.

Экспертные системы, основанные на опыте, по сравнению с ЭСП и ЭСМ, имеют ряд достоинств. При создании БЗ здесь можно непосредственно использовать имеющийся практический опыт или сведения, полученные из литературных источников, т.е. обойтись без привлечения экспертов. В ЭСО значительно сокращаются рассуждения и время решения задач, если известны аналогичные случаи, называемые шаблонами. ЭСО позволяют избежать прошлых ошибок и использовать удачные результаты, в том числе в сложных ситуациях. В них реализуется простая аддитивная модель приобретения знаний и не требуется проведения анализа знаний о предметной области. Кроме того, имеется возможность использования стратегий индексирования для выбора соответствующих случаев из БД.

К недостаткам ЭСО следует отнести то, что в них не учитываются глубокие знания о предметной области, это может приводить к ошибочному применению опыта; использование излишне большой БД в ЭСО приведет к снижению производительности их работы. Кроме того, имеются трудности в определении хороших критериев для индексирования и сравнения случаев.

Создание гибридных экспертных систем, в которых сочетаются возможности систем, основанных, например, на правилах и опыте, или моделях и правилах, или моделях и опыте, позволяет значительно повысить эффективность ЭС.

Так, ГЭС, основанные на правилах и опыте, имеют возможность производить предварительный просмотр известных случаев до начала рассуждений на основе правил, что существенно снижает затраты на поиск решения. Результаты выполненного поиска здесь можно сохранять в БД для будущего использования, чтобы избегать повторного поиска. В БД можно сохранять как положительные примеры, так и исключения.

К достоинствам ГЭС, основанных на моделях и правилах, можно отнести следующие. Они имеют возможность дополнять объяснения теоретическими знаниями. В этих системах повышается устойчивость (робастность) за счет использования в рассуждениях исходных теоретических принципов при отсутствии эвристических правил. Кроме того, поиск на основе модели здесь дополняется эвристическим поиском, за счет этого появляется возможность решения из числа альтернативных.

В ГЭС, основанных на моделях и опыте, увеличиваются возможности объяснения ситуаций; повышается производительность решения задач за счет проверки аналогичных случаев до начала более экстенсивного поиска посредством рассуждений на основе модели; обеспечивается внесение примеров и исключений в БД случаев, которые могут быть использованы для управления выводом на основе модели; имеется возможность записи результатов вывода на основе моделей для будущего применения.

Применение и создание экспертных систем для решения задач моделирования наиболее оправдано в следующих случаях:

- 1) у пользователей отсутствует достаточный опыт в разработке моделей систем;
- 2) проблемная область является хорошо структурированной, определены необходимая терминология и методология построения моделей;
- 3) имеются эксперты, работающие в соответствующей области, способные взаимодействовать между собой и умеющие четко выражать свои мысли, поделиться своими знаниями с пользователями;
- 4) проблема моделирования имеет приемлемые размеры и границы, например, четко определен класс используемых моделей;
- 5) проблема не может быть решена традиционными хорошо известными методами, реализуемыми в пакетах прикладных программ;
- 6) стоимость и усилия по разработке экспертной системы (временные затраты исчисляются человеко-годами) оправдываются важностью и необходимостью решения задач моделирования. Например, экспертная система обеспечит значительную экономию материальных и временных затрат, повысит безопасность человеческой жизни.

Компьютерное моделирование ХТС к настоящему времени полностью доказало свою актуальность и перспективность. С его помощью удастся повысить эффективность и качество функционирования производства и управления технологическими процессами. Но особенно большое значение компьютерное моделирование и оптимизация имеют для проектирования работоспособных и экономичных ХТС, а также для сокращения сроков их проектирования. Можно выделить два этапа в развитии компьютерного моделирования ХТС, и сейчас мы находимся в преддверии третьего. Этот третий этап отличается от первых двух, в первую очередь тем, что при моделировании учитывается неполнота исходной информации, и в постановку задачи включается требование обеспечения работоспособности (гибкости) проектируемой ХТС.

ХТС, как правило, имеет рециклы, ее структура является замкнутой. В связи с этим расчет материальных и тепловых балансов сводится к решению системы нелинейных уравнений и представляет из себя сложную итерационную процедуру, обычно весьма трудоемкую ввиду нелинейности уравнений математической модели.

Первый этап компьютерного моделирования был связан с переводом расчета материальных и тепловых балансов ХТС с ручного на компьютерный. Начало этого этапа датируется 1958 г. и знаменуется появлением первой моделирующей системы Flexible Flowsheet. Его становление и развитие шло на протяжении 1960 – 1970 гг. Была выработана общая концепция универсальной моделирующей программы, состоящей из организующей программы, библиотеки модулей для расчета химико-технологических аппаратов, банка физико-химических свойств и библиотеки математических модулей. Было создано несколько десятков универсальных моделирующих программ: CHEOPS, CHEVRON, SPEED-UP, MACSIM, NETWORK 67, CHESS, PACER 245, FLOWTRAN, FLOWPACK, PROCESS и др.

Настоящий расцвет компьютерного моделирования начался с появлением персональных компьютеров. К этому времени в результате длительного процесса из общего числа выделились четыре универсальные моделирующие программы, которые заняли лидирующее положение в мире: ASPEN PLUS, HYSIS, CHEMCAD и PRO/II. Указанные программы обладают обширными библиотеками технологических модулей, большими банками физико-химических свойств и удобным для пользователя интерфейсом. Их широко используют при перепрофилировании действующих и проектировании новых ХТС. При этом можно отметить как общую тенденцию стремление переходить на использование в расчетах все более сложных и, соответственно, более адекватных математических моделей технологических аппаратов.

В целом в мировой практике в 1960 – 1980 гг. был накоплен большой опыт использования универсальных моделирующих программ для расчета стационарных режимов (а также, в меньшей степени, и динамических режимов) ХТС. Но на определенном этапе совершенствования универсальных моделирующих программ было осознано, что наибольшие возможности компьютерного моделирования ХТС связаны не с передачей компьютеру традиционных для проектировщиков функций

по расчету материальных и тепловых балансов, а с реализацией с помощью компьютера функций оптимизации и управления действующими ХТС, а также оптимального проектирования новых ХТС.

Начало второго этапа в развитии компьютерного моделирования можно условно отнести ко второй половине 1980-х гг., когда в течение короткого времени произошел переход к персональным компьютерам и появились первые прототипы универсальных моделирующих программ ASPEN PLUS, HYSIS, CHEMCAD и PRO/II, в которые были введены оптимизационные процедуры, и их стали применять не только для расчета отдельных технологических схем, но и для оптимизации стационарных режимов ХТС. Однако, вплоть до настоящего времени универсальные моделирующие программы гораздо чаще применяют для расчета материальных и тепловых балансов с использованием наиболее полных и совершенных модулей для расчета аппаратов и банка физико-химических свойств, снабженного новейшими данными. Причина здесь – и в значительно большей математической трудности оптимизационного расчета по сравнению с балансовым, и в непривычности целевой функции оптимизации для проектировщиков. Но главное и принципиальное затруднение, на наш взгляд, связано с неопределенностью исходной информации, которой мы располагаем при решении задачи оптимизации. Неопределенность практически всегда имеет место на этапе проектирования и часто – на этапе эксплуатации ХТС. Учет неопределенности информации требует как разработки новых математических постановок задач, так и новых подходов и методов их решения.

Разберем этот вопрос более подробно. Ограничившись далее задачей оптимизации стационарных режимов ХТС на этапе проектирования, запишем систему уравнений модели ХТС в общем виде:

$$f_i(a, d, z, x) = 0, \quad i = 1, \dots, p, \quad (1.1)$$

где функции  $f_i(a, d, z, x)$  представляют стационарные модели отдельных аппаратов ХТС и соотношений связи между ними;  $a$  – тип аппаратного оформления технологического процесса;  $d$  – вектор конструктивных переменных;  $z$  – вектор режимных (управляющих) переменных;  $x$  – вектор состояний (вектор концентраций, расходов потоков, их энтальпий и т.п.). Обычно из уравнений (1.1)  $x$  определяется как однозначная функция  $a, d, z: x = x(a, d, z)$ .

Далее при проектировании должен быть соблюден ряд требований-ограничений в форме равенств и неравенств:

$$\bar{g}_j(a, d, z, x) = 0 \quad \text{и/или} \quad \bar{g}_j(a, d, z, x) \leq 0, \quad j = 1, \dots, m. \quad (1.2)$$

Ограничения могут быть технологическими, экономическими, экологическими и регламентными.

Наконец при постановке задачи оптимизации должна быть задана целевая функция, зависящая от переменных ХТС:  $\bar{I}(a, d, z, x)$  и подлежащая минимизации или максимизации. Целевая функция может быть технологическим или чаще экономическим критерием (приведенные затраты, прибыль и т.д.).

Если вектор состояний  $x$  выразить из системы (1.1) в виде  $x(a, d, z)$  и подставить в зависимости  $f(a, d, z, x)$  и  $\bar{g}(a, d, z, x)$ , то математически задачу оптимизации ХТС (для случая минимизации критерия) можно записать в виде:

$$\min_{a \in A, d \in D, z \in Z} I(a, d, z) \quad (1.3)$$

$$g_j(a, d, z) = 0 \quad \text{и/или} \quad g_j(a, d, z) \leq 0, \quad j = 1, \dots, m, \quad (1.4)$$

где  $I(a, d, z) = \bar{I}(a, d, z, x(a, d, z))$ ,  $g_j(a, d, z) = \bar{g}_j(a, d, z, x(a, d, z))$ .

В научной литературе задачу (1.3), (1.4) принято называть задачей нелинейного программирования (НЛП). Решив задачу (1.3), (1.4), получим оптимальные значения  $a, d, z$ , обеспечивающие минимум критерия  $I$  при соблюдении ограничений (1.4).

Так, однако, дело обстоит в идеальном случае. В реальности на этапе проектирования в математическом описании ХТС всегда присутствуют неопределенности двух родов. Одни из них, такие как параметры сырья и температура окружающей среды, могут изменяться во время работы ХТС, оставаясь в пределах некоторого диапазона изменений. Для них принципиально невозможно указать единственное значение. Другие могут быть в реальности постоянными для данной ХТС, но их значения известны с точностью до определенного интервала, например некоторые коэффициенты в кинетических уравнениях и уравнениях тепло- и массопереноса. Чтобы учесть неопределенности в математическом описании ХТС, достаточно их выделить в зависимостях для  $I$  и  $g_j$ , считая, что

$$I = I(a, d, z, \xi), \quad g_j = g_j(a, d, z, \xi), \quad j = 1, \dots, m,$$

где  $\xi$  – вектор неопределенных параметров, принимающих любые значения из заданной области  $\Xi$ , которую обычно считают прямоугольной:

$$\Xi = \{\xi: \xi^L \leq \xi \leq \xi^U\}. \quad (1.5)$$

Таким образом, решение задачи (1.3), (1.4) на самом деле зависит от значения, которое принял вектор  $\xi$ , и само оказывается неопределенным.

Традиционный путь преодоления данного затруднения состоит в следующем. Вектору неопределенных параметров приписывают некое «номинальное» значение:  $\xi = \xi^N$  и решают задачу (1.3), (1.4) при номинальном  $\xi^N$  с получением номинального значения вектора конструктивных переменных  $d^N$  при заданном типе аппаратного оформления  $a$ . После этого волевым образом (на основе имеющихся знаний о проектируемом объекте и интуиции) вводят так называемые «запасы»  $k_i (k_i > 1)$  и принимают при проектировании

$$d_i = k_i d_i^N,$$

где  $d_i$  –  $i$ -я компонента вектора  $d$ ,  $i = 1, \dots, n$  (длина и диаметр реактора, поверхность теплообмена в теплообменнике, число тарелок в ректификационной колонне и т.п.).

Недостатки данного подхода очевидны, так как он не гарантирует ни оптимальности полученного решения, ни того, что все ограничения будут выполнены во время эксплуатации ХТС. Если запасы окажутся недостаточными, то ограничения будут нарушены, если слишком большими, то будет перерасход затрат.

Существенно более правильным и научно обоснованным является подход, когда неопределенность в параметрах процесса учитывается в самой постановке оптимизационной задачи. Суть данного подхода заключается в следующем. Вместо ограничений (1.4) вводят единое ограничение

$$\chi(a, d) = \max_{\xi \in \Xi} \min_{z \in Z} \max_{j \in J} g_j(a, d, z, \xi) \leq 0, \quad (1.6)$$

где  $J = \{1, \dots, m\}$  – множество индексов для функций-ограничений.

Это ограничение называют ограничением гибкости, а функцию  $\chi(a, d)$  – функцией гибкости. Если ХТС с вектором  $d$ , получившим определенное значение удовлетворяет ограничению (1.6), то ХТС называют работоспособной (гибкой). Гибкая ХТС сохраняет работоспособность при любых значениях  $\xi$  из области неопределенности  $\Xi$ . В этом случае в качестве критерия оптимизации применяют некоторую усредняющую величину. В точной формулировке это будет математическое ожидание  $I$  по переменной  $\xi$  в  $\Xi$ , что приводит к необходимости чрезвычайно сложных вычислений многомерного интеграла.

Используя прием дискретизации задачи и аппроксимации многомерного интеграла взвешенной суммой с небольшим числом членов (высокой точности вычисления интеграла здесь не требуется) можно записать:

$$\sum_{i=1}^s w_i I(a, d, z^i, \xi^i), \quad (1.7)$$

где  $\xi^i$  – «аппроксимационные» точки (представительные точки области  $\Xi$ , участвующие в операции усреднения критерия);  $s$  – число аппроксимационных точек;  $w_i$  – весовые коэффициенты, значения которых определяют теоретически (если известны вероятностные распределения неопределенных параметров  $\xi$ ), либо принимают субъективно с учетом знаний о процессе

$$w_i > 0, \sum_{i=1}^s w_i = 1.$$

В результате получаем следующую задачу:

$$\min_{a \in A, d \in D, z^i \in Z} \sum_{i=1}^s w_i I(a, d, z^i, \xi^i) \quad (1.8)$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad i=1, \dots, s, \quad j=1, \dots, m, \quad (1.9)$$

$$\chi(a, d) \leq 0. \quad (1.10)$$

Постановка задачи (1.8) – (1.10) наиболее часто используется при аппаратно-технологическом оформлении ХТП. С некоторых позиций эту постановку можно трактовать как оптимальный выбор запасов технологического оборудования, обеспечивающих работоспособность (гибкость) ХТС независимо от изменения неопределенных параметров  $\xi$  в заданной области  $\Xi$ .

Задачами анализа гибкости ХТС принято называть задачи, которые так или иначе связаны с учетом требования выполнения условия гибкости (1.6). К настоящему времени анализ гибкости оформился как целое научное направление. Задачи анализа гибкости ХТС математически значительно сложнее, чем обычные оптимизационные задачи (1.3), (1.4), и данное научное направление пока что не достигло уровня, при котором возможна массовая передача его результатов в производство. Ведущие универсальные программы – ASPEN PLUS, CHEMCAD и др. – не содержат программных средств для анализа гибкости ХТС.

Задача оптимизации с интервальной неопределенностью – основная задача анализа гибкости. Постановка (1.8) – (1.10) является для нее сейчас наиболее распространенной. Ее формулировка в свое время означала «прорыв», но теперь уже ясно, что эта постановка не лишена недостатков, она нуждается в совершенствовании и уточнении.

Наши исследования показали, что для правильной постановки задачи надо учитывать целый ряд моментов, важнейшими из которых являются следующие: вид неопределенности (модельная или параметрическая), тип неопределенности (неопределенность только на первом этапе, этапе проектирования, или на обоих этапах (проектирование и управление действующим процессом) и тип ограничений (жесткие или мягкие). Модельная неопределенность означает, что отсутствует полная уверенность в том, какой математической моделью следует воспользоваться для некоторых аппаратов ХТС (сложности могут возникнуть, например, при выборе для реактора с псевдооживленным слоем), а параметрическая – что общая структура уравнений для каждого аппарата известна, но есть неопределенность в параметрах (коэффициентах) модели. Параметрическая неопределенность на этапе проектирования практически неизбежна, а на втором этапе (управление ХТС) тоже не всегда удается избежать неопределенности. В первом приближении параметры модели ХТС можно разбить на две группы: *первая* – параметры можно непосредственно измерить на действующем процессе, *вторая* – параметры измерить нельзя, но их значения могут быть уточнены по данным измерения параметров первой группы. При этом чтобы уточнить средние значения параметров второй группы и интервалы неопределенности для них, в свою очередь, приходится решать оптимизационную задачу идентификации.

Если ограничение задачи должно быть выполнено всегда, при любых значениях  $\xi$  из заданной области  $\Xi$ , то такие ограничения называют *жесткими*. Жесткими должны быть ограничения по взрывобезопасности процесса и (часто) регламентные ограничения. Но в целом ряде случаев достаточно, чтобы ограничение выполнялось не всегда, а «почти всегда», т.е. с некоторой заданной большой вероятностью (например, с вероятностью 0,95). Такие ограничения принято называть *мягкими*. Ослабление требований к ограничениям позволяет в ряде случаев получить лучшее значение критерия оптимизации.

В настоящее время проблемам оценки гибкости ХТС уделяется большое внимание, потому что без учета неопределенности в большинстве случаев невозможна реальная, практически полезная оптимизация. Реализация третьего этапа (на котором широкое распространение получит решение оптимизационных задач с учетом требования гибкости) выведет компьютерное моделирование ХТС на тот уровень, когда с его помощью можно будет проектировать значительно более экономичные и надежные ХТС и более эффективно управлять работой действующих.

#### 1.4. МЕТОДОЛОГИЯ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ ХТС

Изучая сложные химико-технологические процессы, аппараты и физико-химические явления, мы не можем учесть все факторы: какие-то оказываются существенными, а какими-то можно пренебречь. При этом выдвигается система допущений (гипотез), которая тщательно обосновывается и позволяет выявить и учесть при математическом описании наиболее характерные черты исследуемой ХТС.

Схема, отражающая основные виды работ при построении ММ, компьютерном моделировании и оптимизации ХТС, включает последовательность определенных шагов (рис. 1.5).

На первом шаге производится описание проблемной области, формулируются цели и задачи, для решения которых требуются ММ. Мы рассматриваем, в основном, задачи оптимизации конструкций аппаратов или режимов его работы, оптимального управления и проектирования ХТС.

На втором шаге составляется полный перечень требований к ММ. По существу, этот перечень представляет собой техническое задание на разработку модели и включает сведения о входных и выходных переменных, области их изменения, допустимых погрешностях, требованиях к быстродействию и т.п.

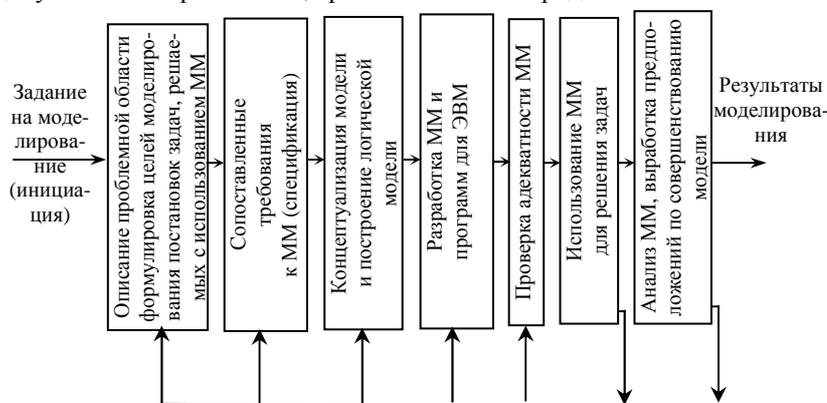


Рис. 1.5. Основные этапы построения математической модели

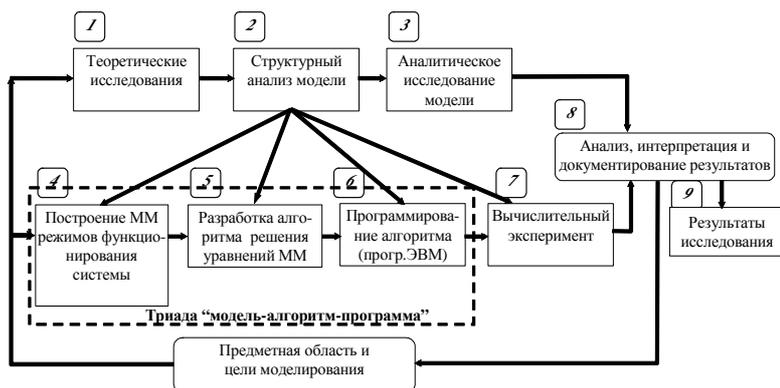
На основе выполнения первых двух шагов разрабатывается концепция модели (третий шаг), дается ее вербальное описание, строится схема, учитывающая логические связи между переменными, делаются предположения о методах построения модели, формировании команды для выполнения работ по компьютерному моделированию.

Четвертый шаг является основным по трудоемкости. В зависимости от применяемого метода построения ММ этот шаг разбивается на несколько стадий: планирование эксперимента, проведение эксперимента, обработка результатов эксперимента и др. Следует заметить, что на данном шаге наряду с системным аналитиком к работе привлекаются программисты.

На пятом шаге осуществляется проверка адекватности разработанной ММ в виде компьютерной программы. ММ тестируется во всем диапазоне изменения переменных. По результатам тестирования и сопоставления с экспериментальными данными делается вывод об адекватности модели или необходимости внесения изменений, т.е. коррекции ранее выполненных этапов.

Окончательная проверка о пригодности модели делается на шестом (завершающем) шаге на основе использования ММ для решения задач, сформулированных на первом шаге.

В процессе компьютерного моделирования и оптимизации исследователь имеет дело с тремя объектами: системой (реальной, проектируемой, воображаемой), математической моделью и программами ЭВМ, реализующими алгоритмы решения уравнений модели и оптимизации. Традиционная схема компьютерного моделирования и оптимизации как единого процесса построения и исследования модели, имеющего соответствующую программную поддержку, может быть представлена в следующем виде (рис. 1.6).



**Рис. 1.6. Схема организации процесса компьютерного моделирования:**

*1 – 9* – блоки процесса; ММ – математическая модель

Исходя из того, что компьютерное моделирование применяется для исследования, оптимизации и проектирования ХТС, можно выделить следующие этапы этого процесса:

- 1) определение объекта – установление границ, ограничений и измерителей эффективности функционирования объекта;
- 2) формализация объекта (построение модели) – переход от реального объекта к некоторой логической схеме (абстрагирование);
- 3) подготовка данных, необходимых для построения и применения модели, и представление их в соответствующей форме;
- 4) разработка моделирующего алгоритма и программы ЭВМ;
- 5) оценка адекватности – повышение до приемлемого уровня степени уверенности, с которой можно судить относительно корректности выводов о реальном объекте, полученных на основании обращения к модели;
- 6) стратегическое планирование – планирование вычислительного эксперимента, который должен дать необходимую информацию;
- 7) тактическое планирование – определение способа проведения каждой серии испытаний, предусмотренных планом вычислительного эксперимента;
- 8) вычислительный эксперимент – процесс осуществления имитации на модели с целью получения желаемых данных о реальном объекте и анализа его чувствительности;
- 9) интерпретация – построение выводов по данным, полученным путем имитации на модели;
- 10) реализация – практическое использование модели и результатов моделирования для решения конкретной задачи;
- 11) документирование – регистрация хода осуществления вычислительного эксперимента и его результатов, а также документирование процесса создания и использования модели.

Перечисленные этапы создания и использования модели определены в предположении, что задача может быть решена наилучшим образом с помощью компьютерного моделирования. Однако, это может быть не самый эффективный способ. В том случае, если задача может быть сведена к аналитическому решению уравнений математической модели, нет никакой нужды в компьютерном моделировании и имитации. Следует изыскивать все возможные средства, подходящие для решения данной конкретной задачи, стремясь при этом к оптимальному сочетанию стоимости и желаемых результатов. Прежде чем приступить к оценке возможностей численной имитации, следует убедиться, что аналитическое решение для данного случая невозможно.

В представленной на рис. 1.6 схеме организации процесса компьютерного моделирования (имитации) основная цепочка (ХТС – математическая модель – моделирующий алгоритм – программа ЭВМ – вычислительный эксперимент) соответствует традиционной схеме, но во главу угла теперь ставятся следующие моменты: понятие триады «*модель–алгоритм–программа*» (блоки 4 – 6), стратегическое и тактическое планирование вычислительного эксперимента (блок 7), интерпретация и документирование его результатов (блок 8).

Можно выделить следующие укрупненные процедуры моделирования и оптимизации ХТС: описание предметной области и построение концептуальной модели; постановка задачи оптимизации; выбор или разработка методов и вычислительных алгоритмов решения задач компьютерного моделирования (исследования) и оптимизации.

Основным назначением первой процедуры является переход от содержательного описания ХТС к ее математической модели. Содержательное описание является исходным материалом для построения формализованной схемы процесса функционирования системы и полной математической модели этого процесса. При построении модели целесообразно придерживаться блочного принципа с выделением основных блоков, описывающих процессы в самой ХТС и имитирующих воздействия внешней среды, и вспомогательных блоков, обеспечивающих компьютерную реализацию и обработку результатов моделирования. Разработка концептуальной модели предполагает четкую формулировку задачи исследования, выбор методики решения задачи с учетом имеющихся ресурсов, анализ масштаба задачи и возможности декомпозиции ее на подзадачи.

Вторая процедура связана с разработкой методов расчета сформулированных математических задач или, как говорят, вычислительных или *моделирующих алгоритмов*. Фактически он представляет собой совокупности алгебраических формул, по которым ведутся вычисления, и логических условий, позволяющих установить нужную последовательность применения этих формул. Вычислительные алгоритмы не должны искажать основные свойства модели и, следовательно, исходной ХТС, быть экономичными и адаптирующимися к особенностям решаемых задач и используемых компьютеров. Как правило, для одной и той же математической модели можно предложить множество вычислительных алгоритмов. Однако, требуется построение эффективных вычислительных алгоритмов, которые позволяют получить решение поставленной задачи с заданной точностью за минимальное количество действий (арифметических, логических), т.е. с минимальными затратами машинного времени. Эти вопросы весьма существенны и составляют предмет теории численных методов.

*Вычислительный эксперимент* имеет "многовариантный" характер. Действительно, решение любой прикладной задачи зависит от многочисленных входных переменных и параметров. Например, если рассчитывается технологическая установка, то имеется множество различных режимных переменных и конструктивных параметров, среди которых нужно определить их оптимальный набор, обеспечивающий эффективное функционирование этой установки. Получить решение соответствующей математической задачи в виде формулы, содержащей явную зависимость от режимных переменных и конструктивных параметров, для реальных задач, как говорилось выше, не удастся. При проведении вычислительного эксперимента каждый конкретный расчет проводится при фиксированных значениях переменных и параметров. Проектируя оптимальную ХТС, т.е., определяя в пространстве переменных и параметров точку, соответствующую оптимальному режиму, приходится проводить большое число расчетов однотипных вариантов задачи, отличающихся значениями некоторых переменных или параметров, поэтому очень важно опираться на эффективные численные методы.

*Вычислительный эксперимент* — это собственно проведение расчетов на ЭВМ и получение информации, представляющей интерес для исследователя. Точность этой информации определяется достоверностью, прежде всего модели, моделирующего алгоритма, программы ЭВМ и исходных данных. Именно по этой причине в серьезных прикладных исследованиях полномасштабным вычислениям предшествует период проведения тестовых расчетов. Они необходимы не только для того, чтобы «отладить» программу, т.е. отыскать и исправить все ошибки и опечатки, допущенные как при создании алгоритма, так и при его программной реализации. В этих предварительных расчетах тестируется также сама математическая модель, выясняется ее адекватность исследуемому объекту. Для этого проводится расчет некоторых контрольных экспериментов, по которым имеются достаточно надежные измерения. Сопоставление опытных данных с результатами расчетов позволяет уточнить математическую модель, обрести уверенность в правильности предсказаний, которые будут получены с ее помощью.

Только после проведения длительной работы в вычислительном эксперименте наступает фаза прогноза (имитации) — с помощью компьютерной модели предсказывается поведение исследуемого объекта в условиях, где натурные эксперименты пока не проводились или где они вообще невозможны.

Важное место в вычислительном эксперименте занимают обработка результатов расчетов, их всесторонний анализ и, наконец, выводы. Выводы бывают в основном двух типов: или становится ясна необходимость уточнения модели, или результаты, пройдя проверку, передаются заказчику. При оптимизации или проектировании ХТС из-за сложности и высокой размерности математической модели проведение расчетов по описанной выше схеме может оказаться чересчур дорогим. Тогда идут на упрощение модели, на построение своего рода инженерных методик (формул), опирающихся на сложные модели и расчеты и дающих возможность получить необходимую информацию значительно более дешевым способом. При этом проводится огромная предварительная работа по анализу сложных моделей, квинтэссенцией которой и являются простые на первый взгляд формулы.

Удобной формой представления логической структуры моделей процессов функционирования ХТС и моделирующих алгоритмов являются схемы. При этом используются:

- *обобщенная (укрупненная) схема*, которая задает общий порядок действий при моделировании и оптимизации ХТС;
- *детальная схема*, она содержит уточнения, отсутствующие в обобщенной схеме, и показывает, что и как следует выполнить на очередном шаге моделирования и оптимизации системы;
- *логическая схема* представляет собой логическую структуру модели процесса функционирования системы, она показывает упорядоченную во времени последовательность логических операций при моделировании.

Третья процедура — создание программы для реализации разработанного моделирующего алгоритма на ЭВМ (создание компьютерной модели). Применение большинства языков высокого уровня порождает ряд проблем, из которых главными являются трудоемкость и недостаточная гибкость. В процессе исследования реальных систем часто приходится уточнять модели, что влечет за собой перепрограммирование моделирующего алгоритма. Ясно, что процесс моделирования в этом случае не будет эффективным, если не обеспечить его гибкости. Для этой цели можно использовать формальные схемы, описывающие классы математических моделей из определенной предметной области, поскольку программировать тогда нужно функционирование данной схемы, а не описываемые ею частные модели.

*Схема программы ЭВМ* представляет собой интерпретацию логической схемы моделирующего алгоритма и отображает порядок программной реализации моделирующего алгоритма с использованием конкретного математического обеспечения. Логическая схема алгоритма и схема программы могут быть выполнены как в укрупненной, так и в детализированной формах. В результате разработки схемы модели, ее алгоритмизации и программирования, т.е. создания триады «модель–алгоритм–программа», исследователь получает в руки универсальный, гибкий и сравнительно недорогой инструмент, который вначале отлаживается, тестируется в «пробных» вычислительных экспериментах. После проверки адекватности триады исходной ХТС с моделью можно проводить разнообразные «опыты», которые позволяют оценить требуемые свойства и характеристики системы. Процесс компьютерного моделирования сопровождается улучшением и уточнением, по мере необходимости, всех звеньев триады.

Чтобы эффективно проанализировать выходные данные, полученные в результате расчетов на ЭВМ, необходимо знать, что делать с результатами рабочих расчетов и как их надо интерпретировать. Эта задача может быть решена на основании предварительного анализа результатов выполнения первых двух процедур моделирования ХТС. Планирование вычислительного эксперимента с моделью позволяет предварительно оценить необходимое количество выходных данных и определить метод их анализа. При этом следует полнее использовать возможности ЭВМ с точки зрения обработки результатов моделирования и представления этих результатов в наиболее наглядном виде, например, в виде таблиц, графиков, диаграмм, схем, средств мультимедиа и т.п.

При массовом использовании методов компьютерного моделирования в технических проектах следует добиваться резкого сокращения сроков разработки моделей, обеспечивающих различные этапы проектирования. Решение этой задачи возможно при соответствующем уровне развития технологии компьютерного моделирования, которая является основой целенаправленной деятельности в обеспечении возможности эффективного выполнения на ЭВМ исследований функционирования сложных систем. С ее помощью организуются действия исследователя на всех этапах его работы с моделями, начиная от изучения предметной области, выделения моделируемой проблемной ситуации и кончая построением и реализацией компьютерных экспериментов для анализа поведения системы.

При решении задач моделирования следует учитывать два важных аспекта:

1) методологическую составляющую технологии как науки, занимающейся выявлением закономерностей, применение которых на практике позволяет находить наиболее эффективные и экономичные приемы компьютерного моделирования ХТС на ЭВМ;

2) прикладные цели и задачи технологии как искусства, мастерства, умения достигать в ходе компьютерного моделирования сложных ХТС практически полезных результатов.

## 1.5. ПОСТАНОВКИ ЗАДАЧ МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ ХТС В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

При анализе и синтезе ХТС должны быть удовлетворены регламентные и проектные ограничения, связанные: 1) с *безопасностью ХТС* (например, температура в каталитическом реакторе должна быть меньше некоторой заданной величины или концентрации некоторых веществ не должны превышать максимально допустимых значений); 2) с *экологической безопасностью* (ограничения на максимальную величину выходных потоков вредных веществ) 3) с *обеспечением заданных производительности, качественных и технико-экономических показателей* выпускаемой продукции и ХТС, соответственно.

Удовлетворение регламентных ограничений осложняется наличием неопределенности (неточности) в математических моделях или в исходных данных задач моделирования и оптимизации. Источниками неопределенности, как правило, являются:

1. Первоначальная неточность математических моделей, используемых для целей анализа и проектирования ХТС. Она порождается: а) неточностью эксперимента, с помощью которого были получены коэффициенты в математических моделях (константы скоростей реакций, коэффициенты межфазного обмена, тепло- и массопереноса и т.д.); б) неточностью химических и физических закономерностей, положенных в основу математических моделей.

2. Изменение внутренних факторов ХТС на этапе функционирования, что приводит к изменению некоторых коэффициентов в математических моделях во время эксплуатации ХТС. Так, изменение активности катализатора приводит к изменению констант скорости реакций, а загрязнение поверхности теплообмена в теплообменнике – к изменению коэффициента теплоотдачи и, соответственно, теплопередачи.

3. Случайное изменение внешних факторов функционирования ХТС во время ее эксплуатации.

4. Конструктивная неточность, т.е. неточность в реализации некоторых размеров оборудования при его изготовлении.

Обычно неполнота наших знаний о ХТС сводится к тому, что некоторые параметры в математических моделях и исходных данных при решении задач моделирования и оптимизации известны неточно. О них известно только, что они принадлежат некоторой области неопределенности  $\Xi$ .

Таким образом, при моделировании и проектировании ХТС мы вынуждены использовать неточные математические модели и поэтому возникает вопрос: как мы можем гарантировать выполнение всех регламентных и проектных ограничений на этапе функционирования ХТС, несмотря на использование неточных математических моделей?

Задача статической оптимизации ХТС при проектировании формулируется при следующих предположениях:

1) в жизненном цикле ХТС можно выделить несколько стадий: стадию проектирования, стадию функционирования и стадию утилизации;

2) имеются регламентные и проектные требования, связанные с экономикой производства, взрывобезопасностью, экологией, качеством выпускаемой продукции, которые записываются в форме (1.4) или (1.9);

3) имеются два типа переменных – конструктивные переменные  $a, d$  (тип аппарата, размеры оборудования, переменные, задающие структуру ХТС) и управляющие переменные  $z$  (температура, давление, расход и др.). На стадии функционирования конструктивные переменные остаются постоянными, а управляющие переменные, вообще говоря, могут изменяться. Поэтому на стадии проектирования можно использовать возможность настройки управляющих переменных для удовлетворения регламентных и проектных ограничений.

На стадии функционирования ХТС можно выделить три группы неопределенных параметров. К первой группе относятся параметры, значения которых могут быть определены (измерены) достаточно точно. Другими словами, на стадии функционирования имеется достаточное число экспериментальных данных, позволяющих определить «точные» значения этих неопределенных параметров. Ко второй группе относятся параметры, которые не могут быть скорректированы на стадии функционирования. Другими словами, область неопределенности для этих параметров остается такой же, как и на стадии проектирования. К третьей группе относятся параметры, значения которых могут быть уточнены на стадии функционирования, однако при этом некоторая ошибка при определении этих параметров остается.

При определении области неопределенности на стадии проектирования мы будем различать два случая. В первом случае мы не знаем плотности распределения вероятностей неопределенных параметров. В этом случае интервалы неопределенности измеряемых параметров могут быть найдены, если известны максимальные ошибки измерения используемых приборов. Существенно труднее проблема отыскания области неопределенности неизмеряемых переменных. Рассмотрим подробнее этот вопрос. Пусть мы провели  $N$  экспериментов. Обозначим через  $[x^j, \xi^j]$  измеренные значения  $[x, \xi]$  в  $j$ -м эксперименте. Пусть

$$x^j = \bar{x}^j + \delta x^j; \quad \xi^j = \bar{\xi}^j + \delta \xi^j,$$

где  $[\bar{x}^j, \bar{\xi}^j]$  – неизвестные точные значения величин  $[x^j, \xi^j]$ , а  $[\delta x^j, \delta \xi^j]$  – ошибки измерения. Из характеристик приборов мы знаем максимальные значения ошибок  $\delta x^j, \delta \xi^j$ :

$$|\delta x^j| \leq \delta_1, \quad |\delta \xi^j| \leq \delta_2.$$

Наиболее целесообразно предполагать в качестве области неопределенности наименьший многомерный прямоугольник, включающий область  $\Omega = \{ \xi(\delta x^1, \dots, \delta x^N, \delta \xi^1, \dots, \delta \xi^N) : |\delta x^j| \leq \delta_1, |\delta \xi^j| \leq \delta_2 \}$ . Для этого мы должны найти  $n_\xi$  пар  $[\xi_i^L, \xi_i^U]$  решением  $2n_\xi$  следующих задач:

$$\xi_i^L = \min_{\delta x^j, \delta \xi^j} \xi(\delta \xi^1, \dots, \delta \xi^N; \delta x^1, \dots, \delta x^N),$$

$$\xi_i^U = \max_{\delta x^j, \delta \xi^j} \xi(\delta \xi^1, \dots, \delta \xi^N; \delta x^1, \dots, \delta x^N),$$

$$|\delta x^j| \leq \delta_1, |\delta \xi^j| \leq \delta_2.$$

Эти задачи относятся к классу двухуровневых задач оптимизации, которые требуют использования методов недифференцируемой оптимизации.

Пусть теперь известны плотности распределения вероятностей неопределенных параметров. Рассмотрим вначале случай, когда все параметры  $\xi_j$  независимы и каждый из них имеет плотность распределения вероятности  $\rho_j(\xi_j)$ . Тогда для каждого параметра  $\xi_j$  можно найти интервал  $\Xi_j^{\alpha_j}$ , удовлетворяющий условию

$$\Pr[\xi_j \in \Xi_j^{\alpha_j}] = \alpha_j,$$

где  $\Pr[\xi_j \in \Xi_j^{\alpha_j}]$  – вероятность принадлежности параметра  $\xi_j$  интервалу  $\Xi_j^{\alpha_j}$ . Это условие может быть записано в виде

$$\int_{\Xi_j^{\alpha_j}} \rho_j(\xi_j) d\xi_j = \alpha_j.$$

В этом случае область неопределенности есть  $n_\xi$  – мерный прямоугольник  $\Xi^\alpha$  со сторонами  $\Xi_j^{\alpha_j}$ ; вероятность попадания  $\xi$  в прямоугольник  $\Xi^\alpha$  равна  $\alpha = \alpha_1 \alpha_2 \dots \alpha_{n_\xi}$ . В случае нормального распределения

$$\rho_i(\xi_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{(\xi_i - \mu_i)^2}{2\sigma_i^2}\right],$$

где  $\mu_i = E\{\xi_i\}$  – среднее значение (математическое ожидание) параметра  $\xi_i$ ;  $\sigma_i$  – среднеквадратичное отклонение.

В этом случае интервал  $\Xi_j^{\alpha_j}$  имеет вид

$$\Xi_j^{\alpha_j} = [\xi: \mu_j - k_j \sigma_j \leq \xi_j \leq \mu_j + k_j \sigma_j, j=1, \dots, n_\xi].$$

При формулировании задач оптимизации в условиях неопределенности мы будем рассматривать два случая:

а) одноэтапная формулировка, в которой неопределенные параметры  $\xi$  (или их часть) не могут быть идентифицированы на стадии функционирования ХТС, и в этом случае управляющие переменные определяются одновременно с определением конструктивных переменных для всей области неопределенности  $\Xi$ ;

б) двухэтапная формулировка, в которой неопределенные параметры (или их часть) могут быть идентифицированы на стадии функционирования ХТС, и в этом случае управляющие переменные могут быть использованы для выполнения ограничений на стадии функционирования ХТС.

Ограничения в задаче оптимизации могут быть жесткими, если они должны безусловно выполняться на стадии функционирования ХТС для любых значений  $\xi$ . Нарушение этих условий может привести к аварии, нанести вред окружающей среде и т.д. Мягкие ограничения могут выполняться с некоторой заданной вероятностью или в среднем.

Мы будем называть ХТС *гибкой*, а соответствующую конструкцию *допустимой*, если на стадии функционирования мы можем удовлетворить все ограничения (жесткие и мягкие) при условии, что неопределенные параметры могут принимать любые значения из области неопределенности  $\Xi$ .

При формулировании задачи оптимизации в условиях неопределенности необходимо сформулировать целевую функцию и ограничения. В качестве целевой функции мы будем использовать некоторую оценку будущей работы ХТС, а в качестве ограничений будут использоваться условия, гарантирующие гибкость ХТС на стадии функционирования. В качестве оценки будущей работы ХТС на стадии функционирования мы будем использовать одну из следующих величин:

- 1) среднее значение, которое может принять критерий оптимизации на стадии функционирования;
- 2) наихудшее значение критерия оптимизации, которое он может принять на стадии функционирования (стратегия наихудшего случая). Решив задачу оптимизации с этим критерием, мы сможем гарантировать, что значение критерия на этапе функционирования не может быть хуже величины, полученной из решения задачи оптимизации;
- 3) верхняя граница для критерия, которая не может быть нарушена с заданной вероятностью.

**Одноэтапная задача оптимизации** [6].

Рассмотрим одноэтапную задачу оптимизации с жесткими ограничениями. Поскольку математическое ожидание  $E_\xi\{I(a, d, z, \xi)\}$  дает среднее значение критерия оптимизации на стадии функционирования ХТС, то естественно использовать эту величину как целевую функцию задачи оптимизации в условиях неопределенности. Объединяя эту целевую функцию и условие гибкости ХТС  $\max_{\xi \in \Xi} g_j(a, d, z, \xi) \leq 0, j=1, \dots, m$ , мы можем сформулировать одноэтапную задачу с жесткими ограничениями в условиях неопределенности:

$$\min_{a, d, z} E_\xi\{I(a, d, z, \xi)\}; \quad \max_{\xi \in \Xi} g_j(a, d, z, \xi) \leq 0, j=1, \dots, m.$$

Можно упростить задачу, переходя к дискретному аналогу критерия оптимизации:

$$\min_{a, d, z} \sum_{i \in I_1} w_i I(a, d, z, \xi^i); \quad \max_{\xi \in \Xi} g_j(a, d, z, \xi) \leq 0, j=1, \dots, m,$$

где  $w_i$  – весовые коэффициенты;  $\xi^i (i \in I_1)$  – аппроксимационные точки;  $I_1$  – множество индексов аппроксимационных точек.

Сформулируем одноэтапную задачу оптимизации с мягкими (вероятностными) ограничениями. Предположим, что мы имеем полную информацию относительно функции распределения вероятностей для  $\xi$ . В этом случае одноэтапная задача оптимизации имеет вид

$$\begin{aligned} \min_{a,d,z} F(a, d, z); \\ \Pr\{g_j(a, d, z, \xi) \leq 0\} = \int_{\Omega_j} P(\xi) d\xi \geq \rho_j, j = 1, \dots, m; \\ \Omega_j = \{\xi : g_j(a, d, z, \xi) \leq 0, \xi \in \Xi\}, \end{aligned}$$

где  $F(\bullet)$  – некоторая целевая функция;  $P(\xi)$  – функция плотности вероятности. В качестве функции  $F(a, d, z)$  мы можем использовать среднее значение первоначальной целевой функции  $I(a, d, z, \xi)$  на этапе функционирования ХТС:  $F(a, d, z) = E_{\xi}\{I(a, d, z, \xi)\}$ .

Главная трудность решения сформулированной одноэтапной задачи оптимизации состоит в необходимости вычисления многомерных интегралов  $E_{\xi}\{I(a, d, z, \xi)\}$  и  $\int_{\Omega_j} P(\xi) d\xi \geq \rho_j$ . Рассмотрим теперь другую формулировку, в которой в качестве критерия будет использоваться его верхняя граница, которая не может быть нарушена с заданной вероятностью:

$$\begin{aligned} \min_{a,d,z,\alpha} \alpha \\ \Pr\{I(a, d, z, \xi) - \alpha \leq 0\} \geq \rho_0, \\ \Pr\{g_j(a, d, z, \xi) \leq 0\} \geq \rho_j, j = 1, \dots, m. \end{aligned}$$

Решение последней задачи позволяет найти конструкцию  $a^*, d^*$  и режим функционирования  $z^*$ , которые гарантируют, что в процессе эксплуатации ХТС целевая функция  $I^*(a^*, d^*, z^*, \xi)$  будет меньше, чем  $\alpha^*$  с вероятностью  $\rho_0$ . Используя ту же целевую функцию мы можем свести задачу с жесткими ограничениями к следующей задаче

$$\begin{aligned} \min_{a,d,z,\alpha} \alpha \\ \Pr\{I(a, d, z, \xi) - \alpha \leq 0\} \geq \rho_0, \\ \max_{\xi \in \Xi} g_j(a, d, z, \xi) \leq 0, j = 1, \dots, m. \end{aligned}$$

Предположим, что мы решили одноэтапную задачу и получили решение  $[a^*, d^*, z^*]$ . Чтобы реализовать это решение, мы должны поддерживать выполнение условий  $z = z^*$  на этапе функционирования ХТС. Это означает, что мы не можем настраивать управляющие переменные на стадии функционирования. Ясно, что использование одноэтапной задачи на стадии проектирования приводит к не вполне экономичным конструкциям аппаратов ХТС, так как не используется настройка управляющих переменных ХТС. С другой стороны, одноэтапная задача оптимизации соответствует случаю, когда на этапе функционирования невозможно уточнить некоторые параметры математической модели и, следовательно, невозможна подстройка управляющих переменных  $z$ .

#### Двухэтапная задача оптимизации [6].

Все формулировки двухэтапных задач оптимизации будут учитывать возможность идентификации неопределенных параметров и настройки управляющих переменных на стадии функционирования ХТС. В этом состоит принципиальная разница между двухэтапной задачей оптимизации и одноэтапной задачей оптимизации. В двухэтапной задаче переменные  $a, d$  по-прежнему постоянны на стадии функционирования ХТС, в то время как переменные  $z$  могут изменяться. В частности, это свойство позволяет настраивать управляющие переменные  $z$  для удовлетворения ограничений задачи.

При формулировке двухэтапной задачи оптимизации мы будем использовать следующее предположение – на стадии функционирования ХТС в каждый момент времени:

- выполняется уточнение всех или части неопределенных параметров на основе доступной экспериментальной информации;
- решается задача оптимизации ХТС с использованием математической модели с уточненными неопределенными параметрами, и найденный оптимальный режим реализуется на функционирующей ХТС. Мы будем называть эту задачу *внутренней задачей оптимизации*.

#### Случай 1. На стадии функционирования ХТС известны точные значения всех неопределенных параметров.

Здесь мы предполагаем, что в каждый момент времени на стадии функционирования ХТС значения неопределенных параметров либо могут быть точно измерены, либо оценены с использованием доступной экспериментальной информации. Рассмотрим условие работоспособности (гибкости) ХТС: *ХТС является гибкой, если для каждого  $\xi \in \Xi$  можно найти такие значения управляющих переменных, что все ограничения задачи будут удовлетворены.*

Условие гибкости ХТС получено в работе [7] и имеет вид

$$\chi_1(a, d) = \max_{\xi \in \Xi} \min_z \max_{j \in J} g_j(a, d, z, \xi) \leq 0.$$

Это необходимое и достаточное условие гибкости ХТС. В работе [8] впервые было введено понятие индекса гибкости  $\gamma$ , определяемого следующим образом:

$$\begin{aligned}\gamma(a, d) &= \max \delta \\ \chi_1(a, d, \delta) &\leq 0, \\ \Xi(\delta) &= \{\xi: \xi^N - \delta \cdot \Delta\xi^- \leq \xi \leq \xi^N + \delta \Delta\xi^+\}, \\ \Delta\xi^- &= \xi^N - \xi^L, \quad \Delta\xi^+ = \xi^U - \xi^N.\end{aligned}$$

Здесь область  $\Xi(\delta)$  зависит от скалярной величины  $\delta$ . Для заданных значений  $a, d$  индекс гибкости  $\gamma$  определяет максимальный прямоугольник, для каждой точки которого можно найти такие значения управляющих переменных, что все ограничения будут удовлетворены. Если  $\gamma > 1$ , то конструктивные переменные  $a, d$  допустимы (ХТС является гибкой), в противном случае ХТС не является гибкой.

Сформулируем двухэтапную задачу двухэтапной оптимизации для случая 1. Выше мы предположили, что в каждый момент времени на стадии функционирования решается внутренняя задача оптимизации с использованием математической модели с уточненными параметрами  $\xi$ . В данном случае она имеет вид

$$\begin{aligned}I^*(a, d, \xi) &= \min_z I(a, d, z, \xi); \\ g_j(a, d, z, \xi) &\leq 0, \quad j = 1, \dots, m.\end{aligned}$$

Предположим, что функция плотности распределения вероятностей  $P(\xi)$  известна. Поскольку в каждый момент времени на стадии функционирования значение критерия оптимизации будет равно  $I^*(a, d, \xi)$ , то на этапе проектирования мы можем оценить будущую работу ХТС, подсчитав математическое ожидание  $E_\xi\{\dots\}$  величины  $I^*(a, d, \xi)$ :

$$E_\xi\{I^*(a, d, \xi)\} = \int_{\Xi} I^*(a, d, \xi) P(\xi) d\xi.$$

Эта величина будет использоваться как целевая функция в задаче оптимизации в условиях неопределенности.

Двухэтапная задача оптимизации есть задача стохастического программирования с рекурсией, которая формулируется следующим образом [7, 9]:

$$\min_{a, d} E_\xi\{I^*(a, d, \xi)\} = \min_{a, d} \int_{\Xi} \min_z \{I(a, d, z, \xi) | g_j(a, d, z, \xi) \leq 0\} P(\xi) d\xi.$$

Поскольку интеграл есть бесконечная сумма и переменные  $z$ , соответствующие различным  $\xi$ , независимы друг от друга, то мы можем изменить порядок операторов интегрирования и минимизации:

$$\begin{aligned}\min_{a, d} \min_{z(\xi)} \int_{\Xi} I(a, d, z(\xi), \xi) P(\xi) d\xi, \\ g_j(a, d, z(\xi), \xi) \leq 0, \quad j = 1, \dots, m, \quad \xi \in \Xi.\end{aligned}$$

Объединяя оба оператора минимизации по  $a, d$  и  $z(\xi)$ , мы получим

$$\begin{aligned}\min_{a, d, z(\xi)} \int_{\Xi} I(a, d, z(\xi), \xi) P(\xi) d\xi, \\ g_j(a, d, z(\xi), \xi) \leq 0, \quad j = 1, \dots, m, \quad \xi \in \Xi.\end{aligned}$$

Задача имеет бесконечное число ограничений и поисковых переменных (одна многомерная функция  $z(\xi)$  эквивалентна бесконечному числу обычных поисковых переменных). Заменяем многомерный интеграл некоторой конечной суммой с помощью той или иной квадратурной формулы [10]:

$$\int_{\Xi} I(a, d, z(\xi), \xi) d\xi = \sum_{i \in L} w_i I(a, d, z^i, \xi^i),$$

где  $\xi^i (i \in L)$  – аппроксимационные (узловые) точки;  $z^i = z(\xi^i)$  – вектор управляющих переменных, соответствующий аппроксимационной точке  $\xi^i$  и  $w_i$  – весовые коэффициенты, удовлетворяющие условиям  $w_i \geq 0, \sum_{i \in I_1} w_i = 1$ . Кроме того, заме-

ним бесконечное число ограничений конечным числом ограничений в аппроксимационных точках  $\xi^i (i \in L)$ . Таким образом, мы получим дискретный вариант двухэтапной задачи оптимизации для случая 1

$$\begin{aligned}I_1 = \min_{d, z^i} \sum_{i \in L} w_i I(a, d, z^i, \xi^i), \\ g_j(a, d, z^i, \xi^i) \leq 0, \quad j = 1, \dots, m, \quad i \in L,\end{aligned}$$

Пусть  $a^*, d^*$  – решение задачи. Однако, мы не можем гарантировать, что эта задача имеет решение для каждого набора  $a, d$  и  $\xi \in \Xi$ . Поэтому ее необходимо дополнить условием гибкости  $\chi_1(a, d) \leq 0$ .

К сожалению, часто функция плотности распределения вероятностей для точек  $\xi^i (i \in L)$  неизвестна, тогда множество аппроксимационных точек  $S_1 = \{\xi^i : i \in L\}$  и весовые коэффициенты  $w_i$  задаются из инженерных соображений. Например,

аппроксимационные точки должны выбираться таким образом, чтобы они попадали в область наиболее вероятных значений, которые параметры  $\xi$  могут принимать при функционировании ХТС.

Если аппроксимационные точки достаточно плотно покрывают область  $\Xi$ , то задача оптимизации в условиях неопределенности решается без ограничения  $\chi_1(a, d) \leq 0$ . Это требует большого числа аппроксимационных точек даже для сравнительно малой размерности  $n_\xi$  вектора  $\xi$ . Если число узловых точек по каждой компоненте вектора  $\xi$  равно  $p$ , то число аппроксимационных точек будет равно  $p^{n_\xi}$ . В этом случае размерность задачи будет равна  $n_a + n_d + p^{n_\xi} n_z$ . В случае, когда число аппроксимационных точек невелико, использование ограничения  $\chi_1(a, d) \leq 0$  совершенно необходимо, так как это гарантирует выполнение ограничений задачи не только в аппроксимационных точках, но и во всех точках области  $\Xi$ .

*Случай 2. На стадии функционирования ХТС известны точные значения только части неопределенных параметров.*

В этом случае вектор неопределенных параметров  $\xi$  состоит из двух подвекторов  $\xi^1$  и  $\xi^2$ . Пусть при этом  $\xi^1 \in \Xi^1$  и  $\xi^2 \in \Xi^2$  ( $\Xi = \Xi^1 \cup \Xi^2$ ). На стадии функционирования ХТС значения компонентов вектора  $\xi^1$  могут быть определены достаточно точно. В то же время компоненты вектора  $\xi^2$  не могут быть уточнены.

Условие гибкости для данного случая имеет вид [11]:

$$\chi_2(a, d) = \max_{\xi^1 \in \Xi^1} \min_z \max_{\xi^2 \in \Xi^2} \max_{j \in J} g_j(a, d, z, \xi^1, \xi^2) \leq 0.$$

Рассмотрим формулировку двухэтапной задачи оптимизации. Мы предполагаем, что векторы  $\xi^1$  и  $\xi^2$  независимы. Для фиксированного момента времени на стадии функционирования ХТС значение  $\xi^1$  известно, а  $\xi^2$  может принимать любые значения из области  $\Xi^2$ . Формулировка внутренней задачи оптимизации имеет вид:

$$\begin{aligned} I^*(a, d, \xi^1) &= \min_z E_{\xi^2} \{I(a, d, z, \xi^1, \xi^2)\} \\ \max g_j(a, d, z, \xi^1, \xi^2) &\leq 0, \quad j = 1, \dots, m, \\ E_{\xi^2} \{I(a, d, z, \xi^1, \xi^2)\} &= \int_{\Xi^2} I(a, d, z, \xi^1, \xi^2) P_2(\xi^2) d\xi^2, \quad \xi^2 \in \Xi^2, \end{aligned}$$

где  $E_{\xi^2} \{I(a, d, z, \xi^1, \xi^2)\}$  – математическое ожидание функции  $I(a, d, z, \xi^1, \xi^2)$  по переменной  $\xi^2$  при фиксированном  $\xi^1 \in \Xi^1$ . Условие  $\chi_2(a, d) \leq 0$  гарантирует существование решения внутренней задачи оптимизации для всех  $\xi^1 \in \Xi^1$ . Следовательно, оно должно быть использовано в постановке двухэтапной задачи оптимизации как ограничение. Так как величина

$$E_{\xi^1} \{I^*(a, d, \xi^1)\} = \int_{\Xi^1} I^*(a, d, \xi^1) P_1(\xi^1) d\xi^1$$

характеризует будущую работу ХТС, то она может служить в качестве целевой функции в двухэтапной задаче оптимизации для случая 2. Сформулируем эту задачу в виде:

$$\begin{aligned} I_2 &= \min_{a, d} E_{\xi^1} \{I^*(a, d, \xi^1)\} \\ \chi_2(a, d) &\leq 0. \end{aligned}$$

Используя квадратурную для аппроксимации интеграла в целевой функции задачи, получим следующую задачу:

$$\begin{aligned} I_2 &= \min_{a, d, z^i} \sum_{i \in L} w_i \sum_{k \in K} v_k I(a, d, z^i, \xi^{1,i}, \xi^{2,k}), \\ \max_{\xi^2 \in \Xi^2} g_j(a, d, z^i, \xi^{1,i}, \xi^2) &\leq 0, \quad j = 1, \dots, m, \quad i \in L, \\ \chi_2(a, d) &\leq 0, \end{aligned}$$

где  $\xi^{1,i}, \xi^{2,k}$  – аппроксимационные точки для внешнего и внутреннего интегралов в критерии оптимизации,  $z^i = z(\xi^{1,i})$ ,  $w_i (i \in L)$  и  $v_k (k \in K)$  – весовые коэффициенты, удовлетворяющие условию

$$w_i \geq 0; \quad \sum_{i \in L} w_i = 1; \quad v_k \geq 0; \quad \sum_{k \in K} v_k = 1.$$

Если функция плотности распределения неизвестна, весовые коэффициенты  $w_i (i \in L)$  и  $v_k (k \in K)$ , а также множества аппроксимационных точек  $S_1 = \{\xi^{1,i} : i \in L\}$ ,  $S_2 = \{\xi^{2,k} : k \in K\}$  должен назначать исследователь.

Следует заметить, что вектор  $z^i$ , соответствующий точке  $\xi^{1,i}$ , используется во всех аппроксимационных точках  $\xi^{1,i}, \xi^{2,j}$  из области

$$\Xi(\xi^{1,i}) = \{\xi : \xi^{2,k} \in \Xi^2; \xi^1 = \xi^{1,i}\}.$$

Можно показать, что справедливо следующее неравенство:

$$I_2 \geq I_1.$$

Сравнивая задачи двухэтапной оптимизации для случая 1 и для случая 2, мы получаем, что полнота экспериментальной информации на стадии функционирования влияет на конструкции (экономичность) аппаратов ХТС, и учет доступной информации позволяет строить более экономичную ХТС.

*Случай 3.* На этапе функционирования ХТС можно определить точные значения неопределенных параметров, все ограничения являются мягкими и должны быть удовлетворены с вероятностью  $\rho$ . Для данного случая двухэтапная задача имеет следующий вид:

$$\begin{aligned} \min_{a,d,z(\xi)} \int_{\Xi} I(a, d, z(\xi), \xi) P(\xi) d\xi, \\ g_j(d, z(\xi), \xi) \leq 0, \quad j = 1, \dots, m, \quad \xi \in \Omega_1, \\ \int_{\Omega_1} P(\xi) d\xi \geq \rho, \quad \Omega_1 = \{\xi: g_j(a, d, z(\xi), \xi) \leq 0, \quad j = 1, \dots, m\}. \end{aligned}$$

Введем штрафную функцию

$$\bar{I}(a, d, \xi, \gamma) = I(a, d, z, \xi) + \gamma \sum_{j=1}^m p(g_j(a, d, z, \xi)),$$

$$\text{где } p(g_j) = \begin{cases} 0, & \text{если } g_j \leq 0; \\ g_j^2, & \text{если } g_j > 0. \end{cases}$$

Здесь  $\gamma$  – штрафной коэффициент.

Далее мы будем использовать следующую внутреннюю задачу оптимизации для всех значений  $\xi$ :

$$\bar{I}^*(a, d, \xi, \gamma) = \min_z \bar{I}(a, d, \xi, \gamma).$$

Ясно, что оптимальное значение критерия этой задачи зависит от  $a, d, \xi, \gamma$ . Пусть  $z(a, d, \xi, \gamma)$  – решение этой задачи. Переменные  $a, d, \gamma$  должны быть выбраны таким образом, чтобы ограничения удовлетворялись с вероятностью  $\rho$  и среднее значение величины  $\bar{I}^*(a, d, \xi, \gamma)$  принимало минимальное значение. Двухэтапная задача оптимизации будет иметь вид:

$$\begin{aligned} F^* = \min_{a,d,\gamma} E_{\xi} \{ \bar{I}^*(a, d, \xi, \gamma) \}, \\ \int_{\Omega_2} P(\xi) d\xi \geq \rho, \\ \Omega_2 = \{ \xi: g_j(a, d, z(a, d, \xi, \gamma), \xi) \leq 0 \}. \end{aligned}$$

Заметим, что неравенство гарантирует удовлетворение ограничений с вероятностью не меньшей  $\rho$ . В определении области  $\Omega_2$  используются значения  $z(a, d, \xi, \gamma)$  управляющих переменных, полученных из решения внутренней задачи оптимизации. В задаче штрафной коэффициент  $\gamma$  используется как дополнительная поисковая переменная.

*Случай 4.* Имеются жесткие и мягкие ограничения. Пусть ограничения с номерами  $j \in J_1 = \{1, \dots, m_1\}$  являются жесткими ограничениями, а с номерами  $j \in J_2 = \{m_1 + 1, \dots, m\}$  – мягкими ограничениями. Ограничения с номерами  $j \in J_2 = \{m_1 + 1, \dots, m\}$  должны удовлетворяться с вероятностью  $\rho$ .

В этом случае двухэтапная задача оптимизации будет иметь вид

$$\begin{aligned} \min_{a,d} F(a, d) = F_1(a, d) + F_2(a, d), \\ \Pr[\xi \in \Omega_1] \geq \rho, \\ \chi_1(a, d; J_1) = \max_{\xi \in \Xi} \min_z \max_{j \in J_1} g_j(a, d, z, \xi) \leq 0, \end{aligned}$$

где

$$\begin{aligned} F_1(a, d) &= \int_{\Omega_1} I^*(a, d, \xi) P(\xi) d\xi, \\ F_2(a, d) &= \int_{\Xi \setminus \Omega_1} I^*(a, d, \xi; J_1) P(\xi) d\xi; \end{aligned}$$

$I^*(a, d, \xi)$  – решение задачи

$$\begin{aligned} I^*(a, d, \xi) &= \min_z I(a, d, z, \xi), \\ g_j(a, d, z, \xi) &\leq 0, \quad j = 1, \dots, m; \end{aligned}$$

и  $I^*(a, d, \xi; J_1)$  – решение задачи

$$I^*(a, d, \xi; J_1) = \min_z I(a, d, z, \xi),$$

$$g_j(a, d, z, \xi) \leq 0, \quad j \in J_1.$$

Заметим, что первая задача решается только для  $\xi \in \Omega_1$ , а вторая задача – в остающейся части области  $\Xi$ . Ограничение  $\Pr[\xi \in \Omega_1] \geq \rho$  гарантирует удовлетворение всех ограничений с вероятностью  $\rho$ . Кроме того, дополнительное ограничение  $\chi_1(a, d; J_1) = \max_{\xi \in \Xi} \min_z \max_{j \in J_1} g_j(a, d, z, \xi) \leq 0$  гарантирует жесткое удовлетворение ограничений с номерами  $j = 1, \dots, m_1$  для всех точек области  $\Xi$ .

Анализ постановок задач оптимизации в условиях неопределенности показывает, что вычисление функции гибкости  $\chi$  является очень важной задачей. Для оценки гибкости ХТС достаточно знать знак функции гибкости.

Рассмотрим вначале свойства функции гибкости  $\chi_1(d)$ . Вычисление функции гибкости  $\chi_1(d)$  сводится к нахождению максимума функции  $h(d, \xi) = \min_z \max_{j \in J} g_j(d, z, \xi)$  в области  $\Xi$ . Функция  $h(d, \xi)$  может быть представлена в виде

$$h(d, \xi) = \min_{z, \alpha} \alpha,$$

$$g_j(d, z, \xi) \leq \alpha, \quad j \in J = (1, \dots, m).$$

Пусть имеется следующая система ограничений:

$$g_1 = -z + \xi \leq 0,$$

$$g_2 = z - 2\xi + 2 - d \leq 0,$$

$$g_3 = -z + 6\xi - 9d \leq 0.$$

Пусть  $\Xi = \{1 \leq \xi \leq 3\}$ . Найдем вид функции  $\chi_1(d)$ . В данном случае задача принимает вид

$$h(d, \xi) = \min_{z, \alpha} \alpha,$$

$$-z + \xi \leq \alpha,$$

$$z - 2\xi + 2 - d \leq \alpha,$$

$$-z + 6\xi - 9d \leq \alpha.$$

Обозначим прямую  $-z + \xi = \alpha$  через  $a$ , прямую  $z - 2\xi + 2 - d = \alpha$  через  $b$  и прямую  $-z + 6\xi - 9d = \alpha$  через  $c$ , точку пересечения прямых  $a$  и  $b$  – буквой  $A$  (рис. 1.7).

Вначале найдем значение  $\chi_1(d)$  при  $d = 1$ . В этом случае задача принимает следующий вид

$$h(1, \xi) = \min_{z, \alpha} \alpha,$$

$$-z + \xi \leq \alpha,$$

$$z - 2\xi + 1 \leq \alpha,$$

$$-z + 6\xi - 9 \leq \alpha.$$

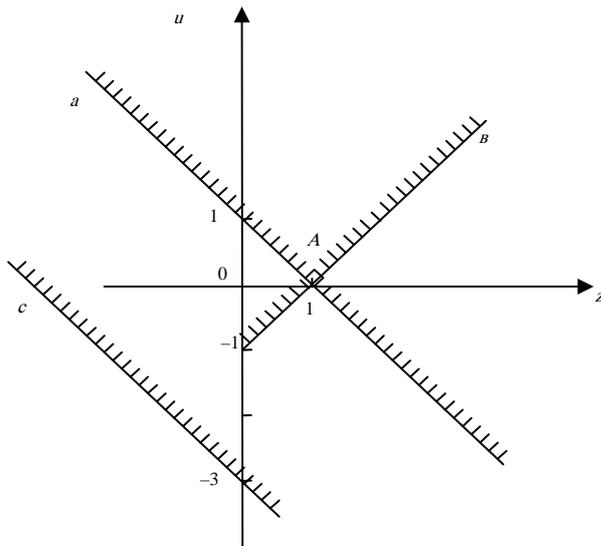


Рис. 1.7. Геометрическая интерпретация задачи

Для построения функции  $h(1, \xi)$  необходимо решить задачу линейного программирования с двумя переменными  $\{z, \alpha\}$  для каждого значения значения  $\xi$  в интервале  $[1; 3]$ . При  $\xi = 1$  задача принимает вид

$$\begin{aligned}
h(1, 1) &= \min_{z, \alpha} \alpha, \\
-z + 1 &\leq \alpha, \\
z - 1 &\leq \alpha, \\
-z - 3 &\leq \alpha.
\end{aligned}$$

Допустимая область, определяемая первым неравенством в задаче, находится выше прямой  $a$  (заштрихованная область на рис. 1.7); допустимая область, определяемая вторым неравенством, находится выше прямой  $b$ , и допустимая область, определяемая третьим неравенством, находится выше прямой  $c$ . Легко видеть, что точка  $A(1, 0)$  пересечения прямых  $a$  и  $b$  соответствует решению последней задачи, поскольку в этой точке переменная  $\alpha$  принимает наименьшее значение в допустимой области  $h(1, 1) = 0$ . Таким образом, в точке решения активными являются первое и второе ограничения.

Найдем вид функции  $h(1, \xi)$  для  $\xi > 1$ . Ясно, что прямые  $a$  и  $b$  будут активными, пока точка  $A$  их пересечения будет выше прямой  $c$ . В этом случае координаты точки  $A$  будут решением исходной задачи. Они могут быть получены решением уравнения  $-z + \xi = z - 2\xi + 1$ .

Отсюда следует

$$z = \frac{3}{2}\xi - \frac{1}{2}, \quad \alpha = -\frac{1}{2}\xi + \frac{1}{2}.$$

Следовательно,

$$h(1, \xi) = -\frac{1}{2}\xi + \frac{1}{2}.$$

Это уравнение будет верно в некотором интервале  $1 \leq \xi \leq \bar{\xi}$ , пока прямые  $a$  и  $b$  являются активными. Следовательно, значение  $\bar{\xi}$  соответствует моменту, когда изменится множество активных ограничений. Это случится, когда точка  $A$  окажется на прямой  $c$ . В этом случае точка  $A$  будет находиться на пересечении трех прямых  $a$ ,  $b$  и  $c$ . В этой точке

$$\bar{\xi} = \frac{9}{5}, \quad \alpha = -\frac{2}{5}.$$

При  $\xi > \frac{9}{5}$  активными ограничениями будут прямые  $c$  и  $b$ . Координаты точки пересечения этих прямых будут найдены из следующего уравнения:

$$z - 2\xi + 1 = -z + 6\xi - 9.$$

Таким образом, мы получаем  $z = 4\xi - 5$ ,  $\alpha = 2\xi - 4$ . Следовательно,  $h(1, \xi) = 2\xi - 4$ .

Таким образом,

$$h(1, \xi) = \begin{cases} -\frac{1}{2}\xi + \frac{1}{2}, & \text{если } 1 \leq \xi \leq \frac{9}{5}; \\ 2\xi - 4, & \text{если } \xi > \frac{9}{5}. \end{cases}$$

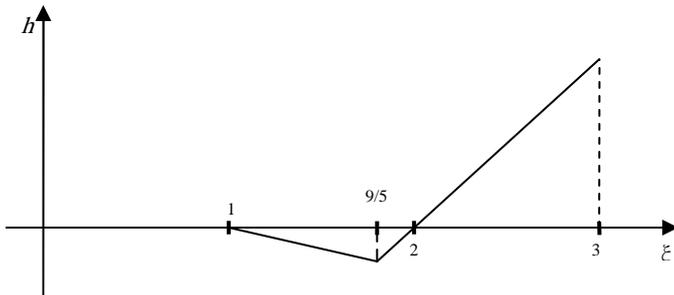


График функции  $h(1, \xi)$  приведен на рис. 1.8.

**Рис. 1.8. Функция гибкости  $h(1, \xi)$**

Найдем теперь вид функции  $h(d, \xi)$  при произвольном  $d$ . В этом случае мы будем использовать те же самые рассуждения, с помощью которых была получена формула для  $h(1, \xi)$ .

Прямая  $a$  лежит выше прямой  $c$ , если выполняется неравенство

$$-z + \xi > z + 6\xi - 9d, \quad \text{или } d > \frac{5}{9}\xi.$$

Пусть это условие выполняется. В этом случае прямые  $a$  и  $b$  будут активными в точке решения задачи. Координаты точки могут быть получены решением уравнения

$$-z + \xi = z - 2\xi + 2 - d.$$

Отсюда получаем

$$z = \frac{3}{2}\xi + \frac{1}{2}d - 1, \quad \alpha = -\frac{1}{2}\xi - \frac{1}{2}d + 1.$$

Следовательно,

$$h(d, \xi) = -\frac{1}{2}\xi - \frac{1}{2}d + 1.$$

Последнее уравнение будет верно в некотором интервале  $1 \leq \xi \leq \bar{\xi}$ , пока точка  $A$  пересечения этих прямых не окажется на прямой  $c$ . В этом случае точка  $A$  будет находиться на пересечении трех прямых  $a$ ,  $b$  и  $c$ . В этой точке координаты  $z$  и  $\alpha$  должны удовлетворять уравнению прямой  $c$ . Подставляя выражения для  $z$  и  $\alpha$  в уравнение прямой  $c$ , получим

$$\frac{3}{2}\bar{\xi} + \frac{1}{2}d - 1 + 6\bar{\xi} - 9d = -\frac{1}{2}\bar{\xi} - \frac{1}{2}d + 1,$$

или  $\bar{\xi} = \frac{9}{5}d$ .

Итак, при  $\xi \leq \frac{9}{5}d$  будут активными первое и второе ограничения исходной задачи, а при  $\xi \geq \frac{9}{5}d$  будут активными второе и третье ограничения задачи. В последнем случае координату точки пересечения прямых  $b$  и  $c$  находим из условия  $z - 2\xi + 2 - d = -z + 6\xi - 9d$ . Отсюда следует  $z = 4\xi - 4d - 1$ ,  $\alpha = 2\xi - 5d + 1$ .

Таким образом,

$$h(d, \xi) = \begin{cases} -\frac{1}{2}\xi - \frac{1}{2}d + 1, & \text{если } 1 \leq \xi \leq \frac{9}{5}d, \\ 2\xi - 5d + 1, & \text{если } \xi > \frac{9}{5}d. \end{cases}$$

Обозначим через  $\xi_{\max}$  значение  $\xi$  в точке максимума функции  $h(d, \xi)$ . Тогда

$$\xi_{\max} = \arg \max_{\xi \in \Xi} h(d, \xi) = \begin{cases} 1, & \text{если } 1 \leq \xi \leq \frac{9}{5}d, \\ 3, & \text{если } \xi \geq \frac{9}{5}d; \end{cases}$$

$$\max_{\xi \in \Xi} h(d, \xi) = \begin{cases} \frac{1}{2} - \frac{1}{2}d, & \text{если } 1 \leq \xi \leq \frac{9}{5}d, \\ 7 - 5d, & \text{если } \xi \geq \frac{9}{5}d. \end{cases}$$

Согласно последнему соотношению в точке  $\xi_{\max} = 3$  должно выполняться условие  $d \leq \frac{15}{9}$ .

Поскольку  $\chi_1(d) = \max_{\xi \in \Xi} h(d, \xi)$ , то  $\chi_1(d) = \max\left(\frac{1}{2} - \frac{1}{2}d, 7 - 5d\right)$ .

Точка  $\bar{d}$ , в которой происходит переход от выражения  $7 - 5d$  к выражению  $0,5 - 0,5d$ , определяется условием  $7 - 5\bar{d} = 0,5d - 0,5\bar{d}$ .

Отсюда  $\bar{d} = \frac{13}{9}$ . Следовательно,

$$\chi_1(d) = \begin{cases} 7 - 5d, & \text{если } d \leq \frac{13}{9}, \\ 0,5 - 0,5d, & \text{если } d \geq \frac{13}{9}. \end{cases}$$

Вид графика функции  $\chi_1(d)$  в интервале  $0 \leq d \leq 3$  приведен на рис. 1.9.

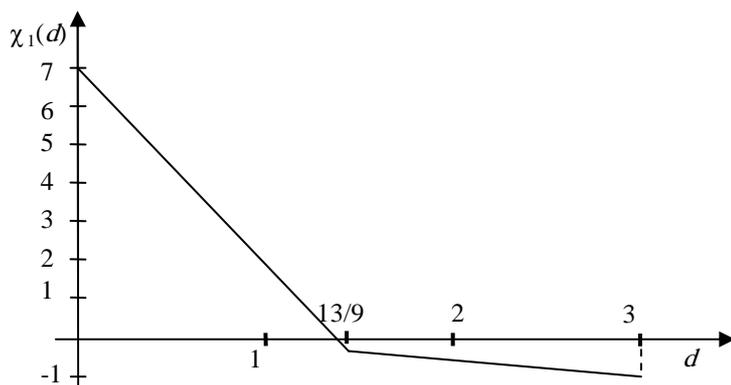


Рис. 1.9. Функция гибкости

Проанализируем вид функций  $h(1, \xi)$  и  $\chi_1(d)$ . Из рис. 1.8 видно, что функция  $h(1, \xi)$  имеет два локальных максимума в точках  $\xi = 1$  и  $\xi = 3$ . Локальный максимум в точке  $\xi = 3$  является глобальным. Поэтому  $\chi_1(d) = 2$ . Поскольку значение  $\chi_1(d)$  положительно, то конструкция  $d = 1$  недопустима для области неопределенности  $\Xi = \{1 \leq \xi \leq 3\}$ . С другой стороны, в каждой точке интервала  $1 \leq \xi \leq 2$  функция  $h(d, \xi)$  неположительна. Следовательно, конструкция  $d = 1$  допустима для области неопределенности  $\Xi = \{1 \leq \xi \leq 2\}$ .

Заметим, что функция  $h(1, \xi)$  не имеет производной в точке  $\xi = 9/5$ . Это точка, в которой изменяется множество активных ограничений. Рассмотрим теперь функцию  $\chi_1(d)$  (рис. 1.9).

Можно видеть, что эта функция не имеет производной в точке  $d = 13/9$ . Таким образом, из анализа задачи вытекает, что функции  $h(d, \xi)$  и  $\chi_1(d)$  являются, вообще говоря, недифференцируемыми функциями, а функция  $h(d, \xi)$  может быть многоэкстремальной. В связи с этим можно рекомендовать следующие методы вычисления функции  $\chi_1(d)$ : 1) метод перебора; 2) метод множества активных ограничений; 3) метод смешанного дискретно-непрерывного, нелинейного программирования; 4) метод ветвей и границ [6].

### Контрольные вопросы

1. Охарактеризуйте сущность системного подхода к моделированию систем.
2. Назовите принципы системного подхода, относящиеся к проблеме моделирования и оптимизации систем.
3. Дайте определение термина «система».
4. Назовите основные этапы моделирования систем.
5. Что понимается под видом моделирования систем?
6. Приведите примеры типовых математических схем моделирования?
7. Что понимается под термином «вычислительный эксперимент»?
8. В чем заключается проверка адекватности математической модели?
9. Рассмотреть задачу

$$\begin{aligned} \min_{z, d} d \\ g_1 = -z + \xi \leq 0, \\ g_2 = z - 2\xi + 2 - d \leq 0, \\ g_3 = -z + 6\xi - 9d \leq 0. \end{aligned}$$

Параметр  $\xi$  имеет однородное распределение:

$$p(\xi) = 0,5, \quad \Xi = \{\xi : 1 \leq \xi \leq 3\}.$$

Пусть имеются ограничения

$$0 \leq d \leq 3, \quad 0 \leq z \leq 5.$$

Требуется:

- а) сформулировать одноэтапную задачу оптимизации с жесткими ограничениями;
- б) сформулировать одноэтапную задачу оптимизации с мягкими ограничениями, при этом ограничения должны удовлетворяться в среднем;
- в) сформулировать одноэтапную задачу оптимизации с мягкими ограничениями, при этом ограничения должны удовлетворяться с вероятностью  $\alpha \geq 0,5$ .

12. Рассмотреть задачу

$$\begin{aligned} \min_{z, d} (d + 0,5z) \\ -d - z\xi + 3 \leq 0, \\ d \geq 0, \quad z \geq 0, \end{aligned}$$

где  $d, z, \xi$  – скаляры. Мы будем предполагать, что параметр  $\xi$  имеет однородное распределение:  $P(\xi) = 1, \quad \Xi = \{\xi : 0 \leq \xi \leq 1\}$ .

Требуется:

- а) сформулировать одноэтапную задачу с жесткими ограничениями;
- б) сформулировать одноэтапную задачу оптимизации с мягким ограничением, при этом мягкое ограничение должно удовлетворяться в среднем;
- в) сформулировать одноэтапную задачу оптимизации с мягким ограничением, при этом мягкое ограничение должно удовлетворяться с вероятностью  $\alpha \geq 0,5$ ;
- г) сформулировать двухэтапную задачу для случая, когда ограничение жесткое;
- д) сформулировать двухэтапную задачу для случая, когда ограничение мягкое, при этом мягкое ограничение должно удовлетворяться с вероятностью  $\alpha \geq 0,5$ .

## 2. МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ

### 2.1. КЛАССИФИКАЦИЯ МЕТОДОВ ПОСТРОЕНИЯ МОДЕЛЕЙ СИСТЕМ

В зависимости от методов получения соотношений связи между переменными и параметрами системы выделяют *формальные* и *неформальные* математические модели. Если структура оператора  $\Sigma$  модели может быть задана на основе некоторых формальных соображений, не имеющих связи с типом объекта, его конструкцией, механизмами протекающих процессов, то такие модели будем называть *формальными*. Задание структуры оператора  $\Sigma$  в формальных ММ производится, как правило, с учетом удобства последующего использования уравнений или простоты определения параметров модели по экспериментальным данным. Под удобством использования ММ понимается возможность получения аналитического или численного нахождения приближенного решения уравнений модели на ЭВМ.

Примерами формальных ММ служат уравнения статики объекта с  $m$  входными  $x_i$  и одной выходной  $y$  переменными:

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m a_{ii} x_i^2 + \sum_{i,j=1}^{m_1} a_{ij} x_i x_j + \dots,$$

или динамики объекта с одним входом

$$a_k y^{(k)}(t) + a_{k-1} y^{(k-1)}(t) + \dots + a_0 y(t) = x(t),$$

где  $m_1$  – число сочетаний из  $m$  по два;  $i \neq j$ ;  $y^{(k)}$  – производная переменной  $y$   $k$ -го порядка;  $a_i$  – параметры модели.

Структура ММ статики выбрана из условия удобства использования ММ в расчетах и сравнительной простоты определения параметров. Формальная ММ динамики объекта также удобна для применения, так как допускает построение аналитического решения  $y(x, a, t)$ . Кроме того, она позволяет анализировать поведение системы методами качественной теории динамических систем, преобразовывать дифференциальное уравнение в алгебраическое (например, с помощью преобразования Лапласа) и т.п.

Следует отметить, что формальные ММ применяют для описания стационарных и нестационарных объектов только с сосредоточенными координатами (переменными). При этом модели динамики всегда выбираются линейными, а уравнения статики задаются в таком виде, чтобы решение  $y(x, a)$  было линейным по  $a$ . Одинаковые (с точностью до параметра  $a$ ) формальные ММ могут описывать различные объекты (системы).

Для составления формальных ММ не требуется глубокого изучения систем, требуемая точность описания достигается увеличением размерности вектора  $a$ . Простота составления таких моделей способствует тому, что они широко применяются в тех задачах, в которых исследование статики и динамики носит массовый характер или производится многократно. Сюда относится исследование характеристик технологических аппаратов с целью построения систем автоматического управления, уточнение параметров ММ нестационарных объектов в процессе их функционирования и т.п.

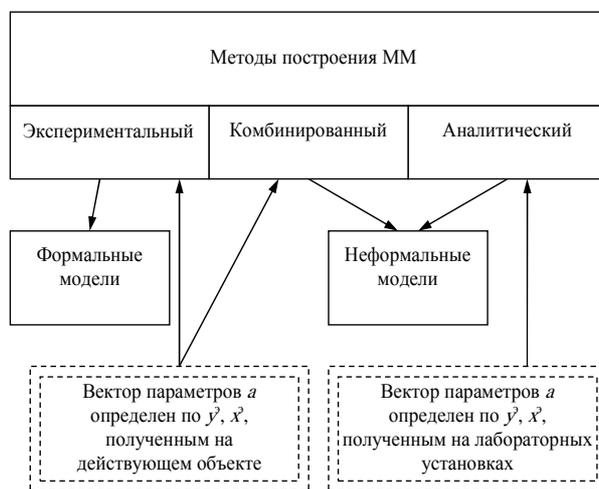
При построении *неформальных* математических моделей операторы  $\Sigma$  выводят на основе теоретического анализа физико-химических процессов, происходящих в системе. Так, например, при выводе уравнений ММ технологических аппаратов учитывают: гидродинамические режимы перемещения веществ, скорости химических реакций, диффузии, передачи тепла, хемосорбции и т.д., уравнения материального и энергетического (теплового) баланса, уравнения фазовых превращений и др. В качестве параметров в модель могут входить (в явной или косвенной форме) основные конструктивные размеры аппарата (поверхности теплообмена, диаметры и длины труб реакторов и т.п.). Чем детальнее и полнее неформальная ММ, тем сложнее структура оператора  $\Sigma$  и выше размерность вектора параметров  $a$ , компонентами которого являются параметры уравнений кинетики (константы скоростей, энергии активации, коэффициенты диффузии и т.п.) и характеристики веществ (теплоемкости, плотности и т.д.).

В процессе вывода уравнений неформальных ММ приходится принимать ряд допущений, например об учете или неучете некоторых физико-химических процессов, протекающих в аппарате. Вследствие этого составлению ММ предшествует трудоемкий этап экспериментального исследования процессов на лабораторных установках с целью определения уравнений кинетики и оценки значимости скоростей процессов. В зависимости от принимаемых допущений ММ одного и того же аппарата могут иметь существенно различный вид. Тем более могут различаться структуры оператора  $\Sigma$  неформальных ММ в тех задачах, в которых требуется массовое или многократное изучение статики и динамики ТП.

Так как неформальные ММ технологических систем содержат разнообразную и обширную информацию о конструкциях элементов, механизмах и скоростях протекающих в них физико-механических процессов, то это позволяет использовать неформальные ММ для оптимального конструирования аппаратов, оптимизации режимов их работы, оптимального управления и т.п.

В зависимости от способа построения моделей статики и динамики и определения вектора параметров  $a$  можно указать три метода построения ММ технологических систем (рис. 2.1): экспериментальный, аналитический и комбинированный.

При экспериментальном методе построения формальных ММ параметры определяются по опытным данным  $y^3, x^3$ , полученным на



**Рис. 2.1. Классификация методов построения математических моделей**

действующем объекте. Построенные этим методом ММ (будем называть их экспериментальными) справедливы только для того объекта, на котором проводились опыты.

Аналитический метод построения ММ заключается в теоретическом расчете или определении параметров неформальных уравнений статики и динамики по опытным данным  $y^3, x^3$ , которые получены при исследовании отдельных физико-химических процессов, происходящих в объекте, на лабораторных установках. В аналитических ММ параметры имеют четкую физическую трактовку и представляют собой самостоятельную ценность, так как могут быть использованы в других задачах. Поэтому к задаче определения вектора  $a$  параметров модели предъявляют следующие требования: единственности, устойчивости к ошибкам измерения и расчета, обеспечения адекватности ММ исследуемому объекту (системе).

Комбинированный (экспериментально-аналитический) метод построения ММ заключается в нахождении параметров неформальных уравнений статики и динамики по сигналам  $y^3, x^3$  полученным на действующем объекте. Модели, полученные таким методом, называют комбинированными. Параметры в таких ММ имеют физическую трактовку, поэтому к задаче определения вектора  $a$  предъявляют те же требования, что и при аналитическом методе.

Математические модели, построенные экспериментальным и комбинированным методами, используются для оптимизации статических режимов действующего объекта (системы) и расчета систем автоматического регулирования. Аналитические модели, кроме того, можно применять для оптимального проектирования технологических объектов (систем) и конструирования систем автоматического управления ими.

## 2.2. ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ МОДЕЛЕЙ СТАТИКИ

Экспериментальный метод построения моделей статики заключается в проведении на действующем объекте эксперимента, т.е. подаче входных сигналов  $x^3$ , записи реакций в виде выходных координат  $y^3$  и аппроксимации опытных данных  $y^3, x^3$  некоторой формальной математической зависимостью в виде оператора  $\Sigma$ . Структура  $\Sigma$  не зависит явно от свойств, осуществляемых в системе физико-химических процессов и выбирается из условий удобства определения вектора  $a$  и применения ММ (обычно структура оператора  $\Sigma$  задается линейной по  $a$ ).

В зависимости от способа задания  $x^3$  различают активные и пассивные экспериментальные методы. В активных методах экспериментатор сам создает испытательный сигнал  $x^3$  желаемой формы, в пассивных методах используются естественные случайные изменения входных и соответственно выходных переменных объекта (системы).

Очевидным достоинством пассивного метода получения  $y^3, x^3$  (случайных функций или случайных величин) является отсутствие нарушения естественного режима функционирования объекта, недостатком – невозможность создания испытательных сигналов желаемого вида, большая длительность «эксперимента» и иногда увеличение объема вычислений при определении параметров модели. Пассивные методы широко применяются для построения адаптивных ММ нестационарных объектов, а также при создании моделей крупномасштабных объектов (систем), искусственное возмущение режимов работы которых связано с большими экономическими потерями.

В большинстве случаев экспериментальный метод построения ММ базируется на трех допущениях: 1) объект есть система с сосредоточенными параметрами (переменными); 2) статические и динамические свойства системы неизменны во времени, т.е. ММ стационарна; 3) уравнения статики и динамики линеаризуемы в малом, т.е. при небольших отклонениях  $y$  от установившегося состояния выполняется принцип суперпозиции. Справедливость второго и третьего допущений проверяется экспериментальным путем.

Под статикой промышленного объекта (системы) обычно понимается совокупность его установившихся состояний, т.е. такие режимы работы, когда все входные и выходные переменные с определенной точностью постоянны во времени. Статической характеристикой системы называют функциональную зависимость между выходной и входными переменными в установившемся режиме. Определение зависимости  $y = f(x)$  является первым этапом при составлении математического описания системы. Статическая характеристика используется для выбора наиболее выгодного режима ее работы, проверки линейности системы, определения передаточного коэффициента и т.п.

При исследовании статики систем наиболее часто встречаются системы со следующими типами структурных схем (рис. 2.2):  $O_{1,1}$  – системы с одной входной  $x$  и одной выходной  $y$  переменными (*SISO-системы, Single Input Single Output*);  $O_{m,1}$  – системы с  $m$  независимыми входными и одной выходной переменными (*MISO-системы, Multi Input Single Output*);  $O_{m,n}$  – системы с  $m$  независимыми входными и  $n$  выходными переменными (*MIMO-системы, Multi Input Multi Output*).



$$\|r\|_E = \left( \sum_{i=0}^m r_i^2 \right)^{1/2} = \left[ \sum_{i=0}^m (\varphi(x_i) - y(i))^2 \right]^{1/2}$$

или

$$\|r\|_C = \max_{0 \leq i \leq m} |r_i| = \max_{0 \leq i \leq m} |\varphi(x_i) - y(i)|.$$

Задача о наилучшем приближении экспериментальных данных  $y(i)$  состоит в нахождении коэффициентов  $c_0, c_1, \dots, c_m$ , минимизирующих норму вектора  $r$ . В зависимости от выбора нормы получим решение задачи. Так, норма  $\|r\|_E$  соответствует задаче о наилучшем среднеквадратичном приближении, а норма  $\|r\|_C$  – задаче о наилучшем равномерном приближении экспериментальных данных.

*Пример 2.1.* Построим наилучшее среднеквадратичное приближение для случая линейного приближения  $n = 1, m = 2$ , когда заданы  $f(x_i) = y(i), i = 0, 1, 2$ . Обозначим  $h_0 = h_0 = x_1 - x_0, h_1 = x_2 - x_1$  и будем искать обобщенный многочлен  $\varphi(x)$  в виде  $\varphi(x) = c_0 + c_1(x - x_1)$ .

Тогда для  $r(x) = \varphi(x) - y$  получим, что  $\|r\|_E^2 = \Phi(c_0, c_1)$ , где

$$\Phi(c_0, c_1) = (c_0 - c_1 h_0 - y(0))^2 + (c_0 - y(1))^2 + (c_0 + c_1 h_1 - y(2))^2.$$

Коэффициенты  $c_0, c_1$  определяются из условия

$$\min_{c_0, c_1} \Phi(c_0, c_1). \quad (2.1)$$

Метод определения коэффициентов обобщенного многочлена из условия (2.1) называют *методом наименьших квадратов*.

Точку минимума  $\Phi(c_0, c_1)$  найдем из условия (в данном случае необходимого и достаточного)

$$\frac{\partial \Phi}{\partial c_0} = 0, \quad \frac{\partial \Phi}{\partial c_1} = 0,$$

которое приводит к системе линейных алгебраических уравнений:

$$\begin{aligned} 3c_0 + (h_1 - h_0)c_1 &= y(0) + y(1) + y(2), \\ (h_1 - h_0)c_0 + (h_0^2 + h_1^2)c_1 &= h_1 y(2) - h_0 y(0). \end{aligned}$$

Отсюда получаем:

$$\begin{aligned} c_0 &= \alpha_0 y(0) + (1 - \alpha_0 - \alpha_2) y(1) + \alpha_2 y(2), \\ c_1 &= \beta \frac{y(2) - y(1)}{h_1} + (1 - \beta) \frac{y(1) - y(0)}{h_0}, \end{aligned} \quad (2.2)$$

$$\text{где } \alpha_0 = \frac{h_1(h_0 + h_1)}{2(h_0^2 + h_1^2 + h_1 h_0)}, \quad \alpha_2 = \frac{h_0(h_0 + h_1)}{2(h_0^2 + h_1^2 + h_1 h_0)} \text{ и } \beta = \frac{h_1(2h_1 + h_0)}{2(h_0^2 + h_1^2 + h_1 h_0)}.$$

Если  $h_0 = h_1 = h$ , то

$$c_0 = \frac{1}{3}(y(0) + y(1) + y(2)), \quad c_1 = \frac{y(2) - y(0)}{2h}. \quad (2.3)$$

Погрешность полученного приближения на равномерной сетке имеет второй порядок по  $h$ , т.е.

$$\|f - \varphi\| = \frac{h^2}{\sqrt{6}} |f''(\xi)|,$$

где  $\xi \in (x_0, x_2)$ .

*Многомерные задачи.* Решение многомерных задач часто сводится к следующему. В некоторой области пространства  $X$  заданы точки  $x_1, x_2, \dots, x_n$  и значения функции  $f(x)$  в этих точках. Требуется получить приближение к значению функции  $f(x)$ .

Простейшим способом решения этой задачи является способ неопределенных коэффициентов. Пусть из каких-то соображений нам известно, что функция  $f(x)$  хорошо приближается линейными комбинациями:

$$\sum_{i=1}^n c_i \varphi_i(x).$$

Потребуем, чтобы такая линейная комбинация совпадала с  $f(x) = y$  в заданных узлах:

$$\sum_{i=1}^n c_i \varphi_i(x_j) = y(j), j = 1, \dots, n.$$

Предположим, что для матрицы  $\|\varphi_i(x_j)\| \det\|\varphi_i(x_j)\| \neq 0$ . Тогда матрица  $\|\varphi_i(x_j)\|$  имеет обратную  $A = \|a_{ij}\|$  и решение записывается в виде:

$$c_i = \sum_{j=1}^n a_{ij} y(j).$$

Функция  $g(x) = \sum_{i=1}^n c_i \varphi_i(x)$  совпадает с  $f(x)$  в точках  $x_j$ . Подставляя  $c_i$  в последнее выражение, получим

$$g(x) = \sum_{i=1}^n z_i(x) y(i),$$

где  $z_i(x) = \sum_{j=1}^n a_{ij} \varphi_j(x)$ .

Такая форма записи интерполяционной функции является аналогом записи интерполяционного многочлена в форме Лагранжа.

Повышение качества приближения может достигаться различными способами, например методом наименьших квадратов. В этом случае приближающая функция ищется в виде

$$g(x) = \sum_{i=1}^m c_i \varphi_i(x), \quad (2.4)$$

где  $m \ll n$ .

Коэффициенты  $c_i$  будем определять из условия:

$$\min_{c_1, c_2, \dots, c_m} \Phi(c_1, c_2, \dots, c_m),$$

где  $\Phi(g) = \Phi(c) = \sum_{j=1}^n \gamma_j (g(x_j) - y(j))^2 = \sum_{j=1}^n \gamma_j \left( \sum_{i=1}^m c_i \varphi_i(x_j) - y(j) \right)^2$ ;  $\gamma_j$  – весовые множители.

В основе метода регуляризации лежат соображения о сглаживании аппроксимирующей функции. Наиболее распространенной формой метода регуляризации является следующая. Приближение отыскивается в виде (2.4), а коэффициенты  $c_i$  выбираются из условия минимума выражения:

$$\Phi(\beta, g) = \Phi(g) + \beta \psi(g), \quad (2.5)$$

где  $\beta$  – параметр регуляризации,  $\beta > 0$ .

Функционал  $\psi(g)$  подбирается из следующего условия: если значение этого функционала невелико, то функция  $g$  обладает определенной гладкостью. Например,  $\psi(g)$  может быть некоторым приближением к интегралу  $\int_x |\text{grad}(g(x))|^2 dx$ .

Пусть минимум  $\Phi(\beta, g)$  достигается при некоторых  $c_1^\beta, c_2^\beta, \dots, c_n^\beta$  и  $g^\beta(x) = \sum_{i=1}^n c_i^\beta \varphi_i(x)$ .

Рассмотрим крайние случаи  $\beta = 0$  и  $\beta \rightarrow \infty$ . Для  $\Phi(0, g)$  имеет место равенство

$$\Phi(0, g) = \sum_{j=1}^n \gamma_j \left( \sum_{i=1}^n c_i \varphi_i(x_j) - y(j) \right)^2. \quad (2.6)$$

Если  $\det \|\varphi_i(x_j)\| \neq 0$ , то система  $\sum_{i=1}^n c_i \varphi_i(x_j) = y(j)$ ,  $j = 1, 2, \dots, n$  имеет решение, и правая часть равенства (2.6) обращается

в нуль. Тогда  $g^0(x)$  совпадает с интерполяционным многочленом в узлах интерполяции  $x_j$ . При больших  $\beta$  в функционале (2.5) определяющим является второе слагаемое, нижняя грань которого достигается на гладкой функции. Следовательно, есть все основания утверждать, что при промежуточных значениях  $\beta$  функции  $g^\beta(x)$  будут гладкими и в то же время не очень сильно отличающимися от приближаемой функции в заданных узлах.

*Приближение функций с помощью нейронных сетей.* В последние годы появился новый алгоритмический аппарат приближения функций многих переменных с помощью линейных операций и суперпозиций функций одного переменного. Такое приближение осуществляется специальными формальными устройствами – *нейронными сетями*, состоящими из формальных *нейронов*.

Нейрон получает на входе вектор сигналов  $x$ , вычисляет его скалярное произведение на вектор весов  $\alpha$  и некоторую функцию одного переменного  $\varphi(z)$ , где  $z$  – скалярное произведение  $x$  на  $\alpha$ . Результат рассылается на входы других нейронов или передается на выход. Таким образом, нейронные сети вычисляют суперпозиции простых функций одного переменного и их линейных комбинаций.

Для описания алгоритмов и устройств в теории нейронных сетей выработана специальная схемотехника, в которой элементарные устройства – сумматоры, синапсы, нейроны и т.п. – объединяются в сети, предназначенные для решения различных задач. Наиболее важные элементы нейросистем – *адаптивный сумматор* и *нелинейный преобразователь*. Адаптивный сумматор вычисляет скалярное произведение входного сигнала  $x$  на вектор параметров  $\alpha$  (рис. 2.3).

Адаптивным его называют из-за наличия вектора настраиваемых параметров  $\alpha$ . Нелинейный преобразователь получает скалярный входной сигнал  $z$  и переводит его в  $\varphi(z)$  (рис. 2.4).

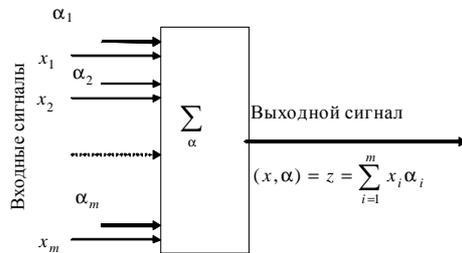


Рис. 2.3. Схема адаптивного сумматора

Стандартный формальный нейрон преобразователя и точки ветвления (рис. 2.5). Точка ветвления служит для рассылки скалярный входной сигнал  $z$  и можно выделить две базовые

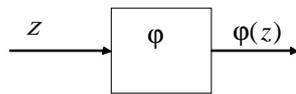


Рис. 2.4. Схема нелинейного преобразователя

составлен из входного сумматора, нелинейного преобразователя и точки ветвления (рис. 2.5). Она передает его выходам. Среди нейронных сетей архитектуры: *слоистые и полносвязные сети*.

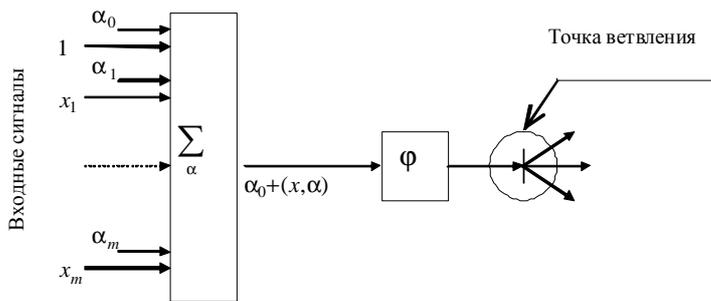


Рис. 2.5. Схема формального нейрона

В слоистых сетях нейроны расположены в несколько слоев (рис. 2.6).

Нейроны первого слоя получают входные сигналы, преобразуют их и через точки ветвления передают нейронам второго слоя. Далее срабатывает второй слой и так далее до  $k$ -го слоя, который выдает выходные сигналы для пользователя. Если не оговорено противное, то каждый выходной сигнал  $i$ -го слоя подается на вход всех нейронов  $(i + 1)$ -го

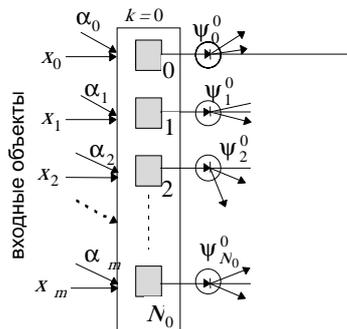


Рис. 2.6. Схема слоистой сети

слоя. Число нейронов в каждом слое может быть любым и никак заранее не связано с количеством нейронов в других слоях. Стандартный способ подачи входных сигналов: все нейроны первого слоя получают каждый входной сигнал. Особое распространение получили трехслойные сети, в которых каждый слой имеет свое наименование: первый – входной, второй – скрытый, третий – выходной.

В полносвязных сетях каждый нейрон передает свой выходной сигнал остальным нейронам, включая самого себя. Выходными сигналами сети могут быть все или некоторые выходные сигналы нейронов после нескольких тактов функционирования сети. Все выходные сигналы подаются всем нейронам.

Таким образом, нейронные сети вычисляют все возможные суперпозиции – функции от функций, получаемые при каскадном соединении сетей.

Рассмотрим более подробно слоистую сеть (рис. 2.6). Ее структура характеризуется числом  $K$  и количеством нейронов в каждом слое. Заметим, что в слоистой сети связи между нейронами в слое отсутствуют.

Введем новые обозначения: вход  $i$ -го нейрона  $k$ -го слоя –  $z_i^k$ , выход  $i$ -го нейрона –  $\psi_i^k$ , количество нейронов в  $k$ -м слое –  $N_k$ ,  $k = 1, 2, \dots, K$ . Тогда суперпозиция входных сигналов  $i$ -го нейрона имеет вид:

$$z_i^k = \sum_{j=0}^{N_{k-1}} \alpha_{ij}^k \psi_j^{k-1}, \quad i = \overline{1, N_k}, \quad k = \overline{1, K}.$$

Здесь  $\alpha_{ij}^k$  – весовые коэффициенты, являющиеся настраиваемыми параметрами и характеризующими связь  $j$ -го нейрона  $(k-1)$ -го слоя с  $i$ -м нейроном  $k$ -го слоя.

Для нулевого слоя имеем  $\psi_j^0 = x_j$ ,  $j = \overline{1, m}$ . С учетом принятых обозначений аппроксимирующая функция  $g_i$ ,  $i = \overline{1, N_k}$ , представляет собой персептрон и может быть записана в виде:

$$\begin{aligned} g_i &= \psi_i^k, i = \overline{1, N_k}; \\ \psi_i^k &= \varphi(z_i^k), i = \overline{1, N_k}, k = \overline{1, K}; \\ \psi_0^k &= 1, k = \overline{0, K-1}. \end{aligned}$$

В качестве функций активации нейронов (нелинейного преобразователя нейронов  $\varphi$ ) часто используют гладкие функции вида:

$$\varphi_1(z) = z; \quad \varphi_2(z) = \frac{1}{1 - \exp(-z)}; \quad \varphi_3(z) = \frac{\exp(z) - \exp(-z)}{\exp(z) + \exp(-z)}.$$

Приближение функций с помощью нейронных сетей сводится к их обучению. При этом входные сигналы  $x$  подаются обучающей сети на обработку, задаются значения весовых коэффициентов  $\alpha$ , а получаемые выходные сигналы  $g$  сравниваются с экспериментальными данными  $y$ . Затем строится оценка работы сети, например, как критерий максимального правдоподобия

$$E(\alpha) = \frac{1}{2} \sum_{\lambda=1}^P \sum_{i=1}^{N_K} (y_{i\lambda} - g_i(x^{(\lambda)}, \alpha))^2,$$

где  $g_i(x^{(\lambda)}, \alpha)$  –  $i$ -й выход сети, соответствующий векторам входных сигналов  $x^{(\lambda)}$  и весовых коэффициентов  $\alpha$ ;  $P$  – объем обучающей выборки  $(x^{(\lambda)}, y_{\lambda})$ .

Поиск оптимальных значений весовых коэффициентов  $\alpha$ , при которых критерий  $E(\alpha)$  минимален, производится с помощью известных методов решения экстремальных задач.

При обучении нейронных сетей целесообразно использовать метод регуляризации, позволяющий получить сглаженные функции  $g_i(x^{(\lambda)}, \alpha)$ . При этом оценка работы сети выбирается в виде:

$$\hat{E}(\beta, \alpha) = E(\alpha) + \beta \Omega(\alpha),$$

где  $\beta$  – параметр регуляризации;  $\Omega(\alpha)$  – равномерно выпуклая функция, например,  $\Omega(\alpha) = \frac{1}{2} \alpha^T \alpha$ .

Оптимальное значение параметра регуляризации  $\beta$  подбирается итерационным методом.

### 2.3. ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ МОДЕЛЕЙ ДИНАМИКИ

Традиционный метод снятия динамической характеристики системы заключается в нанесении искусственного возмущения регулярной формы по входу  $x(t)$  и в регистрации изменений выходной переменной  $y(t)$ . Переходный процесс  $y(t)$ , как правило, аппроксимируется решением линейного дифференциального уравнения с постоянными коэффициентами или по установившимся колебаниям находят значения амплитудно-фазовой характеристики (АФХ).

#### А. Подготовка и планирование эксперимента.

Работа по подготовке эксперимента начинается с изучения конструкции и режимов эксплуатации технологической системы, выбора основных выходных и входных переменных, составления структурной схемы объекта. Экспериментальное исследование динамики систем проводится по каждому из каналов  $x_1(t) \rightarrow y_1(t)$ ,  $x_2(t) \rightarrow y_1(t)$ ,  $x_2(t) \rightarrow y_2(t)$  и т.д. при стабилизированных значениях остальных входных воздействий. Это позволяет структурную схему системы преобразовать в схему с одним входом  $x(t)$  и одним выходом  $y(t)$ .

Планирование эксперимента сводится к выбору вида возмущающего воздействия, количества опытов и величины амплитуды испытательного сигнала. Испытательные воздействия делятся на аperiодические и периодические. К первым относятся ступенчатая функция, прямоугольный импульс, прямоугольная волна. Эти воздействия применяют для снятия переходных функций с промышленных объектов.

Периодические испытательные воздействия типа синусоиды и прямоугольной волны применяют для снятия АФХ. При планировании экспериментов следует учитывать длительность их проведения.

Б. Проведение эксперимента. Перед началом каждого опыта на объекте устанавливается рабочий режим, стабилизируются основные источники возмущений и проверяется правильность включения регистрирующей аппаратуры. При снятии переходных функций  $h(t)$  испытательный сигнал  $x(t) = A$  наносят вручную или с помощью исполнительного механизма

резким изменением положения регулирующего органа. Возмущение наносится в момент времени, принимаемый за нулевой, когда  $y(0) = \text{const}$ ,  $\dot{y}(0) = 0$  и  $\ddot{y}(0) \approx 0$ . Регистрация изменений переменной  $y(t)$ , т.е. запись функции  $h(t)$ , прекращается после того, как, начиная с некоторого момента времени  $t = T_{\text{уст}}$ , установятся значения  $h(t) = \text{const}$ ,  $\dot{h}(t) = 0$  или, для объектов с интегрирующими свойствами,  $h(t) = a_0 + a_1 t$  и  $\dot{h}(t) = \text{const}$ .

На этом этапе частично проверяются предположения о линейности и стационарности динамических свойств системы. Для этого по всем экспериментальным кривым  $h(t)$  вычисляются коэффициенты усиления  $k = h(T_{\text{уст}})/A$  и сравниваются между собой. Помимо этого, сравниваются соответствующие ординаты (точнее, их абсолютные значения) переходных функций, полученных при воздействиях  $+A$  и  $-A$ , а также проверяется выполнение принципа суперпозиции для кривых  $h(t)$ , снятых при сигналах  $\pm A$  и  $\pm 1,5A$ . При существенном различии коэффициентов усиления или при существенном нарушении принципа суперпозиции требуется уменьшить амплитуду испытательного воздействия и провести опыт повторно.

Для проверки справедливости предположения о стационарности динамических свойств объекта требуется постановка еще нескольких серий аналогичных экспериментов по снятию  $h(t)$  через достаточно большие (по сравнению с  $T_{\text{уст}}$ ) промежутки времени.

При снятии АФХ с использованием генератора синусоидальных сигналов первые 3 – 5 периодов колебаний координаты  $y(t)$  носят неустановившийся характер и не пригодны для обработки. В процессе проведения опытов рекомендуется тщательно наблюдать за дрейфом оси колебаний и соблюдением норм допустимых отклонений выходной переменной. С увеличением частоты опыта следует увеличивать и амплитуду испытательного сигнала.

Периодические колебания  $x(t)$  формы прямоугольной волны обычно получают вручную или с помощью исполнительного механизма. Возможны два способа генерирования  $x(t)$ . В первом случае переключения  $x(t)$  с  $+A$  на  $-A$  и наоборот осуществляются в моменты времени  $lT/2$ ,  $l = 0, 1, 2, \dots$ . Такой метод прост, но применять его можно лишь тогда, когда в процессе эксперимента не наблюдается дрейф оси колебаний.

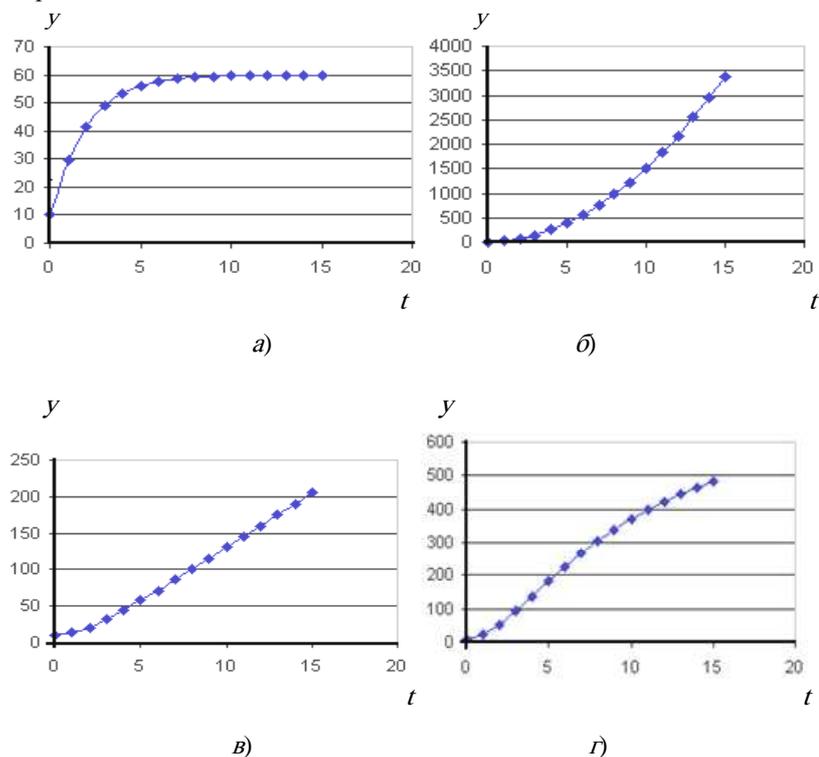
Второй метод генерирования прямоугольной волны предполагает выбор моментов переключения  $x(t)$  в зависимости от значения выходной переменной  $y(t)$ . Переменная  $x(t)$  изменяется с  $+A$  на  $-A$  в момент выполнения равенства

$$y(t) = \pm M,$$

где  $M \geq 0$  – некоторое постоянное для данной частоты число. В этом случае осуществляется своеобразное двухпозиционное регулирование выходной переменной, и опасность дрейфа оси колебаний  $y(t)$  устраняется. Однако этот метод трудоемок и требует больших затрат времени на получение установившихся колебаний.

Проверка предположения о линейности в малом динамических свойств объекта осуществляется путем сравнения частот входных и выходных колебаний. Помимо этого сравниваются ординаты амплитудной характеристики  $A(\omega) = A_{\text{вых}} / A_{\text{вх}}$ , полученные при входных воздействиях  $A_{\text{вх}}$  и  $1,5A_{\text{вх}}$ , где  $A_{\text{вых}}$  – максимальное значение выходных колебаний.

Примеры изменения выходной переменной  $y(t)$  при ступенчатом изменении управляющего воздействия на входе для ряда объектов, динамика которых описывается дифференциальными уравнениями первого и второго порядков, приведенных на рис. 2.7.



**Рис. 2.7.** Траектории выходной переменной при скачкообразном изменении управляющего воздействия для моделей динамики в виде дифференциальных уравнений первого (а) и второго (б–г) порядков:

$$a - \dot{y} = ay(t) + bu(t); \quad \bar{b} - \begin{cases} \dot{z}_1 = z_2(t), \\ \dot{z}_2 = bu(t), \end{cases} \quad y(t) = z_1(t); \quad \bar{B} - \begin{cases} \dot{z}_1 = z_2(t), \\ \dot{z}_2 = z_2(t) + bu(t), \end{cases} \quad y(t) = z_1(t); \quad \bar{\Gamma} - \begin{cases} \dot{z}_1 = z_2(t), \\ \dot{z}_2 = a_1 z_1(t) + a_2 z_2(t) + bu(t), \end{cases} \quad y(t) = z_1(t)$$

В. Обработка результатов эксперимента. *Основной задачей этого этапа является нахождение вида и параметров дифференциальных уравнений по экспериментальным переходным функциям  $h_j(t)$ ,  $j = 1, 2, \dots, n$ .*

Обработка экспериментальных кривых  $h_j(t)$ , снятых при различных испытательных сигналах  $A_j$ , начинается с их усреднения:

$$\bar{h}(t) = \frac{1}{n} \sum_{j=1}^n h_j(t) / A_j.$$

Далее из графика переходной функции  $\bar{h}(t)$  определяется время «чистого» запаздывания  $\tau$ , определяемое как отрезок времени, во всех точках которого выполняется неравенство

$$0 \leq \bar{h}(t) \leq \Delta,$$

где  $\Delta$  – погрешность измерения переменной  $y(t)$ .

Переходная функция  $\bar{h}(t)$  может быть задана непрерывно или дискретно в моменты времени  $t_i, i = 1, 2, \dots, d$  на отрезке времени  $[0, T_{уст}]$ . Коэффициент усиления системы равен  $k = \bar{h}(T_{уст})$ . Во многих случаях экспериментальные переходные функции искажены помехами. Для сглаживания  $h(t)$  применяют усреднение по множеству или по времени. Чаще сглаживание осуществляется по формуле скользящего среднего или методом четвертых разностей.

Так как по сделанным предположениям функция  $h(t)$  снята с линейного объекта с сосредоточенными переменными, динамические свойства которого неизменны во времени, то ее допустимо аппроксимировать решением линейного дифференциального уравнения в обыкновенных производных с постоянными коэффициентами при нулевых начальных условиях. Порядок дифференциального уравнения обычно выбирают не выше 2 – 3 [12].

## 2.4. ПОЛУЧЕНИЕ НЕФОРМАЛЬНЫХ МОДЕЛЕЙ СИСТЕМ

Аналитический метод построения ММ состоит в применении фундаментальных законов природы к конкретной ситуации. Эти законы общепризнанны, многократно подтверждены опытом, служат основой множества научно-технических достижений. Поэтому их обоснованность не вызывает сомнений, что, помимо всего прочего, обеспечивает исследователю мощную психологическую поддержку. На первый план выдвигаются вопросы, связанные с тем, какой закон (законы) следует применять в данном случае и как это делать. Наиболее часто при построении ММ используется следующий закон.

*Сохранение энергии.* Этот закон известен почти двести лет и занимает, пожалуй, наиболее почетное место среди великих законов природы.

Например, инженер оценивает время  $t_k$  сверления слоя металла толщины  $L$  лазером мощностью  $W$ , излучение которого перпендикулярно поверхности материала. Если энергия лазера полностью идет на испарение столбика металла массы  $LSp$  ( $S$  – облучаемая площадь;  $LS$  – объем столбика;  $\rho$  – плотность вещества), то закон сохранения энергии выражается равенством

$$E_0 = Wt_k = hLS\rho,$$

где  $h$  – энергия, требуемая для испарения единицы массы; величина  $h$  имеет составную структуру:  $h = (T_{пл} - T)h_1 + h_2 + h_3$ , поскольку материал необходимо последовательно нагреть до температуры плавления  $T_{пл}$ , а затем расплавить и превратить в пар ( $T$  – исходная температура,  $h_1$  – удельная теплоемкость,  $h_2$  и  $h_3$  – соответственно удельная теплота плавления и парообразования).

Изменение глубины выемки  $l(t)$  со временем определяется из детального баланса энергии в промежутке времени от  $t$  до  $t + dt$ . На испаренную за это время массу  $[l(t + dt) - l(t)]Sp = dlSp$  тратится энергия  $dlhSp$ , равная энергии  $Wdt$ , сообщаемой веществу лазером:  $dlhSp = Wdt$ , откуда получается дифференциальное уравнение  $dl/dt = W/hSp$ . Его интегрирование дает

$$l(t) = \frac{W}{hSp} t = \frac{E(t)}{hSp},$$

где  $E(t)$  – вся энергия, выделенная лазером к моменту времени  $t$ .

Следовательно, глубина выемки пропорциональна затраченной энергии. В действительности процесс сверления гораздо сложнее рассмотренной схемы – энергия тратится на нагрев вещества, на удаление паров из выемки, которая может иметь неправильную форму, и т.д. Вопрос о соответствии объекта и его модели – один из центральных в математическом моделировании, и в дальнейшем мы будем неоднократно к нему возвращаться.

*Сохранение материи.* Пусть имеется небольшое количество радиоактивного вещества (урана), окруженного толстым слоем свинца. Длина свободного пробега продуктов распада  $\lambda_j$  в уране значительно больше характерных размеров самого

материала  $L_I$ , т.е.  $\lambda_I \gg L_I$ . Выражение «толстый слой» означает, что продукты деления полностью поглощаются свинцом. Это гарантируется при выполнении условия  $\lambda_{II} \ll L_{II}$ , где  $\lambda_{II}$  – длина пробега продуктов распада в свинце,  $L_{II}$  – характерный размер свинцового слоя.

Итак, все, что вылетает из области  $I$ , поглощается в области  $II$ , и суммарная масса обоих веществ со временем не меняется. Это и есть закон сохранения материи, примененный к данной ситуации. Если в начальный момент времени  $t = 0$  массы веществ были равны  $M_I(0)$  и  $M_{II}(0)$ , то в любой момент времени справедлив баланс

$$M_I(0) + M_{II}(0) = M_I(t) + M_{II}(t).$$

Одного уравнения, очевидно, недостаточно для определения текущих значений двух масс –  $M_I(t)$  и  $M_{II}(t)$ . Скорость распада (число атомов, распадающихся в единицу времени) пропорциональна общему числу атомов радиоактивного вещества. За время  $dt$  всего распадется  $N_I(t+dt) - N_I(t) = -\alpha N_I(t + \xi dt)$ ,  $\alpha > 0$ ,  $0 < \xi < 1$  атомов. Здесь вторично использован закон сохранения вещества, но применительно к отрезку времени  $dt$ . В этом уравнении, описывающем баланс атомов, в правой части стоит знак минус (вещество убывает), а величина  $N_I(t + \xi dt)$  равна некоторому среднему значению числа атомов за рассматриваемое время. Перепишем его в дифференциальной форме

$$\frac{dN_I(t)}{dt} = -\alpha N_I(t).$$

Учитывая, что

$$M_I(t) = \mu_I N_I(t),$$

где  $\mu_I$  – атомный вес вещества  $I$ , получаем

$$\frac{dM_I(t)}{dt} = -\alpha M_I(t).$$

При самопроизвольной радиоактивности любой атом имеет некоторую не зависящую от состояния окружающего вещества вероятность распада. Поэтому чем больше (меньше) самого радиоактивного вещества, тем больше (меньше) выделяется продуктов распада в единицу времени. Коэффициент пропорциональности  $\alpha > 0$  (постоянная распада) определяется конкретным веществом.

Интегрируя последнее уравнение, получаем, что масса делящегося материала убывает по экспоненциальному закону  $M_I(t) = M_I(0)e^{-\alpha t}$ , и при  $t \rightarrow \infty$  в области  $I$  вещество полностью исчезает.

Так как суммарная масса остается постоянной, то в области  $II$  количество вещества растет:

$$M_{II}(t) = M_{II}(0) + M_I(0) - M_I(0)e^{-\alpha t} = M_{II}(0) + M_I(0)(1 - e^{-\alpha t}),$$

и при  $t \rightarrow \infty$  продукты распада полностью переходят из области  $I$  в область  $II$ .

*Сохранение импульса.* неподвижно стоящая в безветренную погоду на поверхности озера лодка начнет двигаться вперед, если сделать несколько шагов от ее носа к корме. Так проявляет себя закон сохранения импульса, утверждающий: полный импульс системы, не испытывающей действия внешних сил, сохраняется.

Принцип реактивного движения положен в основу многих замечательных технических устройств, например, ракеты, выводящей на орбиту вокруг Земли искусственный спутник, для чего ей требуется развить скорость примерно 8 км/с. Простейшая математическая модель движения ракеты получается из закона сохранения импульса в пренебрежении сопротивлением воздуха, гравитацией и другими силами, исключая, конечно, тягу реактивных двигателей.

Пусть продукты сгорания ракетного топлива покидают расположенные в кормовой части выхлопные сопла со скоростью  $u$  (для современных топлив величина  $u$  равна 3 – 5 км/с). За малый промежуток времени  $dt$  часть топлива выгорела, и масса ракеты изменилась на величину  $dm$ . Изменился также импульс ракеты, однако суммарный импульс системы «ракета + продукты сгорания» остался тем же, что и в момент  $t$ , т.е.

$$m(t)v(t) = m(t+dt)v(t+dt) - dm[v(t+\delta dt) - u],$$

где  $v(t)$  – скорость ракеты;  $v(t+\delta dt) - u$ ,  $0 < \delta < 1$  – средняя за промежуток времени  $dt$  скорость истекающих из сопел газов (обе скорости берутся относительно Земли). Первый член в правой части равенства – импульс ракеты в момент  $t + dt$ , второй – импульс, переданный истекающим газом за время  $dt$ .

Учитывая, что  $m(t+dt) = m(t) + (dm/dt)dt + O(dt^2)$ , закон сохранения импульса можно переписать в виде дифференциального уравнения

$$m \frac{dv}{dt} = - \frac{dm}{dt} u,$$

в котором член  $-(dm/dt)u$ , очевидно, не что иное, как сила тяги ракетных двигателей, и который преобразованный к виду

$$\frac{dv}{dt} = -u \frac{d(\ln m)}{dt},$$

легко интегрируется

$$v(t) = v_0 + u \ln \left( \frac{m_0}{m(t)} \right),$$

где  $v_0, m_0$  – соответственно скорость и масса ракеты в момент времени  $t = 0$ . Если  $v_0 = 0$ , то максимальная скорость ракеты, достигаемая при полном сгорании топлива, равна

$$v = u \ln \left( \frac{m_0}{m_p + m_s} \right),$$

где  $m_p$  – полезная масса (масса спутника);  $m_s$  – масса собственно ракетной конструкции – топливных баков, двигателей, систем управления и т.д. Простая формула Циолковского позволяет сделать фундаментальный вывод о конструкции ракеты для космических полетов. Введем величину  $\lambda = \frac{m_s}{m_0 - m_p}$ , которая характеризует при  $m_p = 0$  отношение структурной и начальной масс ракеты. Тогда для практически реальных значений  $\lambda = 0,1$ ,  $u = 3$  км/с получаем при  $m_p = 0$

$$v = u \ln(1/\lambda) = 7 \text{ км/с.}$$

Отсюда следует, что даже в самой идеальной ситуации (полезная масса равна нулю, отсутствуют гравитация и сопротивление воздуха и т.д.) ракета рассматриваемого типа не способна достичь первой космической скорости. Необходимо использовать многоступенчатые ракеты – вывод, к которому пришли основоположники космонавтики.

*Вариационный принцип* Гамильтона для механической системы. Введем понятие обобщенных координат  $Q(t)$ , полностью определяющих положение механической системы в пространстве. Величину  $dQ/dt$  естественно назвать обобщенной скоростью механической системы в момент времени  $t$ . Набор величин  $Q(t)$  и  $dQ/dt$  определяет состояние механической системы во все моменты времени.

Для описания механической системы вводится функция Лагранжа, которая записывается в виде

$$L(Q, dQ/dt) = E_k - E_{II},$$

где  $E_k, E_{II}$  – кинетическая и потенциальная энергии системы соответственно.

Введем далее величину  $S[Q]$ , называемую действием

$$S[Q] = \int_{t_1}^{t_2} L \left( Q, \frac{dQ}{dt} \right) dt.$$

Интеграл, очевидно, является функционалом от обобщенной координаты  $Q(t)$ , заданной на отрезке  $[t_1, t_2]$ . Он ставит в соответствие некоторое число  $S$  (действие).

Принцип Гамильтона для механической системы гласит: если система движется по законам механики, то  $Q(t)$  – стационарная функция для  $S[Q]$ , или

$$\frac{dS[Q + \epsilon\phi]_{\epsilon=0}}{d\epsilon} = 0.$$

Фигурирующая в *принципе наименьшего действия* функция  $\phi(t)$  – некоторая пробная функция, обращающаяся в ноль в моменты  $t_1, t_2$  и удовлетворяющая тому условию, что  $Q(t) + \epsilon\phi(t)$  – возможная координата данной системы.

Смысл принципа наименьшего действия состоит в том, что из всех априори мыслимых (допускаемых) траекторий (движений) системы между моментами  $t_1, t_2$  выбирается движение, доставляющее минимум функционалу действия. Функция  $\epsilon\phi(t)$  называется *вариацией* величины  $Q(t)$ .

Итак, схема применения принципа Гамильтона для построения моделей механических систем состоит в следующем: определяются обобщенные координаты  $Q(t)$  и обобщенные скорости  $dQ/dt$  системы, строятся функция Лагранжа

$L(Q, dQ/dt)$  и функционал действия  $S[Q]$ , минимизация которого на вариациях  $\varepsilon\varphi(t)$  координаты  $Q(t)$  и дает искомую модель.

Вспользуемся принципом Гамильтона для построения модели движения шарика, соединенного с пружиной. В качестве обобщенной координаты системы естественно выбрать обычную эйлерову координату шарика  $x(t)$ . Тогда обобщенная скорость  $dx/dt = v(t)$  – обычная скорость шарика. Функция Лагранжа равна:

$$L = \frac{m(dx/dt)^2}{2} - k \frac{x^2}{2},$$

где  $k$  – коэффициент, характеризующий упругие свойства пружины;  $x$  – величина растяжения или сжатия пружины относительно нейтрального положения.

Выражение для величины действия имеет вид

$$S[x] = \int_{t_1}^{t_2} L\left(x, \frac{dx}{dt}\right) dt = \int_{t_1}^{t_2} \left[ \frac{m}{2} \left(\frac{dx}{dt}\right)^2 - \frac{k}{2} x^2 \right] dt.$$

Теперь вычислим действие на вариациях  $\varepsilon\varphi(t)$  координаты  $x(t)$ :

$$S[x + \varepsilon\varphi] = \int_{t_1}^{t_2} \left[ \frac{m}{2} \left(\frac{d(x + \varepsilon\varphi)}{dt}\right)^2 - \frac{k}{2} (x + \varepsilon\varphi)^2 \right] dt.$$

Последнюю формулу необходимо продифференцировать по  $\varepsilon$ , учитывая, что функции  $x, \varphi, dx/dt$  от  $\varepsilon$  не зависят:

$$\begin{aligned} \frac{dS[x + \varepsilon\varphi]}{d\varepsilon} &= \\ &= \frac{d}{d\varepsilon} \frac{1}{2} \int_{t_1}^{t_2} \left[ m \left\{ \left(\frac{dx}{dt}\right)^2 + 2\varepsilon \frac{dx}{dt} \frac{d\varphi}{dt} + \varepsilon^2 \left(\frac{d\varphi}{dt}\right)^2 \right\} - k \{ x^2 + 2\varepsilon x\varphi + \varepsilon^2 \varphi^2 \} \right] dt = \\ &= \int_{t_1}^{t_2} \left[ m \left\{ \frac{dx}{dt} \frac{d\varphi}{dt} + \varepsilon \left(\frac{d\varphi}{dt}\right)^2 \right\} - k \{ x\varphi + \varepsilon\varphi^2 \} \right] dt, \end{aligned}$$

и положить в ней  $\varepsilon = 0$ :

$$\left. \frac{dS[x + \varepsilon\varphi]}{d\varepsilon} \right|_{\varepsilon=0} = \int_{t_1}^{t_2} \left[ m \frac{dx}{dt} \frac{d\varphi}{dt} - kx\varphi \right] dt = 0.$$

Правая часть этого выражения с помощью интегрирования ее первого члена по частям и с учетом того, что  $\varphi=0$  в моменты  $t_1, t_2$  преобразуется к виду

$$\left. \frac{dS[x + \varepsilon\varphi]}{d\varepsilon} \right|_{\varepsilon=0} = - \int_{t_1}^{t_2} \varphi \left[ m \frac{d^2x}{dt^2} + kx \right] dt = 0.$$

Поскольку пробная функция  $\varphi(t)$ , фигурирующая в формулировке принципа наименьшего действия, произвольна, то часть выражения, стоящая под знаком интеграла в квадратных скобках, должна быть равна нулю во все моменты времени  $t_1 < t < t_2$ , т.е.

$$m \frac{d^2x}{dt^2} = -kx.$$

Таким образом, движение системы будет описываться дифференциальным уравнением второго порядка, которое также может быть получено из закона Ньютона и закона сохранения энергии.

Методика составления аналитических моделей систем включает несколько этапов.

1. *Изучение системы.* Проводится анализ конструкции системы и протекающих в ней физико-химических процессов.
2. *Составление структурной схемы системы.* Исследуемая система условно разделяется на ряд *элементов* или *элементарных областей*, в качестве которых обычно выделяют или самостоятельные части системы или область заданных размеров, рассматриваемую в течение заданного интервала времени.

Следует отчетливо понимать, что «глубина» расслоения системы на элементы и элементарные области не имеет предела, поэтому выбор числа таких элементов или элементарных областей должен производиться с учетом уровня наших знаний о процессах, реальной возможности определения неизвестных параметров (коэффициентов модели, конст-

руктивных параметров, входных переменных, неопределенных параметров и возмущающих воздействий и т.д.), а также возможности решения полученных систем уравнений, целевого назначения модели системы.

В общем случае под элементарной областью понимается область заданных размеров, рассматриваемая в течение заданного интервала времени и охватывающая как конструкционные элементы технологической системы (ТС), влияющие на поля потенциалов переноса, так и некоторые объемы заполняющих технологическую систему сред (продукты, теплоносители и хладагенты, насадки, сорбенты, катализаторы и т.п.). Полная совокупность элементарных областей составляет весь рабочий объем ТС за конечный период времени.

С задачей рационального расчленения ТС на элементы и элементарные области тесно связана задача принятия системы допущений. В общем случае обсуждаются и затем принимаются или отвергаются следующие важнейшие допущения: о стационарности процессов в элементе или элементарной области; о сосредоточенности или распределенности его переменных; об учете тех или иных физико-химических явлений, имеющих место в данном элементе или элементарной области и т.п. Допущения представляют компромисс между требуемой или желаемой точностью описания свойств объекта и возможностью как количественной оценки физико-химических явлений, так и решения получающихся уравнений.

3. *Составление математического описания процессов, протекающих в отдельных элементах (элементарных областях).* Для бесконечно малого объема элементарной области технологической системы с распределенными переменными выписываются уравнения теплового и материального балансов в интегральной форме. Затем эти уравнения с помощью теорем о среднем и конечных приращениях преобразуются в дифференциальную форму [13]. Если уравнения имеют аналитические решения, то математическое описание процессов, протекающих в элементарной области, задается системой конечных соотношений.

Для ТС с сосредоточенными параметрами (переменными) уравнения материального и теплового балансов записываются в конечной форме.

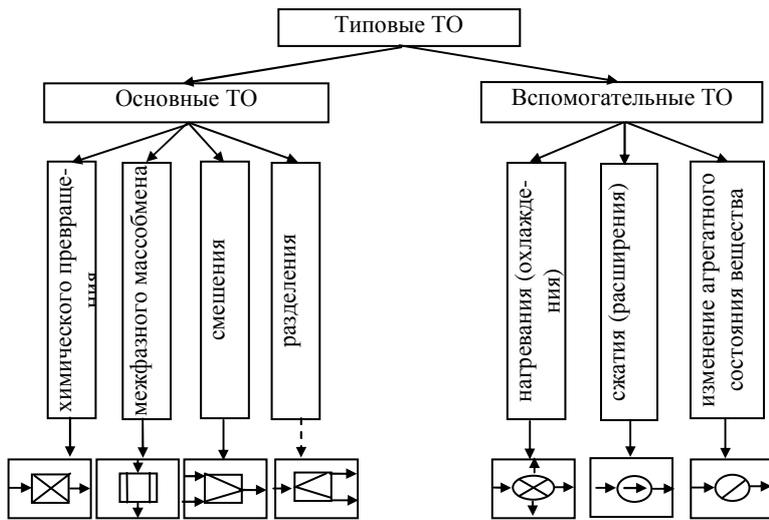
Следует помнить, что в математическое описание элемента входят граничные условия для дифференциальных уравнений и связи с другими, соседними, элементами или областями.

4. *Определение параметров (коэффициентов) уравнений процессов, протекающих в элементах или элементарных областях.* Для определения коэффициентов и других параметров уравнений модели ТС необходимо знать физико-химические свойства перерабатываемых веществ, константы скоростей химических реакций, процессов тепло- и массопереноса и т.д. Разумеется, необходимо знать также все определяющие геометрические размеры элементов и элементарных областей, на которые условно разделяется ТС.

Часть интересующей нас информации можно найти в соответствующей технической и научной литературе, для определения же отдельных коэффициентов и констант требуется постановка специальных лабораторных экспериментов. Результаты экспериментов чаще всего представляют в критериальной форме, что позволяет распространять их на подобные (в определенном смысле) элементы и объекты.

Определение коэффициентов и других параметров модели очень часто является исключительно трудоемкой и кропотливой работой. Реальная возможность определения численных значений тех или иных параметров всегда должна учитываться при составлении структурной схемы объекта и принятии системы допущений. Погрешность определения параметров существенно влияет на точность и адекватность математического описания.

5. *Составление и анализ уравнений аналитической модели ТС.* Процессы, протекающие в ТС (например, в химико-технологических аппаратах), представляют на схемах в виде типовых технологических операторов (ТО), которые подразделяются на основные и вспомогательные (рис. 2.8).



Структурные схемы технологических систем представляют следующие виды потоков: последовательный (а), параллельный (б) и обратный (рециркуляционный) (в) (рис. 2.9).

Обозначения типовых ТО

Рис. 2.8. Классификация типовых технологических операторов (ТО)

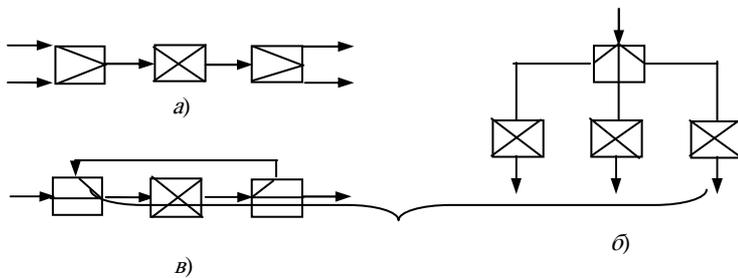


Рис. 2.9. Типы технологических потоков:

а – последовательный; б – параллельный; в – обратный (рецикл)

Структурные схемы ТС почти всегда можно представить в комбинации рассмотренных типовых технологических потоков. Зная, например, модели статики отдельных технологических аппаратов, можно аналитически построить математическую модель статики (статическую характеристику) всей технологической системы.

С использованием описанной выше методики рассмотрим технику вывода (получения) аналитических моделей процессов теплообмена, протекающих в технологических аппаратах, из фундаментальных законов природы.

Теплообменом называется любой процесс переноса теплоты, в котором она в соответствии со вторым началом термодинамики самопроизвольно переходит от более нагретой среды к менее нагретой. Движущей силой процесса переноса теплоты является разность температур.

Перенос теплоты возможен тремя различными способами: теплопроводностью, конвекцией и излучением.

В технике указанные способы теплообмена редко встречаются в обособленном виде: чаще всего приходится иметь дело с сочетанием двух или даже всех трёх способов при их последовательном или одновременном действии. Особое место занимает теплообмен, сопровождающийся изменением агрегатного состояния тел, участвующих в этом процессе (испарение жидкости, конденсация паров). Различают два случая теплообмена: теплоотдачу и теплопередачу. Теплоотдачей называется процесс теплообмена между твёрдым телом (например, стенкой аппарата) и соприкасающейся с ней жидкостью или газом. Теплообмен между жидкостями, газами, жидкостью и газом, разделёнными стенкой, называется теплопередачей.

Количество тепла, проходящее через данную поверхность за время  $\tau$ , называют тепловым потоком, обозначим его через  $Q$ . Тепловой поток, проходящий в единицу времени через единицу поверхности назовём удельным тепловым потоком.

Сформулируем физические закономерности, определяющие процессы, связанные с распространением тепла.

1. Закон Фурье: количество тепла  $dQ$  [Дж], переданного посредством теплопроводности через площадь сечением  $S$  изотермической поверхностью тела за время  $d\tau$ , выражается основным уравнением теплопроводности

$$dQ = -\lambda S \frac{dT}{dn} d\tau.$$

Отрицательный знак в правой части уравнения является следствием падения температуры в направлении переноса теплоты. Коэффициент пропорциональности  $\lambda$  в уравнении называется коэффициентом теплопроводности. Величина  $\lambda$  для различных веществ сильно различается, а для одного и того же вещества зависит от температуры, плотности, структуры, влажности и других факторов.

2. Уравнение Ньютона: тепловой поток  $\frac{dQ}{d\tau}$  при конвективном теплообмене пропорционален разности температур  $(\theta - T)$ , площади поверхности  $S$  твердого тела и соприкасающейся с ним среде:

$$\frac{dQ}{d\tau} = \alpha(\theta - T)S,$$

где  $\theta, T$  – температура поверхности твердого тела и потока соответственно;  $\alpha$  – коэффициент теплоотдачи, Вт/(м<sup>2</sup>·К).

Коэффициент теплоотдачи  $\alpha$  выражает количество теплоты, отданное единицей поверхности ( $S = 1$  м<sup>2</sup>) в единицу времени ( $\tau = 1$  с) при разности температур  $(\theta - T)$ . Заметим, что  $\alpha$  не является постоянной величиной, а зависит от многих параметров.

3. Количество теплоты, которое необходимо сообщить однородному телу, чтобы повысить его температуру на величину  $\Delta T$  равно

$$Q = c_p m \Delta T = c_p \rho V \Delta T,$$

где  $c_p$  – удельная теплоемкость, Дж/кг·К;  $m$  – масса тела, кг;  $\rho$  – плотность тела, кг/м<sup>3</sup>;  $V$  – объем тела, м<sup>3</sup>.

4. Внутри потока может возникать или поглощаться теплота. Выделение теплоты может быть описано плотностью тепловых потоков  $F(x, y, z, \tau)$  в точке  $(x, y, z)$  в момент времени  $\tau$ . В результате действия этих источников за промежутков времени  $(\tau, \tau + \Delta\tau)$  выделится количество теплоты

$$dQ = F(x, y, z, \tau) dx dy dz d\tau,$$

или в интегральной форме:

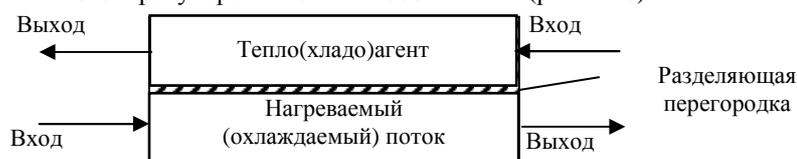
$$Q = \int_{\tau_1}^{\tau_2} \iiint_{XYZ} F(x, y, z, \tau) dx dy dz d\tau,$$

где  $X, Y, Z$  – области изменения независимых пространственных переменных  $x, y, z$ .

Процессы теплообмена осуществляют в теплообменных аппаратах (теплообменниках). По назначению теплообменные аппараты бывают подогревателями, холодильниками, испарителями, конденсаторами, дистилляторами, сублиматорами, плавителями и т.п.

По способу передачи теплоты различают теплообменники поверхностные и смешительные. В первом случае передача тепла происходит через разделяющие твердые стенки, во втором – непосредственный контакт (смешение) нагретых и холодных сред (жидкостей, газов, твердых веществ).

*Пример 2.2.* Рассмотрим вывод уравнений ММ процесса нагрева потока жидкости конденсирующимся паром, осуществляемого в рекуперативном теплообменнике (рис. 2.10).



**Рис. 2.10.** Схема процесса теплообмена через разделяющую перегородку в рекуперативном теплообменнике

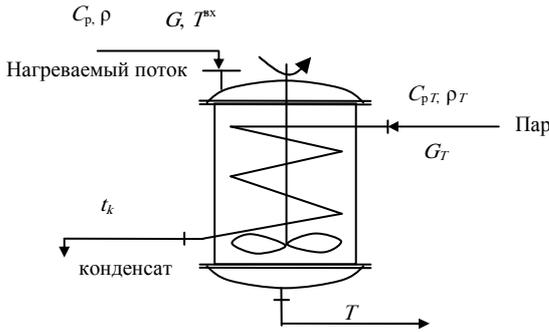
Предположим, что нагреваемый поток жидкости подается в трубное (реакционное) пространство, пар – в межтрубное пространство (змеевик или теплообменную рубашку). Движение потока жидкости соответствует определенной типовой модели гидродинамики: идеальному смешению, идеальному вытеснению, диффузионному режиму и др.

1. Гидродинамика нагреваемого потока соответствует модели идеального смешения, а пар подается в змеевик, находящийся внутри реакционного пространства аппарата (рис. 2.11).

При сделанных допущениях температура жидкости в каждой точке реакционного пространства теплообменника и на его выходе одинакова, а пар конденсируется при температуре  $t_k$ . Составим уравнение теплового баланса по потоку нагреваемой жидкости за промежуток времени  $(\tau_1, \tau_2)$ :

$$c_p \rho G \int_{\tau_1}^{\tau_2} [T(\tau) - T^{\text{BX}}(\tau)] d\tau + \int_{\tau_1}^{\tau_2} k_T S [T(\tau) - t_k] d\tau = c_p \rho V [T(\tau_2) - T(\tau_1)],$$

где  $c_p, \rho, G$  – теплоемкость, плотность и расход потока жидкости соответственно;  $S$  – поверхность теплообмена;  $V$  – объем реакционной зоны теплообменника;  $k_T$  – коэффициент теплопередачи.



**Рис. 2.11. Схема аппарата для нагрева потока жидкости конденсирующимся паром**

Пользуясь теоремой о среднем, получим равенство

$$c_p \rho G [T(\tau) - T^{\text{BX}}(\tau)]_{\tau=\tau_3} \Delta\tau + k_T S [T(\tau) - t_k(\tau)]_{\tau=\tau_4} \Delta\tau = c_p \rho V [T(\tau_2) - T(\tau_1)],$$

которое при помощи теоремы о конечных приращениях можно преобразовать к виду

$$c_p \rho G [T(\tau) - T^{\text{BX}}(\tau)]_{\tau_3 \in [\tau_1, \tau_2]} \Delta\tau + k_T S [T(\tau_4) - t_k]_{\tau_4 \in [\tau_1, \tau_2]} \Delta\tau = c_p \rho V \frac{dT}{d\tau} \Big|_{\tau=\tau_5 \in [\tau_1, \tau_2]},$$

где  $\tau_3, \tau_4, \tau_5$  – промежуточные точки интервала  $[\tau_1, \tau_2]$ .

Отсюда после сокращения на  $\Delta\tau$  находим:

$$c_p \rho G [T(\tau) - T^{\text{BX}}(\tau)]_{\tau=\tau_3} + k_T S [T(\tau_4) - t_k]_{\tau=\tau_4} = c_p \rho V \frac{dT}{d\tau} \Big|_{\tau=\tau_5}.$$

Наши рассуждения относятся к произвольному промежутку времени  $(\tau_1, \tau_2)$ . Переходя к пределу при  $(\tau_1, \tau_2) \rightarrow \tau$ , получим уравнение динамики процесса теплообмена:

$$c_p \rho G [T(\tau) - T^{\text{BX}}(\tau)] + k_T S [T(\tau) - t_k] = c_p \rho V \frac{dT}{d\tau},$$

из которого следует уравнение (модель) статики процесса при  $\frac{dT}{d\tau} = 0$ :

$$c_p \rho G [T - T^{\text{BX}}] + k_T S [T - t_k] = 0 \text{ или } T = \frac{(c_p \rho G T^{\text{BX}} + k_T S t_k)}{c_p \rho G + k_T S}.$$

2. Гидродинамика нагреваемого потока жидкости соответствует модели идеального вытеснения. В этом случае теплообмен осуществляется в аппарате типа «труба в трубе», причем пар подается в межтрубное пространство. Составим уравнение теплового баланса на участке трубы  $(l_1, l_2)$  за промежуток времени  $[\tau_1, \tau_2]$ :

$$c_p \rho G \int_{\tau_1}^{\tau_2} [T(l_2, \tau) - T(l_1, \tau)] d\tau + \int_{\tau_1}^{\tau_2} \int_{l_1}^{l_2} k_T \pi d [T(l, \tau) - t_k] dl d\tau = c_p \rho G \int_{\tau_1}^{\tau_2} [T(l, \tau_2) - T(l, \tau_1)] dl,$$

где  $d$  – внутренний диаметр трубы,  $S$  – площадь поперечного сечения.

Пользуясь теоремой о среднем, получим равенство

$$\begin{aligned} & c_p \rho G [T(l_2, \tau) - T(l_1, \tau)] \Big|_{\tau_3 \in [\tau_1, \tau_2]} \Delta\tau + k_T \pi d [T(l_4, \tau) - t_k] \Big|_{\substack{l_4 \in [l_1, l_2] \\ \tau_4 \in [\tau_1, \tau_2]}} \Delta\Delta\tau = \\ & = c_p \rho S [T(l, \tau_2) - T(l, \tau_1)] \Big|_{\substack{l_4 \in [l_1, l_2] \\ \tau_4 \in [\tau_1, \tau_2]}} \Delta\Delta\tau, \end{aligned}$$

которое при помощи теоремы о конечных приращениях можно преобразовать к виду:

$$\begin{aligned} c_p \rho G \frac{\partial T(l, \tau)}{\partial l} \Big|_{\substack{l=l_5 \in [l_1, l_2] \\ \tau=\tau_3 \in [\tau_1, \tau_2]}} \Delta\Delta\tau + k_T \pi d [T(l_4, \tau) - t_k] \Delta\Delta\tau = \\ = c_p \rho S \frac{\partial T(l, \tau)}{\partial \tau} \Big|_{\substack{l=l_4 \in [l_1, l_2] \\ \tau=\tau_4 \in [\tau_1, \tau_2]}} \Delta\Delta\tau, \end{aligned}$$

где  $\tau_3, \tau_4, \tau_5$  и  $l_3, l_4, l_5$  – промежуточные точки интервалов  $(\tau_1, \tau_2)$  и  $(l_1, l_2)$ .

Отсюда после сокращения на произведение  $\Delta\Delta\tau$  находим:

$$c_p \rho G \frac{\partial T(l, \tau)}{\partial l} \Big|_{\substack{l=l_5 \\ \tau=\tau_3}} + k_T \pi d [T(l, \tau) - t_k] \Big|_{\substack{l=l_4 \\ \tau=\tau_4}} = c_p \rho S \frac{\partial T}{\partial \tau} \Big|_{\substack{\tau=\tau_4 \\ l=l_3}}.$$

Наши рассуждения относятся к произвольным промежуткам  $(l_1, l_2)$  и  $(\tau_1, \tau_2)$ . Переходя к пределу при  $(l_1, l_2) \rightarrow l$  и  $(\tau_1, \tau_2) \rightarrow \tau$ , получим уравнение динамики процесса теплообмена в аппарате типа «труба в трубе»:

$$c_p \rho G \frac{\partial T}{\partial l} + k_T S [T(l, \tau) - t_k] = c_p \rho S \frac{\partial T}{\partial \tau},$$

из которого можно получить уравнение статики при  $\frac{dT}{d\tau} = 0$ :

$$\vartheta \frac{dT}{dl} = \frac{k_T \pi d}{G \rho S} [t_k - T(l)]; \quad T(0) = T^{\text{вх}},$$

где  $\vartheta$  – скорость движения потока жидкости.

3. Гидродинамика нагреваемого потока жидкости соответствует *диффузионной модели*. Уравнение диффузионной модели теплообмена можно получить при подсчете баланса теплоты на отрезке  $(l_1, l_2)$  за некоторый промежуток времени  $(\tau_1, \tau_2)$ :

$$\begin{aligned} & c_p \rho G \int_{\tau_1}^{\tau_2} [T(l_2, \tau) - T(l_1, \tau)] d\tau + c_p \rho S \int_{\tau_1}^{\tau_2} \left[ D_T \frac{\partial T}{\partial l}(l_2, \tau) - D_T \frac{\partial T}{\partial l}(l_1, \tau) \right] d\tau + \\ & + \int_{\tau_1}^{\tau_2} \int_{l_1}^{l_2} k_T \pi d [T(l, \tau) - t_k] dl d\tau = c_p \rho S \int_{l_1}^{l_2} [T(l, \tau_2) - T(l, \tau_1)] dl. \end{aligned}$$

Используя аналитические выкладки, аналогичные вышеприведенным, можно получить уравнение диффузионной модели динамики теплообмена:

$$c_p \rho G \frac{\partial T}{\partial l} + c_p \rho S \frac{\partial}{\partial l} \left( D_T \frac{\partial T}{\partial l} \right) + k_T \pi d (T - t_k) = c_p \rho S \frac{\partial T}{\partial \tau},$$

из которого легко получить уравнение (модель) статики при  $\frac{dT}{d\tau} = 0$ :

$$\vartheta \frac{\partial T}{\partial l} + \frac{\partial}{\partial l} \left( D_T \frac{\partial T}{\partial l} \right) + \frac{k_T \pi d}{c_p \rho S} (T - t_k) = 0, \quad T(0) = T^{\text{вх}}, \quad \frac{dT}{dl}(L) = 0.$$

4. Гидродинамика нагреваемого потока жидкости в аппарате соответствует *ячеечной модели*. В этом случае поток жидкости, например, в аппарате с мешалкой представляется разделенным на  $m$  последовательно соединенных ячеек идеального смешения. Тогда для каждой ячейки и в целом для аппарата можно записать:

$$\begin{aligned} & c_p \rho G (T_{(1)}(\tau) - T_{(1)}^{\text{вх}}) + k_T S_{(1)} (T_{(1)}(\tau) - t_k) = c_p \rho V_{(1)} \frac{dT_{(1)}}{d\tau}; \\ & c_p \rho G (T_{(2)}(\tau) - T_{(1)}(\tau)) + k_T S_{(2)} (T_{(2)}(\tau) - t_k) = c_p \rho V_{(2)} \frac{dT_{(2)}}{d\tau}; \\ & c_p \rho G (T_{(m)}(\tau) - T_{(m-1)}(\tau)) + k_T S_{(m)} (T_{(m)}(\tau) - t_k) = c_p \rho V_{(m)} \frac{dT_{(m)}}{d\tau}. \end{aligned}$$

Если режим движения потока жидкости в аппарате описывается ячейчной моделью с обратными потоками (рис. 2.12), то математическая модель процесса теплообмена принимает вид:

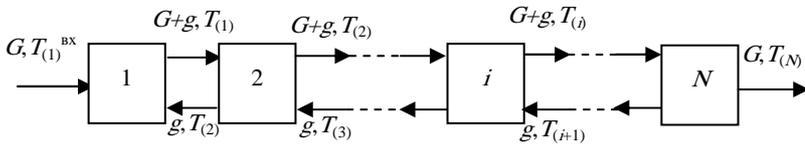


Рис. 2.12. Схема процесса теплообмена по ячеечной модели

с обратными потоками

$$c_p \rho [GT_{(1)}^{BX}(\tau) + gT_{(2)}(\tau) - (G - g)T_{(1)}(\tau)] + k_T S_{(1)}(T_{(1)}(\tau) - t_k) =$$

$$= c_p \rho V_{(1)} \frac{dT_{(1)}}{d\tau};$$

$$c_p \rho [(G + g)(T_{(1)} - T_{(2)}) + g(T_{(3)} - T_{(2)})] + k_T S_{(2)}(T_{(2)}(\tau) - t_k) =$$

$$= c_p \rho V_{(2)} \frac{dT_{(2)}}{d\tau};$$

$$\dots$$

$$c_p \rho [(G + g)(T_{(i-1)} - T_{(i)}) + g(T_{(i+1)} - T_{(i)})] + k_T S_{(i)}(T_{(i)}(\tau) - t_k) =$$

$$= c_p \rho V_{(i)} \frac{dT_{(i)}}{d\tau};$$

$$\dots$$

$$c_p \rho [(G + g)T_{(N-1)} - (G + g)T_{(N)}] + k_T S_{(N)}(T_{(N)}(\tau) - t_k) =$$

$$= c_p \rho V_{(N)} \frac{dT_{(N)}}{d\tau}.$$

Зададим начальные условия для записанных выше уравнений динамики процесса теплообмена:

$$T_{(1)}(0) = T_{10}, \dots, T_{(i)}(0) = T_{i0}, \dots, T_{(N)}(0) = T_{N0}.$$

Для получения уравнений модели статики процесса теплообмена необходимо принять  $\frac{dT_{(i)}}{d\tau} = 0, i = 1, 2, \dots, N$ :

$$c_p \rho [GT_{(1)}^{BX} + gT_{(2)} - (G - g)T_{(1)}] + k_T S_{(1)}(T_{(1)} - t_k) = 0,$$

$$c_p \rho [(G + g)(T_{(1)} - T_{(2)}) + g(T_{(3)} - T_{(2)})] + k_T S_{(2)}(T_{(2)} - t_k) = 0,$$

$$\dots$$

$$c_p \rho [(G + g)(T_{(i-1)} - T_{(i)}) + g(T_{(i+1)} - T_{(i)})] + k_T S_{(i)}(T_{(i)} - t_k) = 0,$$

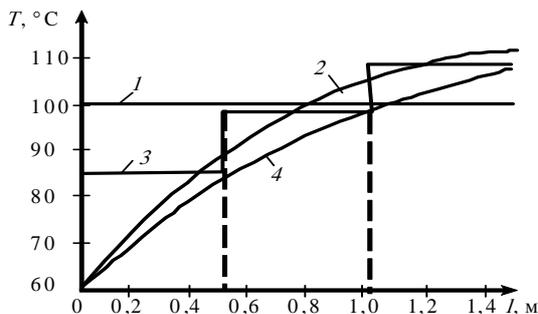
$$\dots$$

$$c_p \rho [(G + g)T_{(N-1)} - (G + g)T_{(N)}] + k_T S_{(N)}(T_{(N)} - t_k) = 0.$$

Пример 2.3. Оценим профиль температуры нагреваемого потока жидкости, исходя из различных гидродинамических моделей движения этого потока. Пусть

$$G = 1000 \text{ кг/ч}; c_p = 2520 \text{ Дж/кг}\cdot\text{К}; \rho = 1200 \text{ кг/м}^3.$$

Обогрев осуществляется насыщенным водяным паром, имеющим температуру  $t_k = 120 \text{ }^\circ\text{C}$ . Диаметр цилиндрической



поверхности теплообмена равен  $d = 0,5 \text{ м}$ . Коэффициент теплопередачи составляет

Рис. 2.13. Результаты расчета температурного профиля по различным моделям:

1 – идеальное смешение; 2 – идеальное вытеснение; 3 – ячеечная модель; 4 – диффузионная модель

$k_T = 600 \text{ Вт/м}^2 \cdot \text{К}$ , длина теплообменника – 1,5 м, параметры ячеечной и диффузионной модели:  $n = 3$ ,  $D_T = 3,54 \cdot 10^{-4} \text{ м}^2/\text{с}$ . На рис. 2.13 приведены результаты расчета температурного профиля по длине теплообменника.

Они свидетельствуют о значительном разбросе температур для различных моделей гидродинамики. Более реальный характер изменения температуры по длине теплообменника отражают ячеечная и диффузионная модели. При этом конечные температуры для данных моделей практически совпадают, но, тем не менее, профили температур различаются существенно.

Приведенный пример подчеркивает важность учета реальной структуры потока в аппарате и его адекватного описания гидродинамическими моделями.

*Вывод уравнения теплопроводности.* Для простоты будем рассматривать одномерные процессы теплопроводности. Они имеют место, например, в длинном тонком металлическом стержне, нагреваемом с одного из торцов при условии, что стержень изотропен. Его начальная температура в любом поперечном сечении не зависит от  $y, z$  (это условие должно выполняться и на торцах стержня), а потерями тепла с боковой поверхности можно пренебречь.

Рассмотрим произвольное сечение стержня с координатой  $x$ . Пусть  $\rho(x), c_p(x), \lambda(x)$  – соответственно плотность, удельная теплоемкость и коэффициент теплопроводности в точках этого сечения. Запишем уравнение распространения этого тепла в стержне (уравнение теплопроводности) на некотором отрезке  $(x_1, x_2)$  за некоторый промежуток времени  $(t_1, t_2)$ , применяя закон сохранения энергии (в интегральной форме):

$$\int_{t_1}^{t_2} \left[ \lambda \frac{\partial T}{\partial x}(\xi, \tau) \Big|_{x=x_2} - \lambda \frac{\partial T}{\partial x}(\xi, \tau) \Big|_{x=x_1} \right] d\tau + \int_{x_1}^{x_2} \int_{t_1}^{t_2} F(\xi, \tau) d\xi d\tau = \int_{x_1}^{x_2} c_p \rho [T(\xi, t_2) - T(\xi, t_1)] d\xi.$$

Предположим, что функция  $T(x, t)$  имеет непрерывные производные  $T_t = \frac{\partial T}{\partial t}$  и  $T_{xx} = \frac{\partial^2 T}{\partial x^2}$ .

Пользуясь теоремой о среднем, получаем равенство

$$\left[ \lambda \frac{\partial T}{\partial x}(\xi, \tau) \Big|_{\xi=x_2} - \lambda \frac{\partial T}{\partial x}(\xi, \tau) \Big|_{\xi=x_1} \right]_{\tau=t_3} \Delta t + F(\xi_4, t_4) \Delta x \Delta t \Big|_{\xi_4=x_4, \tau=t_4} = \left\{ c_p \rho [T(\xi, t_2) - T(\xi, t_1)] \right\}_{\xi=x_3} \Delta x,$$

которое при помощи теоремы о конечных приращениях можно преобразовать к виду:

$$\frac{\partial}{\partial x} \left[ \lambda \frac{\partial T}{\partial x}(\xi, \tau) \right]_{\xi=x_5, \tau=t_3} \Delta x \Delta t + F(x_4, t_4) \Delta x \Delta t = \left[ c_p \rho \frac{\partial T}{\partial t} \right]_{\xi=x_3, \tau=t_5} \Delta x \Delta t,$$

где  $t_3, t_4, t_5$  и  $x_3, x_4, x_5$  – промежуточные точки интервалов  $(x_1, x_2)$  и  $(t_1, t_2)$ .

Отсюда после сокращения на произведение  $\Delta x \Delta t$  получим:

$$\frac{\partial}{\partial x} \left( \lambda \frac{\partial T}{\partial x} \right) \Big|_{x=x_5, t=t_3} + F(x, t) \Big|_{x=x_4, t=t_4} = c_p \rho \frac{\partial T}{\partial t} \Big|_{x=x_3, t=t_5}.$$

Все эти рассуждения относятся к произвольным промежуткам  $(x_1, x_2)$  и  $(t_1, t_2)$ . Переходя к пределу при  $x_1, x_2 \rightarrow x$  и  $t_1, t_2 \rightarrow t$ , получим уравнение

$$\frac{\partial}{\partial x} \left( \lambda \frac{\partial T}{\partial x} \right) + F(x, t) = c_p \rho \frac{\partial T}{\partial t},$$

которое называется уравнением теплопроводности.

*Частные случаи.*

1. Если стержень однороден, то  $\rho, c_p, \lambda = \text{const}$ , и мы получаем линейное уравнение теплопроводности

$$T_t = a^2 T_{xx} + f(x, t),$$

где  $a^2 = \frac{\lambda}{c_p \rho}$  – коэффициент температуропроводности;  $f(x, t) = \frac{F(x, t)}{c_p \rho}$ .

Если источники отсутствуют, т.е.  $F(x, t) = 0$ , то уравнение теплопроводности примет вид:

$$T_t = a^2 T_{xx},$$

2. В случае теплообмена с окружающей средой, подчиняющегося закону Ньютона, количество тепла, теряемого стержнем, рассчитываемого на единицу длины и времени, равно  $F_0 = \alpha(T - \theta)$ , где  $\theta(x, t)$  – температура окружающей среды;  $\alpha$  – коэффициент теплообмена.

Поскольку в нашем приближении не учитывается распределение температуры по сечению, то действие поверхностных источников эквивалентно действию объемных источников тепла. Таким образом, плотность тепловых источников в точке  $x$  в момент времени  $t$  равна

$$F = F_1(x, t) - \alpha(T - \theta),$$

где  $F_1(x, t)$  – плотность других источников тепла.

Если стержень однороден, то уравнение теплопроводности с боковым теплообменом имеет следующий вид:

$$T_t = a^2 T_{xx} - h_1 T + f(x, t),$$

где  $h_1 = \frac{\alpha}{c_p \rho}$ ;  $f(x, t) = h_1 \cdot \theta(x, t) + \frac{F_1(x, t)}{c_p \rho}$  – известная функция.

3. Коэффициенты  $c_p$  и  $\lambda$ , как правило, являются медленно меняющимися функциями температуры, поэтому сделанное выше предположение о постоянстве этих коэффициентов возможно лишь при рассмотрении небольших интервалов изменения температуры.

Изучение процесса теплопроводности в большом интервале изменения температур приводит к нелинейному уравнению теплопроводности, которое для неоднородной среды запишется в виде

$$\frac{\partial}{\partial x} \left( \lambda(T, x) \frac{\partial T}{\partial x} \right) + F(x, t) = c_p(T, x) \rho(T, x) \frac{\partial T}{\partial t}. \quad (2.7)$$

Для получения единственного решения уравнения теплопроводности необходимо к уравнению присоединить начальные и граничные условия.

Начальное условие состоит в задании значений функции  $T(x, t)$  в начальный момент  $t_0$ , т.е.  $T(x, t) = T_0(x)$ ,  $0 \leq x \leq l$ .

Граничные условия могут быть различными в зависимости от температурного режима на торцах стержня. Рассматривают три основных типа граничных условий.

1. На торцах стержня в любой момент времени задается температура:

$$T(0, t) = T_1(t); \quad T(l, t) = T_2(t); \quad t > 0.$$

2. На торцах стержня задаются потоки теплоты как функции времени:

$$\frac{\partial}{\partial x} T(0, t) = v_1(t); \quad \frac{\partial}{\partial x} T(l, t) = v_2(t).$$

К этим условиям мы приходим, если задана величина теплового потока  $Q(l, t)$ , протекающего через торцевое сечение стержня

$$Q(l, t) = -\lambda \frac{\partial}{\partial x} T(l, t),$$

откуда  $\frac{\partial}{\partial x} T(l, t) = v(t)$ , где  $v(t)$  – известная функция, выражающаяся через заданный поток  $Q(l, t)$  по формуле

$$v(t) = -\frac{Q(l, t)}{\lambda}.$$

3. На торцах стержня задаются линейные соотношения между производной и функцией, например, для  $x = l$ :

$$\frac{\partial}{\partial x} T(l, t) = -h_2 [T(l, t) - \theta(t)].$$

Это граничное условие соответствует теплообмену по закону Ньютона на поверхности тела с окружающей средой, температура которой  $\theta$  известна.

Пользуясь двумя выражениями для теплового потока, вытекающего через сечение  $x = l$ :  $Q = \alpha(T - \theta)$  и  $Q = -\lambda \frac{\partial T}{\partial x}$  получаем математическую формулировку третьего граничного условия в виде:

$$\frac{\partial}{\partial x} T(l, t) = -h_2 [T(l, t) - \theta(t)],$$

где  $h_2 = \alpha/\lambda$  – коэффициент теплообмена,  $\theta(t)$  – некоторая заданная функция.

Возможны также и иные виды краевых условий, соответствующие иным физическим ситуациям. Разумеется, допустимы различные комбинации условий, например, на левом конце стержня известна температура, а на правом – поток тепла и т.д.

Более сложный (нелинейный) вариант условий на торцах отвечает сильно нагретому и поэтому излучающему энергию стержню, не контактирующему с какими-либо телами. Тогда в единицу времени стержень теряет на своих границах (торцах) энергию, равную  $\sigma T^4(0, t)$  и  $\sigma T^4(l, t)$  соответственно. В результате получаются условия:

$$\sigma T^4(0, t) = \lambda(T) \frac{\partial T}{\partial x} \Big|_{x=0}, \quad \sigma T^4(l, t) = -\lambda(T) \frac{\partial T}{\partial x} \Big|_{x=l}, \quad t > 0.$$

*Вывод уравнения диффузии.* Если среда неравномерно заполнена газом, то имеет место диффузия его из мест с более высокой концентрацией в места с меньшей концентрацией. Это же явление имеет место и в растворах, если концентрация растворенного вещества в объеме не постоянна.

Рассмотрим процесс диффузии в полой трубке или в трубке, заполненной пористой средой, предполагая, что во всякий момент времени концентрация газа (раствора) по сечению трубки одинакова. Тогда процесс диффузии может быть описан функцией  $c(x, t)$ , представляющей концентрацию в сечении  $x$  в момент времени  $t$ .

Согласно закону Нернста, масса газа, протекающая через сечение  $x$  за промежуток времени  $(t, t + \Delta t)$ , равна

$$dc = -D \frac{\partial}{\partial x} c(x, t) S dt = W S dt, \quad W = -D \frac{\partial c}{\partial x},$$

где  $D$  – коэффициент диффузии;  $S$  – площадь сечения трубки;  $W(x, t)$  – плотность диффузионного потока, равная массе газа, протекающего в единицу времени через единицу площади.

По определению концентрации, количество газа в объеме  $V$  равно

$$G = cV.$$

Отсюда получаем, что изменение массы газа на участке трубки  $(x_1, x_2)$  при изменении концентрации на  $\Delta c$  равно

$$\Delta G = \int_{x_1}^{x_2} \varepsilon(x) \Delta c S dx,$$

где  $\varepsilon(x)$  – коэффициент пористости.

При выводе уравнения диффузии будем считать, что в трубке нет источников вещества, и диффузия через стенки трубки отсутствует.

Составим уравнение баланса массы газа на участке  $(x_1, x_2)$  за промежуток времени  $(t_1, t_2)$ :

$$S \int_{t_1}^{t_2} \left[ D(x_2) \frac{\partial}{\partial x} c(x_2, \tau) - D(x_1) \frac{\partial}{\partial x} c(x_1, \tau) \right] d\tau = S \int_{x_1}^{x_2} \varepsilon(\xi) [c(\xi, t_2) - c(\xi, t_1)] d\xi.$$

Отсюда, подобно выводу уравнения теплопроводности, получим уравнение

$$\frac{\partial}{\partial x} \left( D \frac{\partial c}{\partial x} \right) = \varepsilon \frac{\partial c}{\partial t}, \quad (2.8)$$

являющееся уравнением диффузии. Оно вполне аналогично уравнению теплопроводности.

Если коэффициент диффузии постоянен  $D = \text{const}$ , то уравнение диффузии принимает вид:

$$c_t = a^2 c_{xx},$$

где  $a^2 = D/\varepsilon$ .

Если коэффициент пористости  $\varepsilon = 1$ , а коэффициент диффузии постоянен, то уравнение диффузии имеет вид  $c_t = D c_{xx}$ .

Для получения единственного решения уравнения диффузии необходимо к уравнению присоединить начальные и граничные условия.

При переходе к дискретным моделям теплопроводности и диффузии область непрерывного изменения аргументов  $(x$  и  $t)$  заменяется конечным (дискретным) множеством точек (узлов), называемым сеткой. Вместо функции непрерывного аргумента рассматриваются функции дискретного аргумента, определенные в узлах сетки и называемые сеточными функциями. Производные, входящие в дифференциальные уравнения, заменяются (аппроксимируются) при помощи соответствующих разностных соотношений. В результате дифференциальное уравнение заменяется системой алгебраических (разностных) уравнений. Начальные и краевые условия также заменяются разностными начальными и краевыми условиями.

Естественно требовать, чтобы полученная таким образом разностная краевая задача была разрешима и ее решение при увеличении числа  $N$  узлов сетки приближалось (сходилось) к решению исходной задачи.

Пусть область изменения аргументов  $(x, t)$  есть прямоугольник  $\Pi$  ( $0 \leq x \leq 1, 0 \leq t \leq T$ ). Построим на отрезке  $0 \leq x \leq 1$  сетку  $\{x_i = ih, i = 0, 1, \dots, N_1\}$  с шагом  $h = 1/N_1$  и сетку  $\{t_j = j\tau, j = 0, 1, \dots, N_2\}$  с шагом  $\tau = T/N_2$  на отрезке  $0 \leq t \leq T$ .

Множество узлов  $(x_i, t_j)$  с координатами  $x_i = ih$  и  $t_j = j\tau$  назовем сеткой в прямоугольнике  $\Pi$  и обозначим через  $\omega_{ht}$  сетку  $\{(x_i = ih, t_j = j\tau), i = \overline{0, N_1}; j = \overline{0, N_2}\}$ . Эта сетка равномерна по каждой из переменных  $x$  и  $t$ .

Пусть  $Y$  – сеточная функция, заданная на  $\omega_{ht}$ . Будем обозначать  $Y_i^j = Y(x_i, t_j)$  значение сеточной функции  $Y$  в узле  $(x_i, t_j)$  сетки  $\omega_{ht}$ . Непрерывной функции  $T(x, t)$  или  $c(x, t)$ , где  $(x, t) \in \Pi$ , будем ставить в соответствие сеточную функцию  $Y_i^j = T(x_i, t_j) \vee c(x_i, t_j)$ .

Рассмотрим теперь производную  $\vartheta'_x$  функции  $\vartheta_x$ . Заменить ее разностным выражением можно бесчисленным множеством способов. Простейшими являются замены вида:  $\vartheta'_x \sim \frac{\vartheta_i - \vartheta_{i-1}}{h} = L_h^- \vartheta_i$  – левая разностная производная (левое разностное отношение);  $\vartheta'_x \sim \frac{\vartheta_{i+1} - \vartheta_i}{h} = L_h^+ \vartheta_i$  – правая разностная производная  $\vartheta'_x \sim \frac{\vartheta_{i+1} - \vartheta_{i-1}}{2h} = L_h^0 \vartheta_i$  – центральная разностная производная, где знак  $\sim$  означает соответствие или аппроксимацию.

Обращаясь к формулам для  $L_h^\pm$ , видим, что  $L_h^- \vartheta$  и  $L_h^+ \vartheta$  аппроксимируют  $L \vartheta = \vartheta'$  с первым порядком точности. Выражения для  $L_h^- \vartheta$  содержат значения  $\vartheta$  в двух узлах  $x = x_i$  и  $x = x_{i-1}$  сетки. Говорят, что оператор  $L_h^- \vartheta$  является двухточечным или оператором первого порядка.

Множество узлов, значения сеточной функции в которых входят в выражение  $L_h \vartheta$ , называют шаблоном оператора  $L_h$  в точке  $x_i$ . Очевидно, что шаблон оператора  $L_h^-$  состоит из узлов  $x_i$  и  $x_{i-1}$ , а шаблон  $L_h^+ \vartheta$  – из узлов  $x_i$  и  $x_{i+1}$ .

Возьмем теперь трехточечный оператор, определенный на шаблоне  $x_{i-1}, x_i, x_{i+1}$ :

$$L_h^{(\sigma)} \vartheta_i = \sigma L_h^+ \vartheta_i + (1-\sigma) L_h^- \vartheta_i = \frac{\sigma \vartheta_{i+1} + (1-2\sigma) \vartheta_i - (1-\sigma) \vartheta_{i-1}}{h},$$

где  $\sigma$  – произвольное число. В частности при  $\sigma = 1/2$  получаем центральную разностную производную  $L_h^0$ , которая аппроксимирует  $\vartheta'(x)$  со вторым порядком.

Рассмотрим теперь вторую производную  $L\vartheta = \vartheta''$ . Выберем трехточечный шаблон, состоящий из узлов  $x_{i-1}, x_i, x_{i+1}$ , и рассмотрим разностный оператор

$$L_h \vartheta_i = \frac{1}{h} (L_h^+ \vartheta_i - L_h^- \vartheta_i) = \frac{\vartheta_{i+1} - \vartheta_i}{h} - \frac{\vartheta_i - \vartheta_{i-1}}{h} = \frac{\vartheta_{i+1} - 2\vartheta_i + \vartheta_{i-1}}{h^2}.$$

На практике аппроксимация производных на многоточечных шаблонах используется редко, так как при увеличении шаблона обычно увеличивается объем вычислительной работы и ухудшается качество получающихся разностных операторов (в смысле устойчивости).

Рассмотрим более сложный оператор  $Lu = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial t^2}$ , где  $u = u(x, t)$  – функция двух аргументов  $x$  и  $t$ , меняющаяся в области  $\Pi(0 \leq x \leq 1, 0 \leq t \leq T)$ . Введем сетку  $\{(x_i = ih, t_j = j\tau), i = \overline{0, N_1}; j = \overline{0, N_2}\}$  с шагами  $h = 1/N_1$ ;  $\tau = T/N_2$ . Произведем замену:

$$\frac{\partial u}{\partial t} \sim \frac{u_i^{j+1} - u_i^j}{\tau} = u_{i,i}^{j+1}; \quad \frac{\partial^2 u}{\partial t^2} \sim \frac{u_{i-1}^j - 2u_i^j + u_{i+1}^j}{h^2} = u_{xx,i}^j.$$

В результате получим разностный оператор

$$L_{h\tau} u_i^{j+1} = \frac{u_i^{j+1} - u_i^j}{\tau} - \frac{u_{i-1}^j - 2u_i^j + u_{i+1}^j}{h^2}.$$

Этот оператор определен на шаблоне, состоящем из четырех точек  $(x_i, t_{j+1}), (x_i, t_j), (x_{i-1}, t_j), (x_{i+1}, t_j)$  (рис. 2.14, а).

Оператор  $L_{h\tau}$  определен не во всех узлах  $\omega_{h\tau}$ , а только при  $0 < i < N$  и  $j > 0$ , т.е. во внутренних узлах. В остальных узлах, называемых граничными, должны быть заданы начальные и краевые условия. Оператор  $L_{h\tau}$  имеет первый порядок аппроксимации по  $\tau$  и второй по  $h$ :

$$\max_{\omega_{h\tau}} |L_{h\tau} u_i^j - (Lu)_i^j| = O(h^2 + \tau).$$

Аппроксимируем этот же оператор  $Lu$  на шаблоне, показанном на рис. 2.14, б.

В результате получим оператор

$$L_{h\tau} u_i^{j+1} = \frac{u_i^{j+1} - u_i^j}{\tau} - \frac{u_{i-1}^{j+1} - 2u_i^{j+1} + u_{i+1}^{j+1}}{h^2},$$

аппроксимирующий  $Lu$  с тем же порядком точности, что и предыдущий оператор.

*Пример 2.4.* Рассмотрим постановку разностной задачи для уравнения теплопроводности:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + f(x, t), \quad 0 < x < l, \quad 0 < t < t_k;$$

$$T(x, 0) = T_0(x), \quad 0 \leq x \leq l; \quad T(0, t) = \mu_1(t); \quad T(l, t) = \mu_2(t), \quad 0 \leq t \leq t_k.$$

Введем равномерную сетку  $\omega_{h\tau} = \{(x_i = ih, t_j = j\tau), i = \overline{0, N_1}; j = \overline{0, N_2}\}$  и запишем соответствующую разностную краевую задачу:

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{y_{i-1}^j - 2y_i^j + y_{i+1}^j}{h^2} + \varphi_i^{j+1}, \quad 0 < i < N, \quad j > 0;$$

$$y_i^0 = T_0(x_i); \quad y_0^j = \mu_1(t_j); \quad y_N^j = \mu_2(t_j),$$

где  $\varphi_i^{j+1} = f(x_i, t_j)$ .

Определим

$$y_i^{j+1}: y_i^{j+1} = (1-2\gamma)y_i^j + \gamma(y_{i-1}^j + y_{i+1}^j) + \tau\varphi_i^{j+1},$$

где  $\gamma = \tau/h^2$ .

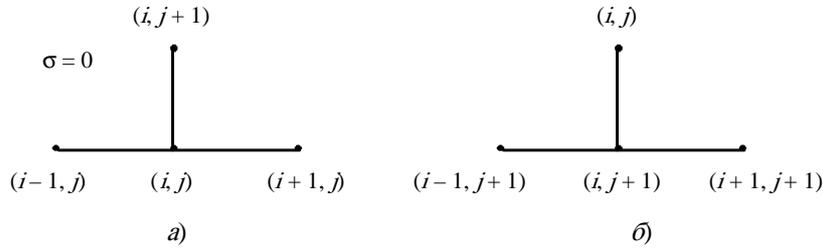


Рис. 2.14. Различные варианты (а, б) четырехточечного шаблона

Если  $y_i^j$  известно, то по этой формуле можно определить  $y_i^{j+1}$  во всех узлах  $i = 1, 2, \dots, N_1 - 1$  (на слое  $j + 1$ ). Так как при  $j = 0$  задано начальное условие  $y_i^0 = T_0(x_i)$ , то последняя формула позволяет определить от слоя к слою значения  $y_i^{j+1}$  во всех внутренних узлах сетки  $\omega_{ht}$ , используя при этом краевые условия. В этом случае полученная разностная схема называется *явной*.

Если выбрать другой шаблон, то разностная краевая задача примет вид

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{y_{i-1}^{j+1} - 2y_i^{j+1} + y_{i+1}^{j+1}}{h^2} + \varphi_i^{j+1}.$$

В этом случае для определения  $y_i^{j+1}$  на новом слое  $j + 1$  получаем систему алгебраических уравнений:

$$\gamma y_{i-1}^{j+1} - (1 + 2\gamma)y_i^{j+1} + \gamma y_{i+1}^{j+1} = -y_i^j - \tau \varphi_i^{j+1}, \quad 0 < i < N_1.$$

Такая схема называется неявной или схемой с опережением.

После того, как разностная схема написана, возникает, прежде всего, вопрос о разрешимости полученной алгебраической системы уравнений. Если эта система неразрешима, то такую схему следует признать непригодной.

Пусть разностная задача разрешима. Тогда естественно требовать, чтобы при неограниченном измельчении сетки решение разностной задачи стремилось к решению исходной задачи для дифференциального уравнения (схема сходилась). При этом предполагаем, что разностная задача решается точно, и решение может быть найдено с любым числом знаков. Практически же все вычисления ведутся с конечным числом знаков и на каждом этапе вычислений допускаются ошибки округления. Если малые ошибки округления, допускаемые на промежуточных этапах вычислительного процесса, при сгущении сетки приводят к большим искажениям решения, то такую схему называют неустойчивой. Она непригодна для практики.

Ошибки вычисления можно рассматривать как возмущения начальных данных или правой части уравнения. Отсюда следует, что от схемы надо требовать, чтобы решение разностной задачи мало менялось при малом изменении входных данных задачи (правой части краевых и начальных условий) или, иными словами, чтобы решение непрерывно зависело от входных данных при измельчении сетки. Если это требование выполняется, то такая схема называется *устойчивой*, в противном случае схема *неустойчива*.

*Разностные схемы для нелинейных уравнений теплопроводности (диффузии)*. При написании разностных уравнений естественно исходить из уравнения баланса, которое содержит интегралы от функций и ее производных:

$$\int_{x_1}^{x_2} c[T(x, t_2) - T(x, t_1)] dx = \int_{t_1}^{t_2} [W(x_1, t) - W(x_2, t)] dt + \int_{x_1}^{x_2} \int_{t_1}^{t_2} f(x, t) dx dt,$$

где  $T(x, t)$  – температура;  $c$  – объемная теплоемкость;  $f(x, t)$  – плотность источников тепла;  $W(x, t) = -\lambda(x, t) \frac{\partial T}{\partial x}(x, t)$  – тепловой поток;  $\lambda(x, t, T)$  – коэффициент теплопроводности.

Если существуют непрерывные производные  $\frac{\partial T}{\partial t}$  и  $\frac{\partial}{\partial x} \left( \lambda \frac{\partial T}{\partial x} \right)$ , то из уравнения баланса следует дифференциальное уравнение теплопроводности:

$$c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( \lambda(x, t) \frac{\partial T}{\partial x} \right) + f(x, t).$$

Пусть дана сетка ( $x_i = ih$ ,  $t_j = j\tau$ ). Для каждой элементарной ячейки (прямоугольника) этой сетки пишется уравнение баланса, которое содержит интегралы от функции и ее производных вдоль границы ячейки. Для их вычисления необходимо предположение о профиле функций. В зависимости от выбора локальной интерполяции как по  $x$ , так и по  $t$  мы получим различные схемы. Вопрос о выборе интерполяции подчинен требованиям устойчивости, точности и простоты реализации.

*Пример 2.5.* Рассмотрим стационарное уравнение теплопроводности

$$\frac{\partial}{\partial x} \left( \lambda(x, t) \frac{\partial T}{\partial x} \right) - q(x)T = -f(x, t), \quad 0 < x < 1, \quad k > 0, \quad q \geq 0,$$

где  $q(x)T$  – мощность стоков тепла (при  $q \leq 0$  – источников), пропорциональная температуре  $T(x)$ .

Выберем на отрезке  $0 \leq x \leq 1$  сетку  $\omega_h = \{x_i = ih, i = \overline{0, N_1}\}$  с шагом  $h$ .

Напишем уравнение баланса тепла на отрезке  $x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}$ ,  $x_{i-\frac{1}{2}} = \frac{1}{2}(x_{i-1} + x_i) = x_{i-1} + \frac{h}{2}$ :

$$W_{i-\frac{1}{2}} - W_{i+\frac{1}{2}} - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)T(x)dx + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)dx = 0.$$

Возьмем простейшую аппроксимацию  $T = \text{const} = T_i$  при  $x_{i-\frac{1}{2}} \leq x \leq x_{i+\frac{1}{2}}$ :

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)T(x)dx \approx h d_i T_i, \quad d_i = \frac{1}{h} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)dx.$$

Проинтегрируем равенство  $\frac{dT}{dx} = -\frac{W}{\lambda}$  на отрезке  $x_{i-1} \leq x \leq x_i$ :

$$T_{i-1} - T_i = \int_{x_{i-1}}^{x_i} \frac{W}{\lambda} dx.$$

Полагая  $W = \text{const} = W_{i-\frac{1}{2}}$  при  $x_{i-1} \leq x \leq x_i$ , будем иметь:

$$T_{i-1} - T_i = W_{i-\frac{1}{2}} \int_{x_{i-1}}^{x_i} \frac{dx}{\lambda(x)} \quad \text{или} \quad W_{i-\frac{1}{2}} = -a_i \frac{T_{i-1} - T_i}{h}; \quad a_i = \frac{1}{\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{\lambda(x)}}.$$

Отметим, что  $\int_{x_{i-1}}^{x_i} \frac{dx}{\lambda(x)}$  есть тепловое сопротивление отрезка  $[x_{i-1}, x_i]$ .

Заменяя интеграл по одной из формул  $\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{\lambda(x)} \approx \frac{1}{\lambda_{i-\frac{1}{2}}}$ ,  $\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{\lambda(x)} \approx \frac{1}{2} \left( \frac{1}{\lambda_{i-1}} + \frac{1}{\lambda_i} \right)$ , получим  $a_i = \lambda_{i-\frac{1}{2}}$ ,  $a_i = \frac{2\lambda_{i-1}\lambda_i}{\lambda_{i-1} + \lambda_i}$  и

т.д.

В результате получим разностную схему вида:

$$\frac{1}{h} \left[ \frac{a_{i+1}(y_{i+1} - y_i)}{h} - \frac{a_i(y_i - y_{i-1})}{h} \right] - d_i y_i = -\varphi_i,$$

$$\text{где } d_i = \frac{1}{h} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} q(x)dx, \quad \varphi_i = \frac{1}{h} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x)dx.$$

Метод баланса, таким образом, позволяет получать схемы, коэффициенты которых во всех узлах сетки вычисляются по одним и тем же формулам как средние значения коэффициентов дифференциального уравнения в окрестности узла сетки.

Все разностные схемы пишутся одинаково во всех узлах сетки и для любых  $\lambda(x)$ ,  $q(x)$ ,  $f(x)$ . Такие схемы называются однородными. Для практических целей целесообразно находить коэффициенты схемы  $a$ ,  $d$ ,  $\varphi$  по более простым формулам, используя значения  $\lambda$ ,  $q$ ,  $f$  в отдельных точках. Обычно используют шаблоны из одной или из двух точек, полагая, например:

$$a_i = \lambda_{i-\frac{1}{2}}, \quad d_i = q_i, \quad \varphi_i = f_i,$$

если  $\lambda$ ,  $q$ ,  $f$  непрерывны. Если  $\lambda$ ,  $q$ ,  $f$  разрывны, то в этих формулах следует брать полусумму предельных значений слева и справа.

Рассмотрим одномерное с начальными и граничными условиями параболическое уравнение в частных производных, описывающее процессы теплопроводности (2.7) и диффузии (2.8), и обсудим алгоритм его решения.

*Пример 2.6.* Запишем в общем виде одномерное параболическое уравнение с краевыми условиями.

$$c(x, t, u) \frac{\partial u}{\partial t} = \mu(x) \frac{\partial}{\partial x} \left( v(x) \lambda(x, t, u) \frac{\partial u}{\partial x} \right) + f(x, t, u), \quad (2.9)$$

$$u(x, t_0) = u_0(x), \quad (2.10)$$

$$r_1 \lambda(x, t, u) \frac{\partial u}{\partial x} \Big|_{x=l} = z_1(u) u \Big|_{x=l} - \theta_1(t, u) \Big|_{x=0}, \quad (2.11)$$

$$r_2 \lambda(x, t, u) \frac{\partial u}{\partial x} \Big|_{x=l} = z_2(u) u \Big|_{x=l} - \theta_2(t, u) \Big|_{x=l}, \quad (2.12)$$

где  $u(x, t)$  – температура (концентрация);  $\lambda(x, t, u)$  – коэффициент теплопроводности (диффузии);  $c(x, t, u)$  – объемная теплоемкость;  $f(x, t, u)$  – плотность мощности источников вещества и тепла.

Постановка (2.9) – (2.12) допускает рассмотрение цилиндрических и сферических областей при наличии радиальной симметрии за счет выбора функций  $\mu(x), \nu(x)$  (например, при  $\mu(x) = \frac{1}{x}, \nu(x) = x$  получаем цилиндрическую область, а при  $\mu(x) = \frac{1}{x^2}, \nu(x) = x^2$  – сферическую область).

Представление решения нелинейной краевой задачи (2.9) – (2.12) в аналитической форме возможно лишь в исключительных случаях. Универсальным методом приближенного решения является метод конечных разностей.

Приступим теперь к построению разностной схемы для краевой задачи (2.9) – (2.12). Пусть  $n$  и  $m$  – фиксированные натуральные числа. Введем на отрезке  $[0, l]$  и  $[0, T]$  сетки  $\Sigma_{nx}$  и  $\Sigma_{mt}$  с узлами  $0 = x_0 < x_1 < \dots < x_n = L, 0 = \tau_0 < \tau_1 < \dots < \tau_m = T$ . Эти сетки  $\Sigma_{nx}^m = \Sigma_{nx} \times \Sigma_{mt}$  определяют сетку в области  $\Pi$ . Положим  $h_i = x_i - x_{i-1}, i = \overline{1, n}, \tau_j = t_j - t_{j-1}, j = \overline{1, m}$ . На этой сетке функция  $u(x, t)$  аппроксимируется сеточной функцией  $y_{ij}$ , определённой в узлах сетки  $(x_i, t_j)$ .

Примем следующие обозначения:

$$C_i = C(x_i, t_{j+1}, y_i); \quad y_i = y_i^{j+1}; \quad \hat{y}_i = y_i^j;$$

$$\lambda_{i \pm \frac{1}{2}} = \frac{1}{2} [\lambda(x_i, t_{j+1}, y_i) + \lambda(x_{i \pm 1}, t_{j+1}, y_{i \pm 1})];$$

$$\mu_i = \mu(x_i); \quad \nu_{i \pm \frac{1}{2}} = \nu\left(\frac{x_i + x_{i \pm 1}}{2}\right); \quad \Phi_i = f(x_i, t_{j+1}, y_i).$$

Для построения неявной разностной схемы будем использовать четырехточечный шаблон. В результате получим неявную разностную схему с точностью  $O(h^2 + \tau)$ :

$$C_i \frac{y_i - \hat{y}_i}{\tau_{j+1}} = \frac{2}{h_i + h_{i+1}} \mu_i \left[ \nu_{i+\frac{1}{2}} \lambda_{i+\frac{1}{2}} \frac{y_{i+1} - y_i}{h_{i+1}} - \nu_{i-\frac{1}{2}} \lambda_{i-\frac{1}{2}} \frac{y_i - y_{i-1}}{h_i} \right] + \Phi_i; \quad (2.13)$$

$$y_i(t_0) = u_0(x), \quad i = \overline{1, n}; \quad j = 0; \quad (2.14)$$

$$\frac{r_1 h_i}{2} \left( C_0 \frac{y_0 - \hat{y}_0}{\tau_{j+1}} + \Phi_0 \right) = r_1 \mu_0 \nu_{\frac{1}{2}} \lambda_{\frac{1}{2}} \frac{y_1 - y_0}{h_1} - \mu_0 \nu_0 (z_1(y_0) y_0 - \theta_1(t_{j+1}, y_0)), \quad (2.15)$$

$$0 \leq j \leq m;$$

$$\frac{r_2 h_{n-1}}{2} \left( C_n \frac{y_n - \hat{y}_n}{\tau_{j+1}} + \Phi_n \right) = r_2 \mu_n \nu_{n-\frac{1}{2}} \lambda_{n-\frac{1}{2}} \frac{y_n - y_{n-1}}{h_{n-1}} - \mu_n \nu_n (z_2(y_n) y_n - \theta_2(t_{j+1}, y_n)). \quad (2.16)$$

Упростим разностную схему (2.13) – (2.16), предполагая  $h = \text{const}, \tau = \text{const}$ :

$$C_i \frac{y_i^{j+1} - y_i^j}{\tau} = \frac{\mu_i}{h^2} \left[ \nu_{i+\frac{1}{2}} \lambda(y^{j+1}) (y_{i+1}^{j+1} - y_i^{j+1}) - \nu_{i-\frac{1}{2}} \lambda(y^{j+1}) (y_i^{j+1} - y_{i-1}^{j+1}) \right] + \Phi(y_i^{j+1}). \quad (2.17)$$

Далее, полагая, что  $\lambda = \lambda(y^j), \Phi = \Phi(y^j)$ , получим схему, линейную относительно  $y_i^{j+1}$ :

$$C_i \frac{y_i^{j+1} - y_i^j}{\tau} = \frac{\mu_i}{h^2} \left[ \nu_{i+\frac{1}{2}} \lambda(y^j) (y_{i+1}^{j+1} - y_i^{j+1}) - \nu_{i-\frac{1}{2}} \lambda(y^j) (y_i^{j+1} - y_{i-1}^{j+1}) \right] + \Phi(y_i^{j+1}). \quad (2.18)$$

Решение разностной краевой задачи для  $y_i^{j+1}$  находится методом прогонки [5].

Разностная схема (2.17) нелинейна относительно  $y_i^{j+1}$ . Для решения получающейся системы нелинейных уравнений применяются итерационные методы.

Преобразуем разностную схему (2.17), вводя обозначения:

$$A_i^{(s)} = \frac{\tau\mu_i}{h^2} v_{i-\frac{1}{2}} \lambda_{i-\frac{1}{2}}(y_i^{(s)}); \quad B_i^{(s)} = c_i + \frac{\tau\mu_i}{h^2} v_{i+\frac{1}{2}} \lambda_{i+\frac{1}{2}}(y_i^{(s)}) +$$

$$+ \frac{\tau\mu_i}{h^2} v_{i-\frac{1}{2}} \lambda_{i-\frac{1}{2}}(y_i^{(s)});$$

$$D_i^{(s)} = \frac{\tau\mu_i}{h^2} v_{i+\frac{1}{2}} \lambda_{i+\frac{1}{2}}(y_i^{(s)}); \quad F_i^{(s)} = c_i y_i^{(s)} + \tau\varphi(y_i^{(s)});$$

где  $S = 0, 1, 2, \dots$  – номер итерации.

В результате получим систему нелинейных алгебраических уравнений вида

$$A_i^{(s)} y_{i-1}^{(s+1)} - B_i^{(s)} y_i^{(s+1)} + D_i^{(s)} y_{i+1}^{(s+1)} = -F_i^{(s)}. \quad (2.19)$$

В качестве нулевого приближения обычно берут значение  $y_i^j$  с предыдущего временного слоя  $y_i^{(0)} = y_i^j$ .

Решение уравнения (2.19) относительно  $y_i^{(s+1)}$  с краевыми условиями при  $i = 0, i = N$  находится итерационным методом. Для окончания итераций используется условие  $\max_{1 \leq i \leq N-1} |y_i^{(s+1)} - y_i^{(s)}| < \varepsilon$  или же задается определенное число итераций.

Обычно уже две-три итерации заметно повышают точность. Неявные схемы вида (2.17) позволяют для обеспечения заданной точности использовать более крупный шаг по времени по сравнению с линейными (безытерационными) схемами (2.18), что зачастую приводит к значительному уменьшению объема вычислительной работы.

*Решение систем разностных уравнений методом прогонки.* Неявные схемы (2.17, 2.18) для уравнения теплопроводности приводят к системе алгебраических уравнений относительно искомой функции  $y_i = y_i^j$  на новом временном слое  $t = t_{j+1}$ . Эта система уравнений имеет вид:

$$A_i y_{i-1} - B_i y_i + D_i y_{i+1} = -F_i, \quad i = \overline{1, n-1}; \quad (2.20)$$

$$y_0 = \chi_1 y_1 + \vartheta_1, \quad (2.21)$$

$$y_n = \chi_2 y_{n-1} + \vartheta_2,$$

где

$$A_i = \frac{2\tau_{j+1}}{(h_i + h_{i+1})h_i} v_{i-\frac{1}{2}} \lambda_{i-\frac{1}{2}}; \quad D_i = \frac{2\tau_{j+1}}{(h_i + h_{i+1})h_{i+1}} v_{i+\frac{1}{2}} \lambda_{i+\frac{1}{2}}; \quad F_i = \varphi_i \tau_{j+1} + c_i y_i;$$

$$\chi_1 = \frac{\frac{r_1 \mu_0 v_{\frac{1}{2}} \lambda_{\frac{1}{2}}}{h_1}}{r_1 \mu_0 v_{\frac{1}{2}} \lambda_{\frac{1}{2}} C_0 + \frac{r_1 h_1}{2} + \mu_0 v_{\frac{1}{2}} z_1(y_0)};$$

$$\vartheta_1 = \frac{\mu_0 v_{\frac{1}{2}} \theta_1(t_{j+1}, y_0) + \frac{r_1 C_0 h_1}{2\tau_{j+1}} \hat{y}_0 - \frac{r_1 h_1}{2} \varphi_0}{r_1 \mu_0 v_{\frac{1}{2}} \lambda_{\frac{1}{2}} C_0 + \frac{r_1 h_1}{2} + \mu_0 v_{\frac{1}{2}} z_1(y_0)};$$

$$\chi_2 = \frac{\frac{r_2 \mu_n v_{n-\frac{1}{2}} \lambda_{n-\frac{1}{2}}}{h_1}}{r_2 \mu_n v_{n-\frac{1}{2}} \lambda_{n-\frac{1}{2}} C_n + \frac{r_2 h_{n-1}}{2} + \mu_n v_n z_2(y_n)};$$

$$\vartheta_2 = \frac{\mu_n v_n \theta_2(t_{j+1}, y_n) + \frac{r_2 C_n h_{n-1}}{2\tau_{j+1}} \hat{y}_n - \frac{r_2 h_n}{2} \varphi_n}{r_2 \mu_n v_{n-\frac{1}{2}} \lambda_{n-\frac{1}{2}} C_n + \frac{r_2 h_{n-1}}{2} + \mu_n v_n z_2(y_n)}.$$

Задача (2.20)–(2.21) разрешима, если  $A_i > 0, D_i > 0, B_i > A_i + D_i, \quad 0 \leq \chi_{1,2} < 1$ . Для нахождения ее решения можно применить обычные методы линейной алгебры или методы итераций. Однако наиболее выгодным или экономичным по объему затрачиваемой работы является метод прогонки или метод факторизации, учитывающий специальный вид матрицы системы уравнений (2.20) – ее трехдиагональность.

Будем искать решение задачи (2.20) – (2.21) в виде

$$y_i = \eta_{i+1} y_{i+1} + \gamma_{i+1}, \quad i = \overline{0, n-1}; \quad (2.22)$$

где  $\eta_i, \gamma_i$  – неизвестные пока функции. Подставляя  $y_{i-1} = \eta_i y_i + \gamma_i$  в (2.20), исключим  $y_{i-1}$  и получим  $(A_i \eta_i - B_i) \eta_{i+1} + D_i y_{i+1} + (A_i \gamma_i + F_i) = 0$ , после чего при помощи (2.22) исключим  $y_i$ :

$$[(A_i \eta_i - B_i) \eta_{i+1} + D_i] y_{i+1} + [(A_i \eta_i - B_i) \gamma_{i+1} + (A_i \gamma_i + F_i)] = 0.$$

Уравнение будет удовлетворено, если выражения в квадратных скобках равны нулю. Из этих двух равенств находим рекуррентные формулы для определения  $\eta_{i+1}, \gamma_{i+1}$ :

$$\eta_{i+1} = \frac{D_i}{B_i - A_i \eta_i}, \quad \gamma_{i+1} = \frac{A_i \gamma_i + F_i}{B_i - A_i \eta_i}, \quad i = \overline{1, n-1}. \quad (2.23)$$

Сравнивая формулу  $y_0 = \eta_1 y_1 + \gamma_1$  с краевым условием  $y_0 = \chi_1 y_1 + \vartheta_1$ , находим

$$\eta_1 = \chi_1, \quad \gamma_1 = \vartheta_1. \quad (2.24)$$

Далее, решая (2.23) с начальными условиями (2.24), найдем  $\eta_i, \gamma_i, i = 2, 3, \dots, n$ .

Определим  $y_n$  через  $\eta_n, \gamma_n$  из краевого условия (2.21) при  $i = n$ . Исключая  $y_{n-1}$  из формул  $y_{n-1} = \eta_n y_n + \gamma_n$  и  $y_n = \chi_2 y_{n-1} + \vartheta_2$ , находим

$$y_n = \frac{\vartheta_2 + \chi_2 \gamma_n}{1 - \chi_2 \eta_2} \quad (2.25)$$

при условии, что  $1 - \chi_2 \eta_2 \neq 0$ .

Из условий разрешимости системы (2.20), (2.21) следует, что  $0 \leq \eta_i < 1$  для всех  $i = \overline{1, n}$ .

Алгоритм решения задачи (2.20), (2.21):

1) по начальным данным (2.24) и формулам (2.23) последовательно определяются  $\eta_i$ , затем  $\gamma_n$  для  $i = 1, 2, 3, \dots, n$  (счет идет слева направо – от  $i$  к  $i+1$ );

2) Из (2.25) находится  $y_n$  и затем по формуле (2.22) последовательно (справа налево – от  $i+1$  к  $i$ ) определяются  $y_{n-1}, y_{n-2}, \dots, y_1, y_0$ .

Счёт по формулам (2.22) устойчив, так как  $0 \leq \eta_i < 1$ .

Существует еще один вариант формул прогонки:

$$\alpha_i = \frac{A_i}{B_i - D_i \alpha_{i+1}}, \quad i = 1, 2, \dots, n-1, \quad \alpha_n = \chi_2; \quad (2.26)$$

$$\beta_i = \frac{D_i \beta_{i+1} + F_i}{B_i - D_i \alpha_{i+1}}, \quad i = 1, 2, \dots, n-1, \quad \beta_n = \vartheta_2; \quad (2.27)$$

$$y_{i+1} = \alpha_{i+1} y_i + \beta_{i+1}, \quad i = 1, 2, \dots, n-1, \quad y_0 = \frac{\vartheta_1 + \chi_1 \beta_1}{1 - \chi_1 \alpha_1}. \quad (2.28)$$

Алгоритм: 1) по формулам (2.26) и (2.27) последовательно от  $i+1$  к  $i$  (справа налево) определяются сначала  $\alpha_i$ , затем  $\beta_i$  для  $i = n-1, n-2, \dots, 1, 0$ ; 2) по формулам (2.28) последовательно от  $i$  к  $i+1$  (слева направо) находятся  $y_1, y_2, \dots, y_n$ .

Нетрудно убедиться в том, что число арифметических операций, производимых при решении задачи (2.20), (2.21), пропорционально числу уравнений.

## 2.5. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ НЕСТАЦИОНАРНЫХ ПРОЦЕССОВ ПЕРЕНОСА

Наибольшую сложность для математического моделирования представляют периодические режимы работы аппаратов и переходные процессы. Поля определяющих характеристик при этом являются нестационарными. Необходимость рассмотрения нестационарных и переходных процессов принципиально усложняет технологические расчеты, так как инженерные (т.е. упрощенные) методики расчета таких процессов дают еще менее надежные, а то и просто неприемлемые результаты.

В настоящее время для выполнения технологических расчетов локальных единиц оборудования повсеместно продолжают использоваться рекомендованные нормативными документами (ГОСТ, РСТ, ОСТ, РД, РТМ) методики расчета тепловых, массообменных и гидродинамических процессов.

Общие недостатки этих расчетных методик следующие:

- каждый процесс рассматривается обособленно от остальных процессов, составляющих единое производство, что затрудняет учет их взаимовлияния;
- методики реализуют решение локальных оптимизационных задач;
- использование в расчетах усредненных и неизменных во времени или пространстве характеристик процессов не позволяет оценить погрешность каждого отдельного расчета, поскольку она будет зависеть от конкретных условий протекания процессов;

– методики расчета времени протекания процессов, как правило, позволяют оценить время только регулярных процессов.

Наилучшее представление о закономерностях процесса, его особенностях может быть получено на основе математического моделирования полей определяющих параметров (температур, концентраций, скоростей и др.).

Технологические расчеты с использованием полей определяющих параметров имеют ряд принципиальных преимуществ, а именно:

- расчеты стационарных, квазистационарных, нестационарных, периодических и переходных процессов могут выполняться по унифицированным методикам;
- поля определяющих параметров могут быть рассчитаны как для отдельных единиц оборудования, так и для групп совместно работающих аппаратов;
- возможен учет локальных значений характеристик исследуемых процессов;
- при известных полях определяющих параметров возможен расчет всех остальных параметров процессов, таких, как потоки, градиенты, амплитуды, интегральные характеристики определяющих параметров.

Значения тепловых потоков определяют интенсивность и продолжительность тепловых процессов. Значения температурных градиентов в продукте, как правило, определяют его качественные показатели. Интегральные теплоты и потери определяют основную составляющую эксплуатационных затрат на осуществление процесса.

Использование современных средств компьютерной техники делает такие расчеты не только практически реализуемыми, но и необходимыми при решении задач разработки и оптимизации современного промышленного химического производства.

Для математического моделирования полей определяющих параметров целесообразно использовать аналитические методы везде, где они применимы.

Преимущества использования аналитических методов очевидны:

- независимость объема вычислений от значений пространственных и временных координат (как следствие – отсутствие накопления систематических расчетных погрешностей);
- не вызывающий дополнительных затруднений расчет потоков, средних и локальных значений, балансных соотношений, присутствующих в любой модели;
- возможность использования решений частных задач, полученных ранее;
- возможность использования унифицированного набора задач для моделирования класса процессов в соответствующем оборудовании;
- возможность анализа и упрощения решений для характерных и предельных значений параметров процесса;
- наглядность и «физичность» промежуточных и конечных расчетных результатов.

В настоящее время для класса процессов в реализующем их оборудовании разработана унифицированная методика расчета полей определяющих параметров на основе аналитических решений систем линейных дифференциальных уравнений в частных производных с соответствующими условиями однозначности.

Методика предназначена для использования как при решении задачи разработки оптимального аппаратного оформления новых или реконструируемых производств, так и для решения локальных задач определения оптимальных конструктивных и режимных параметров единиц оборудования.

Разработанная методика позволяет осуществлять комплексный расчет как совместно работающих, так и отдельно взятых единиц оборудования. В частном случае разработанная методика может быть использована для определения длительности процессов, протекающих в нестационарных или переходных режимах.

Рассмотрим возможности аналитических подходов к математическому моделированию полей определяющих параметров на примере температурных полей. Пространственные нестационарные температурные поля в производственном промышленном оборудовании могут моделироваться решениями класса задач теплопроводности, рассматриваемыми в данной работе.

В самом общем случае пространственное нестационарное температурное поле может быть описано дифференциальным уравнением Фурье – Кирхгофа. Без учета переноса тепла диффузионной теплопроводностью, которым обычно пренебрегают вследствие его малости по сравнению с другими составляющими, уравнение имеет вид:

$$c_p \rho \frac{dt}{d\tau} = \text{div}(\lambda \nabla t) + Q_v + \frac{dp}{d\tau} + \eta \Phi_v + S_v,$$

где  $t = t(\alpha, \beta, \gamma, \tau)$  – определяемая температура, как функция пространственных координат  $\alpha, \beta, \gamma$  и времени  $\tau$ ;  $c_p$  – удельная теплоемкость;  $\rho$  – плотность;  $\tau$  – время;  $\lambda$  – коэффициент теплопроводности;  $Q_v$  – суммарная удельная мощность объемных источников тепла;  $dp/d\tau$  – работа сил давления;  $\eta$  – коэффициент вязкости;  $\Phi_v$  – диссипативная функция Релея;  $S_v$  – суммарная удельная работа внешних сил в процессе диффузионного переноса.

Это уравнение описывает температурное поле на основе фундаментальных законов переноса тепла в пространстве с учетом всех тепловых эффектов, которые встречаются при эксплуатации промышленного производственного оборудования.

К ним относятся следующие составляющие:

- теплоты разбавления и концентрирования растворов ( $Q_v$ );
- теплота фазовых переходов ( $Q_v$ );
- тепловые эффекты химических превращений ( $Q_v$ );
- теплота, приносимая перемешивающими устройствами ( $\Phi_v$ );
- теплота внутреннего трения в потоках продуктов и теплоносителей ( $\Phi_v$ );
- теплота, приносимая внешними электромагнитными, электрическими, акустическими и другими воздействиями ( $Q_v, S_v$ );
- теплота, приносимая работой сил давления ( $dp/d\tau$ ).

Знание температурного поля позволяет определить тепловые потоки, температурные градиенты, интегральные теплоты,

а также как локальные, так и интегральные тепловые потери в окружающую среду, в том числе через монтажные элементы.

Значения тепловых потоков определяют интенсивность и продолжительность тепловых процессов. Значения температурных градиентов в продукте, как правило, определяют его качественные показатели. Интегральные теплоты и потери определяют основную составляющую эксплуатационных затрат на осуществление процесса.

Очевидно, что непосредственное использование дифференциального уравнения Фурье – Кирхгофа для прикладных расчетов практически невозможно по ряду причин. Во-первых, наличие конвективных членов исключает возможность его прямого решения в общем виде не только аналитическими, но и численными методами. Во-вторых, уравнение является нелинейным, т.е. ряд коэффициентов и членов уравнения являются функциями либо температуры, либо только координат и времени. Это также исключает возможность получения прямого аналитического решения и значительно затрудняет использование численных методов.

Рассмотрим возможность использования аналитических подходов к математическому моделированию тепловых процессов в промышленном оборудовании, которые могут быть описаны дифференциальным уравнением Фурье – Кирхгофа.

Избежать использования нелинейных постановок задач теплопроводности, не допускающих в общем случае аналитических решений и значительно усложняющих численные, позволяет очевидный прием дискретизации пространственных и временных координат совместно с допущением о постоянстве характеристик теплового процесса внутри локальной пространственно-временной области.

Только допущение о постоянстве коэффициента теплопроводности внутри рассматриваемой области позволяет упростить вид слагаемого, учитывающего диффузионный перенос тепла в пространстве.

Такой подход, позволяющий свести решение нелинейной задачи теплопроводности к совокупности решений соответствующих линейных задач, имеет свои особенности и налагает ряд специфических требований на постановки линейных задач.

Пространственная и временная дискретизация области осуществления теплообменного процесса как способ линеаризации постановок задач теплопроводности и диффузии лишь внешне напоминает пространственную и временную дискретизацию, лежащую в основе численных методов решения дифференциальных уравнений.

Во-первых, локальная область объединяет области с различными механизмами переноса тепла и массы в пространстве. Во-вторых, поля определяющих параметров локальной области моделируются аналитическими функциями, являющимися решениями соответствующих дифференциальных уравнений в частных производных, тогда как при использовании численных схем для каждой узловой точки пространственно-временной области решается система алгебраических уравнений, приближенно отражающая закономерности протекания процессов переноса, использующая известные значения исследуемого параметра в соседних узловых точках и позволяющая найти фиксированное значение исследуемого параметра в текущей узловой точке.

При использовании аналитического решения дифференциального уравнения теплопроводности для каждой локальной области определяется не только температурное поле области, но и все остальные характеристики теплового процесса. Учет тепловых потоков через границы локальных областей и сохранение температурного профиля в конце локального временного интервала при переходе от одной области к другой позволяет исключить систематическую погрешность расчетной методики и значительно снизить погрешности компьютерного счета. В то же время, ряд приемов, характерных для численных методов, может быть с успехом использован и при аналитической дискретизации. Так, конкретные размеры локальных областей по каждой из координат практически могут быть выбраны методом Рунге.

Другой принципиально важный вопрос – исследование возможности отказа в общем случае от рассмотрения конвективных членов, входящих уравнение Фурье – Кирхгофа при моделировании тепловых полей в промышленном производственном оборудовании. Только при отсутствии конвективных составляющих появляется возможность получения аналитических (да и численных тоже) решений данного уравнения. Исключения составляют случаи движения жидкости в режиме идеального вытеснения по каналам постоянного сечения, для которых возможны постановка и решение задач переноса, содержащих конвективные члены.

Таким образом, моделирование температурных полей в оборудовании возможно осуществлять на основе решения сопряженных систем дифференциальных уравнений, включающих одномерные уравнения переноса тепла жидкостью, движущейся в режиме идеального вытеснения, и уравнений теплопроводности для теплопередающих элементов оборудования (в общем случае многомерных).

Вышеизложенные предпосылки положены в основу методологии математического моделирования полей, определяющих параметров оборудования.

Сущность методологии состоит в представлении как стационарных, так и нестационарных полей определяющих параметров производственного оборудования как совокупности полей пространственно-временных локальных областей, моделируемых аналитическими решениями систем сопряженных линейных дифференциальных уравнений в частных производных с соответствующими условиями однозначности.

В общем случае под локальной областью понимается область ограниченных размеров, рассматриваемая в течение короткого интервала времени, внутри которой присутствуют все виды целевого переноса, характерные для текущего процесса, и охватывающая как участки конструктивных элементов аппарата, влияющих на поля определяющих параметров, так и некоторые объемы заполняющих аппарат сред (сырье и продукты, теплоносители и хладагенты, инертные компоненты, насадки, сорбенты, катализаторы и др.) Полная совокупность локальных областей составляет весь рабочий объем аппарата за конечный период времени.

В частных случаях локальная область может охватывать как все поперечное сечение аппарата, так и все внутреннее пространство аппарата.

Общая структура локальной области аппаратов, в которых осуществляется взаимосвязанный теплообмен, представлена на рис. 2.15. Здесь к теплоносителям относятся жидкие и газообразные продукты и инертные компоненты, а также окружающая среда. К областям, в которых осуществляется перенос тепла теплопроводностью, относятся гранулированные компоненты, сорбенты, высушиваемые материалы и катализаторы. К областям, в которых осуществляется перенос вещества молекулярной диффузией, относятся гранулированные, сорбенты и высушиваемые материалы. Учитываются лишь те конструктивные элементы, которые участвуют в теплообмене между теплоносителями.

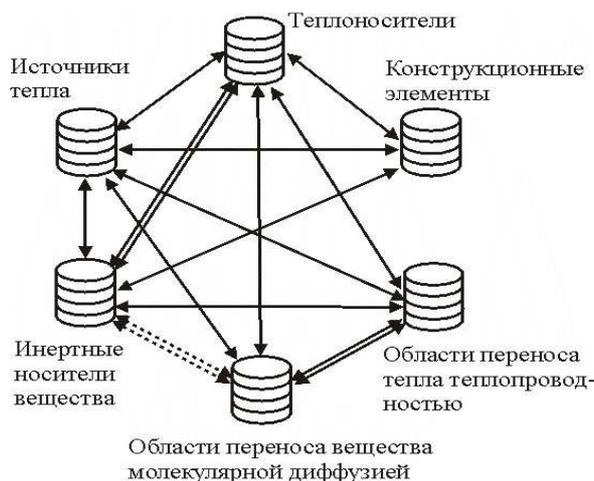


Рис. 2.15.

Конкретный вид локальной области и соответствующая ей математическая модель определяются геометрическими характеристиками оборудования и кинетическими закономерностями протекающих в нем процессов.

### 2.5.1. Математическое моделирование температурного и концентрационного полей локальной области сорбционного оборудования

Рассмотрим методику расчета сорбционного оборудования на основе моделирования температурного и концентрационного полей локальной области адсорбционного аппарата колонного типа (рис. 2.16, 2.17).

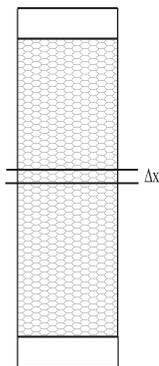


Рис. 2.16. Расположение локальной области сорбционного оборудования

В качестве локальной области выделим тонкий слой толщиной  $\Delta x$ , охватывающий все поперечное сечение колонного аппарата, включая обечайку, и рассматриваемый в течении времени  $\Delta t$  (рис. 2.16). Толщина слоя может быть соизмерима с эквивалентным диаметром гранулы сорбента.

Процессы тепло- и массообмена при сорбции протекают взаимосвязанно.

В процессах переноса тепла и массы много общего. И диффузия, и теплопроводность являются результатом молекулярного движения. Этим объясняется аналогия в их математическом описании.

Температурное поле локальной области сорбционного колонного аппарата, являющееся совокупностью температурных полей гранул сорбента, газового потока и стенки корпуса описывается сопряженной нелинейной задачей теплообмена, прямое аналитическое решение которой не представляется возможным.

Поле концентраций локальной области сорбционного колонного аппарата, которое является совокупностью полей концентраций гранул сорбента и газового потока, также описывается сопряженной нелинейной задачей диффузии, которая также аналитически

не решается.

Кроме того, коэффициенты задачи теплопроводности зависят от текущих концентраций, а диффузионные характеристики зависят от температур.

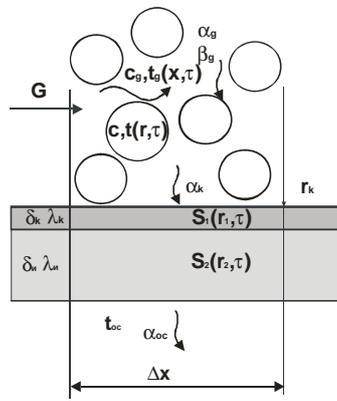


Рис. 2.17. Локальная область сорбционного оборудования

Таким образом, методика моделирования температурного и концентрационного полей локальной области внутреннего пространства сорбционного оборудования схожа с методикой моделирования температурного и концентрационного полей локальной области при сушке гранулированных материалов.

Примем допущение о том, что температурное и концентрационное поле газового потока в колонном аппарате одномерно, т.е. температура и концентрация газового потока меняются только вдоль продольной оси колонны и остаются постоянными по сечению аппарата, перпендикулярному его продольной оси.

Примем допущение о постоянстве теплофизических характеристик газового потока и гранул сорбента внутри локальной области. Значения теплофизических характеристик определяются средними температурами и концентрациями в потоке и гранулах сорбента в локальной области.

Данное допущение позволяет использовать для расчета температурных и концентрационных полей линейные дифференциальные уравнения, допускающие аналитические решения.

Температура и концентрация газового потока, а так же температура стенки корпуса принимаются постоянными по длине локальной области.

Введем следующие обозначения.

$G_n$  – массовый расход газовой смеси на входе в локальную область;

$x_n$  – концентрация поглощаемого компонента газовой смеси на входе в локальную область;

$t_n$  – температура газовой смеси на входе в локальную область;

$D_g, D_c^2$  – соответственно коэффициенты диффузии поглощаемого компонента в газе-носителе и гранулах сорбента;

$\alpha_c, \alpha_k, \alpha_{oc}$  – коэффициенты теплоотдачи соответственно от поверхности гранул и внутренней поверхности корпуса к газовому потоку, а так же от наружной поверхности корпуса (или теплоизоляции, если она есть) в окружающую среду;

$c_1, \rho_1, \lambda_1$  – соответственно теплоемкость, плотность и теплопроводность газовой смеси;

$c_s, c_k, \rho_s, \rho_k, \lambda_s, \lambda_k$  – соответственно теплоемкости, плотности и теплопроводности гранул сорбента и материала корпуса;

$r_k$  – внутренний радиус кожуха аппарата;

$d_c$  – эквивалентный диаметр гранулы сорбента;

$\delta_k$  – толщина стенки корпуса;

$t_{oc}$  – температура окружающей среды.

При известных температурных и концентрационных полях гранул сорбента и температурном поле стенки корпуса с учетом принятых допущений расчет температуры и концентрации газового потока внутри локальной области и на выходе из нее не представляет трудностей, поэтому рассмотрим возможности расчета этих полей.

Расчет температурного поля стенки корпуса, представляющего собой полый неограниченный цилиндр, в частном случае однослойный, подробно рассматривался при моделировании температурного поля локальной области кожухотрубчатого теплообменника.

Предполагается, что гранула сорбента имеет каноническую форму: сфера, ограниченный цилиндр, параллелепипед.

Пусть гранула сорбента имеет сферическую форму.

В этом случае температурное и концентрационное поле локальной области описывается следующими функциями:

- $t_1(x, \tau)$  – температурное поле газового потока;
- $t_{1s}$  – средняя температура газового потока;
- $t(r, \tau)$  – температурное поле гранулы сорбента;
- $t_k(r_1, \tau)$  – температурное поле стенки корпуса;
- $c_1(x, \tau)$  – концентрационное поле газового потока;
- $c_{1s}$  – средняя концентрация поглощаемого компонента в газовом потоке;
- $\alpha(r, \tau)$  – концентрационное поле гранулы сорбента.

Функции  $t(r, \tau)$ ,  $t_k(r_1, \tau)$  и  $t_1(x, \tau)$  являются решениями соответствующих задач теплопроводности, функции  $\alpha(r, \tau)$  и  $c_1(x, \tau)$  – решением задач диффузии.

Температурное поле газового потока описывается следующей системой уравнений:

$$\frac{\partial t_1(x, \tau)}{\partial \tau} + W \frac{\partial t_1(x, \tau)}{\partial x} + K t_1(x, \tau) = F(x, \tau),$$

$$t_1(0, \tau) = t_{10}(\tau); \quad t_1(x, 0) = f_1(x),$$

где  $K = \frac{\alpha_c I_1 + \alpha_k I_2}{S \rho c}$ ;  $F(x, \tau) = \frac{\alpha_c I_1 t_{F1}(x, \tau) + \alpha_k I_2 t_{F2}(x, \tau)}{S \rho c}$ .

Для локального интервала времени  $d\tau$  решение является функцией продольной координаты и имеет вид:

$$t_1(x) = \exp(-Px) \left( t_{10}(0) + \int_0^x V(x) \exp(Px) dx \right),$$

где  $P = \frac{K d\tau + 1}{W d\tau}$ ;  $V(x) = \frac{1}{W} \left( F(x, d\tau) + \frac{f_1(x)}{d\tau} \right)$ .

Температурное поле гранулы сорбента описывается системой уравнений:

$$\frac{\partial t(r, \tau)}{\partial \tau} = a_c^2 \left( \frac{\partial^2 t(r, \tau)}{\partial r^2} + \frac{2}{r} \frac{\partial t(r, \tau)}{\partial r} \right) + \frac{q}{c_c \rho_c}, \quad 0 \leq r \leq R, \quad \tau > 0; \quad t(r, 0) = f(r); \quad \frac{\partial t(0, \tau)}{\partial r} = 0; \quad \lambda_c \frac{\partial t(R, \tau)}{\partial r} + \alpha_c (t(R, \tau) - t_c) = 0.$$

Решение данной задачи имеет вид:

$$t(r, \tau) = \sum_{n=1}^{\infty} \frac{U(\mu_n, \tau) W(r, \mu_n)}{Z_n},$$

где  $W(r, \mu_n) = \frac{1}{r} \sin\left(\frac{\mu_n r}{a_c}\right)$ ;  $\mu_n$  –  $n$ -й положительный корень уравнения  $\frac{\mu}{a_c} \cos\left(\frac{\mu R}{a_c}\right) + \left(\frac{\alpha_c}{\lambda_c} - \frac{1}{R}\right) \sin\left(\frac{\mu R}{a_c}\right) = 0$ ;

$$Z_n = \int_0^R r^2 W^2(r, \mu_n) dr = 0,5 \left( R - \frac{a_c}{\mu_n} \sin\left(\frac{\mu_n R}{a_c}\right) \cos\left(\frac{\mu_n R}{a_c}\right) \right); \quad U(\mu_n, \tau) = \left( U(\mu_n, 0) - \frac{Q_n}{\mu_n^2} \right) \exp(-\mu_n^2 \tau) + \frac{Q_n}{\mu_n^2};$$

$$U(\mu_n, 0) = \int_0^R r^2 (f(r) - t_c) W(r, \mu_n) dr;$$

$$Q_n = \frac{q}{c_c \rho_c} \int_0^R r^2 W(r, \mu_n) dr = \frac{q}{c_c \rho_c} \left( \frac{a_c}{\mu_n} \right)^2 \left( \sin\left(\frac{\mu_n R}{a_c}\right) - \frac{\mu_n R}{a_c} \cos\left(\frac{\mu_n R}{a_c}\right) \right).$$

Температурное поле стенки корпуса описывается следующей системой уравнений:

$$\frac{\partial t_k(r_1, \tau)}{\partial \tau} = a_k^2 \left( \frac{\partial^2 t_k(r_1, \tau)}{\partial r_1^2} + \frac{1}{r_1} \frac{\partial t_k(r_1, \tau)}{\partial r_1} \right), \quad r_k \leq r_1 \leq r_k + \delta_k, \quad \tau > 0;$$

$$t_k(r_1, 0) = f_t(r_1); \quad \lambda_k \frac{\partial t_k(r_k, \tau)}{\partial r_1} - \alpha_k (t_k(r_k, \tau) - t_1) = 0;$$

$$\lambda_k \frac{\partial t_k(r_k + \delta_k, \tau)}{\partial r_1} + \alpha_{oc} (t_k(r_k + \delta_k, \tau) - t_{oc}) = 0.$$

Решение этой задачи имеет вид:

$$t_k(r_1, \tau) = A + B \ln(r_1) + \sum_{n=1}^{\infty} \frac{\xi(v_n, \tau) \zeta(r_1, v_n)}{Z_n},$$

где  $B = \frac{t_1 - t_{oc}}{\ln(r_k) - \ln(r_k + \delta_k) + \lambda_k \left( \frac{1}{r_k \alpha_k} + \frac{1}{(r_k + \delta_k) \alpha_{oc}} \right)}$ ;

$$A = t_1 - B \left( \ln(r_k) - \frac{\lambda_k}{\alpha_k r_k} \right);$$

$$\xi(v_n, \tau) = \xi(v_n, 0) \exp(-v_n^2 \tau);$$

$$\zeta(r_1, v_n) = J_0\left(\frac{v_n r_1}{a_k}\right) + H_n Y_0\left(\frac{v_n r_1}{a_k}\right);$$

$$\xi(v_n, 0) = \int_{r_k}^{r_k + \delta_k} r_1 (f_1(r_1) - A - B \ln(r_1)) \zeta(r_1, v_n) dr_1;$$

$$H_n = - \frac{\frac{\lambda_k v_n}{a_k} J_1\left(\frac{v_n}{a_k} r_k\right) + \alpha_k J_0\left(\frac{v_n}{a_k} r_k\right)}{\alpha_k Y_0\left(\frac{v_n}{a_k} r_k\right) + \frac{\lambda_k v_n}{a_k} Y_1\left(\frac{v_n}{a_k} r_k\right)},$$

$v_n$  –  $n$ -й положительный корень уравнения;

$$\begin{aligned}
& J_0\left(\frac{v(r_k+\delta_k)}{a_k}\right) - \frac{v\lambda_k}{a_k\alpha_{oc}} J_1\left(\frac{v(r_k+\delta_k)}{a_k}\right) + \\
& + H\left(Y_0\left(\frac{v(r_k+\delta_k)}{a_k}\right) - \frac{v\lambda_k}{a_k\alpha_{oc}} Y_1\left(\frac{v(r_k+\delta_k)}{a_k}\right)\right) = 0; \\
& Z_n = \int_{r_k}^{r_k+\delta_k} \zeta^2(r_1, v_n) dr_1 = \int_{r_k}^{r_k+\delta_k} \left( J_0\left(\frac{v_n r_1}{a_k}\right) + H_n Y_0\left(\frac{v_n r_1}{a_k}\right) \right)^2 dr_1 = \\
& 0,5(r_k+\delta_k)^2 \left( J_0^2\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) + J_1^2\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) \right) + (r_k+\delta_k)^2 H_n \times \\
& \times \left( J_0\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) Y_0\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) + J_1\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) Y_1\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) \right) + \\
& + 0,5(r_k+\delta_k)^2 H_n^2 \left( Y_0^2\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) + Y_1^2\left(\frac{v_n(r_k+\delta_k)}{a_k}\right) \right) - \\
& - 0,5r_k^2 \left( J_0^2\left(\frac{v_n r_k}{a_k}\right) + J_1^2\left(\frac{v_n r_k}{a_k}\right) \right) - \\
& - r_k^2 H_n \left( J_1\left(\frac{v_n r_k}{a_k}\right) Y_1\left(\frac{v_n r_k}{a_k}\right) + J_1\left(\frac{v_n r_k}{a_k}\right) Y_1\left(\frac{v_n r_k}{a_k}\right) \right) - \\
& - 0,5r_k^2 H_n^2 \left( Y_0^2\left(\frac{v_n r_k}{a_k}\right) + Y_1^2\left(\frac{v_n r_k}{a_k}\right) \right).
\end{aligned}$$

Концентрационное поле газового потока описывается следующей системой уравнений:

$$\frac{\partial c_1(x, \tau)}{\partial \tau} + W \frac{\partial c_1(x, \tau)}{\partial x} + K_c c_1(x, \tau) = F_c(x, \tau),$$

$$c_1(0, \tau) = c_{10}(\tau); \quad c_1(x, 0) = f_{c1}(x),$$

где  $K_c = \frac{\beta \Pi}{S \rho}$ ;  $F_c(x, \tau) = \frac{\beta \Pi c_F(x, \tau)}{S \rho}$ .

Для локального интервала времени  $d\tau$  решение является функцией продольной координаты и имеет вид:

$$c_1(x) = \exp(-P_c x) \left( c_{10}(0) + \int_0^x V_c(x) \exp(P_c x) dx \right),$$

где  $P_c = \frac{K_c d\tau + 1}{W d\tau}$ ;  $V_c(x) = \frac{1}{W} \left( F_c(x, d\tau) + \frac{f_{c1}(x)}{d\tau} \right)$ .

Средняя концентрация поглощаемого компонента в газовом потоке рассчитывается следующим образом

$$c_{1s} = \frac{1}{\Delta x} \int_0^{\Delta x} c_1(x) dx = \frac{1}{\Delta x} \int_0^{\Delta x} \exp(-P_c x) \left( c_{10}(0) + \int_0^x V_c(\xi) \exp(P_c \xi) d\xi \right) dx.$$

Концентрационное поле гранулы сорбента описывается системой уравнений:

$$\begin{aligned}
\frac{\partial c(r, \tau)}{\partial \tau} &= D_c^2 \left( \frac{\partial^2 c(r, \tau)}{\partial r^2} + \frac{2}{r} \frac{\partial c(r, \tau)}{\partial r} \right), \quad 0 \leq r \leq R, \quad \tau > 0; \\
c(r, 0) &= f_c(r); \\
\frac{\partial c(0, \tau)}{\partial r} &< \infty; \quad D_c \frac{\partial c(R, \tau)}{\partial r} + \beta(c(R, \tau) - c^*) = 0.
\end{aligned}$$

Коэффициент массоотдачи  $\beta$  определяется из критериальных уравнений:

$$\text{Re}_3 > 30, \quad \text{Nu}_D = 0,395 \text{Re}_3^{0,64} \text{Pr}_D^{0,333};$$

$$2 < \text{Re}_3 \leq 30, \quad \text{Nu}_D = 0,725 \text{Re}_3^{0,47} \text{Pr}_D^{0,333};$$

$$\text{Re}_3 < 2, \quad \text{Nu}_D = 0,515 \text{Re}_3^{0,85} \text{Pr}_D^{0,333}.$$

Здесь  $\text{Nu}_D = \frac{\beta d_c^2}{D_g}$ ;  $\text{Re}_3 = \frac{w d_c \rho}{\mu}$ ;  $\text{Pr}_D = \frac{v}{D_g}$ .

В неподвижном слое зернистого материала коэффициент теплоотдачи между потоком и поверхностью частицы определяется из уравнения:

$$Nu = \begin{cases} 0,0035 \left( \frac{Re}{\varepsilon} \right)^{1,5} Pr^{0,33}, & \frac{Re}{\varepsilon} < 200; \\ 0,4 \left( \frac{Re}{\varepsilon} \right)^{0,67} Pr^{0,33}, & \frac{Re}{\varepsilon} \geq 200. \end{cases}$$

Коэффициент теплоотдачи между потоком и стенкой аппарата в неподвижном слое зернистого материала определяется из уравнения:

$$Nu = \begin{cases} 0,31 Re^{0,5} \frac{(1-\varepsilon)^{0,5}}{\varepsilon}, & Re \in (1,5; 57); \\ 0,1 Re^{0,8} \frac{(1-\varepsilon)^{0,2}}{\varepsilon}, & Re \in (57; 150). \end{cases}$$

$$\text{В последних соотношениях } Nu = \frac{\alpha d_3}{\lambda}, \quad Re = \frac{\omega d_3}{\nu}, \quad d_3 = \sqrt{\frac{F}{\pi}}.$$

Расчет температурного и концентрационного полей локальной области выполняется по следующей итеративной методике.

1. Используя температуру газового потока и концентрацию поглощаемого компонента в нем на выходе из предыдущей локальной области, а так же температуру зерна сорбента и концентрацию поглощаемого компонента в нем в предыдущий временной интервал текущей локальной области определяются значения теплофизических характеристик газового потока, средних по объему зерна коэффициентов теплопроводности и диффузии зерна сорбента, значения коэффициентов тепло- и массоотдачи, равновесные концентрации в текущей локальной области.

2. По начальным температурам и концентрациям в локальной области рассчитывается концентрационное поле в зерне сорбента, количество поглощенного компонента, изменение концентрации газового потока.

3. Рассчитывается количество тепла, выделяющееся в процессе сорбции.

4. Рассчитывается температурное поле зерна сорбента с учетом теплоотдачи к газовому потоку. Вычисляется температура газового потока на выходе из локальной области с учетом возможного теплообмена газа с окружающей средой через корпус аппарата.

5. Для уточненных значений средних температур повторяются п. 1 – 4 до совпадения с заданной точностью средних температур в соседних итерациях.

Тепловой и материальный балансы локальной области могут быть использованы для независимой проверки и оценки качества расчета (как и балансы по всему аппарату в целом).

$$\Delta Q_1 = \Delta Q_2 + \Delta Q_3 + \Delta Q_k + \Delta Q_n,$$

где  $\Delta Q_1$  – тепловая мощность теплоты адсорбции;  $\Delta Q_2$  – тепловая мощность, затраченная на нагрев гранул сорбента;  $\Delta Q_3$  – тепловая мощность, затраченная на нагрев газового потока;  $\Delta Q_k$  – тепловая мощность, затраченная на нагрев стенки корпуса;  $\Delta Q_n$  – тепловая мощность потерь в окружающее пространство.

### 2.5.2. Методика расчета адсорбционного оборудования

Методика, основанная на моделировании температурных и концентрационных полей локальной области адсорбционного оборудования, может использоваться для расчетов аппаратов с подвижным или неподвижным слоем сорбента.

Рассмотрим вертикальный аппарат с неподвижным слоем сорбента (рис. 2.16).

Длину локальной области в направлении движения потока газа выбираем методом Рунге, а величину локального интервала времени – равной времени прохождения газового фронта через локальную область.

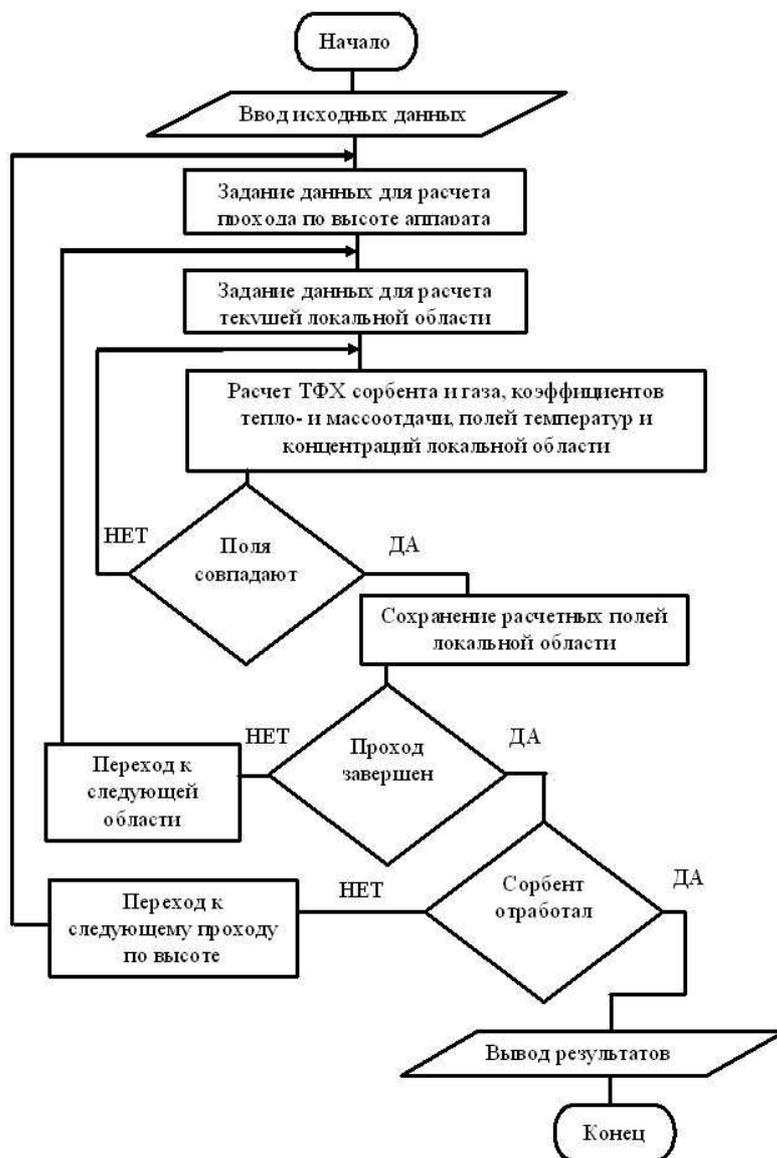
Для каждого локального интервала времени выполняется последовательный расчет температурного и концентрационного полей локальных областей с учетом возможного теплообмена с окружающей средой.

Сохранение расчетных полей температур и концентраций текущей локальной области для использования их в качестве начального условия при переходе к очередному локальному интервалу времени осуществляется путем сохранения массивов коэффициентов и собственных чисел, входящих в решения соответствующих задач теплопроводности и диффузии.

Для того, чтобы избежать совместного решения задач теплопроводности и диффузии в граничных условиях используются постоянные значения температур и концентраций газового потока в локальной области, равные среднеинтегральным значениям.

Расчет продолжается до достижения заданной концентрации газового потока на выходе из аппарата, что соответствует обработке слоя сорбента.

Укрупненная блок-схема расчета нестационарных полей температур и концентраций адсорбционного аппарата с неподвижным слоем сорбента приведена на рис. 2.18.



**Рис. 2.18. Укрупненная схема расчета нестационарных полей температур и концентраций адсорбционного аппарата с неподвижным слоем сорбента**  
**2.5.3. Метод конечных интегральных преобразований**

Для получения аналитического решения задач теплопроводности, включающих уравнение Фурье – Кирхгофа практически наиболее удобным оказался метод конечных интегральных преобразований, теория которого была разработана Н.С. Кошляковым. Дальнейшее развитие теория конечных интегральных преобразований получила в работах Э.М. Карташова.

По сравнению с другими аналитическими методами решения задач математической физики, метод конечных интегральных преобразований имеет ряд практических преимуществ:

- он унифицирован и не требуют изобретательности в технических приемах;
- он позволяет получать решение в случае неоднородных граничных условий без представления задачи в виде совокупности стационарной и нестационарной составляющих;
- он допускает преобразование по нескольким (или всем) пространственным координатам одновременно, а также по координатам, вдоль которых свойства среды изменяются ступенчато.

Использование решений задач теплопроводности для многослойных областей может использоваться для получения приближенных решений нелинейных задач, поскольку увеличение числа слоев в постановке задачи не приводит к принципиальному усложнению решения, так как в решении используются цепочные (или рекуррентные) соотношения, легко реализуемые при программировании.

Смысл метода конечных интегральных преобразований состоит в следующем.

Преобразование, которым функции  $f(x_1, x_2, \dots, x_n)$  ставится в соответствие функция

$$\begin{aligned}
 F(x_1, x_2, \dots, x_{j-1}, \gamma, x_{j+1}, \dots, x_n) &= \\
 &= \int_a^b f(x_1, x_2, \dots, x_{j-1}, \xi, x_{j+1}, \dots, x_n) K(\xi, \gamma) \rho(\xi) d\xi.
 \end{aligned}
 \tag{2.29}$$

является интегральным преобразованием по переменной  $x_j$ , которая в общем случае может быть комплексной. Диапазон  $[a,$

b) является пределами интегрального преобразования, функция  $K(\xi, \gamma)$  является ядром интегрального преобразования и функция  $\rho(\xi)$  – весовой функцией.

Возможно интегральное преобразование по нескольким или сразу по всем пространственным переменным. Интегральное преобразование по нескольким переменным эквивалентно последовательному применению интегрального преобразования по отдельным переменным.

Преобразование, которым функция  $F(x_1, x_2, \dots, x_{j-1}, \gamma, x_{j+1}, \dots, x_n)$  преобразуется в функцию  $f(x_1, x_2, \dots, x_n)$ , является обратным преобразованием.

Интегральное преобразование определено, когда интеграл в правой части (2.29) существует. При практическом использовании интегрального преобразования необходимо существование и обратного преобразования.

Основным отличием интегральных преобразований в конечных пределах от операционного исчисления является использование широкого набора интегральных преобразований, в которых ядра интегральных преобразований и весовые функции определяются индивидуально для каждой конкретной задачи.

Рассмотрим дифференциальное уравнение второго порядка

$$\sum_{i=0}^3 \alpha_i \frac{\partial^2 u}{\partial x_i^2} + \sum_{i=0}^3 \beta_i \frac{\partial u}{\partial x_i} + cu = f.$$

**ВЫБЕРЕМ ПЕРЕМЕННУЮ  $x_j = \xi$ , КОТОРАЯ ИЗМЕНЯЕТСЯ В ПОСТОЯННЫХ КОНЕЧНЫХ ПРЕДЕЛАХ  $[a, b]$  КАЧЕСТВЕ ПЕРЕМЕННОЙ ПРЕОБРАЗОВАНИЯ.**

Ядро интегрального преобразования  $K(\xi, \gamma)$  и весовая функция  $\rho(\xi)$  определяются из условия, чтобы интегральное соотношение

$$\int_a^b \left( \sum_{i=0}^3 \alpha_i \frac{\partial^2 u}{\partial x_i^2} + \sum_{i=0}^3 \beta_i \frac{\partial u}{\partial x_i} + cu \right) K(\xi, \gamma) \rho(\xi) d\xi = \bar{f},$$

(здесь  $\bar{f}$  – интегральное преобразование функции  $f$  являлось дифференциальным уравнением относительно интегрального преобразования

$$\bar{u} = \int_a^b u K(\xi, \gamma) \rho(\xi) d\xi$$

функции  $u$ .

Тогда весовая функция  $\rho$  определяется с точностью до постоянного множителя из условия  $(\alpha_j \rho)' = \beta_j \rho$ .

Если граничные условия по координате  $x_j$  имеют вид

$$\alpha_a u'(a) + \beta_a u(a) = \varphi_a, \quad \alpha_b u'(b) + \beta_b u(b) = \varphi_b,$$

то ядро интегрального преобразования  $K(\xi, \gamma)$  является решением задачи Штурма-Лиувилля

$$(\alpha_j \rho K')' + \rho(c + s^2)K = 0;$$

$$\alpha_a u'(a) + \beta_a u(a) = 0;$$

$$\alpha_b u'(b) + \beta_b u(b) = 0,$$

также полученным с точностью до постоянного множителя. Здесь  $s^2$  – собственные числа задачи.

Обратное преобразование выполняется по формуле

$$u = \sum_{i=0}^{\infty} \frac{\bar{u} K}{N},$$

причем суммирование ведется по собственным числам и

$$N = \int_a^b K^2(\xi, \gamma) \rho(\xi) d\xi.$$

Рассмотрим использование метода конечных интегральных преобразований на примере решения задачи диффузии для сплошного шара.

Запишем задачу относительно равновесной концентрации поглощаемого компонента окружающей среды  $c^*$ :

$$\frac{\partial c(r, \tau)}{\partial \tau} = D_c^2 \left( \frac{\partial^2 c(r, \tau)}{\partial r^2} + \frac{2}{r} \frac{\partial c(r, \tau)}{\partial r} \right), \quad 0 \leq r \leq R, \quad \tau > 0; \quad (2.30)$$

$$c(r, 0) = f_c(r) - c^*; \quad (2.31)$$

$$\frac{\partial c(0, \tau)}{\partial r} < \infty; \quad (2.32)$$

$$D_c \frac{\partial c(R, \tau)}{\partial r} + \beta c(R, \tau) = 0; \quad (2.33)$$

Выполним интегральное преобразование по координате  $r$ .

Переход к изображениям выполняется по формуле:

$$A(\tau) = \int_0^R r^2 c(r, \tau) P(r) dr.$$

Здесь  $P(r)$  – решение вспомогательной задачи

$$\frac{d^2 P(r)}{dr^2} + \frac{2}{r} \frac{dP(r)}{dr} + \frac{\mu^2}{D_c^2} P(r) = 0; \quad (2.34)$$

$$P(0) < \infty; \quad (2.35)$$

$$D_c^2 \frac{dP(R)}{dr} + \beta P(R) = 0. \quad (2.36)$$

Решение задачи (2.34) – (2.36) имеет вид:

$$P(r) = \frac{1}{r} \sin\left(\frac{\mu}{D_c} r\right); \quad (2.70)$$

где  $\mu$  – последовательные положительные корни уравнения, полученного при подстановке (2.37) в (2.36):

$$D_c \mu \cos\left(\frac{\mu}{D_c} R\right) + \left(\beta - \frac{D_c^2}{R}\right) \sin\left(\frac{\mu}{D_c} R\right) = 0.$$

Теперь переходим к изображениям задачи (2.30) – (2.33). В изображениях исходная задача имеет вид:

$$\frac{dA(\tau)}{d\tau} = -\mu^2 A(\tau); \quad (2.71)$$

$$A(0) = \int_0^R r^2 (f(r) - c^*) P(r) dr. \quad (2.72)$$

Задача (2.71), (2.72) имеет решение:

$$A(\tau) = A(0) \exp(-\mu^2 \tau).$$

Остается записать решение исходной задачи (2.30) – (2.33) в соответствии с формулой обратного перехода:

$$c(r, \tau) = c^* + \sum_{n=1}^{\infty} \frac{A(\tau) P(r)}{N_n},$$

(суммирование ведется по значениям  $\mu_n$ ),

$$\text{где } N_n = \int_0^R r^2 P^2(r) dr = \frac{R}{2} - \frac{D_c}{2\mu_n} \sin\left(\frac{\mu_n}{D_c} R\right) \cos\left(\frac{\mu_n}{D_c} R\right).$$

Среднеобъемная концентрация рассчитывается по формуле:

$$\begin{aligned} c(\tau) &= \frac{3}{R^3} \int_0^R r^2 c(r, \tau) dr = \\ &= c^* + \frac{3}{R^3} \sum_{n=1}^{\infty} \frac{A(\tau) D_c^2}{N_n \mu_n^2} \left( \sin\left(\frac{\mu_n}{D_c} R\right) - \frac{\mu_n}{D_c} R \cos\left(\frac{\mu_n}{D_c} R\right) \right). \end{aligned}$$

Таким образом, аналитическое решение получено.

## 2.5. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ В УСЛОВИЯХ ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТИ

Рассмотрим методы решения двухэтапных задач оптимизации. Имеются следующие две трудности решения этих задач: 1) недифференцируемость функции  $h(a, d, \xi)$  по переменной  $\xi$  и функции гибкости  $\chi_1(a, d)$  по переменной  $d$ ; 2) многоэкстремальность задач, связанная с невыпуклостью функций  $I(a, d, z, \xi)$ ,  $g_j(a, d, z, \xi)$  и видом функции  $h(a, d, \xi)$ . Поэтому прямой подход к решению двухэтапных задач оптимизации в условиях неопределенности требует использования методов глобальной недифференцируемой оптимизации. Известно, что эти методы очень трудоемки и существенно менее эффективны, чем хорошо разработанные методы дифференцируемой оптимизации (методы нелинейного программирования). В связи с этим опишем методы решения двухэтапной задачи оптимизации в условиях неопределенности, которые будут использовать только методы дифференцируемой оптимизации. При выполнении некоторых условий выпуклости они будут доставлять глобальное решение двухэтапной задачи оптимизации.

Запишем двухэтапную задачу оптимизации в следующем виде:

$$F_1 = \min_{a \in A, d \in D, z^i \in I_1} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$h(a, d, \xi) \leq 0, \quad \xi \in \Xi.$$

Пусть  $[a^*, d^*, z^{i*}]$  – решение этой задачи. Обозначим бесконечное множество точек, содержащихся в области  $\Xi$ , через  $S_\Xi$  и через  $S_{A,P}$  – множество точек  $\xi^P$ , которым соответствуют активные ограничения  $h(a, d, \xi) = 0$  в точке решения задачи:

$$S_{A,P} = \{\xi^P : h(a^*, d^*, \xi^P) = 0, \xi^P \in \Xi\}.$$

Будем называть эти точки активными.

Решение последней задачи  $[a^*, d^*, z^{i*}]$  есть решение (локальный минимум) задачи

$$F_1 = \min_{a \in A, d \in D, z^i \in I_1} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g(a, d, z^i, \xi^i) = 0, \quad i \in I_1,$$

$$h(a, d, \xi) = 0, \quad \xi \in S_{A,P}.$$

*Нижняя граница для двухэтапной задачи оптимизации.* Введем некоторое произвольное множество  $S_2^{(k)} = \{\xi^l : l \in I_2^{(k)}, \xi^l \in \Xi\}$  точек из области неопределенности, где  $I_2^{(k)}$  – множество индексов точек в  $S_2^{(k)}$  ( $I_1 \cap I_2^{(k)} = \emptyset$ ),  $k$  – номер итерации алгоритма решения задачи. Точки множества  $S_2^{(k)}$  будем называть критическими точками.

Далее будем рассматривать задачу

$$F_1^{L,(k)} = \min_{a \in A, d \in D, z^i \in I_1} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$h(a, d, \xi^l) \leq 0, \quad \xi^l \in S_2^{(k)}.$$

Эта задача имеет следующие свойства.

*Свойство 1.* Величина  $F_1^{L,(k)}$  является нижней границей оптимального значения целевой функции двухэтапной задачи оптимизации. Поскольку  $S_2^{(k)} \subset S_\Xi$ , то область допустимости исходной задачи оптимизации является частью области допустимости последней задачи. Следовательно  $F_1^{L,(k)} \leq F_1$ .

*Свойство 2.* Пусть множество  $S_2^{(k+1)}$  получено добавлением одной или нескольких точек к множеству  $S_2^{(k)}$ , тогда  $F_1^{L,(k+1)} \geq F_1^{L,(k)}$ . Таким образом, добавлением точек к множеству критических точек не ухудшает нижнюю границу, а в большинстве случаев даже ее улучшает.

*Свойство 3.* Если множество критических точек  $\xi^i$ , принадлежащих множеству  $S_2^{(k)}$ , покрывает достаточно плотно область  $\Xi$ , то решение последней задачи достаточно близко к решению исходной задачи. Таким образом

$$F_1 - F_1^{L,(k)} \leq \varepsilon,$$

где  $\varepsilon$  – достаточно малая положительная величина.

Пусть множество  $S_2^{(k)}$  покрывает достаточно плотно область  $\Xi$ . Это означает, что при достаточно малом  $\varepsilon$  в  $\varepsilon$ -окрестности каждой точки из области  $\Xi$  имеется точка множества  $S_2^{(k)}$ . Тогда множество  $S_2^{(k)}$  близко к множеству  $S_\Xi$  и решение последней задачи близко к решению исходной задачи.

*Верхняя граница для двухэтапной задачи оптимизации.* Пусть область  $\Xi$  разбита на  $N_k$  подобластей  $\Xi_i^{(k)}$  ( $\Xi_i^{(k)} = \{\xi : \xi^{L,i(k)} \leq \xi \leq \xi^{U,i(k)}\}$ ),  $\Xi = \Xi_1^{(k)} \cup \dots \cup \Xi_{N_k}^{(k)}$ , где  $k$  – номер итерации алгоритма решения двухэтапной задачи оптимизации. Обозначим через  $\Xi^{(k)}$  совокупность всех подобластей  $\Xi_i^{(k)}$  ( $i = 1, 2, \dots, N_k$ ). Заменим в задаче

$$F_1 = \min_{a, d, z^i \in I_1} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$\chi_1(a, d) \leq 0,$$

последнее ограничение  $N_k$  ограничениями  $\chi_{1,i}^U(a, d) \leq 0$  ( $i = 1, \dots, N_k$ ):

$$F_1^{U,(k)} = \min_{a, d, z^i \in I_1} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad j = 1, \dots, m, \quad i \in I_1,$$

$$\chi_{1,1}^U(a, d) \leq 0, \dots, \chi_{1,N_k}^U(a, d) \leq 0,$$

где функция  $\chi_{1,i}^U(a, d) = \min_{z \in Z} \max_{j \in J} \max_{\xi \in \Xi_i} g_j(a, d, z, \xi)$  – верхняя граница функции гибкости для подобласти  $\Xi_i$ .

В частности, когда  $N_k = 1$ , ограничение  $\chi_1(a, d) \leq 0$  заменяется ограничением  $\chi_1^U(a, d) \leq 0$ . Подставим выражение для  $\chi_{1,i}^U(a, d)$  в последние неравенства, тогда получим

$$\begin{aligned} F_1^{U,(k)} &= \min_{a,d,z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i), \\ g_j(a, d, z^i, \xi^i) &\leq 0, \quad j=1, \dots, m, \quad i \in I_1, \\ \min_z \max_{j \in J} \max_{\xi \in \Xi_l^k} g_j(a, d, z, \xi) &\leq 0, \quad l=1, \dots, N_k. \end{aligned}$$

Последнюю задачу можно переписать в виде

$$\begin{aligned} F_1^{U,(k)} &= \min_{a,d,z^i,z^l} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i), \\ g_j(a, d, z^i, \xi^i) &\leq 0, \quad j=1, \dots, m, \quad i \in I_1, \\ \max_{\xi \in \Xi_l^k} g_j(a, d, z^l, \xi) &\leq 0, \quad j=1, \dots, m, \quad l=1, \dots, N_k. \end{aligned}$$

Пусть  $[a^{(k)}, d^{(k)}, z^{i,(k)}, z^{l,(k)}]$  – решение последней задачи, тогда  $[a^{(k)}, d^{(k)}, z^{i,(k)}]$  – решение исходной задачи. Обозначим через  $\xi^{j,(l)}$  решение задачи

$$\max_{\xi \in \Xi_l^{(k)}} g_j(a^{(k)}, d^{(k)}, z^{l,(k)}, \xi).$$

Будем называть точку  $\xi^{j,(l)}$  активной, если соответствующее неравенство является активным в точке решения последней задачи, т.е.

$$g_j(a^{(k)}, d^{(k)}, z^{l,(k)}, \xi^{j,(l)}) = 0.$$

Аналогично множеству активных точек  $S_{A,P}$  двухэтапной задачи оптимизации введем множество активных точек  $S_{A,P}^{(k)}$  последней задачи:

$$S_{A,P}^{(k)} = \{\xi^{l,j} : g_j(a^{(k)}, d^{(k)}, z^{l,(k)}, \xi^{j,(l)}) = 0, \quad l=1, \dots, N_k, \quad j=1, \dots, m\}.$$

Будем называть область  $\Xi_i^{(k)}$  активной, если ей соответствует хотя бы одно равенство  $g_j(a^{(k)}, d^{(k)}, z^{l,(k)}, \xi^{j,(l)}) = 0$ . Активной области с номером  $l$  соответствует условие  $\chi_{1,l}^U(a, d) = 0$ .

Рассмотрим свойства задачи оптимизации с верхней границей.

*Свойство 1.* Величина  $F_1^{U,(k)}$  является верхней границей оптимального значения целевой функции двухэтапной задачи оптимизации.

*Свойство 2.* Если  $\Xi^{(p+1)}$  получается разбиением некоторых подобластей множества  $\Xi^{(p)}$ , т.е.  $\forall \Xi_i^{(p+1)} \exists \Xi_j^{(p)}$  такое, что  $\Xi_i^{(p+1)} \subseteq \Xi_j^{(p)}$ , то  $F_1^{U,(p+1)} \leq F_1^{U,(p)}$ .

*Свойство 3.* Разобьем область  $\Xi$  на достаточно малые подобласти  $\Xi_i^{(k)}$ , тогда решение задачи

$$\begin{aligned} F_1^{(k)} &= \min_{a,d,z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i), \\ g_j(a, d, z^i, \xi^i) &\leq 0, \quad j=1, \dots, m, \quad i \in I_1, \\ \chi_{1,1}^U(d) &\leq 0, \dots, \chi_{1,N_k}^U(d) \leq 0. \end{aligned}$$

будет достаточно близко к решению двухэтапной задачи оптимизации

$$\lim_{r(\Xi_i^{(k)}) \rightarrow 0} F_1^{U,(k)} = F_1.$$

Рассмотрим теперь алгоритм вычисления верхней границы оптимального значения целевой функции двухэтапной задачи оптимизации. Для ее вычисления можно использовать алгоритм внешней аппроксимации: 1) вычисление нижней границы величины  $F_1^{U,(k)}$ ; 2) проверка критерия окончания итерационной процедуры. В случае необходимости продолжения итерационной процедуры в соответствии с алгоритмом увеличивается множество критических точек для улучшения аппроксимации ограничений.

Приведем алгоритм вычисления верхней границы оптимального значения целевой функции двухэтапной задачи оптимизации [6].

**Алгоритм 1.**

*Шаг 1.* Положить  $p = 1$ . Задать множество подобластей  $T_l^{(k)} (l = 1, \dots, N_k)$ , (число  $k$  фиксировано, следовательно, число подобластей и их размеры остаются постоянными в этом алгоритме). Задать начальные значения  $z^{i,(0)}, z^{l,(0)}, a^{(0)}, d^{(0)} (i \in I_1, l = 1, \dots, N_1)$  соответствующих переменных. Задать достаточно малое число  $\varepsilon > 0$ .

*Шаг 2.* Создать начальное множество  $S_{2l}^{(k,1)}$  критических точек для всех подобластей  $T_l (l = 1, \dots, N_1)$ .

$$S_{2l}^{(k,1)} = \{ \xi^{l,j,(1)}, j = 1, \dots, m : g_j(a^{(0)}, d^{(0)}, z^{i,(0)}, \xi^{l,j,(1)}) > 0 \},$$

где  $\xi^{l,j,(1)}$  есть решение задачи  $\max_{\xi \in T_l^{(k)}} g_j(a^{(0)}, d^{(0)}, z^{i,(0)}, \xi)$ .

*Шаг 3.* Решить задачу

$$F_1^{UL,(k)} = \min_{a,d,z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad j = 1, \dots, m, \quad i \in I_1,$$

$$g(a, d, z^l, \xi^l) \leq 0, \quad \xi^l \in S_{2,l}^{(k,1)}, \quad l = 1, \dots, N_k$$

и определить нижнюю границу  $F_1^{UL,(k)}$  величины  $F_1^{U,(k)}$ . Пусть  $[a^{(p)}, d^{(p)}, z^{i,(p)}, z^{l,(p)}]$  – решение этой задачи.

*Шаг 4.* решить  $N_k$  задач

$$\max_{\xi \in \Xi_l^{(k)}} g_j(a^{(p)}, d^{(p)}, z^{l,(p)}, \xi),$$

где  $\Xi_l^{(k)} = \{ \xi : \xi^{L,l,(k)} \leq \xi \leq \xi^{U,l,(k)} \}$ . Обозначим через  $\xi^{l,j,(p)}$  решение этой задачи при фиксированных  $[l, j]$ .

*Шаг 5.* Проверить выполнение условий

$$g_j(a^{(p)}, d^{(p)}, z^{l,(p)}, \xi^{l,j,(p)}) \leq \varepsilon, \quad j = 1, \dots, m, \quad l = 1, \dots, N_k.$$

Если все неравенства выполняются, решение задачи

$$F_1^{U,(k)} = \min_{a,d,z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad j = 1, \dots, m, \quad i \in I_1,$$

$$\max_j \max_{\xi \in T_l^{(k)}} g_j(a, d, z^l, \xi) \leq 0, \quad l = 1, \dots, N_k,$$

где  $z^l$  – вектор управляющих переменных, соответствующий  $l$ -подобласти; получено, перейти к шагу 9; в противном случае перейти к шагу 6.

*Шаг 6.* Создать множества  $R^{l,(p)} (l = 1, \dots, N_k)$  точек  $\xi^{l,j,(p)}$ , для которых ограничения на шаге 5 нарушаются.

$$R^{l,(p)} = \{ \xi^{l,j,(p)}, j = 1, \dots, m : g_j(a^{(p)}, d^{(p)}, z^{l,(p)}, \xi^{l,j,(p)}) > \varepsilon, \quad l = 1, \dots, N_k \}.$$

*Шаг 7.* Создать новые множества критических точек

$$S_{2l}^{(p+1)} = S_{2l}^{(p)} \cup R^{l,(p)}, \quad l = 1, \dots, N_k,$$

которые будут использоваться на следующей итерации.

*Шаг 8.* Положить  $p := p + 1$ , перейти к шагу 3.

*Шаг 9.* Создать и запомнить множество активных точек

$$S_{A,P}^{(k)} = \{ \xi^{l,j,(p^*)} : g_j(a^{(k)}, d^{(k)}, z^{l,(k)}, \xi^{l,j,(p^*)}) = 0, \quad l = 1, \dots, N_k, \quad j = 1, \dots, m \},$$

где  $p^*$  – номер последней итерации этого алгоритма.

Двухэтапная задача оптимизации

$$F_1 = \min_{a \in A, d \in D, z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$\max_{\xi \in \Xi} h(a, d, \xi) \leq 0$$

имеет вид задачи полубесконечного программирования. Поэтому для ее решения можно использовать алгоритм внешней аппроксимации.

Алгоритм 2 (внешней аппроксимации).

*Шаг 1.* Положить  $k = 1$ . Задать начальные значения  $a^{(0)}, d^{(0)}$ , множество аппроксимационных точек  $S_1 = \{ \xi_i : i \in I_1 \}$ , начальное множество критических точек  $S_2^{(0)}$  и достаточно малое число  $\varepsilon > 0$ .

*Шаг 2.* Решить задачу

$$F_1^{L,(k)} = \min_{a \in A, d \in D, z^i, z^l} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$g_j(a, d, z^l, \xi^l) \leq 0, \quad j=1, \dots, m, \quad \xi^l \in S_2^k$$

и определить нижнюю границу  $F_1^{L,(k)}$  величины  $F_1$ . Пусть  $a^{(k)}, d^{(k)}$  – решение этой задачи.

*Шаг 3.* Найти значение  $\chi_1(a^{(k)}, d^{(k)})$ . Пусть  $\xi^{(k)}$  – решение этой задачи.

*Шаг 4.* Если  $\chi_1(a^{(k)}, d^{(k)}) \leq 0$ , то решение найдено. В противном случае перейти к шагу 5.

*Шаг 5.* Образовать новое множество критических точек  $S_2^{(k+1)} = S_2^{(k)} \cup \{\xi^{(k)}\}$ .

*Шаг 6.* Положить  $k := k + 1$  и перейти к шагу 2.

Для вычисления функции гибкости  $\chi_1(a^{(k)}, d^{(k)})$  можно использовать метод ветвей и границ [6].

Метод разбиений и границ.

Он представляет двухуровневую итерационную процедуру, основанную на разбиении области неопределенности  $\Xi$  на подобласти и состоит из набора следующих алгоритмов и процедур:

- 1) алгоритм вычисления верхней границы  $F_1^{U,(k)}$ ;
- 2) алгоритм вычисления нижней границы  $F_1^{L,(k)}$ ;
- 3) правило (процедура) выбора множества  $S_2^{(k)}$  критических точек в алгоритме вычисления нижней границы.

Пусть на  $k$ -й итерации область  $\Xi$  разбита на подобласти  $\Xi_i^{(k)}, i=1, \dots, N_k$ . Введем множество  $L^{(k)}$  подобластей  $\Xi_i^{(k)}$  следующим образом:

$$L^{(k)} = \{\Xi_i^{(k)}, i=1, \dots, N_k : r(\Xi_i^{(k)}) > \delta_1\},$$

где  $\delta_1$  – заранее заданное число.

Алгоритм 3.

*Шаг 1.* Положить  $k = 1$ . Задать начальное множество подобластей  $\Xi_l^{(1)} (l=1, \dots, N_k)$ , множество аппроксимационных точек  $S_1 = \{\xi^i : i \in I_1\}$ , начальное множество критических точек  $S_1^{(0)}$ . Задать начальные значения  $z^{i,(0)}, z^{l,(0)}, a^{(0)}, d^{(0)}$  ( $i \in I_1, l=1, \dots, N_1$ ) соответствующих переменных, число  $\delta_1 > 0$ , достаточно малые числа  $\varepsilon_1 > 0, \varepsilon_2 > 0, \delta_2 > 0$  ( $\varepsilon_2 > \varepsilon_1, \delta_1 > \delta_2$ ).

Положить

$$F_1^{U,(0)} = a, F_1^{L,(0)} = -a,$$

где  $a$  – достаточно большое число ( $a > -F_1$ ).

*Шаг 2.* Вычислить верхнюю границу задачи

$$F_1^{U,(k)} = \min_{a, d, z^i} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad j=1, \dots, m, \quad i \in I_1,$$

$$\chi_{1,1}^U(a, d) \leq 0, \dots, \chi_{1,N_k}^U(a, d) \leq 0$$

с помощью алгоритма 1.

Пусть  $[a^{(k)}, d^{(k)}, z^{i,(k)}, z^{l,(k)}]$  ( $i \in I_1, l=1, \dots, N_k$ ) – решение этой задачи.

*Шаг 3.* Определить множество  $Q^{(k)} = \{\Xi_l^{(k)} : l \in I_Q^{(k)}\}$  подобластей  $\Xi_l^{(k)}$  которым соответствуют активные ограничения:

$$\chi_{1,l}^U(a^{(k)}, d^{(k)}) = 0, \quad \Xi_l^{(k)} \in Q^{(k)}.$$

*Шаг 4.* Если  $Q^{(k)}$  – пустое множество, то решение двухэтапной задачи оптимизации найдено.

*Шаг 5.* Если выполняется условие  $F_1^{U,(k-1)} - F_1^{U,(k)} \leq \varepsilon_1 |F_1^{U,(k)}|$ , то решение задачи найдено. В противном случае проверить условие

$$F_1^{U,(k-1)} - F_1^{U,(k)} \leq \varepsilon_2 |F_1^{U,(k)}|,$$

и если оно нарушается, то перейти к шагу 8, в противном случае – к шагу 6.

*Шаг 6.* Найти нижнюю границу  $F_1^{L,(k)}$ , решая задачу

$$F_1^{L,(k)} = \min_{a \in A, d \in D, z^i, z^l} \sum_{i \in I_1} w_i I(a, d, z^i, \xi^i),$$

$$g_j(a, d, z^i, \xi^i) \leq 0, \quad i \in I_1,$$

$$g_j(a, d, z^l, \xi^l) \leq 0, \quad j = 1, \dots, m, \quad \xi^l \in S_2^k.$$

Здесь мы полагаем  $S_2^{(k)} = S_{A,P}^{(P)}$ .

*Шаг 7.* Если выполняется условие  $F_1^{U,(k-1)} - F_1^{L,(k)} \leq \varepsilon_1 |F_1^{U,(k)}|$ , то решение задачи получено.

*Шаг 8.* Если выполняется условие  $r(\Xi_i^{(k)}) \leq \delta_2$ ,  $i = 1, \dots, N_k$ , то решение задачи получено.

*Шаг 9.* Если выполняется условие  $r(\Xi_l^{(k)}) \leq \delta_1$ ,  $l = 1, \dots, N_k$ , то перейти к шагу 11.

*Шаг 10.* Создать множество  $L^{(k)}$  и определить множество  $V^{(k)}$  подобластей  $\Xi_l^{(k)}$ , принадлежащих одновременно множествам  $L^{(k)}$  и  $Q^{(k)}$

$$V^{(k)} = L^{(k)} \cap Q^{(k)}.$$

Разбить каждую подобласть  $\Xi_l^{(k)} \in V^{(k)}$  на две подобласти таким образом, что  $\Xi_{l_1}^{(k+1)} \cup \Xi_{l_2}^{(k+1)} = \Xi_l^{(k)}$ . Образовать новое множество  $\Xi^{(k+1)}$  подобластей из старого множества  $\Xi^{(k)}$ , заменяя подобласти  $\Xi_l^{(k)}$  новыми подобластями  $\Xi_{l_1}^{(k+1)}$  и  $\Xi_{l_2}^{(k+1)}$ .  
Перейти к шагу 12.

*Шаг 11.* Положить  $\delta_1 := \delta_1 / 2$  и перейти к шагу 9.

*Шаг 12.* Положить  $k := k + 1$  и перейти к шагу 2.

Рассмотрим пошаговую работу алгоритма 3 на примере задачи

$$\begin{aligned} & \min_{d, z} d, \\ & -z + \xi \leq 0, \\ & z - 2\xi - d + 2 \leq 0, \\ & -z + 6\xi - 9d \leq 0, \\ & 0 \leq d \leq 4. \end{aligned}$$

Пусть  $\varepsilon_1 = 0,01$ ;  $\varepsilon_2 = 0,1$ ;  $\delta_1 = 0,1$ ;  $\delta_2 = 0,01$ ;  $F_1^{U,(0)} = a$ ;  $F_1^{L,(0)} = -a$ , где  $a$  – достаточно большое число. Область неопределенности имеет вид  $\Xi = \{\xi : 1 \leq \xi \leq 3\}$ .

Рассмотрим первую итерацию.

*Шаг 1.* Пусть начальная совокупность подобластей состоит из всей области  $\Xi$ .

*Шаг 2.* Определить верхнюю границу  $F_1^{U,(1)}$ , решая задачу линейного программирования:

$$\begin{aligned} & \min_{d, z} d, \\ & -z + 3 \leq 0, \\ & z - d \leq 0, \\ & -z - 9d + 18 \leq 0, \end{aligned}$$

получим  $[d, z] = [3, 3]$ . Отсюда следует, что  $F_1^{U,(1)} = d^{(1)} = 3$ .

В этой задаче активными являются первое и второе ограничения. Активные точки  $\xi = 1$  и  $\xi = 3$  соответствуют этим ограничениям. Следовательно,  $S_{A,P}^{(1)} = (1, 3)$ .

*Шаг 3.*  $Q^{(1)} = \{\Xi\}$ .

*Шаг 4.* Ясно, что множество  $Q^{(1)} = \{\Xi\}$  не пустое.

*Шаг 5.* Условие  $F_1^{U,(0)} - F_1^{U,(1)} \leq \varepsilon_1 |F_1^{U,(1)}|$  не выполняется.

*Шаг 6.* Определить нижнюю границу  $F_1^{L,(1)}$ , в которой  $S_2^{(1)} = S_{A,P}^{(1)} = \{1, 3\}$ .

Эта задача имеет вид:

$$\begin{aligned} & \min_{d, z^1, z^2} d, \\ & -z^1 + 1 \leq 0, \\ & z^1 - d \leq 0, \\ & -z^1 + 6 - 9d \leq 0, \\ & -z^2 + 3 \leq 0, \\ & z^2 - 4 - d \leq 0, \\ & -z^2 + 18 - 9d \leq 0. \end{aligned}$$

Здесь первые три ограничения соответствуют первой критической точке, а последние три ограничения соответствуют второй критической точке. Решая последнюю задачу, получим  $F_1^{L,(1)} = \bar{d} = 1,4$ .

*Шаг 7.* Условие  $F_1^{U,(1)} - F_1^{L,(1)} \leq \varepsilon_1 |F_1^{U,(1)}|$  не выполняется, поскольку  $3 - 1,4 \geq 0,03$ .

*Шаг 8.* Так как  $r(\Xi) = 3 > 0,01$ , то  $L^{(1)} = \{\Xi\}$ .

*Шаг 9.*  $V^{(1)} = \{\Xi\}$ . Разобьем  $\Xi$  на две подобласти  $\Xi_p = \{\xi : 1 \leq \xi \leq 2\}$  и  $\Xi_q = \{\xi : 2 \leq \xi \leq 3\}$ .

Рассмотрим вторую итерацию.

*Шаг 2.* Решить задачу для случая, когда имеется две подобласти  $(N_2 - 2)$ :  $\Xi_p$  и  $\Xi_q$ . В этом случае задача имеет вид:

$$\begin{aligned} & \min_{d, z^1, z^2} d, \\ & \begin{cases} -z^1 + 2 \leq 0, & z^1 - 9 \leq 0, & -z^1 + 12 - 9d \leq 0, \\ -z^2 + 3 \leq 0, & z^2 - 2 - d \leq 0, & -z^2 + 18 - 9d \leq 0. \end{cases} \end{aligned}$$

Здесь переменные  $[z^1, z^2]$  относятся к областям  $[\Xi_p, \Xi_q]$ , соответственно. Решения первой и второй вспомогательных задач имеют вид:

$$[\bar{d}^1, \bar{z}^1] = [2; 2] \text{ и } [\bar{d}^2, \bar{z}^2] = [1,6; 3,6].$$

Число  $\bar{d} = \max(\bar{d}^1, \bar{d}^2) = 2$  будет решением этой задачи, если для  $d = \bar{d}$  можно найти такие управляющие переменные  $[z^1, z^2]$ , что все ограничения задачи будут выполняться. Все ограничения будут удовлетворяться при  $d = \bar{d} = 2$ ,  $z^1 = 2$  и  $3 \leq z^2 \leq 4$ . Таким образом,  $F_1^{U,(2)} = 2$ .

Можно видеть, что по сравнению с первой итерацией верхняя граница уменьшилась от 3,0 до 2,0. Заметим, что первое и второе ограничения, соответствующие подобласти  $\Xi_p$ , являются активными. Активные точки  $[\xi^1, \xi^2] = [1,0; 2,0]$  соответствуют этим ограничениям. Поскольку любая величина переменной  $z^2$  из интервала  $[3,0; 4,0]$  соответствует решению задачи, то первое и второе ограничения, соответствующие подобласти  $\Xi_q$ , являются активными. Активные точки  $\xi = 2,0$ ;  $\xi = 3,0$  соответствуют этим ограничениям. Таким образом, множество  $S_{A,P}^{(2)}$  состоит из трех критических точек  $\xi^1 = 1$ ,  $\xi^2 = 2$  и  $\xi^3 = 3$ .

*Шаг 3.* Так как активные ограничения соответствуют подобластям  $\Xi_p$  и  $\Xi_q$ , то  $Q^{(2)} = \{\Xi_p, \Xi_q\}$ .

*Шаг 4.* Множество  $Q^{(2)}$  не пусто.

*Шаг 5.* Здесь  $3 - 2 > 0,03$ . Следовательно условие

$$F_1^{U,(1)} - F_1^{U,(2)} \leq \varepsilon_1 |F_1^{U,(1)}|$$

не выполняется.

*Шаг 6.* Решить задачу линейного программирования при  $S_2^{(2)} = S_{A,P}^{(2)}$ :

$$\begin{aligned} & \min_{d, z^1, z^2, z^3} d, \\ & \begin{cases} -z^1 + 1 \leq 0, \\ z^1 - 9 \leq 0, \\ -z^1 + 6 - 9d, \\ -z^2 + 2 \leq 0, \\ z^2 - 2 - d \leq 0, \\ -z^2 + 12 - 9d \leq 0, \\ -z^3 + 3 \leq 0, \\ z^3 - 4 - d \leq 0, \\ -z^3 + 18 - 9d \leq 0. \end{cases} \end{aligned}$$

Здесь переменные  $[z^1, z^2, z^3]$  соответствуют критическим точкам  $[\xi^1, \xi^2, \xi^3]$ . Запишем решение этой задачи  $F_1^{L,(2)} = \bar{d} = 1,4$ . Мы видим, что нижняя граница опять равна 1,4, как и на первой итерации.

*Шаг 7.* Условие  $F_1^{U,(2)} - F_1^{L,(2)} \leq \varepsilon_1 |F_1^{U,(2)}|$  не выполняется.

*Шаг 8.* так как для обеих подобластей мы имеем  $r(\Xi_p) > 0,01$  и  $r(\Xi_q) > 0,01$ , то  $L^{(2)} = \{\Xi_p, \Xi_q\}$ .

*Шаг 9.*  $V^{(2)} = \{\Xi_p, \Xi_q\}$ . Следовательно, мы должны разбить подобласти  $\Xi_p$  и  $\Xi_q$ .

Для получения решения необходимы последующие итерации.

Под неопределенностью будем понимать случайность параметров и нечеткость констант, относящихся к процессу или производству, например: теплопереносу (теплофизические характеристики материала, коэффициенты теплоотдачи, степень черноты, скорость движения среды, теплофизические характеристики движущейся среды), массопереносу (коэффициенты диффузии, кинетические константы, коэффициенты массоотдачи и массопередачи, скорость движения среды), гидродинамике (скорость и направление потоков, вязкость), потокам на входе в систему (гранулометрический состав, температура среды, давление, концентрация примесей).

Для всех этих параметров в справочниках приведены их средние или интервальные значения, а на практике они носят случайный характер, который определяется взаимным влиянием этих величин друг на друга, флуктуацией структуры потоков, особенностью конструкций аппаратов, машин и т.д.

Неопределенность может вызываться также погрешностями датчиков, исполнительных механизмов, информационных каналов, неточностью задания переменных в математической модели объекта, начальных и граничных условиях и т.п.

### 2.7.1. Определение и классификация неопределенности

На рис. 1.29 представлена классификация неопределенностей. Основные типы неопределенностей приведены на рис. 2.19, а.

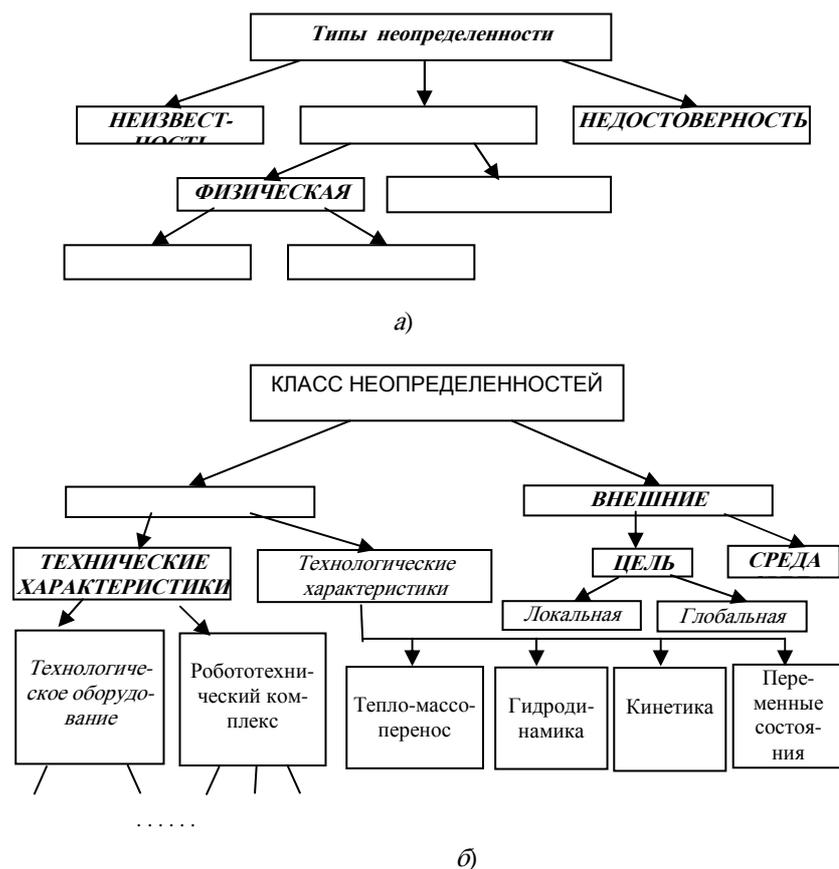


Рис. 2.19. Классификация типов (а) и классов (б) неопределенностей

*Неизвестность* – этот тип неопределенности имеет место в том случае, когда неизвестна информация на выходе объекта исследования, т.е. нет технических средств и способов ее измерения.

*Неоднозначность* – имеет место в том случае, когда при строго определенной информации на входе некоторого технологического объекта информация на выходе имеет различие. Неоднозначность может быть различной природы:

- физическая природа обуславливается случайностью или неточностью информации, что, как правило, определяется техническими измерительными средствами;
- лингвистическая природа имеет место в том случае, когда информация, используемая для анализа технологического объекта, может быть сформулирована на качественном (лингвистическом) уровне.

*Недоуверенность* – этот тип неопределенности может быть в случае, когда информация об исследуемом объекте, явлении или технологических параметрах не является строго определенной (детерминированной).

На рис. 2.21, б представлены классы неопределенностей. Внутренняя неопределенность обуславливается техническими и технологическими характеристиками процессов и производств, а внешняя – определяется целью и состоянием окружающей среды.

Начиная с конца XX в. развитие промышленных производств характеризуется внедрением прогрессивных технологий, обеспечивающих высокий уровень энерго- и ресурсосбережения, ужесточением требований к качеству выпускаемой продукции и экологической чистоты. Это приводит к тому, что ошибки в управлении при использовании детерминированных математических моделей могут приводить к огромным экономическим потерям и способствуют возникновению аварийных ситуаций.

Для решения задач моделирования, управления и оптимизации в условиях неопределенности обычно используют три подхода: стохастический, интервальный анализ и теорию нечетких множеств.

Построение стохастических моделей для решения задач оптимизации, проверки адекватности этих моделей требует проведения большого числа независимых экспериментальных исследований. Последнее сопряжено со значительными трудностями, которые обуславливаются как сложным аппаратным оформлением современных ТП, так и ненаблюдаемостью некоторых из них, а также требует больших материальных затрат. Однако даже в том случае, когда применение теории управления стохастическими объектами строго обосновано и получены необходимые статистические характеристики, можно выделить основной недостаток, присущий большинству задач управления. В этих задачах функционал и ограничения рассматриваются в среднем, при этом не формализуется выполнение технологических и технических требований на тех или иных режимах с заданной вероятностью.

Такой подход имеет ограниченное применение, так как совершенно не исключает возможности аварийных ситуаций, потерь качества продукта, нарушений технологических требований и условий. Причина этого, очевидно, заключается в том, что стремление приблизить работу реальных объектов к оптимальным режимам одновременно сопровождается приближением к предельным значениям технологических требований. В этих условиях учет стохастических свойств объектов лишь «в среднем» является технологически недопустимым, тем более, невозможно использовать стохастические методы для вновь проектируемых производств, так как в этом случае в принципе нельзя провести прямые эксперименты.

Обычно на практике бывает известен лишь приближенно вид распределения, которому принадлежит неизвестная истинная плотность распределения, или неопределенный параметр задается только верхней и нижней границей, и о его поведении на границах и внутри интервала ничего не известно. В этом случае могут быть применены методы интервального анализа, когда неточность формализуется на основе использования интервальных оценок вместо фиксированных чисел.

Однако интервальный анализ также имеет недостатки, которые зачастую могут привести к ошибкам в управлении, потере оптимального решения или к нарушению качественных показателей. Это объясняется, прежде всего, сложностью выбора интервальной оценки технологического параметра, так как при выборе малого интервала увеличивается вероятность нарушения качественных показателей, а при выборе большого интервала возможно нахождение неоптимального технологического режима.

Проблема математической обработки качественной информации включает сбор, оценку достоверности, систематизацию, формализацию, переработку информации качественного характера с применением современных средств вычислительной техники.

Введенное Л. Заде понятие нечеткого множества как математического объекта [16], позволяющего формализовать термины словесного описания особенностей ТП, стимулировало развитие качественного этапа системного анализа и позволило подойти к решению указанной проблемы. При этом стали очевидны следующие достоинства подхода, основанного на аппарате теории нечетких множеств:

- «сжатие» качественной информации, которая определяется целью исследования и осуществляется с использованием методов инженерии знаний;
- наглядность и простота агрегирования и классификации сведений об исследуемом ТП, получаемых из различных источников;
- возможность применения качественной информации при переходе от смысловой к математической постановке задач;
- формирование стратегий управления ТП на основе качественной формализации действий оператора-технолога;
- синтез формальных вычислительных процедур для решения задач оптимизации и управления при нечеткой исходной информации и в нечетко определенных ситуациях задач оптимизации и управления с использованием качественной информации.

Пусть  $X = \{x\}$  – универсальное множество, а  $C$  – определенное свойство. Обычное (четкое) непустое подмножество  $A$  универсального множества  $X$  однозначно определяется характеристическим функционалом

$$Q_A(x) = \begin{cases} 1, & \text{если } x \in X; \\ 0, & \text{если } x \notin X, \end{cases} \quad (2.73)$$

т.е. подмножество  $A$  определяется как совокупность объектов, имеющих некоторое общее свойство  $C$ , наличие или отсутствие которого у любого элемента  $x$  задается характеристическим функционалом, принимающим значение 1, когда  $x$  удовлетворяет свойству  $C$ , и 0 – в противном случае.

Однако такие понятия, как множество «больших», «не очень больших» или «малых величин», уже не являются множествами в классическом смысле, так как не определены границы их степеней малости, которые позволили бы провести классификационную процедуру (2.73) и четко отнести каждый объект к определенному классу. Большинство классов реальных систем и технологических процессов относятся именно к такому нечетко определенному типу. Поэтому возникает необходимость введения понятия о нечетком множестве как о классе с непрерывной градацией степеней принадлежности.

Понятие нечеткого множества – эта попытка математической формализации нечеткой информации для построения математических моделей. В основе этого понятия лежит представление о том, что составляющие данное множество элементы, имеющие общее свойство, могут обладать этим свойством в различной степени и, следовательно, принадлежать к данному множеству с различной степенью.

Для нечеткого подмножества, являющегося расширением понятия множества в классическом смысле, на пространстве объектов  $X = \{x\}$  вводится уже не функционал вида (2.73), а характеристическая функция, задающая для всех элементов степень наличия у них некоторого свойства, по которому они относятся к подмножеству  $A$ . Эта характеристическая функция для нечеткого множества традиционно носит название *функции принадлежности*.

Численное значение функции принадлежности характеризует степень принадлежности элемента некоторому нечеткому множеству, являющемуся в выражении естественного языка некоторой элементарной характеристикой явления (степени эффективности технологического режима, степени загрязнения среды, степени достоверности параметров модели и др.).

*Нечеткое подмножество*  $A$  множества  $X$  характеризуется функцией принадлежности  $\mu_A : X \rightarrow [1,0]$ , которая ставит в соответствие каждому элементу  $x \in X$  число  $\mu_A(x)$  из интервала  $[0, 1]$ , характеризующее степень принадлежности элемента  $x$

подмножеству  $A$ . Причем 0 и 1 представляют собой соответственно низшую и высшую степень принадлежности элемента к определенному подмножеству. Будем обозначать такое нечеткое множество  $\tilde{A}$ .

*Пример.* Пусть  $X = \{0, 1, 2, \dots, 9, 10\}$  целочисленное множество значений давления в аппарате. Тогда нечеткое множество  $\tilde{A}$  «высокое давление» может быть представлено следующим образом:

$$\tilde{A} = \{6|0,2, 7|0,5, 8|0,8, 9|1, 10|1\}.$$

Это следует понимать следующим образом: значения 9 и 10 с абсолютной уверенностью можно отнести к «высокому давлению», 8 – есть «высокое давление» со степенью 0,8 и т.д., а значения 1, 2, ..., 5 абсолютно не являются «высоким давлением». Можно использовать кусочно-линейную аппроксимацию функции принадлежности нечеткого множества, как показано на рис. 2.20.

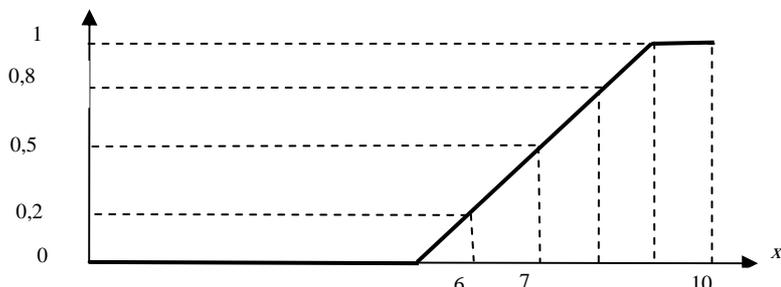


Рис. 2.20. Функция принадлежности нечеткого множества «высокое давление»

Таким образом, значение функции принадлежности  $\mu_{\tilde{A}}(x)$  элемента  $x$  к нечеткому множеству  $\tilde{A}$  можно интерпретировать как определенную экспертом (группой экспертов) степень соответствия элемента  $x$  понятию, формализованному нечетким множеством  $\tilde{A}$ .

*Высотой* нечеткого множества  $\tilde{A}$  называется величина  $\sup_{x \in X} \mu_{\tilde{A}}(x)$ . Нечеткое множество  $\tilde{A}$  *нормально*, если максимальное значение его функции принадлежности  $\sup_{x \in X} \mu_{\tilde{A}}(x) = 1$ . При  $\sup_{x \in X} \mu_{\tilde{A}}(x) < 1$  нечеткое множество называется *субнормальным*.

Непустое субнормальное нечеткое множество можно преобразовать (нормализовать) в нормальное по формуле:

$$\mu_{\tilde{A}}(x) = \frac{\mu'_{\tilde{A}}(x)}{\sup_{x \in X} \mu'_{\tilde{A}}(x)}.$$

*Носителем* нечеткого множества  $\tilde{A}$  называется четкое подмножество универсального множества  $X$ , элементы которого имеют ненулевые значения функции принадлежности, т.е.  $\text{supp}(\tilde{A}) = \{x : \mu_{\tilde{A}}(x) > 0\}$ ,  $\forall x \in X$ . Нечеткое множество называется *пустым*, если его носитель является пустым множеством.

*Ядром* нечеткого множества  $\tilde{A}$  называется четкое подмножество множества  $X$ , элементы которого имеют значения функции принадлежности, равные 1, т.е.  $\text{core}(\tilde{A}) = \{x : \mu_{\tilde{A}}(x) = 1\}$ ,  $\forall x \in X$ . Ядро субнормального нечеткого множества пустое.

Множеством  $\alpha$ -уровня ( $\alpha$ -сечение) нечеткого множества  $\tilde{A}$  называется четкое подмножество множества  $X$ , элементы которого имеют значения функции принадлежности большие или равные  $\alpha$ , т.е.  $A_{\alpha} = \{x : \mu_{\tilde{A}}(x) \geq \alpha\}$ ,  $\forall \alpha \in [0, 1]$ .

*Пример.* Пусть  $\tilde{A} = \{1|1, 2|0,8, 3|0,5, 4|0,1, 5|0\}$ . Тогда  $A_{0,1} = \{1, 2, 3, 4\}$ ,  $A_{0,5} = \{1, 2, 3\}$ .

Иллюстрация определения носителя, ядра,  $\alpha$ -сечения,  $\alpha$ -уровня нечеткого множества приведена на рис. 2.21.

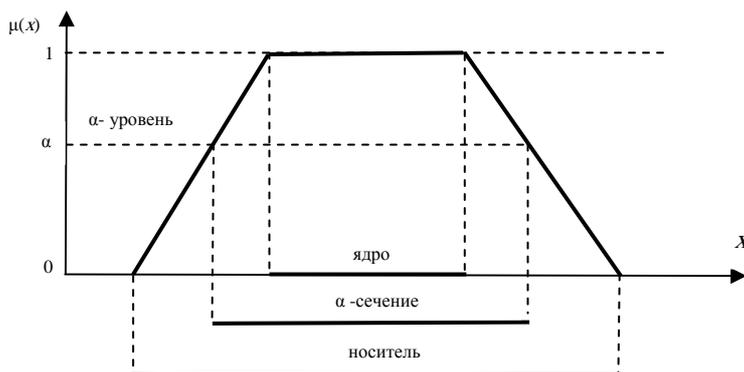


Рис. 2.21. Ядро,  $\alpha$ -сечение,  $\alpha$ -уровень, носитель нечеткого множества

С нечеткими множествами можно производить все алгебраические действия, выполняемые для обычных множеств: получать их сумму (объединение), произведение (пересечение) и дополнение. При этом действия с нечеткими множествами сводятся к алгебраическим операциям с функциями принадлежности.

**Объединение нечетких множеств.** Пусть даны множества  $\tilde{A}$  и  $\tilde{B}$ . Их объединением  $\tilde{A} \cup \tilde{B}$  называется нечеткое множество, функция принадлежности которого вычисляется следующим образом

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \max_x (\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)).$$

Объединение соответствует союзу *или* и его можно записывается как

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \mu_{\tilde{A}}(x) \vee \mu_{\tilde{B}}(x),$$

где символ  $\vee$  соответствует операции взятия  $\max$ .

**Пример.** Пусть нечеткие множества  $\tilde{A}$  и  $\tilde{B}$  заданы в табличной форме:

$\tilde{A}$

$\tilde{B}$

$x_1$	$x_2$	$x_3$	$x_4$
0,1	0,2	0,3	1

$x_1$	$x_2$	$x_3$	$x_4$
0,5	0,9	0,1	0

Тогда

$\tilde{A} \cup \tilde{B}$

$x_1$	$x_2$	$x_3$	$x_4$
0,5	0,9	0,3	1

**Пример.** Пусть  $\tilde{A}$  нечеткий интервал между 5 и 8 и  $\tilde{B}$  нечеткое число приблизительно 4. Объединение нечеткого множества между 5 и 8 с приблизительно 4 показывается на рис. 2.22 (жирная линия).

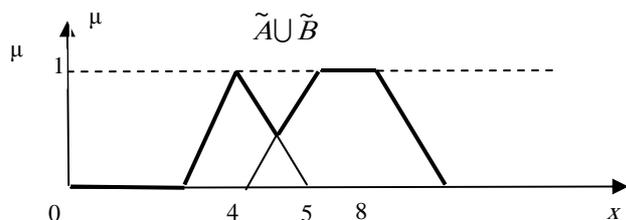


Рис. 2.22. Функция принадлежности  $\mu_{\tilde{A} \cup \tilde{B}}(x)$

**Пересечение нечетких множеств.** Пусть даны множества  $\tilde{A}$  и  $\tilde{B}$ . Их пересечением  $\tilde{A} \cap \tilde{B}$  называется нечеткое множество, функция принадлежности которого вычисляется следующим образом

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \min_x (\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)).$$

Пересечение соответствует союзу *и* и его можно записать как

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \mu_{\tilde{A}}(x) \wedge \mu_{\tilde{B}}(x),$$

где символ  $\wedge$  соответствует операции взятия  $\min$ .

**Пример.** Пусть нечеткие множества  $\tilde{A}$  и  $\tilde{B}$  заданы в табличной форме:

$\tilde{A}$

$\tilde{B}$

$x_1$	$x_2$	$x_3$	$x_4$
0,1	0,4	0,5	1

$x_1$	$x_2$	$x_3$	$x_4$
0,7	0,9	0,3	0

Тогда

$\tilde{A} \cap \tilde{B}$

$x_1$	$x_2$	$x_3$	$x_4$
0,1	0,4	0,3	0

**Пример.** Пусть  $\tilde{A}$  нечеткий интервал между 5 и 8 и  $\tilde{B}$  — нечеткое число, приблизительно 4. Пересечение нечеткого множества между 5 и 8 с приблизительно 4 показывается на рис. 2.23 (жирная линия).  $\tilde{A} \cap \tilde{B}$ .

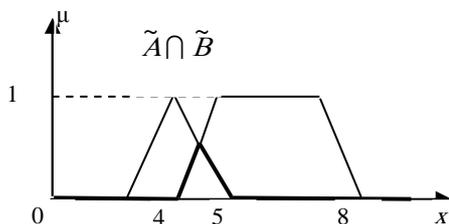


Рис. 2.23. Функция принадлежности  $\mu_{\tilde{A} \cap \tilde{B}}(x)$

**Дополнение нечеткого множества.** Дополнением нечеткого множества  $\tilde{A}$  служит также нечеткое множество  $\tilde{A}'$ , функция принадлежности которого вычисляется следующим образом

$$\mu_{\tilde{A}'}(x) = 1 - \mu_{\tilde{A}}(x).$$

*Пример.* Пусть нечеткое множество  $\tilde{A}$  задано в табличной форме, т.е.

$\tilde{A}$				Тогда $\tilde{A}'$			
$x_1$	$x_2$	$x_3$	$x_4$	$x_1$	$x_2$	$x_3$	$x_4$
0,8	1	0,7	0	0,2	0	0,3	1

Дополнение нечеткого множества  $\tilde{A}$  можно показать графически (рис. 2.24).

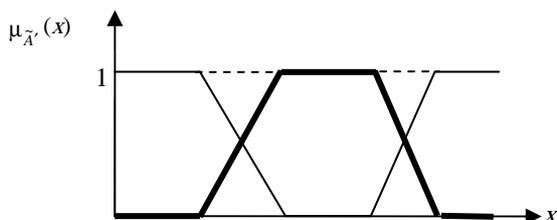


Рис. 2.24. Функция принадлежности  $\mu_{\tilde{A}'}(x)$

В теории нечетких множеств разработан общий подход к выполнению операторов пересечения, объединения и дополнения, реализованный в так называемых треугольных нормах и конормах. Приведенные выше реализации операций пересечения и объединения – наиболее распространенные случаи  $t$ -нормы и  $t$ -конормы.

Существует свыше десятка типовых форм кривых для задания функций принадлежности. Наибольшее распространение получили: треугольная, трапециевидальная и гауссова формы функции принадлежности.

Треугольная функция принадлежности (рис. 2.25) определяется тройкой чисел  $(a, b, c)$ , и ее значение в точке  $x$  вычисляется согласно выражению:

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b; \\ \frac{c-x}{c-b}, & b \leq x \leq c; \\ 0, & x \notin [a, c]. \end{cases}$$

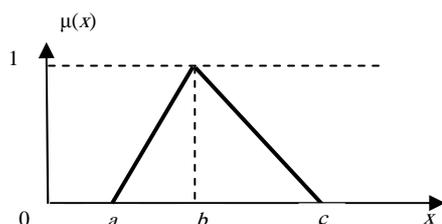


Рис. 2.25. Треугольная функция принадлежности

При  $(b - a) = (c - b)$  имеем случай симметричной треугольной функции принадлежности, которая может быть однозначно задана двумя параметрами из тройки  $(a, b, c)$ .

Аналогично для задания трапециевидальной функции принадлежности (рис. 2.27) необходима четверка чисел  $(a, b, c, d)$ :

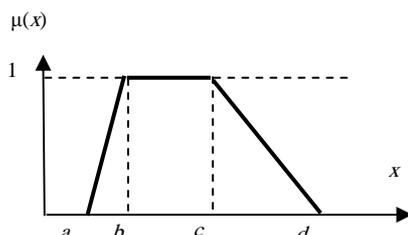


Рис. 2.26. Трапециевидальная функция принадлежности

$$\mu(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b \leq x \leq c; \\ \frac{d-x}{d-c}, & c \leq x \leq d; \\ 0, & x \notin [a; d]. \end{cases}$$

При  $(b-a) = (d-c)$  трапецидальная функция принадлежности принимает симметричный вид. Функция принадлежности гауссова типа (рис. 2.27) описывается формулой

$$\mu(x) = \exp\left[-\left(\frac{x-c}{\sigma}\right)^2\right]$$

и оперирует двумя параметрами. Параметр  $c$  обозначает центр нечеткого множества, а параметр  $\sigma$  отвечает за крутизну функции.

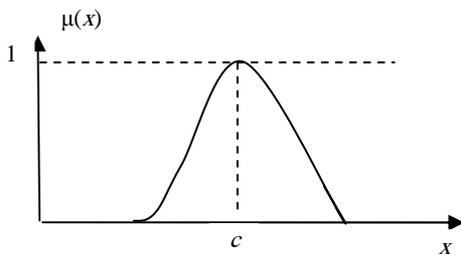


Рис. 2.27. Функция принадлежности гауссова типа

Задание значений переменных технологического режима (производства) для человека более естественно словами, т.е. без использования чисел. Оператор или технолог принимают решения на основе лингвистической информации типа: «очень низкий уровень», «высокая температура», «низкое качество сырья», «длительное перемещение», «скорость больше средней», «красивый цвет» и т.п. Психологи установили, что человеческий мозг почти всю числовую информацию вербально перекодирует и хранит в виде лингвистических термов.

*Лингвистической* называется переменная, принимающая значения из множества слов или словосочетаний естественного или формального языка, служащие, как правило, некоторой элементарной характеристикой явления. Множество допустимых значений лингвистической переменной называется *терм-множеством*. *Термом* называется любой элемент терм-множества, который формализуется нечетким множеством с помощью функции принадлежности. Например, лингвистическая переменная «температура» может иметь следующие термы: «низкая», «высокая», «средняя», которые образуют терм-множество лингвистической переменной «температура». Лингвистическая переменная «клапан» может иметь следующие термы: «открывать быстро», «не трогать», «закрывать медленно», «закрывать быстро».

Отметим, что для большинства практических задач и необходимой точности описания систем вполне достаточно трех термов в лингвистической переменной, а если не хватает словарного запаса в термах, то следует увеличить их число (рекомендуется не более семи).

Понятия нечеткой и лингвистической переменных используются при описании объектов и явлений с помощью нечетких множеств.

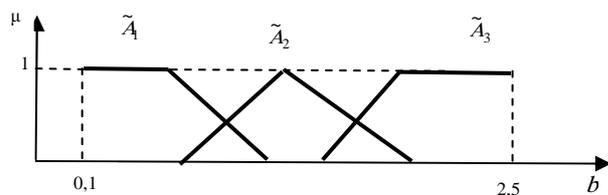
Нечеткая переменная характеризуется тройкой  $\langle a, X, \tilde{A} \rangle$ , где  $a$  – имя переменной;  $X$  – универсальное множество (область определения  $a$ );  $\tilde{A}$  – нечеткое множество на  $X$ , описывающее ограничение (т.е.  $\mu_{\tilde{A}}(x)$ ) на значение нечеткой переменной  $a$ .

*Лингвистической переменной* называется набор  $\langle b, T, X, G, M \rangle$ , где  $b$  – имя лингвистической переменной;  $T$  – множество ее значений (терм-множество), представляющих собой наименования нечетких переменных, областью определения каждой из которых является множество (носитель)  $X$ . Множество  $T$  называется базовым терм-множеством лингвистической переменной;  $G$  – синтаксическое правило, порождающее названия значений лингвистической переменной  $b$ ;  $M$  – семантическая процедура, позволяющая превратить каждое новое значение лингвистической переменной, образуемое процедурой  $G$ , в нечеткую переменную, т.е. сформировать соответствующее нечеткое множество.

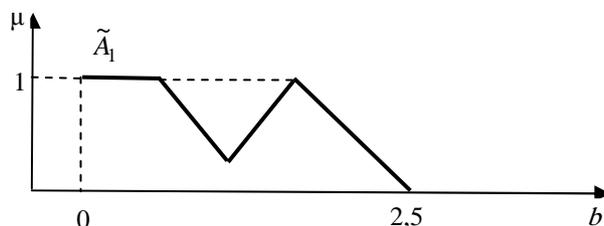
*Идея.* Пусть эксперт определяет уровень жидкости в емкости с помощью выражений «низкий уровень», «средний уровень» и «высокий уровень», при этом минимальный уровень равняется 0,1 м, а максимальный – 2,5 м.

Формализация этого описания может быть проведена с помощью лингвистической переменной  $\langle b, T, X, G, M \rangle$ , для которой  $b$  – уровень жидкости;  $T = \{\text{«низкий уровень»}, \text{«средний уровень»}, \text{«высокий уровень»}\}$ . Заметим, что наряду с базовыми значениями лингвистической переменной возможно определение значений лингвистической переменной «уровень жидкости» как «около 0,2 м», «приблизительно 2 м», т.е. в виде нечетких чисел;  $X = [0,1, 2,5]$ ;  $G$  – процедура образования новых термов с помощью связок «и», «или» и модификаторов типа «очень», «не», «слегка» и др. (например, «не низкий», «очень низкий уровень» и др.);  $M$  – процедура задания на  $X = [0,1, 2,5]$  нечетких подмножеств  $\tilde{A}_1 = \text{«низкий уровень»}$ ,  $\tilde{A}_2 = \text{«средний уровень»}$ ,  $\tilde{A}_3 = \text{«высокий уровень»}$ , а также нечетких множеств для термов из  $G(T)$  соответственно правилам трансляции нечетких связок и модификаторов «и», «или», «не», «очень», «слегка» и другие операции над нечеткими множествами вида:  $\tilde{A} \cap \tilde{B}$ ,  $\tilde{A} \cup \tilde{B}$ .

Следует отметить, что базовое терм-множество и расширенное терм-множество можно характеризовать функциями принадлежности, например как показано на рис. 2.28 и рис. 2.29.



**Рис. 2.28. Функции принадлежности терм-множества:**  
 $\tilde{A}_1$  – «низкий уровень»;  $\tilde{A}_2$  – «средний уровень»;  $\tilde{A}_3$  – «высокий уровень»



**Рис. 2.29. Функция принадлежности:**  
 $\tilde{A}_1 \cup \tilde{A}_2$  – нечетное множество «низкий или средний уровень»

### 2.7.2. Модели нечеткого логического вывода

Простейшие модели нечеткого логического вывода, имеющие практические приложения, основаны на продукционных правилах двух видов:

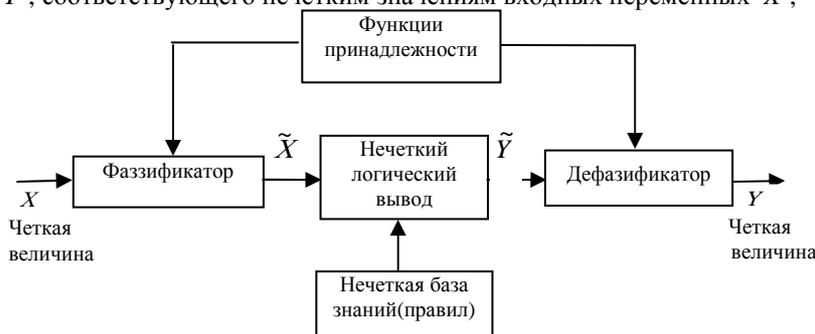
П<sub>i</sub>: Если  $x_1$  есть  $A_{i1}$  ... и ...  $x_n$  есть  $A_{in}$ , то  $Y_k$  есть  $D_{isk}$ ,  $i = \overline{1, m}$ ,  $k = \overline{1, n}$ ;

П<sub>i</sub>: Если  $x_1$  есть  $A_{i1}$  ... и ...  $x_n$  есть  $A_{in}$ , то  $y_k = f_{ik}(x)$ ,  $i = \overline{1, m}$ ,  $k = \overline{1, n}$ , где  $X, Y$  – входные и выходные лингвистические переменные типа «плотность», «производительность», «давление», «температура»,  $A_{ik}, D_{ik}$  – означают термы этих переменных, например, «очень высокая», «большая», «низкая», определенные как нечеткие подмножества соответствующих множеств численных значений переменных, и  $f_i$  – некоторые вещественные функции,  $m$  – число правил,  $n$  – количество переменных.

Нечеткие модели, основанные на правилах первого или второго типа, соответственно называются моделями Мамдани или Сугено. Кроме названных моделей известны модели Цукамото и Ларсена.

Типовая структура модели на основе нечеткого логического вывода содержит следующие блоки (рис. 2.30):

- в блоке «Фаззификатор» определяются степени истинности, т.е. значения функций принадлежности для левых частей каждого правила (предпосылок);
- блок «Нечеткая база знаний» содержит информацию о зависимости  $Y = f(X)$  в виде лингвистических правил типа «если – то»;
- в блоке «Нечеткий логический вывод» на основе правил базы знаний определяется значение выходной величины в виде нечеткого множества  $\tilde{Y}$ , соответствующего нечетким значениям входных переменных  $\tilde{X}$ ;



**Рис. 2.30. Структура модели нечеткого логического вывода**

– блок «Дефаззификатор» преобразует выходное нечеткое множество  $\tilde{Y}$  в четкую величину  $Y$ . Существует несколько методов дефаззификации, например метод центра тяжести, метод центроида, метод наибольшего значения.

В качестве примера рассмотрим, как можно преобразовать выходное нечеткое множество в четкую величину методом центра тяжести. Для этого используется следующая формула:

$$y' = \frac{\int y \mu(y) dy}{\int \mu(y) dy}, \quad (2.74)$$

где  $y'$  – расчетное выходное четкое значение,  $\mu(y)$  – функция принадлежности соответствующего выходного нечеткого множества.

Для кусочно-линейной функции принадлежности формулу (2.74) можно представить в следующем виде:

$$y' = \frac{\sum_i y'_i S_i}{\sum_i S_i}, \quad (2.75)$$

где  $y'_i$  – координата  $u$  центра фигуры под  $i$ -м отрезком прямой функции принадлежности;  $S_i$  – площадь этой фигуры.

Расчет четкого значения по формуле (5.3) представлен на рис. 2.31:

$$y^t = \frac{1,67 \cdot 0,5 + 3,0 \cdot 2,0 + 4,67 \cdot 1,0}{0,5 + 2,0 + 1,0} = 3,29.$$

Рассмотрим нечеткий вывод на примере модели Мамдани, в которой взаимосвязь между входами  $X = (x_1, x_2, \dots, x_n)$  и выходом  $u$  определяется нечеткой базой знаний следующего вида:

если  $(x_1 = a_{1,jl})$  и  $(x_2 = a_{2,jl})$  и ... и  $(x_n = a_{n,jl})$  или

...

или  $(x_1 = a_{1,jk})$  и  $(x_2 = a_{2,jk})$  и ... и  $(x_n = a_{n,jk})$ , то  $y = d_j$ ,  $j = \overline{1, s}$ ,

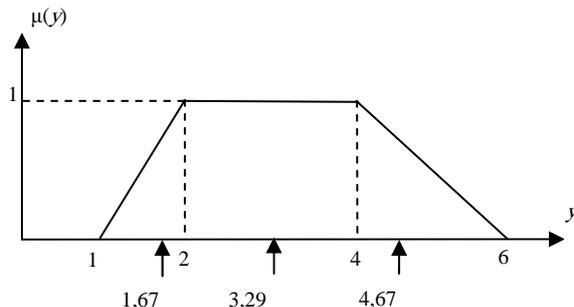


Рис. 2.31. Функция принадлежности для расчета методом центра тяжести

где  $a_{i,jk}$  – терм лингвистической переменной  $x_i$  в строке с номером  $jp$  ( $p = \overline{1, m}$ );  $m$  – число правил, в которых выходная лингвистическая величина  $u$  оценивается термом  $d_j$ ;  $s$  – количество термов, используемых для оценки лингвистической переменной  $u$ .

Используя операции «пересечения» и «объединения» нечеткую базу знаний можно представить в более компактной форме

$$\bigcup_{p=1}^{k_j} \bigcap_{i=1}^n (x_i = a) \rightarrow y = d_j.$$

На рис. 2.32, 2.33 показан нечеткий вывод по Мамдани для двух входных лингвистических переменных  $x_1, x_2$  и двух нечетких правил  $\Pi_1, \Pi_2$  и одной выходной лингвистической переменной  $u$ .

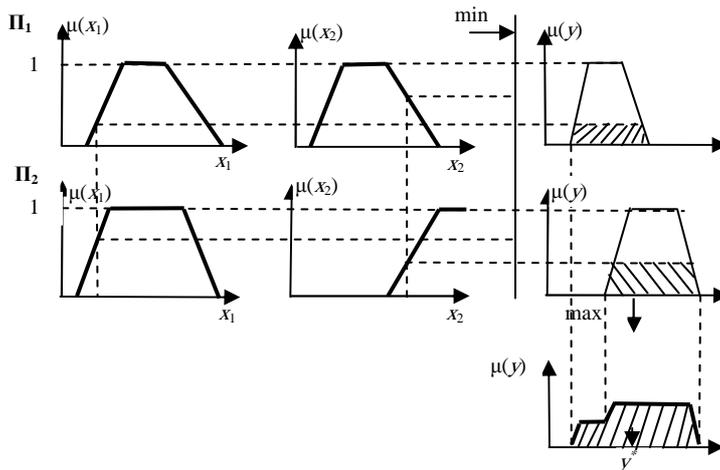


Рис. 2.32. Схема нечеткого вывода Мамдани

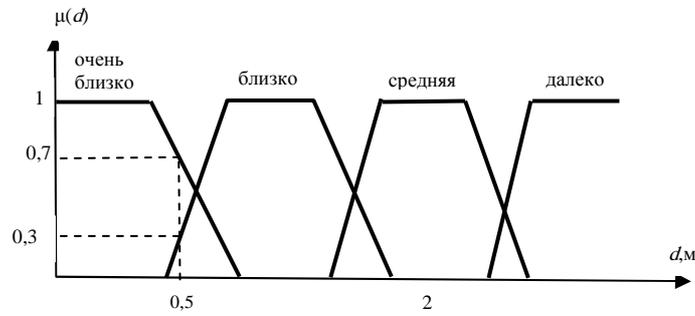


Рис. 2.33. Лингвистическая переменная и функция принадлежности

Математическое описание ХТС, функционирующих в условиях неопределенности в общем виде может быть представлено следующим образом

$$\tilde{y} = M(\tilde{x}, u, \tilde{b}), \quad \tilde{x} \in \tilde{X}, \quad \tilde{b} \in \tilde{B}, \quad \tilde{y} \in \tilde{Y},$$

где  $M$  – оператор нечеткой математической модели;  $\tilde{X}$  – нечеткое множество значений режимных параметров (например, производительность, качество исходного сырья или полупродуктов и др.);  $\tilde{B}$  – нечеткое множество значений настроечных параметров (например, коэффициенты тепло- и массоотдачи, физико-химические константы, фактические характеристики оборудования и др.);  $\tilde{y}$  – нечеткая выходная величина.

Очевидно, что функция принадлежности выходных величин  $\mu_{\tilde{y}}(y)$  зависит от управляющего воздействия  $u$ . Чтобы подчеркнуть эту зависимость будем в дальнейшем обозначать  $\mu_{\tilde{y}}(y|u)$ .

Математическую модель, позволяющую определить функцию принадлежности  $\mu_{\tilde{y}}(y|u)$  в зависимости от детерминированного значения управляющего воздействия  $u$  и функций принадлежности  $\mu_{\tilde{x}}(x)$  и  $\mu_{\tilde{b}}(b)$ , запишем в виде

$$\mu_{\tilde{y}}(y|u) = M(\mu_{\tilde{x}}(x), u, \mu_{\tilde{b}}(b)).$$

Будем называть  $M$  «определяющей моделью» в том смысле, что она рассчитывает (определяет) функцию принадлежности  $\mu_{\tilde{y}}(y|u)$  выходных величин.

Эта модель является функциональным оператором, который ставит в соответствие функциям принадлежности  $\mu_{\tilde{x}}(x)$ ,  $\mu_{\tilde{b}}(b)$  и управляющему воздействию  $u$  функцию принадлежности  $\mu_{\tilde{y}}(y|u)$

$$M: X \times U \times B \rightarrow Y,$$

где  $X, B, Y$  – функциональные пространства соответственно  $\mu_{\tilde{x}}(x)$ ,  $\mu_{\tilde{b}}(b)$  и  $\mu_{\tilde{y}}(y)$ .

При этом решение математической модели  $M$  может быть определено на основе использования принципа расширения Заде, т.е. задать эту модель алгоритмически следующим образом:

$$\mu_{\tilde{y}}(y|u) = \max_{x,b} \min(\mu_{\tilde{x}}(x), \mu_{\tilde{b}}(b)) | y = M(x, u, b),$$

где  $M$  – детерминированная математическая модель,  $\mu_{\tilde{y}}(y|u) = 0$ , если  $\{(x, u, b) | y = M(x, u, b)\} = \emptyset$ .

Таким образом, определяющая математическая модель  $M$  при расчете  $\mu_{\tilde{y}}(y|u)$  многократно использует детерминированную модель ТП.

Алгоритм реализации определяющей математической модели представлен на рис. 2.34.

В блоке 1 блок-схемы (рис. 2.34) вводится значение управляющего воздействия  $u$ , для которого необходимо построить реакцию: функцию принадлежности  $\mu(y)$ .

Блок 2 организует цикл перебора  $x$  и  $b$ .

Блок 3 – для каждого  $x_j$  и  $b_k$  определяет соответствующие значения функции принадлежности  $\mu(x_j)$  и  $\mu(b_k)$  и минимальное значение из этих двух величин  $a_{ijk}$ :  $a_{ijk} = \min[\mu(x_j), \mu(b_k)]$ .

Блок 4 – вычисляет  $y_{ijk}$ , соответствующее заданным значениям  $u, x_j, b_k$  по математической модели  $y = M(u, x, b)$ .

Блок 5 – запоминаются значения  $y_{ijk}$  и  $a_{ijk}$ , формируя таблицы  $Y$  и  $A$ .

Блок 6 – определяет окончание цикла перебора  $x$  и  $b$ .

Таким образом в блоках 2 – 6 рассчитываются и запоминаются все возможные  $y_{ijk}$  для заданного  $u$  и соответствующие им  $a_{ijk}$ .

Блок 7 – организует цикл перебора  $u$  и определяет для каждого из них значение функции принадлежности  $\mu(y_i)$ .

Блок 8 – определяет интервал величиной  $2\Delta_i$ , где  $\Delta_i$  – заданная точность расчета такая, что принадлежность  $u$  этому интервалу идентифицируется как значение  $y = y_i$ .

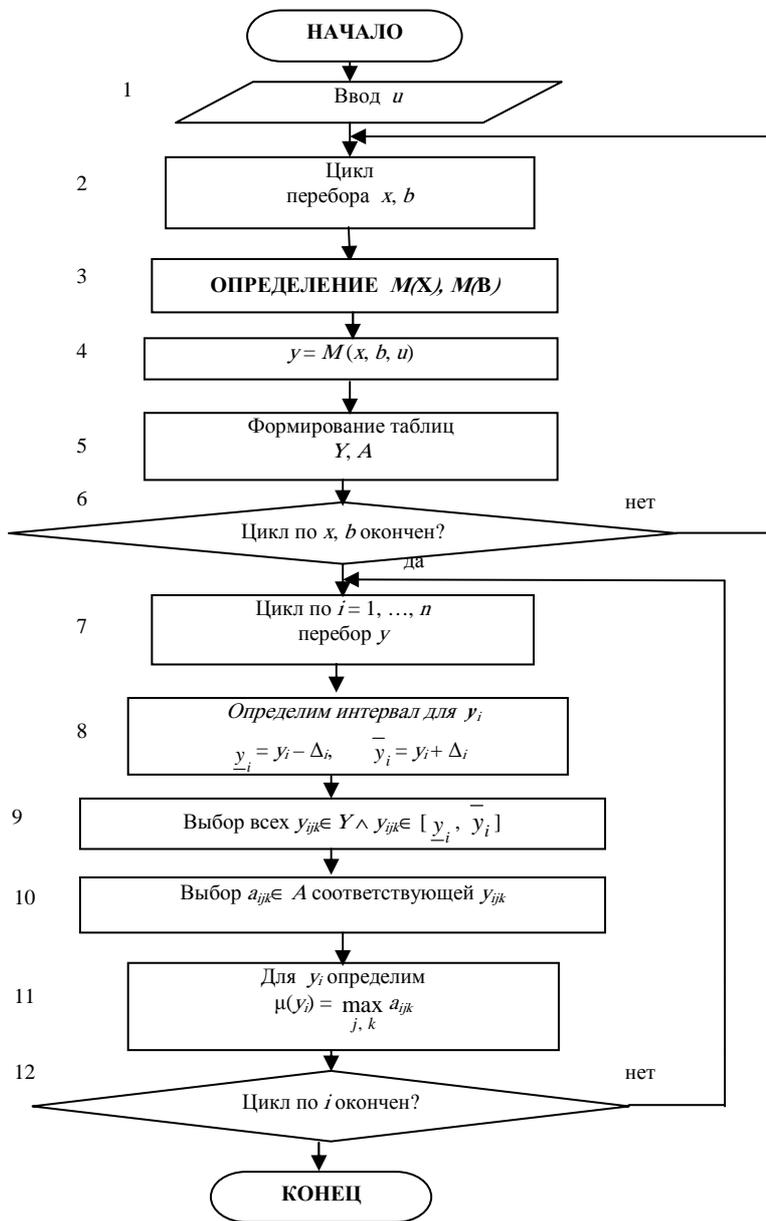
Блок 9 – находит из заполненной таблицы  $Y = \{y_{ijk}\}$ , значения  $y_{ijk} \in [y_i - \Delta_i, y_i + \Delta_i]$ , где  $y_i - \Delta_i$  и  $y_i + \Delta_i$  и идентифицирует их как  $y_i$ .

Блок 10 – для каждого из найденных  $y_{ijk}$  выбирается из таблицы  $A$  соответствующее значение  $a_{ijk}$ .

Таким образом, формируется множество  $a_{ijk}$ , соответствующих  $y_i$ .

Блок 11 – определяет значение функции принадлежности  $\mu(y_i)$ , соответствующее значению  $y_i$  по формуле  $\mu(y) = \max_{j,k} a_{ijk}$ .

Блок 12 – определяет окончание цикла.

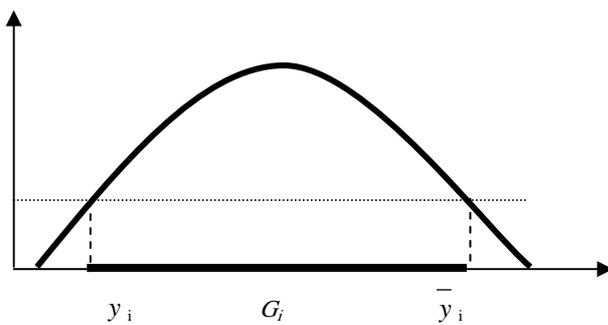


**Рис. 2.34. Схема алгоритма решения уравнений определяющей модели**

Если цикл окончен, то функция принадлежности  $\mu(y)$  для заданного значения  $u$  построена.

Параметрами адекватности будем называть составляющие вектора выходных величин математической модели, т.е.  $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_m)$ , используемые при решении задач моделирования, управления и оптимизации. Вектор выходных величин описывается соответствующей функцией принадлежности  $\mu_{\tilde{y}}(y|u)$ .

Назовем некоторую постоянную величину  $\epsilon_i$  уровнем существенности параметра адекватности  $\tilde{y}_i$  ( $i = \overline{1, m}$ ). Областью существенности  $G_i$  множество  $y_i$  таких, что (рис. 2.35):



**Рис. 2.35. Изображение области существования  $G_i$**

$$G_i = \{ y_i \mid \mu_i(y_i | u) \geq \epsilon_i \}.$$

Назовем границами  $\underline{y}_i, \overline{y}_i$  существенности значений параметров  $y_i$  числа, определяемые по формулам:

$$\underline{y}_i = \min_{G_i} y_i; \quad \bar{y}_i = \max_{G_i} y_i. \quad (2.76)$$

Будем считать, что некоторые экспериментальные значения  $y_i^3$  удовлетворяют условиям адекватности, если

$$|y_i^3 - \hat{y}_i| \leq \frac{\Delta_i}{2}, \quad (2.77)$$

где  $\hat{y}_i = \frac{\underline{y}_i + \bar{y}_i}{2}$ ;  $\Delta_i = \bar{y}_i - \underline{y}_i$ .

Будем считать, что математическая модель адекватна реальному объекту, если при любых значениях управляющих воздействий, принадлежащих множеству допустимых управлений, значения параметра адекватности модели удовлетворяют условиям (2.77).

Таким образом, определение адекватности определяющей математической модели реальному объекту управления заключается в следующем:

1. Определение экспериментального значения управляющего воздействия  $u_j^3$ , где  $j$  – номер эксперимента.
2. Вычисление для  $u_j^3$  функций принадлежности  $\mu_{ij}(y_{ij} | u_j^3)$  параметров адекватности  $\tilde{y}_i$ .
3. Определение значений  $\underline{y}_i$ ,  $\bar{y}_i$ ,  $\hat{y}_i$  и  $\Delta_i$  по формулам (2.76).
4. Измерение экспериментальных значений или вычисление по экспериментальным значениям  $y_i^3 = (y_1^3, \dots, y_j^3)$ .
5. Проверка выполнения условия (2.77).

На рис. 2.36 представлена блок-схема алгоритма проверки адекватности определяющей математической модели.

В том случае, если определяющая математическая модель адекватна реальному объекту управления, т.е. выполняются условия (2.77), то возможно ее применение для решения задач моделирования, управления и оптимизации.

При невыполнении условия (2.77) хотя бы для одного параметра  $y_i$  ( $i = \overline{1, m}$ ), необходимо проводить коррекцию математической модели.

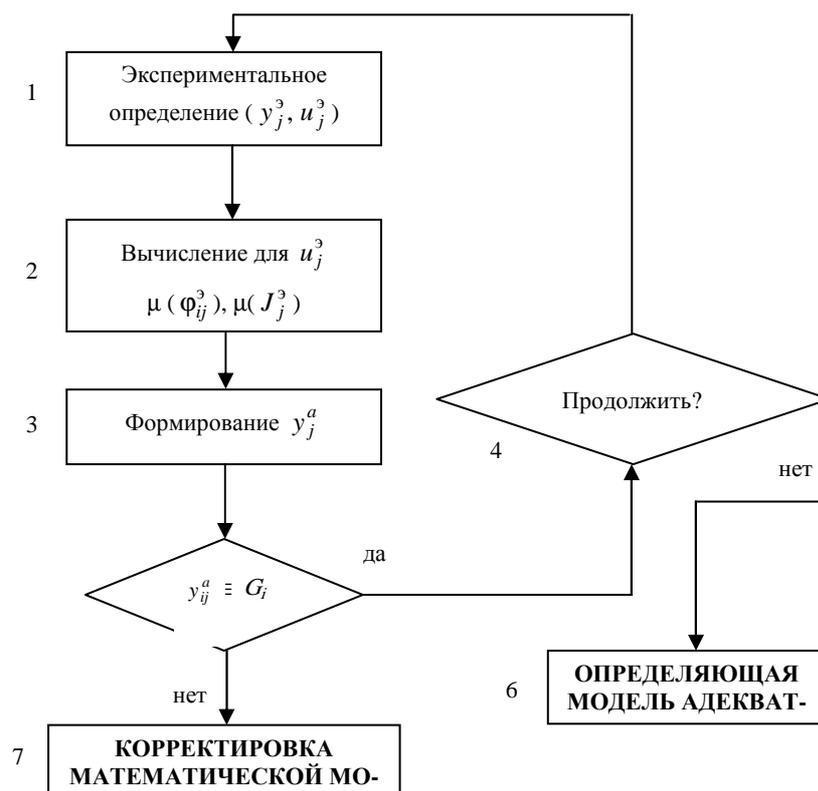
Задача коррекции определяющей математической модели может быть сформирована следующим образом.

Необходимо из множества  $B$  функций принадлежности  $\mu(b)$  настроечных параметров  $b$  модели найти  $\mu^*(b)$ , при которой для всех экспериментальных значений  $(u_j^3, y_j^3)$  ( $j = \overline{1, k}$ , где  $k$  – число экспериментов) выполняются условия адекватности и при этом принимает минимальное значение рассогласование между откорректированной функцией принадлежности  $\mu(b)$  и функцией принадлежности  $\mu^{\text{экс}}(b)$ , полученной на основании экспертных оценок, при выполнении физической реализуемости  $\mu(b)$ .

Математически эта задача формализуется следующим образом: необходимо найти такую функцию принадлежности  $\mu^*(b)$ , при которой достигается минимума функционал

$$L^*(\mu(b)) = \min_{\mu(b) \in B} \|\mu^{\text{экс}}(b) - \mu(b)\|, \quad (2.78)$$

при удовлетворении условий  $\underline{b} \geq \underline{b}_{\text{доп}}$ ,  $\bar{b} \leq \bar{b}_{\text{доп}}$ , где неравенства должны выполняться по соответствующим составляющим (5.5) и



**Рис. 2.36. Алгоритм проверки адекватности определяющей математической модели**

$\forall(j = \overline{1, k}), \forall(i = \overline{1, m}); \underline{b}, \overline{b}$  – соответственно минимальное и максимальное значение соответствующих составляющих областей определения откорректированной функции принадлежности  $\mu(b)$ ;  $\underline{b}_{\text{доп}}, \overline{b}_{\text{доп}}$  – соответственно минимальное и максимальное допустимое значение  $\underline{b}$  и  $\overline{b}$ .

Корректировка математической модели в поставленном виде является чрезвычайно сложной вариационной задачей.

Эта задача значительно упрощается, если функцию принадлежности искать в классе кусочно-линейных функций, например, в треугольной, трапециoidalной (рис. 2.28 – 2.29) и других формах.

При этом функция принадлежности задается несколькими точками, например, функция принадлежности треугольной формы задается тремя точками  $\underline{b}, \overline{b}$  и  $b_c$ , где  $b_c$  – значение, соответствующее максимальному значению функции принадлежности  $\mu(b)$ , функция принадлежности трапециoidalной формы задается четырьмя точками и т.п.

При таком подходе функционал (2.78) становится функцией конечного числа переменных, и задача коррекции значительно упрощается и переходит в разряд минимизации функции от конечного числа переменных.

### Контрольные вопросы

1. Какие подходы применяются при построении моделей систем?
2. В чем заключается экспериментальный метод построения моделей систем?
3. Изложите суть и особенности аналитического и комбинированного методов построения модели системы.
4. Какие фундаментальные законы природы используются при построении аналитических моделей?
5. Приведите примеры моделей статики и динамики систем.
6. Каковы подходы к решению задач моделирования и оптимизации в условиях неопределенности?
7. Что понимается под термином «лингвистическая переменная»?
8. Перечислите основные свойства нечетких множеств.
9. Сформулируйте понятие термина «функция принадлежности».
10. Дайте определение термина «адекватность нечеткой модели».
11. Рассмотреть задачу

$$\begin{aligned} \min_{z, d} d \\ g_1 = -z + \xi \leq 0, \\ g_2 = z - 2\xi + 2 - d \leq 0, \\ g_3 = -z + 6\xi - 9d \leq 0. \end{aligned}$$

Параметр  $\xi$  имеет однородное распределение:

$$p(\xi) = 0,5, \quad \Xi = \{\xi : 1 \leq \xi \leq 3\}.$$

Пусть имеются ограничения

$$0 \leq d \leq 3, \quad 0 \leq z \leq 5.$$

Требуется:

- а) решить одноэтапную задачу оптимизации с жесткими ограничениями;
- б) решить одноэтапную задачу оптимизации с мягкими ограничениями, при этом ограничения должны удовлетворяться в среднем;
- в) решить одноэтапную задачу оптимизации с мягкими ограничениями, при этом ограничения должны удовлетворяться с вероятностью  $\alpha \geq 0,5$ .

12. Рассмотреть задачу

$$\begin{aligned} \min_{z, d} (d + 0,5z) \\ -d - z\xi + 3 \leq 0, \\ d \geq 0, \quad z \geq 0, \end{aligned}$$

где  $d, z, \xi$  – скаляры. Мы будем предполагать, что параметр  $\xi$  имеет однородное распределение:  $P(\xi) = 1, \quad \Xi = \{\xi : 0 \leq \xi \leq 1\}$ .

Требуется:

- а) решить одноэтапную задачу с жесткими ограничениями;
- б) решить одноэтапную задачу оптимизации с мягким ограничением, при этом мягкое ограничение должно удовлетворяться в среднем;
- в) решить одноэтапную задачу оптимизации с мягким ограничением, при этом мягкое ограничение должно удовлетворяться с вероятностью  $\alpha \geq 0,5$ ;
- г) решить двухэтапную задачу для случая, когда ограничение жесткое;
- д) решить двухэтапную задачу для случая, когда ограничение мягкое, при этом мягкое ограничение должно удовлетворяться с вероятностью  $\alpha \geq 0,5$ .

## 3. АДСОРБЦИОННЫЙ МЕТОД ОЧИСТКИ И РАЗДЕЛЕНИЯ СМЕСЕЙ

### 3.1. АДСОРБЦИОННОЕ РАВНОВЕСИЕ И СТРУКТУРА

*Адсорбцией* называют процесс концентрирования веществ на поверхности раздела фаз или в объеме пор твердого тела. В процессе адсорбции участвуют как минимум два агента: тело, на поверхности или в объеме пор которого происходит концентрирование поглощаемого вещества (его называют *адсорбентом*), и поглощаемое вещество (*адсорбтив*).

По типу сил, обуславливающих процесс адсорбции, различают *физическую адсорбцию* и сорбцию, основанную на силах химического взаимодействия, или *хемосорбцию*. Физическая адсорбция вызывается силами молекулярного взаимодействия. В отличие от физической адсорбции при хемосорбции не сохраняется индивидуальность адсорбтива и адсорбента. При сближении молекул адсорбтива с поверхностью происходит перераспределение электронов взаимодействующих компонентов с образованием химической связи. Если физическую адсорбцию можно сравнить с конденсацией, то хемосорбционный процесс должен рассматриваться как химическая реакция, протекающая на поверхности раздела фаз.

Адсорбция является процессом самопроизвольным и экзотермическим. Физическую и химическую адсорбцию можно различить по теплоте адсорбции. Теплота физической адсорбции соизмерима с теплотой конденсации веществ и не превышает 80...120 кДж/моль. Теплота хемосорбции достигает нескольких сотен килоджоулей на моль. Хемосорбция, как правило, протекает с невысокой скоростью. Она может происходить при высоких температурах, когда скорость физической адсорбции пренебрежимо мала.

Твердые адсорбенты имеют поры разного размера: *микро-, мезо- и макропоры*. Микропоры имеют размеры, соизмеримые с адсорбируемыми молекулами. По данным рентгеновского метода их эффективные радиусы преимущественно находятся в интервале 0,5...1,0 нм. В качестве верхней границы размера микропор принимают радиус 1,5 нм. Суммарный объем микропор промышленных адсорбентов обычно не превышает 0,5 см<sup>3</sup>/г.

Адсорбция в микропорах сводится к заполнению пространства микропор адсорбируемыми молекулами. Основным параметром пористой структуры является объем микропор (например, для единицы массы адсорбента) и обобщенная характеристика их размеров. Понятие «удельная поверхность» для микропор не имеет геометрического смысла и ее определение по принятым уравнениям сводится к нахождению формальной константы этих уравнений.

Эффективные радиусы мезопор лежат в интервале от 1,5 до 100...200 нм и во много раз превышают размеры адсорбируемых молекул. Для этого случая приобретает физический смысл понятие о поверхности раздела фаз, т.е. о поверхности пор адсорбента. Обычно поверхность адсорбента относят к единице его массы и называют *удельной поверхностью*. Основными параметрами мезопор являются удельная поверхность, объем пор и функция распределения объема пор по размерам. Для мезопор действие адсорбционных сил проявляется не во всем объеме пор, а практически только на небольшом расстоянии от стенок. Поэтому на поверхности пор происходит мономолекулярная и полимолекулярная адсорбция паров, т.е. образование последовательных адсорбционных слоев, завершающееся заполнением этой разновидности пор по механизму капиллярной конденсации. Удельные поверхности мезопор могут заключаться в интервале 10...400 м<sup>2</sup>/г. Мезопоры являются главными транспортными артериями, по которым осуществляется подвод вещества к микропорам.

Самые крупные поры адсорбентов – макропоры имеют эффективные радиусы больше 100...200 нм. Их удельная поверхность очень мала – 0,5...2 м<sup>2</sup>/г, вследствие чего адсорбцией на поверхности пор этого типа практически можно пренебречь. Объем макропор у активных углей составляет 0,2...0,8 см<sup>3</sup>/г. Макропоры играют роль крупных транспортных артерий в зернах адсорбентов.

Все адсорбенты в соответствии с преобладающим размером пор можно подразделить на три предельных структурных класса: макропористые, мезопористые и микропористые.

При равновесии для выбранной системы адсорбент–адсорбтив количество  $a$  поглощенного газа или пара является функцией давления  $p$  поглощаемого вещества и температуры  $T$ :  $a = f(p, T)$ . При описании адсорбционного равновесия количество поглощенного вещества (*адсорбционную способность, величину адсорбции, адсорбируемость*) выражают обычно в молях на 1 г адсорбента, граммах на 100 г чистого адсорбента, а иногда в граммах на 1 см<sup>3</sup> адсорбента. В качестве характеристики адсорбционных свойств твердых тел используют зависимость адсорбционной способности от давления при постоянной температуре – *изотерму сорбции*:  $a = f(p)$  при  $T = \text{const}$ .

Первым фундаментальным уравнением изотермы адсорбции было уравнение Ленгмюра. Процесс адсорбции Ленгмюр представлял как результат динамического взаимодействия молекул адсорбтива с активными центрами на поверхности адсорбента:



В соответствии с кинетической теорией газов в единицу времени о поверхность, равную единице, ударяется  $\mu$  молекул, причем

$$\mu = p / (2\pi mKT)^{1/2},$$

где  $p$  – давление;  $m$  – масса молекулы;  $K$  – постоянная Больцмана;  $T$  – абсолютная температура.

Из числа ударяющихся молекул только часть фиксируется на поверхности, остальные упруго отражаются в газовую фазу. Доля неупругих столкновений от общего числа столкновений молекул газа с поверхностью была названа коэффициентом конденсации  $\alpha_0$ . Необходимо учитывать, что в процессе адсорбции часть поверхности  $S$  уже покрыта молекулами адсорбата и, следовательно, активной является лишь вторая, свободная часть.

Тогда скорость адсорбции, отнесенная к единице поверхности, равна

$$W_A = \alpha_0 (1 - S) \mu.$$

При десорбции только те молекулы могут снова перейти в газовую фазу, которые обладают достаточной энергией, превышающей теплоту адсорбции молекул  $q_m$ . Число молекул, десорбирующихся с единицы поверхности в единицу времени

$$v = K_0 \exp(-q_m / KT),$$

где  $K_0$  – энтропийный коэффициент.

Скорость десорбции с единицы поверхности равна  $W_d = vS$ . При равновесии скорости адсорбции и десорбции равны:  $W_A = \alpha_0(1-S)\mu = W_d = vS$ . Решая последнее уравнение относительно  $S$ , получим

$$S = \frac{\alpha_0 \mu / v}{1 + \alpha_0 \mu / v} = \frac{bp}{1 + bp},$$

где  $b = \frac{\alpha_0 \exp(q_m / KT)}{K_0 (2\pi mKT)^{1/2}}$ .

Степень заполнения поверхности можно выразить через отношение адсорбционной способности при равновесном давлении  $p$  к предельной адсорбционной способности  $a_m$ , соответствующей оккупации всех активных центров поверхности.

Уравнение изотермы адсорбции Ленгмюра  $a = \frac{a_m bp}{1 + bp}$  охватывает широкий интервал давлений. В начальной области изотермы  $bp \ll 1$ , и уравнение принимает вид  $a \approx a_m bp$ . На этом участке уравнение Ленгмюра идентично уравнению Генри для растворимости газов: адсорбционная способность растет линейно с давлением. При большом давлении, наоборот,  $bp \gg 1$ , наступает заполнение поверхности монослоем молекул и изотерма становится параллельной оси абсцисс  $a \approx a_m$ .

Уравнение  $a = \frac{a_m bp}{1 + bp}$  может быть преобразовано следующим образом:

$$\frac{p}{a} = \frac{1}{a_m b} + \frac{1}{a_m} p \quad \text{или} \quad \frac{1}{a} = \frac{1}{a_m} + \frac{1}{a_m b} \frac{1}{p}.$$

Линейность изотермы является обязательным условием при доказательстве правомерности применения мономолекулярной адсорбции для рассматриваемой системы.

В уравнении локализованной адсорбции Ленгмюра величина  $a_m$  связана с количеством центров адсорбции. Однако опыты свидетельствуют, что уравнение Ленгмюра справедливо и для нелокализованной адсорбции. В этом случае  $a_m$  – емкость монослоя – пропорциональна удельной поверхности адсорбента.

Уравнение Ленгмюра часто применяют в качестве равновесного интерполяционного соотношения при расчете кинетики и динамики процессов поглощения примесей из газовой среды.

Широкое распространение в практических расчетах по адсорбции получило степенное уравнение Фрейндлиха:

$$a = kp^n,$$

где  $a$  – количество адсорбированного вещества;  $p$  – давление адсорбтива;  $k, n$  – постоянные величины.

Достоинство уравнения Фрейндлиха в простоте и удовлетворительном соответствии опыту в широком интервале давлений, исключая, может быть, интервал самых низких равновесных давлений, когда адсорбция строго описывается уравнением Генри. Обычно в расчетах используют линейную форму уравнения:  $\lg a = \lg k + n \lg p$ .

В число основных видов адсорбентов входят активные угли, силикагели, цеолиты, пористые стекла и природные глинистые породы. Активные угли получают из различных видов органического сырья: твердого топлива (торфа, бурого и каменного угля, антрацита), дерева и продуктов его переработки (древесного угля, опилок). Структура угля представлена гаммой пор всех размеров, причем адсорбционная емкость и скорость адсорбции компонентов промышленных газов определяется содержанием микропор в единице массы или объема гранул.

Силикагели получают на основе диоксида кремния, который существует в кристаллической (кварц, тридимит, крстобалит) и аморфной формах. Основные преимущества силикагелей: низкая температура, требуемая для регенерации (110...200 °С) и, как следствие, более низкие энергозатраты, чем при регенерации других промышленных минеральных сорбентов (оксид алюминия, цеолиты); высокая механическая прочность по отношению к истиранию и раздавливанию.

Другим типом неорганических адсорбентов, широко применяющихся в технике для осушки различных сред и для других целей является активный оксид алюминия и алюмогели. Оксид алюминия получают прокаливанием гидроксидов алюминия (тригидратов и моногидратов).

Цеолиты – алюмосиликаты, содержащие в своем составе оксиды щелочных и щелочноземельных металлов, – отличаются строго регулярной структурой пор, которые в обычных температурных условиях заполнены молекулами воды. Эта вода, названная цеолитной, при нагреве выделяется. Свойства природных цеолитов изучены и систематизированы в трудах академиков Вернадского и Ферсмана.

Микропористые и биопористые стекла разной структуры оказалось возможным получить изменением в структуре исходного стекла сплошной борокислородной сетки и отдельных включений боратной фазы. Пористые стекла нашли широкое применение для решения некоторых задач разделения веществ, например, в медицине.

Среди адсорбентов значительное место занимают глинистые породы, в состав которых входят минералы с регулярной структурой. Глинистые материалы как адсорбенты применяют в основном для очистки различных жидких сред от примесей. Часть глинистых природных адсорбентов обладает высокой активностью в естественном виде и их подготовка в промышленности заключается лишь в термической обработке. Другая часть глинистых пород (бентониты) приобретает высокие адсорбционные свойства после химической активации.

Поверхность твердого тела, находящегося при температуре  $T$  в контакте с газом при давлении  $P$ , непрерывно бомбардируется молекулами газа. Число ударяющихся о единицу поверхности за единицу времени молекул равно:

$$n = \frac{N_A P}{\sqrt{2\pi MRT}},$$

где  $N_A$  – число Авогадро;  $M$  – молекулярный вес;  $R$  – газовая постоянная. При этом часть газовых молекул упруго отражается от поверхности, а часть на некоторое время задерживается, т.е. адсорбируется. Относительную долю задерживающихся («прилипших») молекул  $\chi$  часто называют коэффициентом прилипания. Вероятность и продолжительность адсорбции зависят от сил межмолекулярного взаимодействия, места, куда ударились молекула, температуры и т.д. Но даже кратковременная задержка молекул приводит к их концентрированию на поверхности или, как говорят, избыточной концентрации или плотности (по сравнению с концентрацией или плотностью на большом удалении от поверхности). Это избыточное содержание молекул в приповерхностном слое и называется адсорбцией.

В равновесных условиях величина адсорбции определяется равенством скоростей адсорбции и десорбции. При малых покрытиях, когда вероятность соударения адсорбтива с местом, уже занятым адсорбатом, мала, среднее число молекул  $\Gamma$ , адсорбированных на единице поверхности, может быть определено как  $\Gamma = \chi n \tau$ , где параметры  $\chi$  и  $n$  определены выше;  $\tau$  – среднее время пребывания молекулы в адсорбированном состоянии.

Яков Френкель в 1924 г. предложил уравнение, связывающее среднее время пребывания молекулы в адсорбированном состоянии  $\tau$  с теплотой десорбции  $Q_d$  в виде

$$\tau = \tau_0 e^{\frac{Q_d}{RT}},$$

где  $\tau_0$  выражается через статистические суммы для адсорбата и образующегося активированного комплекса, причем в типичных случаях  $\tau_0 \approx 1/\nu_c \approx 10^{-13}$  с, где  $\nu_c$  – величина порядка частоты колебаний химических связей, следовательно  $\tau \approx 10^{-13} e^{\frac{Q_d}{RT}}$ .

Это уравнение применимо для различных оценок, в частности, любопытны оценки среднего времени  $\tau$  пребывания молекулы в адсорбированном состоянии в зависимости от  $Q_d$ . Для таких оценок используем характерные значения теплоты физической адсорбции (He – ~0,4 кДж/моль, Ar, O<sub>2</sub>, N<sub>2</sub> – ~17, NH<sub>3</sub>, CO<sub>2</sub>, этилена – ~30...35, воды на полярной поверхности – 40...60 кДж/моль), величины которых для обратимых процессов совпадают с  $Q_d$ . Во всех этих случаях при 293 К среднее время «жизни» молекул в адсорбированном состоянии – малые доли секунды. Но при переходе к хемосорбции с  $Q_d$  80 – 100 кДж/моль значения  $\tau$  быстро возрастают. Уже при 176 кДж/моль (хемосорбция CO на Ni) величина  $\tau \sim 7,8 \cdot 10^{10}$  лет, т.е. на порядок больше возраста Земли. В свою очередь, снижение температуры существенно увеличивает  $\tau$  даже при слабой физической адсорбции. Так, расчетное время  $\tau$  для гелия при снижении температуры до 77 К увеличивается всего в 1,5 раза, но при 4 К – в  $1,4 \cdot 10^5$  раз (до значений, характерных для адсорбции при 293 К с теплотой в 27,9 кДж/моль).

В общем случае при адсорбции молекула теряет, по крайней мере, степень свободы поступательного движения в направлении, перпендикулярном поверхности. Остальные степени свободы могут сохраняться, в том числе две степени свободы поступательного движения, обеспечивающие миграцию по поверхности.

Простейшее уравнение изотермы адсорбции (закон Генри) в этой трактовке можно записать в виде  $\Gamma = n\chi(\tau_0 e^{Q_d/(RT)}) = K_{\Gamma} P$ . Это уравнение применимо при минимальных заполнениях поверхности и малых давлениях, когда газ можно рассматривать как идеальный, а число соударений его молекул с поверхностью  $n$ , не занятой другими адсорбированными молекулами, пропорционально  $P$ . Величина адсорбции  $\Gamma$  отнесена к единице поверхности, а коэффициент пропорциональности  $K_{\Gamma}$  обычно называют константой Генри. Формально  $K_{\Gamma}$  можно рассматривать как константу термодинамического равновесия:  $K_{\Gamma} = \Gamma/P$ , особенность которой в том, что она является смешанной, поскольку содержит давление  $P$  и величину  $\Gamma$  – эквивалент концентрации.

Это уравнение – предельная форма всех теоретических изотерм адсорбции при  $P \rightarrow 0$  и основа стандартизации термодинамических функций, характеризующих адсорбцию. Но закон Генри выполняется строго только при малых давлениях и малых степенях заполнения (покрытия) поверхности.

Степень заполнения поверхности обычно определяется как  $\theta = \Gamma/\Gamma_{\max}$ , где  $\Gamma_{\max}$  – предельное содержание молекул на единице поверхности в заполненном монослое.

Это уравнение изотермы, описывающее ситуации при больших покрытиях, предложено в 1916 г. выдающимся американским ученым Ирвингом Ленгмюром (1881 – 1957), лауреатом Нобелевской премии 1932 г. По смыслу вывода, это уравнение мономолекулярной адсорбции на однородной поверхности, полученное без учета взаимодействия адсорбат–адсорбат.

Простейший вывод этого уравнения основан на анализе квазихимической реакции между молекулой адсорбата  $M$  и активным центром  $Z$  на поверхности адсорбента  $M + [Z] \leftrightarrow [MZ]$ , где  $[MZ]$  – локализованный адсорбционный комплекс. Если пренебречь взаимодействием адсорбат–адсорбат и влиянием образовавшихся адсорбционных комплексов на соседние активные центры, допустить, что число свободных центров на поверхности составляет долю  $(1 - \theta)$  от общего числа центров, а число занятых центров – долю  $\theta$ , то константа равновесия этой реакции  $K = \theta / (P(1 - \theta))$ , откуда следует

$$\theta = \frac{KP}{(1 + KP)} = \frac{\Gamma}{\Gamma_{\max}}.$$

Чаще последнее уравнение записывают в форме

$$a = a_0 KP / (1 + KP),$$

где  $a$  и  $a_0$  – соответствующие величины адсорбции с произвольной, но одинаковой размерностью.

Изотерма Ленгмюра имеет вид кривой с насыщением (рис. 3.1), где в области больших давлений при  $KP \gg 1$  величина  $a = a_0$ , что соответствует заполнению монослоя.

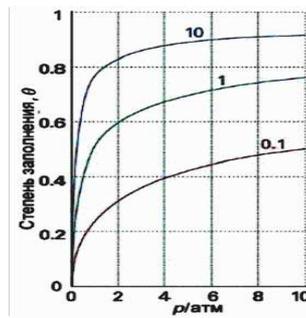


Рис. 3.1 Изотермы адсорбции Ленгмюра с убывающими значениями константы  $K$

Крутизна начального участка изотермы при прочих равных условиях определяется величиной константы равновесия  $K$ , а в области малых давлений при  $KP \rightarrow 0$  уравнение Ленгмюра сводится к уравнению Генри, поэтому константа равновесия  $K$  равна коэффициенту Генри.

Уравнение Ленгмюра широко используется на практике. Для расчета значений  $a_0$  и константы равновесия  $K$  оно записывается в линейном виде как  $P/a - 1/(a_0K) + (1/a)P$ , или  $a/P = Ka_0 - Ka$ , или  $1/a = 1/a_0 + (1/a_0K)(1/P)$ .

Константа адсорбционного равновесия  $K$ , подобно константам равновесия реакций в газах или растворах, связана со стандартным изменением изохорного потенциала уравнением

$$\Delta F^\theta = \Delta U^\theta - T\Delta S^\theta = -RT \ln K,$$

откуда следует  $K = e^{\Delta S^\theta/R} - e^{\Delta U^\theta/RT} = g e^{Q/RT}$ ,

где  $g$  – энтропийный множитель;  $Q$  – теплота адсорбции.

Аналогично можно получить уравнение Ленгмюра для адсорбции смеси компонентов, которые могут адсорбироваться на одних и тех же центрах. Так, адсорбцию бинарной смеси газов  $A$  и  $B$  можно рассматривать как две параллельные реакции взаимодействия газов  $A$  и  $B$  со свободной поверхностью по схеме Ленгмюра, полагая, что молекулы  $A$  и  $B$  адсорбируются на одних и тех же свободных центрах:  $A$  (газ) + свободное место  $[Z] \leftrightarrow [AZ]$  и  $B$  (газ) + свободное место  $[Z] \leftrightarrow [BZ]$ , где  $[AZ]$  и  $[BZ]$  – соответствующие поверхностные комплексы. В условиях равновесия для каждого компонента  $K_A = \theta_A/(P_A[1 - \theta_A - \theta_B])$  и  $K_B = \theta_B/(P_B[1 - \theta_A - \theta_B])$  или  $\theta_A/\theta_B = K_A P_A/(K_B P_B)$ . Дальнейшие простейшие преобразования дают

$$\theta_A = \frac{K_A P_A}{1 + K_A P_A + K_B P_B}.$$

Аналогично получается выражение для  $\theta_B$ . В более общем случае одновременной адсорбции  $i$  компонентов степень покрытия поверхности компонентом  $i$  равна

$$\theta_i = \frac{K_i P_i}{1 + \sum_i K_i P_i},$$

а суммарная степень покрытия поверхности всеми компонентами определяется как

$$\theta_\Sigma = \frac{\sum_i K_i P_i}{1 + \sum_i K_i P_i},$$

где  $P_i$  – парциальное давление компонента смеси  $i$ ;  $K_i$  – его константа равновесия.

Из последнего уравнения следует, что адсорбция каждого компонента смеси возрастает с увеличением его парциального давления  $P_i$ , убывает с ростом парциального давления других компонентов, а также зависит от парциальных значений  $K_i$ . Уравнение Ленгмюра для хемосорбции, сопровождающейся диссоциацией молекул на два фрагмента, каждый из которых

занимает один отдельный центр, записывается как  $\theta = \frac{\sqrt{KP}}{1 + \sqrt{KP}}$ . Уравнение Ленгмюра и его модификации описывают моно-

молекулярную адсорбцию на однородной поверхности, которая не осложнена взаимодействием адсорбированных компонентов между собой. При наличии такого взаимодействия и тех же прочих условиях используется уравнение Фаулера–Гугенгейма:

$$P = \frac{\theta}{K(1-\theta)} e^{-\frac{K_1\theta}{RT}}.$$

В последнем уравнении первый член – уравнение Ленгмюра с константой  $K$ , характеризующей взаимодействие адсорбат–адсорбент, а константа  $K$  в экспоненте характеризует межмолекулярное взаимодействие адсорбат–адсорбат в монослое.

Линеализованная форма этого уравнения имеет вид  $\ln \left[ \frac{\theta}{1-\theta} \right] - \ln P = \ln K + \frac{K_1\theta}{RT}$ .

Среди других уравнений изотерм адсорбции на однородной поверхности, учитывающих взаимодействие адсорбат–адсорбат, отметим уравнение, которое предложено А.В. Киселевым в форме

$$P = \frac{\theta}{K(1-\theta)(1+K_1\theta)}$$

с тем же смыслом констант, что и уравнение Фаулера–Гугенгейма.

Дальнейшее действие сил при физической адсорбции обуславливает образование нескольких слоев адсорбированных молекул. Заполнение второго слоя маловероятно до тех пор, пока на поверхности находятся только отдельные изолированные

молекулы, так как силы взаимодействия между единичной адсорбированной молекулой и ударяющейся об нее молекулой обычно малы по сравнению с силами на поверхности адсорбента, где во взаимодействии одновременно участвует большое число молекул. Но по мере увеличения поверхностной концентрации адсорбента возрастает вероятность одновременного взаимодействия ударяющейся молекулы адсорбтива с тремя, четырьмя и более адсорбированными ранее молекулами. «Время жизни»  $\tau$  возрастает до значений, при которых на отдельных участках поверхности может начинаться полимолекулярная адсорбция.

Простейшее уравнение полимолекулярной адсорбции, полученное независимо на основе разных предпосылок тремя разными авторами и по этому обычно называемое уравнением Френкеля–Хелси–Хилла, имеет вид

$$\theta = B_1 [RT \ln(P/P_0)]^{-n} \quad \text{или} \quad t = B_2 [RT \ln(P/P_0)]^{-n},$$

где  $t$  – среднестатистическая толщина полимолекулярной пленки адсорбата на поверхности адсорбента, получаемая делением объема адсорбата на величину поверхности адсорбента; плотность адсорбата обычно отождествляется с плотностью его нормальной жидкой фазы. Френкель и Хилл вывели это уравнение путем интегрирования выражений для дисперсионного потенциала без учета сил отталкивания, действие которых во втором и последующих монослоях становится малым. В этом случае величина показателя степени  $n = 3$ . Хелси получил подобное уравнение эмпирически, анализируя изотермы адсорбции азота при 77 К на непористых системах, величина показателя степени оказалась близкой  $\sim 2,75$ . Это уравнение обычно используется при анализе распределения пор по размерам по данным капиллярно-конденсационных измерений, а в практике измерений удельной поверхности используется другое уравнение полимолекулярной адсорбции, называемое по имени авторов уравнением Брунауэра, Эмметта, Теллера.

При выводе этого уравнения использованы допущения Ленгмюра: рассматривается адсорбция на однородной поверхности без учета взаимодействия между молекулами в плоскости слоя, но дополнительно вводится учет «вертикальных» межмолекулярных взаимодействий адсорбат–адсорбат. Предполагается, что ударяющиеся молекулы, попадая на уже занятые места, не покидают их немедленно, а образуют адсорбционные комплексы из групп молекул. По мере приближения к давлению насыщения  $P_0$  сначала сокращается число свободных мест, а затем мест, занятых комплексами из одной, двух, трех и т.д. молекул.

Рассмотрим такую полимолекулярную адсорбцию как серию квазихимических реакций образования единичных и кратных комплексов:

адсорбтив + свободная поверхность  $\rightleftharpoons$  [единичные комплексы],

адсорбтив + единичные комплексы  $\rightleftharpoons$  [двойные комплексы],

адсорбтив + двойные комплексы  $\rightleftharpoons$  [тройные комплексы] и т.д.

Пусть  $\theta_1, \theta_2, \theta_3, \theta_4, \dots$  – доля поверхности, занятая единичными, двойными, тройными и т.д. комплексами. Общая величина адсорбции,  $a$  с учетом всех типов комплексов и их кратности, определяется как

$$a = a_m(\theta_1 + 2\theta_2 + 3\theta_3 + 4\theta_4 + \dots),$$

где  $a_m$  – емкость плотного монослоя. Выражения для констант равновесия соответствующих квазихимических реакций записываются в виде  $K_1 = \theta_1/(P\theta_0)$ ;  $K_2 = \theta_2/(P\theta_1)$ ;  $K_3 = \theta_3/(P\theta_2)$  и т.д. ( $\theta_0$  – доля свободной поверхности). Величина константы  $K_1$  обычно намного больше значений  $K_2, K_3, K_4, \dots$ , так как сила взаимодействия быстро убывает по мере увеличения расстояния от поверхности. Константы  $K_2, K_3, K_4$  и т.д. также не равны друг другу, однако разница между ними обычно намного меньше их отличия от  $K_1$ . Поэтому принимается допущение, что все эти константы, кроме  $K_1$ , равны друг другу и близки константе равновесия на границе жидкость–пар, равной

$$K_L = 1/P_0,$$

где  $P_0$  – давление насыщенных паров жидкой фазы адсорбтива при данной температуре. Это допущение позволяет получить блок уравнений типа

$$\theta_1 = K_1 P \theta_0; \quad \theta_2 = K_2 P \theta_1 = K_L P \theta_1 = (P/P_0)^1 \theta_1;$$

$$\theta_3 = K_3 P \theta_2 = (K_L P)^2 = (P/P_0)^2 \theta_1; \quad \text{и т.д.}$$

Суммарная величина сорбции с учетом этих уравнений при введении обозначения  $P/P_0 = h$  записывается как  $a = a_m K_1 P \theta_0 [1 + 2h + 3h^2 + \dots + nh^{n-1}]$ . Баланс заполнения поверхности можно также представить в виде  $\theta_0 + \theta_1 + \theta_2 + \theta_3 + \theta_4 + \dots = \theta_0 \{1 + K_1 P [1 + h + h^2 + \dots + h^n]\} = 1,0$ , причем выражение в квадратных скобках представляет убывающую геометрическую прогрессию  $1 + h + h^2 + h^3 + \dots + h^n = 1/(1 - h)$ . После дальнейших несложных преобразований и введения обозначений  $K_L = 1/P_0$  и  $K_1/K_L = C$ , получим

$$a = a_m C h / \{ [1 - h] [1 + (C - 1)h] \}.$$

Это и есть уравнение Брунауэра, Эмметта, Теллера, которое уже более 60 лет является одним из важнейших соотношений прикладной адсорбции и используется как основа наиболее распространенной методики измерения удельной поверхности катализаторов и адсорбентов.

Уравнение (имеет две константы  $a_m$  и  $C$ , которые легко определяются по линейной форме этого уравнения

$$Y = h/a(1 - h) = 1/(a_m C) + (C - 1)h/(a_m C)$$

из графика зависимости  $Y$  от  $h$ . Константа  $a_m$  равна емкости заполненного монослоя, а константа  $C$ , которую называют энергетической константой уравнения Брунауэра, Эмметта, Теллера и обозначают как  $C_{БЭТ}$ , согласно современной трактовке, равна  $C_{БЭТ} \sim e^{\Delta S^R / (Q - L)/(RT)}$ , где первая экспонента связана с изменениями энтропии при адсорбции, а вторая – с так называемой чистой теплотой адсорбции, равной разнице изостерической теплоты адсорбции  $Q$  и теплоты конденсации чистого адсорбтива  $L$  в жидкость.

Уравнение Брунауэра, Эмметта, Теллера широко используется для расчета емкости монослоя  $a_m$ , по которой далее рассчитывается площадь удельной поверхности катализаторов и адсорбентов. Если величина  $a_m$  выражена в молях на грамм адсорбента, то удельная поверхность  $A$  с размерностью  $\text{м}^2/\text{г}$  определяется по формуле

$$A = a_m N_A \omega,$$

где  $N_A$  – число Авогадро;  $\omega$  – площадка, приходящаяся на одну молекулу в заполненном монослое на поверхности адсорбента.

## 3.2. МЕХАНИЗМЫ АДСОРБЦИИ

### 3.2.1. Квантовомеханические механизмы адсорбции

В теории выделяют два предельных случая адсорбции: хемосорбцию и физическую адсорбцию.

Типичная хемосорбция – это химическая реакция между адсорбатом и поверхностными атомами или группами атомов адсорбента, при которой молекула адсорбата отдает или получает от поверхности электрон (гемолитические реакции) или электронную пару (гетеролитические реакции). При этом молекула адсорбата диссоциирует на радикалы или отдельные атомы. В результате хемосорбированные молекулы теряют индивидуальность и обычно могут десорбироваться только в виде продуктов реакции, участвуют в изотопном обмене с атомами поверхности или другими хемосорбированными молекулами. Природа образующихся при хемосорбции связей в принципе та же, что и в объемных химических соединениях. Однако состояние поверхностных атомов адсорбента отличается от объемного, что может существенно влиять на особенности образующихся связей, распределение электронов и т.д.

Природа сил, вызывающих адсорбцию, может быть различной. В популярном курсе Фейнмановских лекций силы межмолекулярного взаимодействия рассматриваются как результат «неимоверно сложного взаимодействия всех электронов и ядер одной молекулы со всеми электронами и ядрами другой» и отмечается, что их нельзя свести полностью к фундаментальным элементарным взаимодействиям типа кулоновских или гравитационных. Поэтому эти силы относят к нефундаментальным, а их разбиение на составляющие базируется на различных модельных приближениях. Но в общем случае все виды адсорбции имеют общую квантово-механическую основу, природа взаимодействия определяется электронной структурой молекул адсорбата и адсорбента.

Выражение для энергии Ван-дер-ваальсового взаимодействия молекулы с поверхностью твердого тела при больших расстояниях  $z$  молекулы от поверхности можно получить как частный случай из общих выражений макроскопической теории ван-дер-ваальсового взаимодействия двух тел, принимая, что одно из тел является сильно разряженным. При  $a_0 \ll z \ll \lambda_0$  (где  $a_0$  – атомный радиус, а  $\lambda_0$  – длина волн, характерных для спектров поглощения молекулы и твердого тела) из общих выражений макроскопической теории ван-дер-ваальсовых взаимодействий вытекает следующее выражение для потенциальной энергии ван-дер-ваальсового взаимодействия молекулы с твердым телом:

$$\Phi = -\frac{1}{6\pi z^3} \sum_n \int_0^\infty \frac{E_{n0} \left| \vec{q}_{n0} \right|^2}{E_{n0}^2 + \omega^2} \left[ \frac{\varepsilon(i\omega) - 1}{\varepsilon(i\omega) + 1} \right] d\omega,$$

где  $E_{n0}$  – разность энергий в основном и  $n$ -ом возбужденном состояниях;  $\vec{q} = -e \sum_i \vec{r}_i$  – оператор электрического диполя;

$\varepsilon(i\omega)$  – диэлектрическая проницаемость твердого тела, выраженная как функция мнимой частоты  $i\omega$ .

Для потенциальной энергии ван-дер-ваальсового взаимодействия молекулы с полубесконечной изотропной диэлектрической средой при  $a_0 \ll z \ll \lambda_0$  методом электрических изображений Мак-Лечлен получил выражение

$$\Phi = -\frac{h}{4\pi z^3} \sum_n \int_0^\infty \alpha(i\omega) \left[ \frac{\varepsilon(i\omega) - 1}{\varepsilon(i\omega) + 1} \right] d\omega,$$

где  $\alpha(i\omega)$  – электрическая поляризуемость молекулы, выраженная как функция мнимой частоты  $i\omega$ .

Выражение получено при помощи макроскопической модели, можно получить также при помощи микроскопической модели, принимая, что молекула адсорбата и атомы, составляющие твердое тело, состоят из изотропных гармонических осцилляторов.

Оцененная по этому выражению потенциальная энергия взаимодействия атомов благородных газов с графитом заметно меньше определенной эмпирически потенциальной энергии взаимодействия этих атомов с графитированной сажей.

Проблема Ван-дер-ваальсового взаимодействия молекулы с поверхностью твердого тела рассматривалась также с использованием квантовомеханической теории возмущения. Применяя теорию возмущений второго порядка, для энергии  $\Phi$  притяжения между сферической неполярной молекулой и плоской поверхностью любого полубесконечного твердого тела Синаноглу и Питцер получили выражение:

$$\Phi = -\frac{1}{2} \alpha_1 \left| \langle 0 | \vec{F}_1 | 0 \rangle \right|^2 - \frac{1}{2} \alpha_1 \frac{\delta_1}{\delta_1 + \delta_s} \sum_{\tau \neq 0} \left| \langle 0 | \vec{F}_1 | \tau \rangle \right|^2,$$

где  $\alpha_1$  – поляризуемость молекулы;  $\delta_1$  и  $\delta_s$  – средние энергии возбуждения молекулы и твердого тела;  $\vec{F}_1$  – мгновенное электростатическое поле, создаваемое всеми электронами и ядрами твердого тела в центре молекулы адсорбата;  $\tau$  – обозначает различные энергетические уровни твердого тела; первый член равен энергии поляризации молекулы электростатическим полем поверхностью; второй член соответствует энергии дисперсионного взаимодействия.

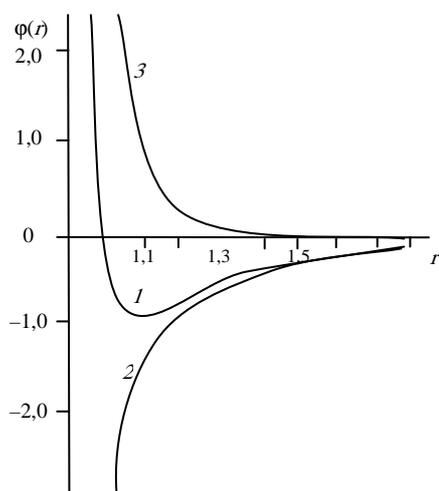


Рис. 3.2. Суммарный потенциал сил взаимодействия (1) и его составляющие: силы притяжения (2) и силы отталкивания (3)

При выводе всех этих выражений для энергии взаимодействия молекулы с поверхностью твердого тела принималось, что расстояние молекулы от поверхности значительно больше атомных размеров. Поэтому в этих теориях фактически учитываются только силы притяжения ван-дер-ваальсового типа; силы отталкивания в них не учитываются. Кроме того, в этих теориях не учитывается атомная структура твердого тела: твердое тело рассматривается как однородная непрерывная среда с математически ровной поверхностью. Однако, как отмечалось выше, адсорбция определяется главным образом потенциальной энергией  $\Phi$  взаимодействия молекулы с поверхностью вблизи потенциального минимума. При таких расстояниях необходимо учесть атомное строение твердого тела и силы отталкивания. Поэтому указанные выше теории не могут быть использованы для определения потенциальной энергии  $\Phi$  вблизи потенциального минимума. При выборе формы потенциала  $\Phi$  и оценке его параметров вблизи потенциального минимума обычно используются результаты приближенных квантово-механических теорий ван-дер-ваальсового взаимодействия двух молекул вблизи равновесного расстояния.

Существует непрерывный спектр форм адсорбции, различающихся по интенсивности взаимодействия, типу сил, характеру образующихся связей Г/Х, изменениям в структуре как адсорбированных молекул, так и поверхности адсорбента.

Качественно картина взаимодействия двух молекул  $A$  и  $B$  выражается показанным на рис. 3.2 графиком зависимости суммарной потенциальной энергии (кривая 1) от расстояния  $r_{12}$  между центрами этих молекул. На большом удалении молекулы не взаимодействуют. При их сближении возникает сила притяжения, которая достигает некоторого максимального значения, а при дальнейшем сближении молекулы начинают отталкиваться друг от друга из-за перекрытия электронных оболочек. Для разделения этих эффектов кривую 1 представляют в виде суммы кривых 2 и 3, соответствующих потенциальной энергии притяжения и отталкивания.

Энергия отталкивания обычно описывается эмпирическими формулами типа

$$\varphi_{от}(r) = Vr^{-n} \text{ или } \varphi_{от}(r) = be^{-\alpha r},$$

где  $r$  – расстояние между центрами взаимодействующих молекул. Силы притяжения в модельном приближении представляются в виде суммы вкладов электростатического, индукционного и дисперсного межмолекулярных взаимодействий.

### 3.2.2. Термодинамические основы поверхностных явлений

Поверхностные явления происходят на границе раздела фаз. Реальная межфазная граница имеет конечную толщину, в пределах которой происходит резкое изменение термодинамических параметров (концентрация, температура, давление и др.). Гиббс предложил допущение, состоящее в следующем: граница раздела имеет нулевую толщину, а термодинамические параметры считаются постоянными вплоть до поверхности раздела.

Свободную поверхностную энергию можно представить в виде изменения энергии Гиббса:

$$dG \leq -S'dT + Vdp + \sigma dS + \sum_{i=1}^m \mu_i dn_i + \zeta dq_i,$$

где  $S'$  – энтропия;  $T$  – температура;  $V$  – объем;  $p$  – давление;  $\sigma$  – удельная свободная поверхностная энергия;  $S$  – площадь поверхности раздела фаз;  $\mu_i$ ,  $n_i$  – химический потенциал и число молей компонента;  $\zeta$ ,  $q$  – электрический потенциал и заряд поверхности. Знак «меньше» в уравнении характеризует самопроизвольное снижение энергии Гиббса, а знак «равно» характеризует равновесие.

Из уравнения следует, что  $\sigma = \left( \frac{\partial G}{\partial S} \right)_{T, p, n_i, q_i}$ , т.е. удельная свободная поверхностная энергия есть частная производная от

энергии Гиббса по площади поверхности раздела фаз при постоянной температуре, давлении, числе молей и заряде.

Для раскрытия физического смысла свободной поверхностной энергии применяется пример, в котором сравнивается поведение молекул на межфазной границе и в объеме тела. Молекулы в объеме находятся в окружении таких же молекул и поэтому их силовые поля скомпенсированы. Молекулы поверхностного слоя взаимодействуют с молекулами обеих граничных фаз, в результате результирующая сил взаимодействия не равна нулю и направлена внутрь той фазы, с которой взаимодействие больше. Свободная поверхностная энергия это результат не скомпенсированного межмолекулярного взаимодействия на границе раздела фаз.

Рассматриваются так же понятия *когезии* и *адгезии*. Эти два понятия описывают два вида межмолекулярных взаимодействий – внутри фазы и между фазами.

Взаимодействие молекул, атомов, ионов внутри одной фазы называется когезия (сцепление). Взаимодействие между контактирующими поверхностями конденсированных тел различной природой называется адгезией (прилипанием).

Когезия обеспечивает существование вещества в конденсированном состоянии. Адгезия обеспечивает соединение определенной прочности между двумя телами.

Рассмотрены количественные характеристики когезионного взаимодействия. Работа когезии  $W_k$  определяется затратой энергии на разрыв тела по сечению. Так как при разрыве образуются две единичные поверхности, то работа когезии равна удвоенному значению поверхностного натяжения на границе с газом:  $W_k = 2\sigma_{2,1}$ . Адгезия – результат стремления системы к уменьшению поверхностной энергии, поэтому адгезия является самопроизвольным процессом. Работа адгезии  $W_a$  – это работа разрыва межмолекулярных связей на поверхности раздела фаз, отнесенная к единице площади. Рассматривается связь между работой адгезии и поверхностным натяжением на примере соединения двух конденсированных взаимно нерастворимых фаз. При совмещении этих поверхностей (нанесение одного на другое) происходит явление адгезии. Так как система остается опять двухфазной, на границе раздела образуется поверхностное натяжение  $\sigma_{2,3}$ . В результате энергия Гиббса уменьшается на величину, равную работе адгезии, т.е.  $\Delta G + W_a = 0$  или  $W_a = -\Delta G$ . Для начального и конечного состояния  $G_{нач} = \sigma_{2,1} + \sigma_{3,1}$  и  $G_{кон} = \sigma_{2,3}$ . Изменение энергии Гиббса в процессе адгезии описывается уравнением Дюпре:  $W_a = -\Delta G = \sigma_{2,1} + \sigma_{3,1} - \sigma_{2,3}$ .

Отсюда следует, что работа адгезии тем больше, чем больше начальное поверхностное натяжение компонентов и меньше поверхностное натяжение после соединения.

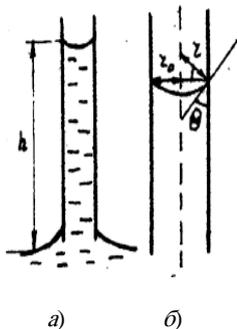
### 3.2.3. Основы капиллярных явлений

Процесс адсорбции часто сопровождается конденсацией паров поглощаемого вещества в капиллярных порах адсорбента, присоединением молекул поглощаемого вещества по месту ненасыщенных валентностей элементов, составляющих кристаллическую решетку адсорбента, и другими процессами.

Рассмотрим поведение жидкости в тонком капилляре, опущенном в сосуд с жидкостью. В этом случае можно считать, что мениск имеет сферическую форму. При условии смачивания поверхности капилляра (острый угол  $\theta$ ) будет образовываться искривление с отрицательным радиусом кривизны (вогнутый мениск). В результате давление жидкости под поверхностью мениска будет пониженным по сравнению с давлением под плоской поверхностью на  $2\sigma/r$ . Жидкость будет подниматься по капилляру до тех пор, пока капиллярное давление не уравнивается гидростатическим давлением столбика жидкости, т.е.

$$\Delta p = (\rho - \rho_0)gh,$$

где  $\rho$  и  $\rho_0$  – плотности жидкости и ее насыщенного пара (или воздуха);  $g$  – ускорение свободного падения;  $h$  – высота подъема столбика жидкости (рис. 3.3, а).



Кривизна поверхности определяется условиями смачивания, т.е. значением краевого угла  $\theta$ . Радиус кривизны мениска  $r$  связан с радиусом капилляра через соотношение  $r = r_0/\cos\theta$  (рис. 3.3, б). Высота поднятия жидкости в капилляре определяется формулой Жюренна:

$$h = \frac{2\sigma \cos \theta}{r_0(\rho - \rho_0)g}.$$

Капиллярные явления возможны и между не полностью погруженными пластинами, расположенными близко друг к другу.

Рассмотренные поверхностные явления вызваны самопроизвольным снижением поверхностной энергии за счет уменьшения поверхностного натяжения.

Рис. 3.3. Поднятие жидкости в смачиваемом капилляре

### 3.2.4. Физическая и химическая адсорбция

Сначала проиллюстрируем различия химической и физической адсорбции на зависимости потенциала взаимодействия  $\zeta(r)$  от межатомного расстояния  $r$  (рис. 3.4).

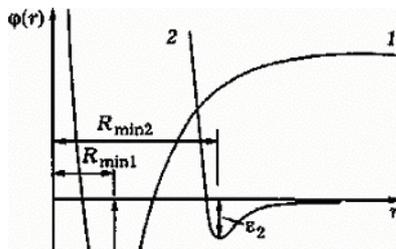


Рис. 1.2 Схема зависимости потенциала межмолекулярного взаимодействия от межатомного расстояния. 1 – при химической и 2 – физической адсорбции

Рис. 3.4. Схема зависимости потенциала межмолекулярного взаимодействия от межатомного расстояния: 1 – при химической адсорбции; 2 – при физической адсорбции

Такой метод наглядной иллюстрации различий между химической и физической адсорбцией был предложен в 1932 г. Леннард-Джонсом. Пусть для конкретности эти графики условно изображают адсорбцию кислорода на платине.

График 2 соответствует рассмотренной выше физической адсорбции. Здесь глубина потенциальной «ямы»  $\varepsilon_2$  связана с теплотой физадсорбции, которая обычно близка теплоте конденсации адсорбтива (в случае кислорода  $\sim 7 \dots 10$  кДж/моль). При ван-дер-ваальсовском радиусе атома кислорода  $\sim 0,140$  нм и платины  $\sim 0,138$  нм суммарный размер  $R_{ww} = R_{\min 2} = 0,278$  нм. Дальнейшему сближению препятствует отталкивание электронных оболочек.

График 1 построен для условной ситуации, когда молекула адсорбтива на бесконечном расстоянии от поверхности диссоциирует на два атома и в результате имеет избыточную энергию, равную энергии диссоциации. Этой энергии обычно достаточно для преодоления сил отталкивания электронных оболочек и сближения с поверхностными атомами вплоть до образования химической связи. Энергии связи соответствует глубокий минимум  $\varepsilon_1$  на кривой 1, пропорциональный теплоте хемосорбции, которая в системе О/Рt составляет  $\sim 280 \dots 300$  кДж/моль. Межатомное расстояние  $R_x = R_{\min 1}$ , оцененное как сумма ионных радиусов (0,064 нм для платины и 0,136 нм для кислорода) близко 0,2 нм, теплота хемосорбции в  $30 \dots 40$  раз больше теплоты физадсорбции, а межатомное расстояние в 1,4 раза меньше.

Различие энергий взаимодействия приводит к столь же существенной разнице значений энтропии адсорбции и, соответственно, степеней свободы адсорбированных молекул. Так, при физадсорбции молекула может терять лишь способность к поступательному движению перпендикулярно к поверхности, а при хемосорбции могут ограничиваться и колебательные степени свободы. Из рис. 1.2 также следует, что переход от кривой 2 к кривой 1 связан с преодолением потенциального барьера, определяющего энергию активации хемосорбции  $E_a$ , причем  $E_a$  перехода от физадсорбции к хемосорбции на поверхности обычно ниже энергии, необходимой для активации в газовой фазе. При наличии нескольких форм хемосорбции, которым соответствуют разные минимумы потенциала и, соответственно, разные межатомные расстояния (например, образование одинарных, двойных и тройных связей), между этими формами также возможны активационные барьеры.

Сильное взаимодействие при хемосорбции сопровождается возмущением приповерхностного слоя атомов адсорбента. В результате поверхность при сильной хемосорбции может радикально реконструироваться, т.е. перестраиваться. Поэтому при хемосорбции адсорбент и адсорбат необходимо рассматривать как единую квантово-механическую систему. В то же время при физадсорбции возмущения в приповерхностном слое относительно слабы, структуру адсорбента в первом приближении допустимо рассматривать как инертную, а адсорбент и адсорбат – как две независимые системы.

Однако существующее многообразие форм адсорбции допускает и достаточно плавный переход от идеальной физадсорбции, обусловленной дисперсионными силами, к хемосорбции, обусловленной химическими поверхностными реакциями. Между этими предельными формами адсорбции возможны различные промежуточные типы специфической адсорбции и слабой хемосорбции. Поэтому часто затруднительно найти экспериментальные критерии, достоверно отличающие специфическую адсорбцию от слабой хемосорбции, хотя проблему идентификации типа адсорбции обычно позволяет решить совокупность используемых критериев. Рассмотрим эти критерии.

1) Теплота адсорбции. Теплота физической адсорбции по порядку величины близка теплоте конденсации паров в жидкость и обычно находится в пределах  $10 \dots 50$  кДж/моль. Однако, например, теплота адсорбции первых порций паров воды на цеолитах может достигать  $80 \dots 100$  кДж/моль из-за взаимодействия диполей молекул воды с катионами в цеолите, причем молекулы воды при такой адсорбции сохраняют индивидуальность и десорбируются в том же виде. Это случай сильной специфической физической адсорбции. В свою очередь, теплота хемосорбции близка теплоте химических реакций и обычно превышает  $80 \dots 100$  кДж/моль. Однако в литературе описаны и случаи эндотермической хемосорбции, например, диссоциативной хемосорбции  $H_2$  на стекле, где эндотермичность адсорбции компенсируется ростом энтропии свободно мигрирующих по поверхности атомов водорода.

2) Зависимость скорости адсорбции от температуры. Физическая адсорбция обычно происходит при сравнительно низких температурах, а повышение температуры, как правило, снижает величины адсорбции. Для хемосорбции характерны более высокие температуры и более сложная зависимость от температуры. Равновесные величины хемосорбции с ростом температуры, как и при физической адсорбции, снижаются. Но в температурном диапазоне, где достижение равновесия лимитируется энергией активации, повышение температуры сопровождается ростом величин хемосорбции. Этот эффект наглядно проявляется при измерении изобар адсорбции.

Физическую адсорбцию и хемосорбцию иногда различают как процессы, происходящие, соответственно, при относительно низких и высоких температурах. Но такое определение без измерений изобар адсорбции в широкой температурной области также не является строгим. Например, в микропористых системах при комнатных температурах физическая адсорбция кислорода и азота происходит в столь заметных количествах, что на этом основаны современные методы разделения воздуха (метод PSA). Но в этом случае температура физической адсорбции  $O_2$  и  $N_2$  на  $\sim 200^\circ$  выше температуры кипения и на  $\sim 150^\circ$  выше критической, и поэтому вряд ли может быть названа относительно невысокой. Кстати, адсорбция воды на цеолитах заметна даже при  $\sim 500$  К. Наконец, известны и случаи диссоциативной хемосорбции при весьма низких температурах.

3) Специфичность адсорбции также часто используется в качестве критерия для установления форм адсорбции. Но, как следует из изложенного выше, хемосорбция практически всегда специфична, а физическая адсорбция может быть полностью неспецифична лишь при соответствующем подборе адсорбента и адсорбтива, на чем и основано широкое использование физической адсорбции для измерения удельной поверхности и других текстурных характеристик катализаторов и адсорбентов.

4) Форма изотерм адсорбции (т.е. зависимостей величин адсорбции от давления паров адсорбтива при постоянной температуре) также в ряде случаев может быть использована в качестве критерия. Хемосорбция обычно завершается монослойным заполнением и имеет в таких случаях форму так называемых ленгмюровских изотерм или изотерм I типа. Для физадсорбции характерен более богатый ассортимент форм изотерм, который включает ту же форму изотерм I типа, которая типична при адсорбции на микропористых адсорбентах.

5) Современные спектроскопические методы (ИКС, ЯМР, ЭПР, ЭСХА и т. д.) существенно расширяют возможности идентификации и исследования типов адсорбционных взаимодействий. Эти методы основаны на взаимодействии излучения или потока электронов с веществом и квантово-механических переходах в молекулярных структурах адсорбата и адсорбента и составляющих их атомов и ядер. В результате методы чувствительны к разного рода изменениям в этих структурах. Изменений интенсивности соответствующих спектров, положения и ширины линий, характерных химических сдвигов и т.д. позволяют детально исследовать механизмы и энергетику межмолекулярных взаимодействий при адсорбции. Новые возможности открывают методы современной сканирующей полевой микроскопии, сочетающие возможности сканирующей туннельной микроскопии с прямыми измерениями межмолекулярного взаимодействия на супрамолекулярном уровне.

### 3.2.5. Тепловые эффекты процессов адсорбции

В процессе адсорбции может выделяться значительное количество тепла. Теплота адсорбции [в кДж/(кмоль)] обычно определяется опытным путем.

Теплота адсорбции дает прямую информацию об энергии связи и природе адсорбционного взаимодействия каждой пары адсорбат/адсорбент, а зависимость теплот адсорбции от заполнения характеризует энергетическую однородность или неоднородность поверхности адсорбента.

Основные методы определения теплот адсорбции  $q$ :

- из изотерм адсорбции путем построения изостер и расчета изостерических теплот адсорбции  $q_{ст}$ ;
- из результатов газохроматографических измерений объемов удерживания  $V_x$  при очень малых концентрациях адсорбтива в виде так называемых дифференциальных теплот адсорбции при нулевом заполнении поверхности  $q_0$  путем построения зависимости  $\ln V_x$  от  $1/T$ , наклон которой равен  $q_0/R$ ;
- из прямых калориметрических измерений в виде дифференциальных теплот адсорбции при разных заполнениях поверхности  $q_d(\theta)$ .

Изостерическая теплота адсорбции  $q_{ст}$  определяется уравнением Клаузиуса–Клапейрона

$$q_{ст} = -RT/(\partial \ln P / \partial T)_a = -R/(\partial \ln P / \partial (1/T))_a ,$$

а дифференциальная теплота адсорбции  $q_d$  численно равна изменению свободной энергии  $\Delta F$  при адсорбции, взятой с обратным знаком, т.е.

$$q_d = -\Delta F = RT^2 / (\partial \ln P / \partial T)_a - RT = -R/(\partial \ln P / \partial (1/T))_a - RT .$$

Следовательно,  $q_{ст} = q_d + RT$ .

Из общих соотношений термодинамики поверхностных явлений следует уравнение для расчета изменений дифференциальной внутренней энергии адсорбции в виде

$$-\frac{\partial \Delta U}{\partial a} = RT^2 \left| \frac{\partial \ln P}{\partial T} \right|_a - RT^2 \frac{d \ln P_0}{dT} .$$

В этом уравнении величина  $RT^2 \left| \frac{\partial \ln P}{\partial T} \right|_a$  равна изостерической теплоте адсорбции  $q_{ст}$ , а величина  $RT^2 \frac{d \ln P_0}{dT}$  – теплоте испарения (конденсации)  $q_L$  жидкой фазы чистого адсорбата в отсутствие адсорбции. Поэтому разность

$$q_{ст} - q_L = -\frac{\partial \Delta U}{\partial a}$$

называется чистой теплотой адсорбции. Эта величина характеризует взаимодействие адсорбат–адсорбент.

Надежное определение изостерических теплот из изотерм возможно только в случаях относительно слабого взаимодействия при легко достижимом обратимом равновесии, т.е. в основном при физической адсорбции, хотя и здесь возникают проблемы точности определения малых давлений. При хемосорбции взаимодействие существенно сильнее, равновесие может долго не устанавливаться и в ряде случаев адсорбционная калориметрия оказывается одним из немногих методов, дающих информацию о количестве и качестве адсорбционных центров с разной энергией на поверхности.

Рассмотрим кратко основные принципы адсорбционной калориметрии. Современный калориметр – это прецизионный научный прибор, имеющий чувствительность порядка  $10^{-4}$  Дж, стабильно работающий в широком температурном диапазоне, включающий дополнительно системы дозирования и измерения величин адсорбции. Существует два основных типа адсорбционных калориметров: теплопроводящий (типа Тиана–Кальве) и адиабатический.

В теплопроводящих калориметрах между адсорбционной ячейкой и окружающей ее изотермической оболочкой поддерживается постоянная разница температур за счет электронагревателя, питаемого импульсами тока. Тепло от оболочки в ампулу с образцом передается через большое число термопар, рабочие спаи которых окружают ампулу с образцом, а спаи сравнения связаны с термостатом. Тепловые эффекты при адсорбции автоматически компенсируются изменением числа импульсов тока в нагревателе, изменения энтальпии определяются по суммарному расходу тока.

Адиабатические калориметры снабжены адиабатической оболочкой, температура которой изменяется так же, как и температура ампулы с образцом. Здесь чувствительным элементом обычно являются платиновые термометры сопротивления, а автоматическая компенсация тепловыделений с помощью импульсных нагревателей минимизирует инерционность системы. Адиабатические калориметры удобны для измерений в изотермическом режиме, а теплопроводящие калориметры – при изменяющейся температуре.

Для исследований физической адсорбции также используется калориметрический метод измерения теплот смачивания пористого тела жидкостью. Такие эксперименты проводят обычно в калориметрах теплопроводящего типа, заливая предварительно подготовленный образец определенным количеством жидкости. Заранее обезгаженная жидкость вводится в калориметр в запаянных ампулах, которые разбивают в вакууме в начале эксперимента. Тепловыделение  $\Delta H$  при смачивании близко интегральной теплоте адсорбции при образовании монослойного покрытия, которая пропорциональна поверхности адсорбента  $A$  и равна

$$\Delta H = hA,$$

где  $h$  – удельная теплота смачивания, отнесенная к единице поверхности. Если величина  $h$  заранее определена в экспериментах с образцами с известной поверхностью, то эта методика позволит определять удельную поверхность. Такая методика была предложена в 1944 г. Гаркинсом и Джура как абсолютный метод измерения удельной поверхности, не связанный с интерпретацией изотерм адсорбции и какими-либо модельными допущениями.

Но все оказалось не так просто. Более сложна и теория метода, которая только в первом приближении сводится к уравнению (3.1.55). Удельные значения  $h$  очень чувствительны к детальному составу и структуре поверхности, а малые значения

$h$  (порядка  $1 \text{ Дж/м}^2$ ) требуют применения прецизионных калориметров и учета тепловыделений, даже связанных с разбиванием ампулы и размешиванием жидкости. Для устранения невоспроизводимого влияния природы поверхности адсорбента Гаркинс и Джура предложили покрывать ее достаточно толстой адсорбционной пленкой той же жидкости до полной экранировки влияния специфики поверхности адсорбента. В этом случае  $h$  определяется только работой когезии жидкости. Но при существовавшей в то время технике оказалось, что толщина покрытия должна достигать пяти – семи мономолекулярных слоев, что делает бессмысленным использование такого метода для измерений, например, поверхности тонкопористых систем.

В более поздних экспериментах, выполненных в 1979 г. Роукероллем с сотрудниками на современном прецизионном калориметре, показано, что для полной экранировки влияния природы поверхности достаточно адсорбционная пленка толщиной в один–полтора монослоя. Ее дальнейшее утолщение уже не влияет на удельные теплоты смачивания, которые совпали в измерениях для 11 адсорбентов разной химической природы по девяти разным жидкостям, включая воду и органические компоненты.

### 3.3. КИНЕТИКА АДСОРБЦИИ

В основе расчета процессов очистки и рекуперации, осуществляемых с помощью твердых адсорбентов, лежат закономерности динамики адсорбции. Динамика сорбционных процессов рассматривает пространственно-временные распределения компонентов между фазами системы (одна из которых – твердая), возникающие при перемещении этих фаз относительно друг друга.

Процессы адсорбции экзотермичны, обратные процессы (десорбция) требуют подвода тепла. Поэтому при анализе некоторых задач динамики сорбции следует учитывать распределение температуры в обеих фазах системы.

Пусть концентрация адсорбата в адсорбенте до начала процесса равна нулю, концентрация адсорбтива во входящем потоке постоянна. Введем поток в колонну и проследим за изменением концентрации в обеих фазах. В этот момент во всех сечениях слоя, кроме лобового, концентрация адсорбата равна нулю. Адсорбент в лобовом слое начинает поглощать адсорбтив, концентрация которого в вошедшей «элементарной» порции газа убывает. Частично истощенная «элементарная» порция адсорбтива вместе с потоком газа поступает в следующий «элементарный» объем слоя адсорбента. В нем также начинается поглощение вещества. Далее в сорбционный процесс вовлекаются третий, четвертый и так далее «элементарные» объемы.

Рассмотрим процесс извлечения пара в гипотетическом горизонтальном адсорбере. В момент времени  $\tau_1$  количество поступившего адсорбтива еще недостаточно для насыщения лобового слоя адсорбента и распределение адсорбата в слое характеризуется кривой 1 (рис. 3.5).

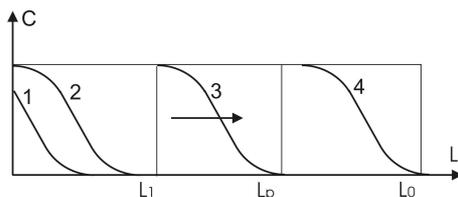


Рис. 3.5. Фронт сорбции в моменты времени  $\tau_1, \tau_2, \tau_3$  и  $\tau_4$

В момент времени  $\tau_2 > \tau_1$  лобовой слой адсорбента насыщается адсорбатом до величины равновесной адсорбционной способности.

Распределение концентрации по длине адсорбера в любой из фаз, достигнутое ко времени  $\tau_2$ , называют *фронтом сорбции*. После этого момента кривая распределения адсорбата часто перемещается с постоянной скоростью  $U$ , определяемой по формуле:

$$U = wc_0 / a_p,$$

где  $w$  – скорость газового потока, см/с;  $c_0$  – концентрация адсорбтива в газе, моль/см<sup>3</sup> газа;  $a_p$  – равновесная адсорбционная способность, моль/см<sup>3</sup> адсорбента.

В момент времени  $\tau_3$  слой адсорбента можно разделить на три зоны: полностью отработавший слой  $L_1$ , работающий слой  $L_p$  и еще не вступивший в работу слой  $L_0$ . В момент времени  $\tau_4$  кривая распределения перемещается в концевой слой адсорбента. С этого момента, момента «проскока», в выходящем потоке газа появляется и начинает прогрессивно возрастать концентрация извлекаемого вещества:

$$\tau_4 = \tau_{пр},$$

где  $\tau_{пр}$  – момент проскока или время защитного действия слоя адсорбента.

Получим уравнение динамики адсорбции, приняв ряд упрощающих допущений:

- 1) адсорбируется лишь один компонент потока;
- 2) подвижная фаза несжимаема, и концентрация адсорбируемого вещества в ней так мала, что можно пренебречь изменением плотности потока вследствие убыли адсорбтива;
- 3) движение потока осуществляется в одном направлении с постоянной скоростью ( $w = \text{const}$ );

В этом случае процесс адсорбции описывается системой уравнений материального и теплового балансов, кинетики адсорбции и теплопередачи, изотермы адсорбции [14].

Пусть  $x$ -координата соответствует направлению перемещения потока со скоростью  $w$  через элемент пористого тела с площадью поперечного сечения  $dS$  и толщиной  $dx$ . Скорость потока при неизотермической адсорбции не является посто-

янной величиной, а подобно концентрациям  $c$  и  $a$ , есть функция координаты и времени. При составлении баланса вещества в выделенном элементарном объеме будем учитывать:

изменение количества вещества в элементарном объеме

$$-\frac{\partial(wcdSdt)dx}{\partial x} = -\frac{\partial wc}{\partial x} dSdxdt,$$

изменение к концентрации вещества в адсорбенте  $(\partial a / dt)d\tau dx dS$

и подвижной фазе  $(\partial c / dt)d\tau dx dS$ .

Отсюда общий материальный баланс в элементарном слое имеет вид:

$$-\frac{\partial(wc)}{\partial x} dSdxdt = \frac{\partial a}{\partial \tau} dSdxdt + \frac{\partial c}{\partial \tau} dSdxdt$$

или

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + \frac{\partial(wc)}{\partial x} = 0.$$

Учет изменения концентрации вещества в элементарном объеме вследствие диффузии приводит к следующему уравнению

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + \frac{\partial(wc)}{\partial x} = D^* \frac{\partial^2 c}{\partial x^2},$$

где  $D^*$  – коэффициент продольной диффузии.

Кинетику нестационарного переноса конвективного (внешнего) массообмена обычно описывают следующим уравнением:

$$da / dt = \beta_{\Gamma}(c - c^*),$$

где  $a$  – величина адсорбции;  $\tau$  – время;  $\beta_{\Gamma}$  – коэффициент внешнего массообмена, отнесенный к единице объема адсорбента;  $c$  – текущая концентрация адсорбтива в потоке;  $c^*$  – концентрация адсорбтива на поверхности раздела фаз, равновесная текущей величине адсорбции.

Другой часто используемой формой кинетического уравнения является формула, в которой движущая сила внутридиффузионного процесса адсорбции записывается как разность концентраций адсорбата в твердой фазе (формула Глюкауфа):

$$da / dt = \beta_{\Gamma}(a^* - a),$$

где  $\beta_{\Gamma}$  – кинетический коэффициент;  $a$  – текущая величина адсорбции;  $a^*$  – величина адсорбции, равновесная текущей концентрации адсорбтива в потоке на внешней поверхности гранул.

При смешанном диффузионном переносе адсорбтива обычно используют ту или иную формулу, понимая под  $\beta$  общий коэффициент массопередачи.

Уравнение теплового баланса для элементарного объема зернистого слоя по форме записывается аналогично уравнению материального баланса. Для адиабатического процесса тепловой баланс может быть записан следующим образом:

$$h_a \frac{\partial T_a}{\partial \tau} = -h_{\Gamma} \frac{\partial T_{\Gamma}}{\partial \tau} - wh_{\Gamma} \frac{\partial T_{\Gamma}}{\partial x} - \frac{\partial(Qa)}{\partial \tau},$$

где  $h_a, h_{\Gamma}$  – теплоемкость адсорбента и газа;  $T_a, T_{\Gamma}$  – температура адсорбента и газа;  $Q$  – тепловой эффект сорбции.

В практических расчетах обычно пренебрегают продольным эффектом диффузии и принимают, что тепловой эффект сорбции  $Q$  является величиной постоянной.

Теплообмен между твердым телом и потоком лимитируется или внешней теплоотдачей, или теплопроводностью гранул (внутренняя задача), или определяется совместно этими двумя механизмами.

Анализ показывает, что в реальных условиях работы зернистых слоев адсорбентов и катализаторов теплообмен, в отличие от массообмена, практически полностью определяется теплоотдачей от ядра потока к внешней поверхности частиц твердого тела. В этом случае уравнение передачи тепла можно записать следующим образом:

$$h_a \frac{\partial T}{\partial \tau} = K_{\Gamma}(T_{\Gamma} - T_a) + Q \frac{\partial a}{\partial \tau}.$$

Дифференциальные уравнения массо- и теплообмена в зернистом слое решаются с учетом начальных и граничных условий. Например, можно принять следующие распределения концентрации и температур в начальный момент времени:  $\tau = 0$ ;  $0 \leq x \leq L$ ;  $c = 0$ ;  $a = 0$ ;  $T_a = T_{\Gamma} = T_n$ . Принятое начальное условие означает, что до начала процесса ни в газовой, ни в твердой фазах на всем протяжении длины(высоты) слоя адсорбат не содержался.

Граничные условия при  $x = 0$  можно задать следующим образом:

$$\tau > 0; x = 0; c = c_0 = \text{const}; a = a(\tau); T_{\Gamma} = T_{\Gamma 0} = \text{const},$$

т.е. в течение всего процесса в слой поступает газовый поток при постоянной концентрации адсорбтива и температуре.

Математическое описание адсорбционных процессов даже с учетом сделанных допущений представляет большие трудности. Рассмотрим постепенно усложняющийся ряд моделей динамики сорбции.

Примем, что между концентрациями адсорбтива и адсорбата в каждый момент времени и в каждой точке слоя соблюдается равновесное соотношение в силу, например, бесконечно большого значения коэффициента массопередачи. Пусть также

эффект продольной диффузии пренебрежимо мал  $D^* = 0$ , тепловыделения в слое бесконечно малы, а температуры потока и адсорбента одинаковы. В соответствии со сделанными допущениями процесс динамики равновесной изотермической адсорбции описывается двумя уравнениями:

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + \frac{\partial (wc)}{\partial x} = 0,$$

$$a = f(c),$$

где  $a = f(c)$  – уравнение изотермы адсорбции.

Система имеет два решения: одно при наличии конечного градиента концентраций в слое (случай 1), другое при так называемом «обрывном» фронте (случай 2). «Обрывным» в динамике адсорбции называют фронт такой гипотетической формы, при которой концентрация адсорбтива скачкообразно меняется от  $c = c_0$  до  $c = 0$ .

$$\text{Случай 1: } U_i = \frac{w}{1 + f'(c_i)};$$

$$\text{Случай 2: } U = \frac{wc_0}{c_0 + a_0},$$

где  $U_i$  – скорость движения концентрационной точки  $i$  по слою адсорбента;  $f'(c_i) = (da/dc)_i$  – производная изотермы адсорбции в точке  $i$ ;  $U$  – скорость движения всех точек «обрывного» фронта;  $c_0$  – концентрация адсорбтива в потоке;  $a_0$  – равновесная ей величина адсорбции.

Уравнение для случая 1 носит название закона Вика, для случая 2 – закона Вильсона.

Модель динамики неравновесной изотермической адсорбции при пренебрежимо малом влиянии продольнодиффузионных эффектов может быть описана следующими уравнениями [15]:

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + w \frac{\partial a}{\partial x} = 0,$$

$$a^* = f(c),$$

$$\partial a / \partial \tau = \varphi(a, c).$$

В этой системе уравнений последняя строка представляет собой уравнение кинетики в общей форме. Данная система уравнений с начальными и граничными условиями общего решения не имеет. Поэтому следует рассматривать частные случаи (выпуклая, вогнутая, линейная изотермы и т.п.).

Рассмотренные выше математические модели динамики сорбции основаны на предположении о пренебрежимо малой величине теплового эффекта адсорбции. Это допущение справедливо при небольших концентрациях адсорбтива в потоке. Однако при повышенных концентрациях тепловой эффект может оказывать существенное влияние на протекание адсорбционного процесса и нуждается в учете. Учет тепловыделений и теплообмена необходим также в тех случаях, когда температуры потока и зернистого материала различны. Опишем динамику адиабатической адсорбции, предполагая в первом приближении, что массообмен и теплообмен между фазами системы происходит бесконечно быстро. Такой анализ впервые был сделан Годесом и Лезиным. Основные уравнения этой модели могут быть записаны следующим образом:

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + w \frac{\partial c}{\partial x} = 0,$$

$$H \frac{\partial T}{\partial \tau} + wh_r \frac{\partial T}{\partial x} + Q \frac{\partial a}{\partial \tau} = 0,$$

$$a^* = f(c),$$

где  $H$  – суммарная теплоемкость адсорбента и адсорбата.

Начальные и граничные условия пусть будут такими же, как и в предыдущих моделях. При изотермической адсорбции, когда тепловым эффектом адсорбции можно пренебречь, скорость движения адсорбционного фронта составит

$$U_{\text{изф}} = w \frac{c_0}{a_0 + c_0} \approx w \frac{c_0}{a_0}.$$

Из теплового баланса следует, что скорость движения тепловой волны, измеренная по ее «центру тяжести», будет равна  $W = wh_r / H$ . При протекании неизотермического процесса сорбционная и тепловая волны будут двигаться взаимно согласованно.

Дальнейшее развитие теории динамики неизотермической сорбции состояло в учете кинетики тепловых и массообменных процессов. Лезин рассмотрел модель, основанную на предположении о бесконечно быстром протекании теплообмена между газом и шихтой при массообмене, описываемом уравнением кинетики нестационарного конвективного массообмена. Основные уравнения этой модели записываются в виде:

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + w \frac{\partial c}{\partial x} = 0,$$

$$-H \frac{\partial T}{\partial \tau} - wh_r \frac{\partial T}{\partial x} + Q \frac{\partial a}{\partial \tau} = \frac{2\alpha_{\text{II}}}{R} T,$$

$$\frac{\partial a}{\partial \tau} = \beta_r [c - \varphi(a, T)],$$

где  $\alpha_{\text{II}}$  – коэффициент теплоотдачи от стенок адсорбера в окружающую среду;  $R$  – радиус адсорбера.

Адсорбционное равновесие  $\varphi(a, T)$  задавалось уравнениями Ленгмюра и Дубинина-Радушкевича. Рассматривались два процесса – адсорбция и десорбция. Начальные условия выбирались равными:

при адсорбции

$$\tau = 0; \quad c(x, 0) = a(x, 0) = T(x, 0) = 0;$$

при десорбции

$$\tau = 0; \quad c(x, 0) = c_0; \quad a(x, 0) = a_0; \quad T(x, 0) = 0.$$

Граничные условия при адсорбции:

$$x = 0; \quad \tau > 0; \quad c(0, \tau) = c_0; \quad a(0, \tau) = a(\tau); \quad T(0, \tau) = 0.$$

Граничные условия при десорбции:

$$x = 0; \quad \tau > 0; \quad c(0, \tau) = 0; \quad a(0, \tau) = a(\tau); \quad T(0, \tau) = 0.$$

Математическая модель адиабатической адсорбции включает уравнения:

$$\frac{\partial a}{\partial \tau} + \frac{\partial c}{\partial \tau} + w \frac{\partial c}{\partial x} = 0,$$

$$\frac{\partial a}{\partial \tau} = \beta_r [c - c^*],$$

$$\frac{\partial I_a}{\partial \tau} + \frac{\partial I}{\partial \tau} + w \frac{\partial I}{\partial x} = 0,$$

$$\frac{\partial I_a}{\partial \tau} = K(T_r - T_a) + \beta_r (c - c^*) Q,$$

$$c^* = f(a, T_a),$$

где  $I_a$  – энтальпия адсорбента;  $I$  – энтальпия потока;  $K$  – коэффициент теплоотдачи, отнесенный к единице объема зернистого слоя.

Начальные и граничные условия принципиально не отличаются от краевых условий, приведенных выше. Значения коэффициентов тепло- и массообмена определяли путем сопоставления результатов одного из опытов с результатами расчетов.

### 3.4. ПРОЦЕССЫ ОЧИСТКИ И РЕГЕНЕРАЦИИ ВОЗДУХА

Регенерация и очистка воздуха помещений являются сложными химико-технологическими процессами, проходящими с применением различного по сложности исполнения и способам применения оборудования. Аппараты и системы регенерации и очистки воздуха занимают ключевую позицию в числе прочих составляющих систем жизнеобеспечения различных объектов.

Здесь и далее под термином «*регенерация воздуха*» будем понимать процесс создания и поддержания в воздухе помещения определенных концентраций кислорода и диоксида углерода с применением различных аппаратов (регенераторов), комплексов и систем регенерации воздуха.

Очистка воздуха, в общем, также может быть отнесена к процессу его регенерации (восстановления). Однако, учитывая то, что термин «регенерация воздуха» мы традиционно будем применять только для таких компонентов воздуха, как кислород и диоксид углерода, возникает необходимость говорить об очистке воздуха (удалении из него вредных примесей), как об отдельном процессе.

Таким образом, под *очисткой воздуха* будем понимать удаление из него вредных примесей, начиная с текущих концентраций и до уровня ниже ПДК (либо ниже следов вредных примесей).

Под *вредной примесью* понимается компонент газового состава воздуха, оказывающий негативное воздействие на организм человека.

Перечень вредных примесей, их допустимые разовые, мгновенные и другие концентрации определены ГОСТ 12.1.007 для промышленных помещений и рядом отраслевых стандартов и нормативов для помещений специальных объектов. Надо сказать, что номенклатура и количественный состав вредных примесей, указанных в ГОСТ 12.1.007 и отдельными нормативными документами может отличаться существенно, что определяется назначением самого объекта и режимами его функционирования.

Процессы очистки воздуха начали изучать гораздо раньше процессов регенерации, что было обусловлено необходимостью практического применения регенерационного и поглотительного (очистительного) оборудования.

Первые масштабные исследования процессов и способов очистки воздуха от вредных примесей начались, по-видимому, в годы первой мировой войны, когда возникла насущная необходимость разработки и применения средств индивидуальной защиты войск при применении химического оружия. В эти годы были проведены глубокие теоретические и уникальные экспериментальные исследования, позволившие сформировать подходы к разработке и производству средств индивидуальной защиты органов дыхания, которые, позднее, были с успехом перенесены и на средства коллективной защиты.

Системы регенерации воздуха известны более 100 лет, практически с момента появления подводного флота. Кроме объектов ВМФ – подводные лодки, батискафы, системы регенерации атмосферы были востребованы, прежде всего, в космосе, нашли широкое применение в узлах связи, пунктах управления, убежищах ГО, в местах укрытия промышленных предприятий, АЭС, в пунктах управления газовой и нефтяной промышленности при возникновении аварийных ситуаций.

Продолжительное непрерывное пребывание персонала объекта в замкнутом объеме носит комплексный характер и выдвигает целый ряд задач и вопросов, относящихся к самым различным областям науки и техники. Это обусловлено тем, что самочувствие человека зависит от многих факторов, их сочетания, характера и времени воздействия, а также от индивидуальных физиологических особенностей организма.

Если на заре подводного кораблестроения использовалась система регенерации на основе сжатого воздуха и щелочного поглотителя диоксида углерода, то в настоящее время существуют разнообразные средства поддержания жизнедеятельности

людей, основанные на различных принципах восстановления атмосферы, различной энергетике, различной длительности работы для обеспечения автономного пребывания.

Системы регенерации воздуха можно разделить:

1. В зависимости от количества потребляемой электроэнергии:

- не требующие электроэнергии, – т.е. работающие в условиях гравитационной конвекции;
- требующие мало электроэнергии, потребление энергии на одного человека не более 10 Вт. К таким средствам обычно относят средства регенерации на запасах химических веществ;
- энергоемкие, потребление энергии на одного человека более 100 Вт. К ним относятся электрохимические системы регенерации, системы регенерируемого типа на твердых химических поглотителях, сорбентах и жидких регенерируемых поглотителях.

2. По методам регенерации атмосферы:

- раздельный метод регенерации. Получение кислорода и удаление диоксида углерода происходит независимо друг от друга;
- совмещенный способ регенерации воздуха, когда и выделение кислорода, и поглощение диоксида углерода производится в одном месте и в одно время. Примером реализации такого метода являются системы регенерации на основе супероксидов щелочных металлов;
- полураздельный (или полусовмещенный), когда в одном аппарате производится поглощение диоксида углерода и выделение кислорода одновременно, а в другом аппарате производится только поглощение диоксида углерода. Все современные системы регенерации на химически связанном кислороде работают на таком принципе регенерации воздуха.

3. По способам использования химических реагентов и поглотителей:

- регенерируемые системы, например, широко известные поглотители ТРП;
- нерегенерируемые системы, например, системы регенерации воздуха в убежищах ГО, содержащие вещества на основе супероксидов калия или натрия.

### 3.4.1. Очистка воздуха изолированных помещений

Рассмотрим процесс очистки воздуха от вредных примесей в условиях, когда вентилирование помещения производится без забора внешнего воздуха (в режиме полной изоляции). Существенных различий в работе систем фильтровентиляции, работающих по замкнутому (полная изоляция) или разомкнутому (с забором внешнего воздуха) режимам нет. Разница заключается в производительности систем фильтровентиляции (расходе очищаемого воздуха) и в составе вредных примесей, от которых осуществляется очистка. Мы будем рассматривать случай, когда необходимо удалять из помещения вредные примеси, при этом использование наружного воздуха для вентилирования помещения по каким-либо причинам невозможно. Эта ситуация соответствует так же случаям использования индивидуальных изолирующих средств защиты дыхания.

По степени воздействия на организм вредные вещества подразделяют на 4 класса опасности:

- 1 – вещества чрезвычайно опасные;
- 2 – вещества высокоопасные;
- 3 – вещества умеренно опасные;
- 4 – вещества малоопасные.

Для указанных классов веществ существуют нормативы по содержанию в помещениях. В общем случае для производственных помещений эти нормативы представлены в ГОСТ 12.1.007.

Вредные примеси в воздухе обладают эффектом сочетаемого действия. В зависимости от сочетания вредных примесей в воздухе их действие может быть суммирующим, потенцирующим и независимым. В тех случаях, когда в воздухе присутствуют одновременно несколько веществ, токсическое действие которых суммируется, руководствуются формулой, согласно которой сумма отношений концентрации каждого вещества к значению его ПДК не должна превышать единицы:

$$\sum_j \frac{C_j}{\text{ПДК}_j} \leq 1.$$

В случаях, когда токсическое действие одновременно действующих токсических веществ усиливается, концентрации таких веществ необходимо снижать в соответствии с формулой:  $\sum_j \frac{C_j}{\text{ПДК}_j} \leq 0,7$ .

Кроме показателя ПДК часто используют показатель максимальной допустимой концентрации МДК. МДК – концентрация вредной примеси, при однократном воздействии которой в течение установленного времени гарантируется сохранение жизни и здоровья.

Вопрос очистки воздуха герметичных и условно-герметичных помещений от вредных примесей начал приобретать все большую актуальность в связи с:

- увеличением времени автономной работы объектов;
- необходимостью обеспечения безопасности персонала объектов после аварийных ситуаций в объектах с ненормируемым выделением вредных примесей в воздух помещений и по ряду других причин.

Для очистки воздуха помещений от вредных примесей используют системы фильтровентиляции, которые, в зависимости от назначения, могут очищать воздух отдельных помещений (локальная очистка) или очищать воздух всего помещения (системы рециркуляционной очистки).

В зависимости от назначения систем очистки различается и их состав (используемые средства очистки). В качестве фильтров систем фильтровентиляции применяют противоаэрозольные, каталитические, шихтовые и совмещенные фильтры.

В настоящее время практически все системы очистки представляют собой набор фильтров-поглотителей, работающих за счет сорбции газов.

Для помещений основную массу вредных примесей в штатном режиме образуют вредные примеси бытового происхождения. К ним относится большое количество вредных и дурнопахнущих веществ, выделяемых людьми, их одеждой, обувью, санитарно-техническими устройствами, продуктами питания в процессе их хранения и приготовления пищи, пищевыми отходами.

Источником неприятных запахов физиологического происхождения являются выделения кожных покровов тела, слизистых оболочек, кишечные газы, а также продукты разложения этих выделений. Например, в запахе пота выделяются ароматические летучие жирные кислоты – муравьиная, пропионовая и другие соединения. Плохое санитарное состояние кожных покровов тела, белья и одежды увеличивает количество вредных примесей, поступающих в воздух помещения.

Содержание вредных примесей, выделяемых человеком, приведено в табл. 3.1.

### 3.1. Вредные примеси, выделяемые человеком в течение 1 часа

Выделяемое вещество	Количество, мг	Распределение, %	
		Воздух	Конденсат
Аммиак	12,5	21,5	78,5
Альдегиды	0,02		100
Кетоны	9,7	40	60
Сероводород и меркаптаны	0,21		100
Жирные кислоты	3,72	25,7	74,3
Оксид углерода:			
– у курящих;	11,6	100	
– у некурящих	17,4	100	
Углеводороды в пересчете на углерод	21,2	100	

Из вредных примесей бытового происхождения следует отметить сероводород, углеводороды, аммиак и амины.

### 3.4.2. Фильтры-поглотители

Любой фильтр-поглотитель содержит твердое вещество (сорбент), поглощающее вредные примеси в результате процесса сорбции.

Качество сорбентов в отношении сорбции какого-либо газа характеризуется величиной сорбции из воздуха. Максимальное количество газа, сорбируемое единицей массы сорбента при равновесии с газом, содержащемся в воздухе при температуре  $T$  и имеющем конденсацию  $C$ , называется равновесной статической активностью сорбента. При неизменном давлении равновесная статическая активность является функцией концентрации газа и температуры.

Динамическая активность слоя сорбента характеризует так называемый динамический метод работы слоя, в котором воздух, содержащий примеси газов, пропускается через слой сорбента, задерживающий хорошо сорбируемые газы. Процесс поглощения ведется до момента появления следов газа за слоем сорбента (момент проскока), после чего слой сорбента считается отработавшим. Понятие динамической активности впервые было введено Н.А. Шиловым на основании исследований в фронтальной лаборатории в 1916 г. Динамическая активность представляет собой сложную функцию, в основном зависящую от статической активности сорбента и скорости сорбции газа.

Кроме понятий динамической и статической активности сорбента с процессом сорбции связана величина времени защитного действия (время от начала процесса сорбции до проскока). Время защитного действия определяется по уравнению Шилова:

$$\tau = kL - \theta_0,$$

где  $\theta_0$  – потеря времени защитного действия, обусловленная конечной скоростью сорбции;  $k$  – коэффициент защитного действия, который представляет собой обратную величину постоянной скорости продвижения фронта газа;  $L$  – длина слоя сорбента.

Коэффициент защитного действия может быть вычислен по формуле:

$$k = \frac{a}{vC_0},$$

где  $v = \frac{V}{S}$  – линейная скорость газа через сорбент;  $a$  – статическая активность сорбента;  $C_0$  – концентрация примеси газа в воздухе.

Выражение для определения времени защитного действия может быть представлено в виде:

$$\tau = \frac{\delta_0}{D} \frac{a}{F\chi C_0} \left( \frac{\chi}{v} \right)^n \left( \ln \frac{C_0}{C_1 - C'} - 1 \right),$$

где  $\delta_0$  – константа для рассматриваемого сорбента, связанная с толщиной адгезированного слоя;  $D$  – коэффициент диффузии газа;  $F$  – поверхность сорбента;  $C'$  – концентрация газа на поверхности сорбента;  $C_1$  – концентрация газа на границе работающего слоя;  $\chi$  – доля площади в сечении слоя, соответствующей промежуткам между зернами сорбента.

Для целей практического применения приведенные выше формулы заменяют на следующие соотношения:

$$\theta = kL - \tau, L > L_0,$$

$$k = \frac{\beta C_0^{\mu-1}}{\nu},$$

$$\tau = b_0 k \frac{d}{\nu^{n-1}} (\lg C_0 - b),$$

$$L_0 = b_0 \frac{d}{\nu^{n-1}} (\lg C_0 - b + 0,434).$$

Приведенные в этих уравнениях константы  $\beta, \mu, n, b_0, b$  характеризуют систему газ-сорбент в интервале изменения параметров  $C_0, \nu, d$ , применимом для вычисления констант уравнений.

Процесс очистки воздуха в замкнутом объеме можно описать следующей системой уравнений:

$$\frac{\partial c}{\partial t} + \frac{\partial a}{\partial t} = -\omega \frac{\partial c}{\partial x},$$

$$\frac{\partial a}{\partial t} = \beta(c - c^*),$$

$$a = \frac{W_0}{\nu^*} \exp \left[ -B \frac{T^2}{\beta^{*2}} 0,43^2 \left( \ln \frac{P_s}{P^*} \right)^2 \right],$$

$$c = \frac{p^* M \cdot 10^3}{18,2}$$

при начальных условиях  $c(x,0) = c_0 \exp(-\beta x / \omega)$ ,  $a(x,0) = 0$  и граничных условиях, получаемых из решения уравнения идеального смешения для объема

$$\frac{dc_V}{dt} = \frac{S\omega}{W} (C|_{x=L} - C_V);$$

$$C_V(0) = C_0, C(0,t) = C_V.$$

где  $\beta$  – коэффициент массообмена,  $1/c$ ;  $\omega$  – линейная скорость движения воздуха в слое шихты, м/с;  $W_0, B, b$  – константы, зависящие от применяемого сорбента;  $P_s$  – давление насыщения, мм рт. ст.;  $\beta^*$  – коэффициент аффинности;  $\nu^*$  – мольный объем, см<sup>3</sup>/ммоль.

В приведенных уравнениях не учитывается взаимное влияние примесей. Следует отметить, что это условие выполняется практически всегда, так как объем шихты сорбента рассчитывается на поглощение всей гаммы вредных примесей.

Кроме того, система уравнений предполагает очистку воздуха от вредных примесей при отсутствии их источников, т.е. после накопления в течение какого-либо времени вредных примесей в объеме помещения и очистке этого при отсутствующем (или несущественном за время очистки) поступлении вредных примесей в объем помещения. При невыполнении этого условия уравнение кинетики должно быть изменено на уравнение

$$\frac{dc_V}{dt} = \frac{S\omega}{W} (C|_{x=L} - C_V) + q,$$

где  $q$  – скорость поступления вредной примеси в объем помещения.

В этой системе в качестве уравнения кинетики использовано обычное уравнение массообмена. Для вычисления равновесной концентрации примеси используется уравнение Дубинина–Радушкевича, позволяющее по молекулярной массе, мольному объему, плотности, давлению насыщения, критической температуре и некоторым константам сорбентов рассчитать величину адсорбции и из величины  $a$  найти  $C^*$ .

Для газов уравнение Дубинина–Радушкевича имеет вид:

$$a = \frac{W_0}{b} \left[ \exp \left( -B \frac{T^2}{\beta^2} 0,43^2 \left( \ln \left( \frac{T}{T_{kp}} \right) \frac{P_s}{P} \right)^2 \right) \right].$$

Представленная система уравнений позволяет рассчитывать динамику процесса очистки воздуха в помещении. Еще одним полезным уравнением, применяемом при проектировании систем очистки, является уравнение вентиляции

$$Q = S\omega = \frac{V}{t} \ln \frac{C_{\text{ПДК}} - C_{\text{Ф}}}{C_0 - C_{\text{Ф}}},$$

которое позволяет определить необходимый расход воздуха через фильтр (систему фильтров)  $Q$  для очистки помещения объемом  $V$  от вредной примеси за время  $t$  от начальной концентрации вредной примеси  $C_0$  до уровня ПДК  $C_{\text{ПДК}}$ . При этом считается, что на выходе фильтра очищаемый воздух имеет концентрацию вредной примеси  $C_{\text{Ф}}$  (проскок).

### 3.4.3. Регенерация воздуха в герметично замкнутом объеме

Замкнутая система (объект), процессы очистки и регенерации воздуха в которой мы будем рассматривать, состоит, в основном, из:

- герметично (условно-герметично) замкнутого объема, системы помещений;
- систем жизнеобеспечения;
- технологического оборудования;

– персонала.

В системы жизнеобеспечения входят:

- системы регенерации и очистки воздуха;
- система терморегулирования;
- система удаления жидких и твердых отходов.

За формирование газового состава воздуха помещения отвечают системы регенерации и очистки воздуха.

Рассмотрим источники и стоки компонент воздушной среды, участвующие в формировании газового состава. Так как объект является изолированным от внешней среды, воздух внешней среды в формировании газового состава воздуха в помещении не участвует.

Попадание воздуха внешней среды в объект исключается организацией подпора (созданием в объекте повышенного относительно внешней среды давления). Ситуации разгерметизации объекта являются экстремальными и в настоящей главе не рассматриваются.

Внутри объекта в формирование газового состава воздуха вносят свой вклад:

- технологическое оборудование (источники и стоки кислорода, диоксида углерода, вредных примесей, тепла и влаги);
- материалы помещений (источники и стоки кислорода, диоксида углерода, вредных примесей);
- персонал (стоки кислорода, источники диоксида углерода, вредных примесей, тепла и влаги);
- системы регенерации и очистки воздуха (источники кислорода, тепла, стоки диоксида углерода и вредных примесей).

Поглощение и выделение кислорода и диоксида углерода технологическим оборудованием, материалами и персоналом зависит от большого числа факторов.

Для каждого объекта поглотительные и выделительные характеристики либо задаются в виде усредненных значений, либо определяются в ходе отдельных исследований.

Так, примеры зависимости выделения диоксида углерода и поглощения кислорода человеком от характера (тяжести) выполняемой работы представлены на рис. 3.6 и 3.7.

$G^{CO_2}, 0,06, [м^3/ч]$

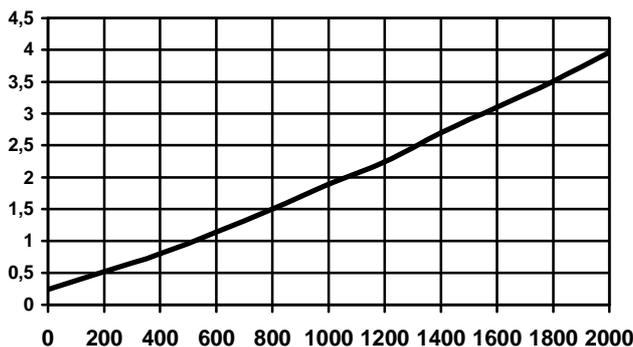


Рис. 3.6. Выделение человеком диоксида углерода в зависимости от нагрузки:

$G^{CO_2}$  – количество выделяемого человеком диоксида углерода в единицу времени;  $d$ , кгм/ч – нагрузка; 1 Вт – 360 кгм/ч

$G^{O_2}, 0,06, м^3/ч$

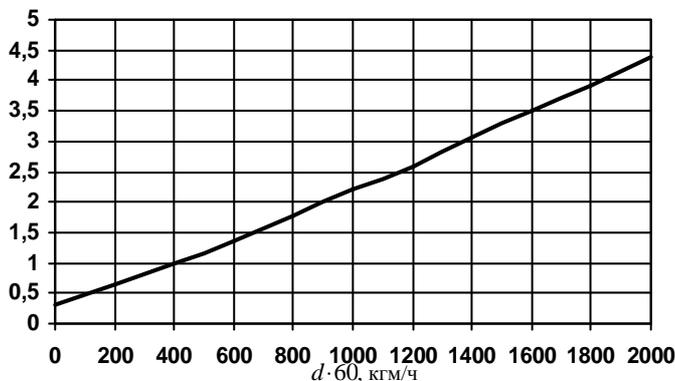


Рис. 3.7. Потребление кислорода человеком в зависимости от нагрузки:

$G^{O_2}$  – количество потребляемого кислорода в единицу времени

Расчет потребления кислорода в зависимости от нагрузки для одного человека можно проводить по формуле, которая является аппроксимацией графика для нагрузок более 200 [кгм/мин]:

$$G^{O_2} = \begin{cases} m < 85, S(0,02154d + 0,0306)/1000, \\ m \geq 85, S(0,002154d + 0,0306 + (m - 85)0,04)/1000 \end{cases}, \text{ м}^3/\text{мин},$$

где  $G^{O_2}$  – количество потребляемого кислорода в единицу времени;  $d$ , кгм/мин, нагрузка 1 Вт = 6 кгм/мин;  $m$  – масса человека, кг;  $S$  – коэффициент, характеризующий пол человека, для мужчин он равен 1, для женщин 0,7...0,8. Можно также учитывать возраст, накладывая соответствующее ограничение на величину максимального потребления кислорода.

Количество выводимого из организма диоксида углерода при дыхании можно рассчитывать по формуле:

$$G^{CO_2} = k(d)G^{O_2}, \text{ м}^3/\text{мин},$$

где  $k(d)$  – параметр, характеризующий соотношение выделенного диоксида углерода к поглощенному кислороду за единицу времени в зависимости от нагрузки (дыхательный коэффициент) (рис. 3.8).

Чаще всего при проведении предварительных расчетов для большого количества людей (>10) и длительного (>10 суток) функционирования систем регенерации и очистки воздуха используются усредненные дыхательные характеристики человека. При этом принимается, что одним человеком выделяется (20...30) л/ч  $CO_2$ .

$k(d)$

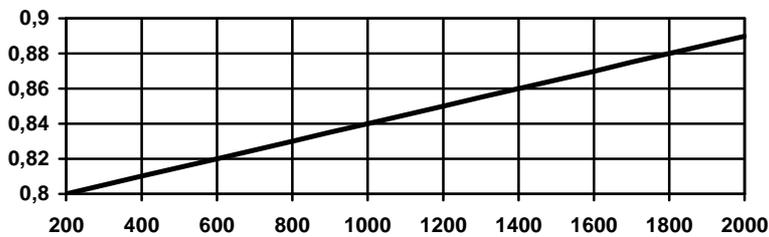


Рис. 3.8. Соотношение диоксида углерода к поглощенному кислороду за единицу времени в зависимости от нагрузки

В любом случае для моделирования процессов регенерации и очистки воздуха будем принимать за  $G^{O_2}$  – суммарное поглощение кислорода, л/ч, а за  $G^{CO_2}$  – суммарное выделение диоксида углерода, л/ч.

Основной составляющей, влияющей на формирование газового состава воздуха в замкнутых помещениях, является система регенерации и очистки воздуха.

В зависимости от условий эксплуатации объекта системы регенерации и очистки воздуха могут быть:

- совмещенными (элементы систем поглощают диоксид углерода с одновременным выделением кислорода);
- полусовмещенными (одни элементы систем поглощают диоксид углерода с одновременным выделением кислорода, а другие – только поглощают диоксид углерода);
- раздельными (поглощение диоксида углерода и выделение кислорода осуществляется различными элементами системы).

Далее будут рассмотрены совмещенные системы регенерации воздуха, элементы которых содержат регенеративные вещества с химически связанным кислородом.

Наиболее часто в качестве регенеративных продуктов используются надперекиси калия и натрия, некоторые свойства которых представлены в табл. 3.2.

Надперекись натрия представляет собой вещество желтого цвета с удельной поверхностью 0,1...0,3 м<sup>2</sup>/г, насыпной плотностью 08...1,0 г/см<sup>3</sup> и истинной плотностью 2,21...2,25 г/см<sup>3</sup>. Чистую надперекись натрия в промышленном производстве получают с 1949 г путем окисления перекиси натрия кислородом при давлении 280 кг/см<sup>2</sup> и температуре 470 °С.

### 3.2. Свойства надперекиси калия и натрия

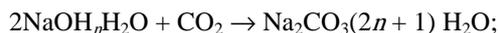
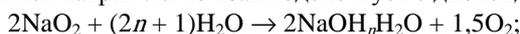
Название	Формула	Молекулярный вес	Содержание кислорода, % вес	Стехиометрическая емкость, л/кг		
				по O <sub>2</sub>	по CO <sub>2</sub>	
					Me <sub>2</sub> CO <sub>3</sub>	MeHCO <sub>3</sub>
Надперекись натрия	NaO <sub>2</sub>	55	43,6	308	203	406
Надперекись калия	KO <sub>2</sub>	71,1	33,6	236	157	314

При нагревании надперекись натрия начинает разлагаться при температуре 100...120 °С с образованием непрерывного ряда твердых растворов до перекиси при 250 °С, до окиси при 540 °С.

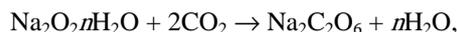
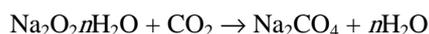
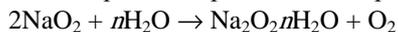
Надперекись натрия может быть переведена в расплав при температуре плюс 370 °С и давлении кислорода 150 кг/см<sup>2</sup>.

Надперекись натрия при комнатной температуре энергично взаимодействует с парами воды с выделением всего активного кислорода. При пониженных температурах (~0 °С) взаимодействие надперекиси натрия с парами воды сопровождается выделением лишь надперекисного кислорода и образованием кристаллогидратов перекиси натрия.

В присутствии паров воды надперекись натрия легко взаимодействует с диоксидом углерода по схеме:



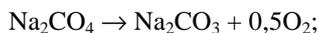
Однако при температурах ниже плюс 10 °С взаимодействие паров воды и диоксида углерода с надперекисью натрия протекает только с выделением надперекисного кислорода и образованием перкарбонатов натрия:



где  $n$  может иметь любые значения от 2 до 8.

Образованию перкарбонатов и гидратов перекиси натрия обязан индукционный период в работе регенеративных продуктов на основе надперекиси натрия, характеризующийся малой скоростью выделения кислорода.

Перкарбонаты и гидраты перекисей натрия начинают разлагаться с выделением кислорода только при температурах выше плюс 25 °С по схеме:



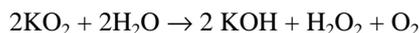
Надперекись натрия не реагирует в отсутствие паров воды с карбонатом, бикарбонатом и гидроокисью натрия.

С органическими веществами (спиртом, олифой, растворителями и синтетическими лаками) надперекись натрия и регенеративные продукты на ее основе реагируют со взрывом. Нагревание надперекиси натрия и регенеративных продуктов на ее основе до температур ~ плюс 100 °С приводит к выделению части кислорода и незначительному увеличению их механической прочности.

Надперекись калия представляет собой мелкокристаллическое вещество желтого цвета с удельной поверхностью 0,8...1,0 м<sup>2</sup>/г, насыпной плотностью 0,3...0,4 г/см<sup>3</sup> и истинной плотностью 2,16 г/см<sup>3</sup>.

При нагревании свыше плюс 145 °С надперекись калия начинает разлагаться до К<sub>2</sub>О с выделением кислорода и образованием твердых растворов начинаясь с + 425 °С и заканчиваясь при + 660 °С.

Надперекись калия энергично взаимодействует с парами воды и при температурах выше плюс 25 °С выделяет весь активный кислород. При температурах близких к 0 °С взаимодействие надперекиси калия с водой может протекать по реакции:



с частичным образованием перекиси водорода. В присутствии паров воды надперекись калия легко реагирует с диоксидом углерода, образуя карбонаты и бикарбонаты и выделяя активный кислород.

При этом в зависимости от температуры пары воды расходуются не только на прямую реакцию выделения кислорода из надперекиси калия, но также на гидратацию продуктов реакции.

По результатам экспериментов 1 моль поглощенной продуктом воды выделяет при:

+ 50 °С – 1,4 моля кислорода;

+ 25 °С – 0,8...1,0 моля кислорода;

+ 4 °С – 0,2 моля кислорода.

При пониженных температурах реакция надперекиси калия с парами воды с диоксидом углерода идет с образованием перкарбонатов, которые при температуре выше + 50 °С разлагаются до бикарбонатов с выделением активного кислорода.

С органическими веществами надперекись калия взаимодействует со взрывом.

Нагревание регенеративных продуктов на основе надперекиси калия при температуре + 100 °С в течении 3 – 5 часов сопровождается уменьшением их объема на 2 % и увеличением их прочности на разлом и истирание на 10 %.

Свойства регенеративных продуктов определяются не только физикохимическими свойствами составляющих их основу надперекисных соединений, но также и пористой структурой этих продуктов. Для компенсации объемных изменений и стабилизации регенеративных свойств надперекисей в состав продуктов на их основе вводят компоненты, обеспечивающие необходимое развитие пористой структуры и сохраняющие форму продуктов в процессе использования.

Таким образом, работа элемента совмещенной системы регенерации и очистки воздуха состоит в поглощении из входящего воздуха паров воды и диоксида углерода с выделением кислорода (рис. 3.9).

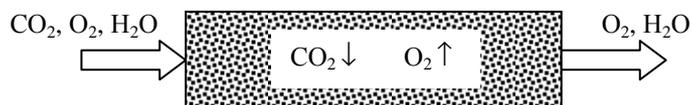


Рис. 3.9. Элемент системы регенерации и очистки воздуха

Воздух на вход регенеративного элемента может поступать за счет принудительной вентиляции (механический или электрический вентилятор) или за счет действия конвективных сил (конвективный режим регенерации воздуха).

### 3.5. МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ОЧИСТКИ И РЕГЕНЕРАЦИИ ВОЗДУХА В ЗАМКНУТОМ ОБЪЕМЕ

#### 3.5.1. Математические модели динамики сорбции в поглотительном патроне

Для описания процесса сорбции в системах очистки воздуха воспользуемся системой следующих уравнений:

1. Система уравнений, описывающая работу поглощающего продукта в патроне;
2. Уравнение, описывающее работу группы поглотительных патронов в замкнутом объеме, при наличии источников диоксида углерода и потребителей кислорода.

Математическая модель включает в себя следующие уравнения:

Уравнение материального баланса

$$\frac{\partial c}{\partial t} + \frac{\partial a}{\partial t} + \nabla(\bar{v}c) = \nabla^2(Dc), \quad (3.1)$$

где  $\frac{\partial c}{\partial t}$  – скорость изменения концентрации в объеме сорбента в газовой фазе;  $\frac{\partial a}{\partial t}$  – скорость поглощения вещества в адсорбенте;  $\bar{v}$  – вектор скорости потока;  $\nabla(\bar{v}c)$  – градиент концентраций;  $\nabla^2(Dc)$  – диффузионный член.

Уравнение неразрывности

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho\bar{v}) = 0, \quad (3.2)$$

где  $\rho$  – плотность вещества потока.

Уравнение движения жидкости (Навье-Стокса)

$$\frac{\partial \bar{v}}{\partial t} = F - \frac{1}{\rho} \text{grad}(p) + \mu \nabla^2 \bar{v}, \quad (3.3)$$

где  $p$  – давление;  $\mu$  – вязкость;  $F$  – напряженность поля массовых сил.

Уравнение состояния

$$\rho = \Psi(c, P, T). \quad (3.4)$$

Уравнение кинетики

$$\frac{\partial a}{\partial t} = \beta c f(a). \quad (3.5)$$

Система уравнений (3.1) – (3.5) позволяет описывать процессы, проходящие внутри патрона. Но ее решение в общем виде проблематично.

Введем следующие допущения:

1. Уравнение материального баланса (3.1) будем рассматривать как одномерное.

2. В уравнении (3.2) предположим, что движение воздушного потока через поглотительный патрон происходит без изменения плотности ( $\frac{\partial \rho}{\partial t} = 0$ ), тогда  $\text{div}(\rho\bar{v}) = 0$ .

3. Изменение температуры в процессах, проходящих в поглотительных патронах, незначительно, т.е. вязкость, плотность и давление можно считать постоянными.

4. В уравнении Навье-Стокса действует только поле гравитационных сил ( $F = g$ ), влияние которого пренебрежимо мало.

5. Скорость с течением времени не меняется ( $v = \text{const}$ ).

С учетом принятых допущений получаем следующую систему:

$$\begin{cases} \frac{\partial c}{\partial t} + \frac{\partial a}{\partial t} + \frac{\partial c}{\partial x} = D \frac{\partial^2 c}{\partial x^2}; \\ \frac{\partial a}{\partial t} = \beta c f(a) \end{cases} \quad (3.6)$$

при начальных условиях

$$\begin{cases} a(x,0) = a_0; \\ c(x,0) = f(c_0, t). \end{cases}$$

В случае хемосорбции уравнение изотермы исключается (в предельном случае прямоугольная изотерма – при любой концентрации реагента в случае химической реакции при статических условиях достигается полное насыщение).

В общей форме уравнение кинетики можно представить в форме:

$$\frac{\partial a}{\partial t} = \beta c f(a). \quad (3.7)$$

В зависимости от физико-химических свойств компонентов наиболее часто используется одно из следующих уравнений кинетики:

$$\frac{\partial a}{\partial t} = \beta c(a - a_0) \text{ – уравнение типа бимолекулярной реакции; } \quad (3.8)$$

$$\frac{\partial a}{\partial t} = \beta c \left(1 - \frac{a}{a_0}\right)^n \text{ – уравнение типа Трепнела } (n = 2); \quad (3.9)$$

$$\frac{\partial a}{\partial t} = \beta c e^{-\frac{a}{a_0} \gamma} \text{ – уравнение Зельдовича–Рогинского–Еловича; } \quad (3.10)$$

$$\frac{\partial a}{\partial t} = \beta c \frac{1 - \left(1 - \frac{a}{a_0}\right)^{1/3}}{\left(1 - \frac{a}{a_0}\right)^{1/3}} - \text{уравнение Бермана и Нагаева}; \quad (3.11)$$

$$\frac{\partial a}{\partial t} = \beta c \frac{1}{\frac{a}{a_0}} - \text{уравнение Мельникова и Дубинина}; \quad (3.12)$$

$$\frac{\partial a}{\partial t} = \beta(c - c^*) - \text{уравнение внешнедиффузионной кинетики}. \quad (3.13)$$

где  $a$ ,  $a_0$  – текущая и предельная емкость сорбента;  $\gamma$ ,  $l$  – константы;  $c^*$  – равновесная концентрация, которая определяется из уравнения  $c^* = \varphi(a)$ ;  $\beta$  – суммарный коэффициент массообмена.

Предложенные кинетические уравнения носят формальный характер.

Уравнение (3.8) получено на основании общих соображений формальной кинетики. Скорость реакции пропорциональна концентрации примеси в зоне непрореагировавшего вещества.

Уравнение кинетики (3.9) может быть интерпретировано как и (3.8), т.е. уравнение бимолекулярной реакции, но  $\beta$  – кинетический коэффициент, уменьшается пропорционально отработке, т.е.  $\beta_{\text{CO}_2}^2 = \beta_0(1 - a/a_0)$ .

Уравнение (3.10) обычно называют уравнением Зельдовича–Рогинского–Еловича. Оно очень часто используется для описания кинетики хемосорбции процессов поглощения  $\text{CO}_2$  и выделения кислорода. Уравнение выведено из предположения об экспоненциальном распределении энергии активации различных участков поверхности хемосорбента.

Уравнение кинетики (3.12) хорошо известно. Уравнение отражает факт лимитирования скорости реакции диффузии через плоский слой продуктов реакции. Уравнение было применено для описания процессов окисления Вагнером, для поглощения диоксида углерода ХП-И А.Х. Мельниковым и М.М. Дубининым, оно приводится в числе других и Д.А. Франк-Каменецким.

Уравнение (3.13) описывает при тех же условиях процесс адсорбции для элементов шихты шарообразной формы.

Замечательным является то, что предлагаемые ниже кинетические уравнения поглощения  $\text{CO}_2$  одинаково хорошо описывают процессы в хемосорбентах как на основе супероксидов щелочных металлов, так и на основе поглотителей.

Наиболее приемлемыми для описания скорости взаимодействия хемосорбентов с диоксидом углерода (подчеркнем, что рассматривается в настоящий момент скорость взаимодействия без учета влияния формы и пористой структуры элемента шихты) являются кинетические уравнения следующего вида

$$\frac{\partial a_{\text{CO}_2}}{\partial t} = \beta_{\text{CO}_2}^1 c_{\text{CO}_2} \left(1 - a_{\text{CO}_2} / a_{0\text{CO}_2}\right); \quad (3.14)$$

$$\frac{\partial a_{\text{CO}_2}}{\partial t} = \beta_{\text{CO}_2}^2 c_{\text{CO}_2} \left(1 - a_{\text{CO}_2} / a_{0\text{CO}_2}\right)^2 \quad (3.15)$$

$$\frac{\partial a_{\text{CO}_2}}{\partial t} = \beta_{\text{CO}_2}^3 c_{\text{CO}_2} \frac{1}{a + a_{\text{CO}_2} / a_{0\text{CO}_2}} \quad \text{или} \quad \frac{\partial a}{\partial t} = \beta_{\text{CO}_2}^3 c_{\text{CO}_2} \frac{1}{a_{\text{CO}_2} / a_{0\text{CO}_2}} \quad (3.16)$$

$$\frac{\partial a_{\text{CO}_2}}{\partial t} = \beta_{\text{CO}_2}^4 c_{\text{CO}_2} \frac{\left(1 - a_{\text{CO}_2} / a_{0\text{CO}_2}\right)^{1/3}}{1 - \left(1 - a_{\text{CO}_2} / a_{0\text{CO}_2}\right)^{1/3}} \quad (3.17)$$

$$\frac{\partial a_{\text{CO}_2}}{\partial t} = \beta_{\text{CO}_2}^5 c_{\text{CO}_2} e^{-\gamma a_{\text{CO}_2} / a_{0\text{CO}_2}}. \quad (3.18)$$

Таким образом, в уравнении (3.1) член  $\frac{\partial a}{\partial t}$  может быть раскрыт в виде уравнений (3.10) – (3.14).

Численные методы решения позволяют получить важные с точки зрения решения задач динамики хемосорбции результаты.

При решении задачи динамики важно знать макроскопическую скорость реакции.

Если обозначить  $F(t)$  как степень отработки, то для частицы шарообразной формы имеем

$$F(t) = \frac{3}{R^3} \int_0^R r^2 \frac{a(t, r)}{a_0} dr, \quad (3.19)$$

а  $\frac{\partial F}{\partial t}$  определяет скорость макроскопической реакции.

Для протекания реакции в чисто кинетической области комплекс  $\frac{k(\beta)R^2}{D} \leq 60$ .

Чисто диффузионная область имеет место при значении комплекса  $\frac{k(\beta)R^2}{D} \geq 500$ .

Переходная область наблюдается при значении  $60 \leq \frac{k(\beta)R^2}{D} \leq 500$ .

Однако наиболее интересным результатом явился вид аппроксимирующих уравнений, описывающих с точностью до 5 % указанные области (т.е. макроскопическую скорость реакции).

$$\text{При } K = \frac{k(\beta)R^2}{D} \leq 60 \quad \frac{\partial F}{\partial t} = \frac{D}{R^2} \frac{F_0}{a_0} c_0 (1-F).$$

$$\text{При } 60 \leq K = \frac{k(\beta)R^2}{D} \leq 500 \quad \frac{\partial F}{\partial t} = \frac{D}{R^2} \frac{1}{a_0} c_0 (F_0 - bF) \text{ при } F \leq F_1$$

$$\frac{\partial F}{\partial t} = \frac{D}{R^2} \frac{1}{a_0} c_0 \frac{3\sqrt[3]{1-F}}{1-\sqrt[3]{1-F}} \text{ при } F > F_1,$$

где  $F_0 = 3 \left( \sqrt{\frac{kR^2}{D}} \operatorname{cth} \left( \sqrt{\frac{kR^2}{D}} \right) - 1 \right)$ ; значения  $F_1, b$  находятся через  $F_0$ .

$$\text{При } K = \frac{k(\beta)R^2}{D} \geq 500; \quad \frac{\partial F}{\partial t} = \frac{D}{R^2} \frac{1}{a_0} c_0 \frac{3\sqrt[3]{1-F}}{1-\sqrt[3]{1-F}}.$$

Таким образом, в практически значимых случаях в широкой области значений комплекса  $K$  (кинетической и переходной) можно воспользоваться аппроксимирующим уравнением вида

$$\frac{\partial a}{\partial t} = \beta c (1 - a/a_0).$$

Вторая подсистема представляет собой уравнение, описывающее процессы, проходящие в замкнутом объеме обычно по одной из следующих моделей:

1. Реактор идеального смешения

$$V_p \frac{dC}{dt} = \sum_{i=1}^n G_i C_i - \sum_{j=1}^m G_j C + \phi(t),$$

где  $V_p$  – объем реактора;  $G$  – расход соответствующего компонента;  $\phi(t)$  – функция, описывающая динамику поглощения (выделения) компонента;  $n, m$  – соответственно количество входящих и выходящих потоков.

2. Реактор идеального вытеснения

$$\frac{\partial c}{\partial t} = -w \frac{\partial c}{\partial x} + \varphi(x, t),$$

где  $c$  – концентрация компонента;  $t$  – время;  $w$  – скорость;  $\varphi(x, t)$  – функция, описывающая динамику поглощения (выделения) компонента.

3. Реактор промежуточного типа

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - w \frac{\partial c}{\partial x} + \varphi(x, t),$$

где  $c$  – концентрация компонента;  $t$  – время;  $w$  – скорость;  $D$  – коэффициент диффузии;  $\varphi(x, t)$  – функция, описывающая динамику поглощения (выделения) компонента.

### 3.5.2. Математическая модель процесса регенерации воздуха в регенеративном патроне

Для вывода уравнений, описывающих процессы, протекающие в регенеративном элементе, примем ряд правомерных допущений:

– в регенеративном элементе происходят только реакции поглощения диоксида углерода и выделение кислорода, поглощение других веществ из воздуха помещений пренебрежимо мало;

– воздух, поступающий на вход регенеративного элемента, является ньютоновской несжимаемой жидкостью (т.е. для плотности воздуха можно принять  $\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial x} = 0$ );

– движение воздуха внутри регенеративного элемента – осевое со средней линейной скоростью  $\omega$ , определяемой по расходу вентилятора и сечению элемента;

– переходные режимы (запуск регенеративного элемента в работу) непродолжительны, а основное время регенеративный элемент работает при малом изменении теплофизических параметров, что позволяет считать параметры процесса независимыми от изменений температуры.

С учетом принятых допущений динамика процесса регенерации воздуха, протекающего в регенеративном элементе, будет описываться системой следующих уравнений: уравнение материального баланса, уравнение кинетики сорбции и уравнение изотермы сорбции.

Для вывода уравнения материального баланса рассмотрим элементарный объем регенеративного вещества длиной  $dx$  с площадью поперечного сечения  $dS$ .

Количество вещества (диоксида углерода или кислорода), вносимого в элементарный объем потоком воздуха через площадь  $dS$  за время  $dt$ , составляет  $\omega C dS dt$ , а потоком, вызванным продольной турбулентной диффузией

$$D \frac{\partial}{\partial x} \left( C + \frac{\partial C}{\partial x} dx \right) dS dt,$$

где  $\omega$  – линейная скорость воздуха;  $C$  – концентрация вещества в газовой фазе;  $D$  – коэффициент продольной диффузии. Количество вещества, вышедшего с воздушным потоком, равно:

$$\omega \left( C + \frac{\partial C}{\partial x} dx \right) dS dt ;$$

с диффузионным потоком –  $D \frac{\partial C}{\partial x} dS dt$ .

Изменение концентрации вещества в элементарном объеме вызовет изменение концентрации вещества в адсорбенте и подвижной фазе. В адсорбенте оно будет равно  $\frac{\partial a}{\partial t} dS dt dx$ , в потоке  $\frac{\partial C}{\partial t} dS dt dx$ , где  $a$  – концентрация вещества в регенеративном веществе.

В соответствии с законом сохранения массы разность между входящими и выходящими потоками будет составлять изменение количества вещества в рассматриваемом элементарном объеме (рис. 3.10), т.е.

Изменение = Приход вещества – Расход вещества

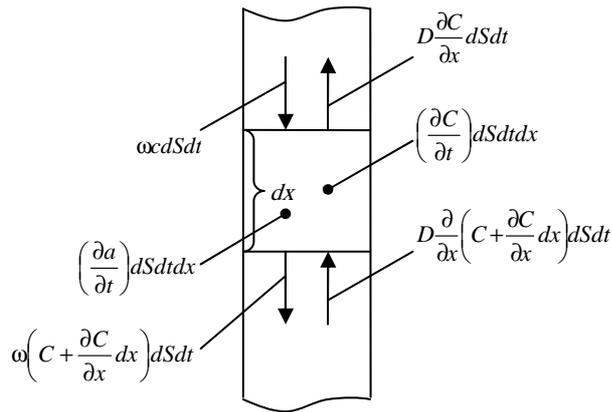


Рис. 3.10. Элементарный объем

$$\frac{\partial a}{\partial t} dS dt dx + \frac{\partial C}{\partial t} dS dt dx = \left( \omega C dS dt + D \frac{\partial}{\partial x} \left( C + \frac{\partial C}{\partial x} dx \right) dS dt \right) - \left( \omega \left( C + \frac{\partial C}{\partial x} dx \right) dS dt + D \frac{\partial C}{\partial x} dS dt \right).$$

В результате преобразований получаем:

$$\frac{\partial a}{\partial t} + \frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - \omega \frac{\partial C}{\partial x}.$$

Принимая во внимание тот факт, что концентрации поглощаемого (выделяемого) вещества в газовой и твердой фазах являются функциями времени и координаты, а скорость потока – функцией времени, запишем уравнение материального баланса в общем виде для каждого интересующего нас компонента воздушной смеси и любого регенеративного элемента:

$$\frac{\partial C_i^j(t, x)}{\partial t} = D_i \frac{\partial^2 C_i^j(t, x)}{\partial x^2} - \omega_i(t) \frac{\partial C_i^j(t, x)}{\partial x} - \frac{\partial a_i^j(t, x)}{\partial t},$$

где  $\omega_i(t) = \frac{W_i(t)}{F_i}$  – скорость потока в регенеративном элементе, которая зависит от расхода через него  $W_i(t)$  и его поперечного сечения  $F_i$ ;  $D_i$  – коэффициент продольной диффузии;  $C_i^j(t, x)$  – концентрация в газовой фазе компонента  $j$  в регенеративном элементе;  $a_i^j(t, x)$  – текущая концентрация компонента  $j$  (1 – кислород, 2 – диоксид углерода) в твердой фазе регенеративного элемента;  $x \in [0, L_i]$  – текущая координата по длине;  $L_i$  – длина слоя регенеративного продукта в  $i$ -м регенеративном элементе.

Начальные и граничные условия для полученных выше уравнений имеют вид:

$$C_i^j(0, x) = C^j(x), a_i^{\text{CO}_2}(0, x) = 0, a_i^{\text{O}_2}(0, x) = a_0^{\text{O}_2},$$

$$C_i^j(t, 0) = C^j(t), \frac{\partial C_i^j(t, L_i)}{\partial x} = 0, j = 1, 2.$$

Скорость сорбции  $\frac{\partial a}{\partial t}$  определяется из уравнения кинетики сорбции, которое показывает количество поглощаемого вещества из газового потока в единицу времени единицей объема поглотителя и в общем случае имеет вид:

$$\frac{\partial a}{\partial t} = \beta C f(a),$$

где  $\beta$  – кинетический коэффициент, означающий, какое количество вещества передается из газового потока объемом  $1 \text{ м}^3$  в слой поглотителя за 1 с, при разности концентраций  $1 \text{ кг/м}^3$ .

Вид функции  $f(a)$  зависит от типа изотермы сорбции.

Выделяют пять основных типов изотерм, которые изображены на рис. 3.11.

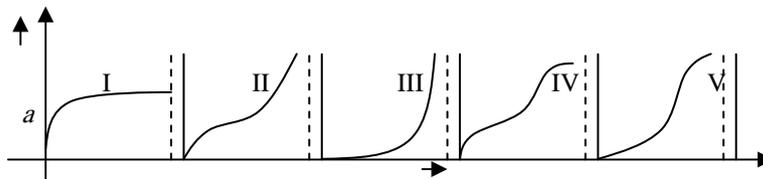


Рис. 3.11. Основные типы изотерм сорбции

Детальное рассмотрение теоретических и практических вопросов, связанных с определением уравнения кинетики сорбции приводится в работах В.М. Рамма, В.Л. Колина, А.А. Кримштейна, О. Левеншпиля, О.М. Годеса, Е.Н. Серпионовой, Л.В. Радужкевича и др.

Будем рассматривать уравнение кинетики сорбции на примере уравнения типа бимолекулярной реакции, что не исключает использования остальных видов. Для диоксида углерода уравнение имеет следующий вид:

$$\frac{\partial a^{\text{CO}_2}}{\partial t} = -\beta^{\text{CO}_2} C^{\text{CO}_2}(t, x) \left( 1 - \frac{a^{\text{CO}_2}(t, x)}{a_0^{\text{CO}_2}} \right),$$

а для кислорода:

$$\frac{\partial a^{\text{O}_2}}{\partial t} = -\beta^{\text{O}_2} C^{\text{CO}_2}(t, x) \left( \frac{a^{\text{O}_2}(t, x)}{a_0^{\text{O}_2}} \right),$$

поскольку показывает, что в отличие от поглощаемого диоксида углерода, кислород – выделяется.

Начальные условия для уравнений имеют вид:

$$a^{\text{CO}_2}(0, x) = 0; \quad a^{\text{O}_2}(0, x) = a_0^{\text{O}_2}.$$

Таким образом, система уравнений динамики сорбции принимает следующий вид:

$$\frac{\partial C_i^j(t, x)}{\partial t} = D_i \frac{\partial^2 C_i^j(t, x)}{\partial x^2} - \omega_i(t) \frac{\partial C_i^j(t, x)}{\partial x} - \frac{\partial a_i(t, x)}{\partial t};$$

$$C_i^j(0, x) = C_0^j(x), \quad C_i^j(t, 0) = C^j(t), \quad \frac{\partial C_i^j(t, L_i)}{\partial x} = 0;$$

$$\frac{\partial a_i(t, x)}{\partial t} = \begin{cases} -\beta_i^{\text{O}_2} C_i^{\text{CO}_2}(t, x) \frac{a_i^{\text{O}_2}(t, x)}{a_0^{\text{O}_2}}; \\ \beta_i^{\text{CO}_2} C_i^{\text{CO}_2}(t, x) \left( 1 - \frac{a_i^{\text{CO}_2}(t, x)}{a_0^{\text{CO}_2}} \right); \end{cases}$$

$$a^{\text{CO}_2}(0, x) = 0, \quad a^{\text{O}_2}(0, x) = a_0^{\text{O}_2};$$

$$i = 1..n, \quad j = 1, 2.$$

Окончательное формирование математического описания процесса регенерации воздуха в герметично замкнутом объеме заключается в сведении в единую модель рассмотренных выше математических описаний процессов.

Дополним модель уравнениями для объема идеального смешения

$$V_0 \frac{dC^j(t)}{dt} = \sum_{i=1}^n G_i (C_i^j(x, L_i) - C^j(t)) - (-1)^j H^j,$$

где  $H^j$  – нагрузка по диоксиду углерода и кислороду ( $j=1$  – кислород,  $j=2$  – диоксид углерода);  $V_0$  – объем помещений.

В случае, когда допущение об идеальном смешении в рассматриваемом объеме не может быть сделано, необходимо описывать процессы изменения концентраций в нем с учетом распределенности параметров, например, системой уравнений Навье-Стокса. При этом возникает необходимость геометрического определения места расположения элементов системы регенерации воздуха, всех источников и стоков компонент воздуха, что существенно усложняет вид математической модели и затрудняет ее решение.

Система уравнений представляет собой математическое описание процесса регенерации воздуха в замкнутом объеме, которая в большинстве случаев позволяет решать задачи проведения имитационных исследований, проектирования регенеративных и поглотительных элементов, систем регенерации воздуха.

### 3.6. РАСЧЕТ ТЕМПЕРАТУРНЫХ ПОЛЕЙ В ПОТОКЕ РАБОЧЕЙ СРЕДЫ И В КОНСТРУКЦИОННЫХ ЭЛЕМЕНТАХ РЕГЕНЕРАТИВНОГО ТЕПЛООБМЕННИКА

#### 3.6.1. Математическая модель температурного поля регенеративного патрона

С целью линеаризации дифференциальных уравнений теплопроводности, позволяющей использовать аналитические решения, предлагается в пределах локального временного интервала рассматривать весь объем активного вещества в регенеративном патроне как совокупность локальных пространственных областей, в каждой из которых возможно использование

постоянных значений теплофизических характеристик активного вещества и газовой среды, определяемых для средней температуры в локальной пространственно-временной области.

Так как для расчета температурного поля регенеративного патрона необходимо взаимосвязанное решение ряда задач, а их параметры являются функциями текущих температур, расчет температурного поля каждой последующей локальной области выполняется по итеративной методике, предусматривающей многократное последовательное решение всех используемых задач с текущим уточнением кинетических и теплофизических характеристик рассматриваемых процессов.

В связи со сложным характером пространственного температурного поля регенеративного патрона предлагается рассмотрение локальной пространственной области в виде кольца прямоугольного сечения, ось которого совпадает с осью патрона (исходя из осесимметричного характера температурного поля) (рис. 3.12). Использование такой формы локальной пространственной области позволяет учесть неравномерность обработки активного вещества по сечению патрона вследствие перегрева центральных областей и пристеночных эффектов.

В связи с этим целесообразно дополнительно использовать решение задачи нестационарной теплопроводности для полного ограниченного цилиндра с функционально меняющейся температурой окружающей среды со стороны боковых поверхностей. Это наиболее общая математическая постановка, поскольку решение задачи с граничными условиями 3 рода может быть использовано для расчета температурных полей при граничных условиях как 1, так и 2 рода путем задания соответствующих значений коэффициентов, входящих в граничные условия 3 рода.

Таким образом, температурное поле выбранной области моделируется решением следующей задачи нестационарной теплопроводности (рис. 3.12).

В приводимой постановке задача записана относительно температуры окружающей среды с торцовых поверхностей.

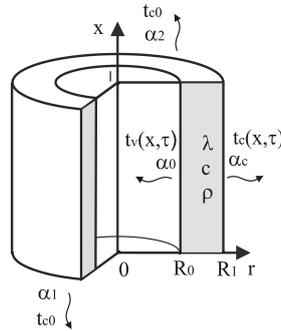


Рис. 3.12. Полный ограниченный цилиндр с переменными температурами окружающей среды

$$\frac{\partial t(x, r, \tau)}{\partial \tau} = a^2 \left( \frac{\partial^2 t(x, r, \tau)}{\partial x^2} + \frac{\partial^2 t(x, r, \tau)}{\partial r^2} + \frac{1}{r} \frac{\partial t(x, r, \tau)}{\partial r} \right), \quad (3.26)$$

$$0 \leq x \leq l, \quad R_0 \leq r \leq R_1, \quad \tau > 0;$$

$$t(x, r, 0) = f(x, r) - t_{c0}; \quad (3.27)$$

$$\lambda \frac{\partial t(0, r, \tau)}{\partial x} - \alpha_1 t(0, r, \tau) = 0; \quad (3.28)$$

$$\lambda \frac{\partial t(l, r, \tau)}{\partial x} + \alpha_2 t(l, r, \tau) = 0; \quad (3.29)$$

$$\lambda \frac{\partial t(x, R_0, \tau)}{\partial r} - \alpha_0 (t(x, R_0, \tau) - t_v(x, \tau) + t_{c0}) = 0; \quad (3.30)$$

$$\lambda \frac{\partial t(x, R_1, \tau)}{\partial r} + \alpha_c (t(x, R_1, \tau) - t_c(x, \tau) + t_{c0}) = 0; \quad (3.31)$$

Для торцевых поверхностей могут быть выбраны иные граничные условия с соответствующей постановкой задачи.

Решение задачи (3.26) – (3.31) так же может быть получено методом конечных интегральных преобразований, примененных последовательно по линейной и цилиндрической координате.

Для исключения координаты  $x$  используем формулу перехода к изображениям

$$T(r, \tau) = \int_0^l t(x, r, \tau) S(\mu, x) dx;$$

где  $S(\mu, x)$  – ядро интегрального преобразования, являющееся решением задачи с однородными граничными условиями:

$$\frac{d^2 S(x)}{dx^2} = -\mu^2 S(x); \quad (3.32)$$

$$\lambda \frac{dS(0)}{dx} + \alpha_1 S(0) = 0; \quad (3.33)$$

$$\lambda \frac{dS(l)}{dx} + \alpha_2 S(l) = 0; \quad (3.34)$$

Задача (3.32) – (3.34) с точностью до постоянного множителя имеет решение:

$$S(x) = \sin(\mu x + \varphi), \quad (3.35)$$

где

$$\varphi = \arctg\left(\frac{\lambda\mu}{\alpha_1}\right) \quad (3.36)$$

$\mu_n$  –  $n$ -й положительный корень уравнения

$$\alpha_3 \sin(\mu l + \varphi) + \lambda \mu \cos(\mu l + \varphi) = 0. \quad (3.37)$$

Обратный переход может быть выполнен по формуле

$$t(x, r, \tau) = \sum_{n=1}^{\infty} \frac{T(r, \tau) S(\mu_n, x)}{N_n},$$

где 
$$N_n = \int_0^l S^2(\mu_n, x) dx = 0,5 \left( l - \frac{1}{\mu_n} (\sin(\mu_n l + \varphi_n) \cos(\mu_n l + \varphi_n) - \sin(\varphi_n) \cos(\varphi_n)) \right) \quad (3.38)$$

В изображениях задача (3.26) – (3.31) имеет вид:

$$\frac{\partial T(r, \tau)}{\partial \tau} = a^2 \left( \frac{\partial^2 T(r, \tau)}{\partial r^2} + \frac{1}{r} \frac{\partial T(r, \tau)}{\partial r} + \mu^2 T(r, \tau) \right), \quad (3.39)$$

$$T(r, 0) = \int_0^l (f(x, r) - t_{c0}) S(\mu, x) dx; \quad (3.40)$$

$$\lambda \frac{\partial T(R_0, \tau)}{\partial r} - \alpha_0 (T(R_0, \tau) - U(\tau)) = 0; \quad (3.41)$$

$$\lambda \frac{\partial T(R_1, \tau)}{\partial r} - \alpha_c (T(R_1, \tau) - W(\tau)) = 0; \quad (3.42)$$

где 
$$U(\tau) = \int_0^l (t_v(x, t) - t_{c0}) S(\mu, x) dx, \quad (3.43)$$

$$W(\tau) = \int_0^l (t_c(x, t) - t_{c0}) S(\mu, x) dx. \quad (3.44)$$

Теперь возможно исключение координаты  $r$  путем использования следующего интегрального преобразования:

$$V(\tau) = \int_{R_0}^{R_1} T(r, \tau) r P(\eta, r) dr; \quad (3.45)$$

с обратным переходом по формуле

$$T(r, \tau) = \sum_{k=1}^{\infty} \frac{P(\eta_k, \tau) V(\tau)}{Z_k},$$

причем  $P(\eta, r)$  является решением задачи

$$\frac{d^2 P(r)}{dr^2} + \frac{1}{r} \frac{dP(r)}{dr} + \eta^2 P(r) = 0;$$

$$\lambda \frac{\partial P(R_0)}{\partial r} - \alpha_0 P(R_0) = 0;$$

$$\lambda \frac{\partial P(R_1)}{\partial r} + \alpha_c P(R_1) = 0.$$

Это решение с точностью до постоянного множителя имеет вид:

$$P(r) = J_0(\eta r) + A Y_0(\eta r);$$

где 
$$A = -\frac{\eta \lambda J_1(\eta R_0) + \alpha_0 J_0(\eta R_0)}{\eta \lambda Y_1(\eta R_0) + \alpha_0 Y_0(\eta R_0)};$$
 а  $\eta_k$  –  $k$ -й положительный корень уравнения

$$\frac{\eta \lambda J_1(\eta R_1) - \alpha_c J_0(\eta R_1)}{-\eta \lambda Y_1(\eta R_1) + \alpha_c Y_0(\eta R_1)} = -\frac{\eta \lambda J_1(\eta R_0) + \alpha_0 J_0(\eta R_0)}{\eta \lambda Y_1(\eta R_0) + \alpha_0 Y_0(\eta R_0)}.$$

Тогда

$$\begin{aligned}
Z_k &= \int_{R_0}^{R_1} r P^2(r) dr = \\
&= \frac{R_1^2}{2} \left( (J_0(\eta_k R_1) + A_k Y_0(\eta_k R_1))^2 + (J_1(\eta_k R_1) + A_k Y_1(\eta_k R_1))^2 \right) - \\
&- \frac{R_0^2}{2} \left( (J_0(\eta_k R_0) + A_k Y_0(\eta_k R_0))^2 + (J_1(\eta_k R_0) + A_k Y_1(\eta_k R_0))^2 \right).
\end{aligned}$$

Применяя преобразование (3.45) к задаче (3.39) – (3.42), переходим к новым изображениям:

$$\frac{dV(\tau)}{d\tau} + \eta_k^2 a^2 V(\tau) = -\frac{\alpha_0 R_0^2}{\lambda} P(R_0) U(\tau) + \frac{\alpha_c R_1^2}{\lambda} P(R_1) W(\tau); \quad (3.46)$$

$$V(0) = \int_{R_0}^{R_1} \int_0^l (f(x, r) - t_{c0}) S(x) r P(r) dx dr. \quad (3.47)$$

Решением задачи (3.46) – (3.47) является функция

$$\begin{aligned}
V(\tau) &= \exp(-\eta_k^2 a^2 \tau) \left( \int_{R_0}^{R_1} \int_0^l (f(x, r) - t_{c0}) S(x) r P(r) dx dr - \right. \\
&- \frac{\alpha_0 a^2 R_0}{\lambda} P(R_0) \int_0^l \int_0^l (t_v(x, \tau) - t_{c0}) S(x) \exp(\eta_k^2 a^2 \tau) dx d\tau + \\
&\left. + \frac{\alpha_c a^2 R_1}{\lambda} P(R_1) \int_0^l \int_0^l (t_c(x, \tau) - t_{c0}) S(x) \exp(\eta_k^2 a^2 \tau) dx d\tau \right). \quad (3.48)
\end{aligned}$$

Возврат к оригиналам выполняется по формуле

$$t(x, r, \tau) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{V(\tau) P(r) S(x)}{N_n Z_k}.$$

При расчете температурного поля локальной области возможно использовать рассмотренную задачу в упрощенной постановке. В этом случае

$$\begin{aligned}
\frac{\partial t(x, r, \tau)}{\partial \tau} &= a^2 \left( \frac{\partial^2 t(x, r, \tau)}{\partial x^2} + \frac{\partial^2 t(x, r, \tau)}{\partial r^2} + \frac{1}{r} \frac{\partial t(x, r, \tau)}{\partial r} \right), \\
0 \leq x \leq l, \quad R_0 \leq r \leq R_1, \quad \tau > 0; \\
t(x, r, 0) &= f(x, r) - t_{c0}; \quad \frac{\partial t(0, r, \tau)}{\partial x} = 0; \quad \frac{\partial t(l, r, \tau)}{\partial x} = 0; \\
\lambda \frac{\partial t(x, R_0, \tau)}{\partial r} - \alpha_0 (t(x, R_0, \tau) - t_v(x, \tau) + t_{c0}) &= 0; \\
\lambda \frac{\partial t(x, R_1, \tau)}{\partial r} + \alpha_c (t(x, R_1, \tau) - t_c(x, \tau) + t_{c0}) &= 0;
\end{aligned}$$

Методика решения остается прежней, но ряд формул приобретает упрощенный вид.

Формула (3.33):  $\frac{\partial S(0)}{\partial x} = 0.$

Формула (3.34):  $\frac{\partial S(l)}{\partial x} = 0.$

Формула (3.35):  $S(x) = \cos \mu x.$

Формула (3.36):  $\varphi = 0.$

Формула (3.37):  $\mu_n = \frac{\pi n}{l}.$

Формула (3.38):  $N_n = \int_0^l S^2(\mu_n, x) dx = 0,5 \left( l + \frac{1}{\mu_n} \sin \mu_n l \cos \mu_n l \right)$

Формула (3.40):  $T(r, 0) = \int_0^l f(x, r) S(\mu, x) dx.$

Формула (3.43):  $U(\tau) = \int_0^l t_v(x, \tau) S(\mu, x) dx.$

Формула (3.44):  $W(\tau) = \int_0^l t_c(x, \tau) S(\mu, x) dx.$

Формула (3.47): 
$$V(0) = \int_{R_0}^{R_1} \int_0^l f(x, r) S(x) r P(r) dx dr.$$

Формула (3.48):

$$V(t) = \exp(-\eta_k^2 a^2 \tau) \left( \int_{R_0}^{R_1} \int_0^l f(x, r) S(x) r P(r) dx dr - \frac{\alpha_0 a^2 R_0}{\lambda} P(R_0) \int_0^{\tau} \int_0^l t_v(x, \tau) S(x) \exp(\eta_k^2 a^2 \tau) dx dt + \frac{\alpha_c a^2 R_1}{\lambda} P(R_1) \int_0^{\tau} \int_0^l t_c(x, \tau) S(x) \exp(\eta_k^2 a^2 \tau) dx dt \right).$$

Благодаря использованию аналитических решений задач теплопроводности, время расчета температурного поля локальной области на современном персональном компьютере составляет, как правило, доли секунды. Кроме того, длительность и пространственные размеры локальной области определяются временем изменения температур газового потока, гранул активного вещества и конструктивных элементов на несколько (5 – 10) градусов, что в реальных условиях может составлять секунды или даже десятки секунд. Поэтому возможно моделирование температурных полей в течение длительных рабочих циклов, начиная с момента включения нагревателей в холодном патроне.

Алгоритм расчета температурного поля локальной области регенеративного патрона:

1. Задаются начальные температурные поля конструктивных элементов, газовых потоков и слоя гранул активного вещества, первоначально – по результатам предварительного прогрева, в последующем – в виде температурных полей конца предыдущего интервала времени для текущей пространственной локальной области.

2. Определяются теплофизические характеристики воздуха и активного вещества, зависящие от температуры, а так же скорости газовых потоков и коэффициенты тепло- и массоотдачи.

Теплофизические характеристики воздуха как функции температуры  $t$ , °С при атмосферном давлении могут быть аппроксимированы следующими зависимостями.

Плотность, кг/м<sup>3</sup>

$$\rho(t) = 1,293 \frac{273}{t + 273}.$$

Динамическая вязкость, Па·с

$$\mu(t) = 17,3 \cdot 10^{-6} \frac{397}{t + 397} \left( \frac{t}{273} + 1 \right)^{1,5}.$$

Теплоемкость, Дж/(кг·К)

$$c(t) = 1 + 0,116 \cdot 10^{-3} t.$$

Теплопроводность, Вт/(м·К)

$$\lambda(t) = \left( 2,56 + 2,318 \cdot 10^{-3} t - \frac{0,32}{\exp t} \right) \cdot 10^{-2}.$$

В неподвижном слое зернистого материала коэффициент теплоотдачи между потоком и поверхностью частицы

$$\text{Nu} = \begin{cases} 0,0035 \left( \frac{\text{Re}}{\varepsilon} \right)^{1,5} \text{Pr}^{0,33}, & \frac{\text{Re}}{\varepsilon} < 200; \\ 0,4 \left( \frac{\text{Re}}{\varepsilon} \right)^{0,67} \text{Pr}^{0,33}, & \frac{\text{Re}}{\varepsilon} \geq 200. \end{cases}$$

Здесь  $\text{Nu} = \frac{\alpha d_3}{\lambda}$ ,  $\text{Re} = \frac{W d_3}{\nu}$ ,  $\text{Pr} = \frac{\nu}{a}$ ,  $a = \frac{\lambda}{c \rho}$ ,  $d_3 = \sqrt{\frac{F}{\pi}}$ ,  $W$  – скорость газа, м/с,  $F$  – площадь поверхности частицы

зернистого материала, м<sup>2</sup>,  $\varepsilon$  – порозность зернистого слоя,  $\lambda$ ,  $c$ ,  $\rho$ ,  $\nu$  – соответственно теплопроводность, Вт/(м·К), теплоемкость, Дж/(кг·К), плотность, кг/м<sup>3</sup>, и кинематическая вязкость, м/с, газовой среды при текущей температуре,  $a$  – определяемый коэффициент теплоотдачи, Вт/(м<sup>2</sup>·К).

Коэффициент теплоотдачи излучением от частиц к газу

$$\alpha_{\text{лг}} = 0,227 \left( \frac{1}{1 + \frac{\varepsilon(1-\xi)}{2\xi(1-\varepsilon)}} \right) \left( \frac{T}{100} \right)^3,$$

здесь  $\xi$  – степень черноты поверхности частицы.

Коэффициент теплоотдачи излучением между частицами

$$\alpha_{\text{лз}} = 0,227 \frac{\xi}{2 - \xi} \left( \frac{T}{100} \right)^3.$$

В неподвижном слое зернистого материала коэффициент массоотдачи между газом и поверхностью частицы

$$Re < 2, \quad Nu_d = 0,515 Re^{0,85} Sc^{0,33},$$

$$30 > Re \geq 2, \quad Nu_d = 0,725 Re^{0,47} Sc^{0,33},$$

$$Re \geq 30, \quad Nu_d = 0,395 Re^{0,64} Sc^{0,33}.$$

Здесь  $Nu_d = \frac{\beta d_p}{D}$ ,  $Sc = \frac{\nu}{D}$ ,  $D$  – коэффициент диффузии (самодиффузии) газа, м/с,  $\beta$  – определяемый коэффициент массоотдачи, 1/с.

3. По средним значениям температур газовых сред и коэффициентов теплоотдачи рассчитываются температуры поверхностей гранул активного вещества.

4. Для найденных температур поверхностей уточняются коэффициенты теплоотдачи и рассчитываются средние температуры газовых компонентов в текущей локальной области.

5. Пункты 2 – 4 повторяются до достаточного совпадения расчетных результатов (практически – не более 3 раз).

6. Вычисляются температурные поля, тепловые потоки и составляющие тепловых балансов для момента окончания локальной временной области.

При известных температурных полях численные значения удельных тепловых потоков определяются по закону Фурье:

$$q = \lambda \frac{\partial t(n)}{\partial n}.$$

Тепловой баланс локальной области в нестационарном температурном режиме

$$\Delta Q_{r1} + \Delta Q_x + \Delta Q_{\phi} + \Delta Q_{r2} + \Delta Q_k + \Delta Q_{ав} + \Delta Q_{гг} + \Delta Q_{п} = 0,$$

где  $\Delta Q_{r1}$  – тепловая мощность, привносимая газовым потоком;  $\Delta Q_x$  – тепловая мощность теплового эффекта химических реакций;  $\Delta Q_{\phi}$  – тепловая мощность, расходуемая на фазовый переход активного вещества;  $\Delta Q_{r2}$  – тепловая мощность, отводимая с отходящими газами;  $\Delta Q_k$  – тепловая мощность, затраченная на нагрев конструктивных элементов;  $\Delta Q_{ав}$  – тепловая мощность, затраченная на нагрев активного вещества;  $\Delta Q_{гг}$  – тепловая мощность, затраченная на нагрев газов в аппарате;  $\Delta Q_{п}$  – тепловая мощность потерь в окружающее пространство.

7. Сохраняются массивы параметров, необходимые для расчета начального распределения температур локальной области в последующем временном интервале.

8. Переход к расчету температурных полей очередной локальной временной области.

Алгоритм расчета нестационарного температурного поля регенеративного патрона.

1. Определяется температурное поле слоя активного вещества в результате предварительного прогрева.

2. Задаются предварительные геометрические размеры пространственных локальных областей по радиальной и продольной координатам.

3. Для локального интервала времени выполняется последовательный расчет температурных полей локальных пространственных областей по слоям с последовательным расчетом по радиальной координате внутри текущего слоя.

4. Определяются составляющие независимого интегрального теплового баланса для локальной временной области. При значительном разбалансе осуществляется переход к п. 2.

5. Возврат к п. 3 при переходе к расчету очередного локального временного интервала. Несмотря на возможность произвольного выбора длительности локального временного интервала, целесообразно выбирать ее кратной времени прохождения газовым потоком пространственной локальной области.

Таким образом, последовательный расчет температурных полей локальных областей во времени и пространстве позволяет получить расчетное нестационарное температурное поле регенеративного патрона за весь рабочий цикл.

### 3.6.2. Моделирование тепловых процессов в потоке рабочей среды и в конструктивных элементах регенеративного теплообменника

Конструкция регенеративного теплообменника индивидуальных изолирующих дыхательных аппаратов (РД ИИДА) представлена на рис. 3.13. Основным конструктивным элементом регенеративного теплообменника является свернутая в спираль гофрированная алюминиевая лента. При вдохе лента нагревается, охлаждая вдыхаемый воздух; при выдохе – охлаждается выдыхаемым воздухом.

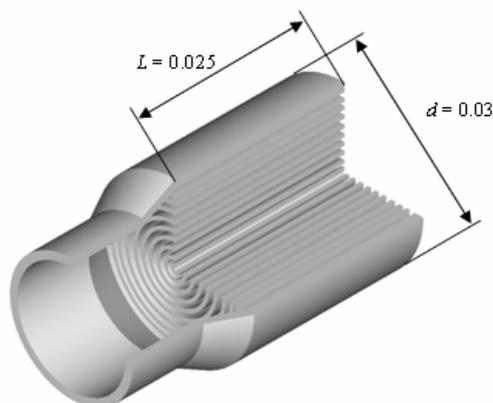


Рис. 3.13. Конструкция РТ ИИДА

Выбор рациональных конструктивных параметров РТ ИИДА обеспечивает комфортные условия работы человека при использовании индивидуальных средств защиты. Это означает, что должен быть выбран разумный баланс следующих параметров: массы, гидравлического сопротивления (которые должны быть минимальны) и снижения температуры вдыхаемого воздуха (которое должно быть максимальным) при различных режимах работы.

Для решения данной задачи разработана математическая модель нестационарных температурных полей РТ ИИДА.

Температурные поля РТ ИИДА моделируются решениями следующих взаимосвязанных задач:

- задача нестационарного переноса тепла в ленте по направлению потока;
- задача переноса тепла потоком газовой смеси.

При выводе уравнений математической модели были приняты следующие допущения:

- температура потока по сечению канала не меняется;
- газ считается несжимаемым;
- канал имеет постоянное сечение;
- теплофизические характеристики среды внутри малой пространственно-временной области остаются постоянными.

Рассмотрим участок канала РТ ИИДА (рис. 3.14). Канал имеет замкнутый периметр, так как образован двойной лентой, уложенной в виде спирали. Кроме того, поверхности, образующие канал, имеют одинаковую температуру в каждом поперечном сечении.

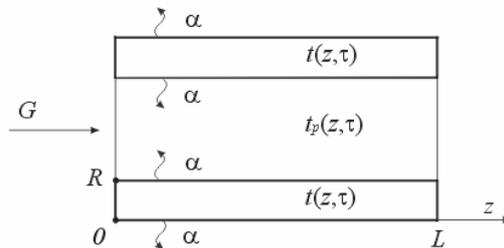


Рис. 3.14. Схема канала, образованного двумя поверхностями с одинаковыми температурами

Следовательно, основной конструктивный элемент РТ ИИДА (тонкая гофрированная алюминиевая лента) может быть рассмотрен как однослойная тонкая пластина. При этом канал, по которому движется воздух, образован одной стенкой замкнутого периметра с температурой  $t(z, \tau)$ .

Задача нестационарной теплопроводности для тонкой ленты имеет вид:

$$\frac{\partial t(z, \tau)}{\partial \tau} = a^2 \frac{\partial^2 t(z, \tau)}{\partial z^2} - H t(z, \tau) + H t_p(z, \tau), \quad (3.49)$$

где  $H = \frac{2\alpha}{R\rho c}$ ,  $\tau > 0$ .

Условия однозначности:

$$t(z, 0) = f(z), \quad (3.50)$$

$$\lambda \frac{\partial t(0, \tau)}{\partial z} - \alpha(t(0, \tau) - t_p(0, \tau)) = 0, \quad (3.51)$$

$$\lambda \frac{\partial t(L, \tau)}{\partial z} + \alpha(t(L, \tau) - t_p(L, \tau)) = 0, \quad (3.52)$$

где  $z$  – пространственная координата по длине пластины;  $\tau$  – время;  $t(z, \tau)$  – текущая температура пластины;  $t_p(z, \tau)$  – температура газа, омывающего пластину;  $R$  – толщина пластины;  $L$  – длина пластины;  $a$  – температуропроводность материала пластины;  $\lambda$  – теплопроводность материала пластины;  $c$  – теплоемкость материала пластины;  $\rho$  – плотность материала пластины;  $\alpha$  – коэффициент конвективной теплоотдачи от поверхности пластины к газу.

Аналитическое решение задачи (3.49) – (3.52) имеет вид:

$$t(z, \tau) = \sum_{n=1}^{\infty} \frac{V(\tau)W(z)}{N}, \quad (3.53)$$

причем суммирование ведется по значениям  $\mu$ , которые определяются как последовательные положительные корни уравнения

$$\alpha \sin \left( \sqrt{\frac{\mu^2 - H}{a^2}} L + \varphi \right) + \lambda \sqrt{\frac{\mu^2 - H}{a^2}} \cos \left( \sqrt{\frac{\mu^2 - H}{a^2}} L + \varphi \right) = 0.$$

В формуле (3.53)

$$W(z) = \sin\left(\sqrt{\frac{\mu^2 - H}{a^2}} z + \varphi\right),$$

$$\varphi = \operatorname{atg}\left(\frac{\lambda\sqrt{\frac{\mu^2 - H}{a^2}}}{\alpha}\right),$$

$$V(\tau) = \left(V(0) + \int_0^\tau Q(\tau) \exp(\mu_n^2 \tau) d\tau\right) \exp(-\mu_n^2 \tau),$$

$$V(0) = \int_0^L f(z) W(z) dz,$$

$$Q(\tau) = \frac{\alpha a^2}{\lambda} (W(L) t_p(L, \tau) + W(0) t_p(0, \tau)) + H \int_0^L t_p(z, \tau) W(z) dz,$$

$$N = \int_0^L W^2(z) dz = \int_0^L \sin^2\left(\sqrt{\frac{\mu^2 - H}{a^2}} z + \varphi\right) dz =$$

$$= \frac{1}{2\sqrt{\frac{\mu^2 - H}{a^2}}} \left(\sqrt{\frac{\mu^2 - H}{a^2}} L + \sin(\varphi) \cos(\varphi) - \right.$$

$$\left. - \sin\left(\sqrt{\frac{\mu^2 - H}{a^2}} L + \varphi\right) \cos\left(\sqrt{\frac{\mu^2 - H}{a^2}} L + \varphi\right)\right).$$

Без учета молекулярного переноса тепла, уравнение переноса тепла потоком жидкости, движущейся в режиме идеального вытеснения по каналу, образованному двумя поверхностями, имеет вид:

$$\frac{\partial t_p(z, \tau)}{\partial \tau} + W \frac{\partial t_p(z, \tau)}{\partial z} + K t_p(z, \tau) = F(z, \tau), \quad (3.54)$$

где  $W = \frac{G}{S\rho}$  – скорость движения потока,

$$K = \frac{\alpha_1 \Pi_1}{Gc}, \quad F(z, \tau) = K t_1(z, \tau). \quad (3.55)$$

Условия однозначности:

$$t_p(z, 0) = f(z); \quad t_p(0, \tau) = t_0(\tau). \quad (3.56)$$

где  $z$  – пространственная координата по направлению движения потока газа;  $\tau$  – время;  $t_p(z, \tau)$  – текущая температура потока газа;  $G$  – массовый расход газа;  $\Pi_1$  – омываемый периметр стенки канала;  $\alpha_1$  – коэффициент конвективной теплоотдачи от стенки канала к потоку газа;  $t_1(z, \tau)$  – температура стенки канала;  $S$  – площадь сечения канала;  $\rho$  – плотность газа;  $c$  – теплоемкость газа.

Для короткого интервала времени  $\Delta\tau$  решение задачи (3.54) – (3.56) имеет вид:

$$t_p(z, \Delta\tau) = \exp(-Pz) \left( t_0 + \int_0^z V(z) \exp(Pz) dz \right).$$

$$\text{Здесь } P = \frac{K\Delta\tau + 1}{W\Delta\tau}, \quad V(z) = \frac{1}{W} \left( F(z, \Delta\tau) + \frac{f(z)}{\Delta\tau} \right).$$

Таким образом, температурные поля РТ ИИДА описываются системой дифференциальных уравнений в частных производных с соответствующими начальными и граничными условиями.

### 3.6.3. Алгоритм расчета нестационарных температурных полей РТ ИИДА

На основе аналитических решений уравнений математической модели РТ ИИДА с использованием экспериментально полученных значений коэффициентов теплоотдачи разработан алгоритм расчета нестационарных температурных полей потока газозвушной смеси и конструкционного элемента (тонкой гофрированной ленты, компактно уложенной в виде спирали в цилиндрическом кожухе) РТ ИИДА.

Расчет температурных полей гофрированной ленты и потока газозвушной смеси может быть осуществлен на основе последовательного теплового расчета соответствующих локальных областей. При последовательном просчете температурных полей локальных областей выходные температуры ленты являются входными для последующей временной области, а выходные температуры потока вещества являются входными для последующей пространственной области.

Расчет нестационарных температурных полей РТ ИИДА и газозвушной смеси включает последовательный расчет температурных полей для каждого локального временного интервала. При этом для каждой локальной пространственной области сохраняются текущие значения температуры газозвушной смеси и температуры ленты. Эти данные используются

в качестве исходных для теплового расчета той же локальной пространственной области в очередной локальной временной интервал.

Расчет температурных полей РТ ИИДА осуществляется для последовательности дыхательных циклов, причем температуры газовой смеси и ленты на выходе из теплообменника при выдохе являются входными температурами при последующем вдохе.

Алгоритм расчета нестационарных температурных полей потока газовой смеси и ленты РТ ИИДА (рис. 3.15) предусматривает выполнение следующей последовательности действий.

- Ввод исходных данных для расчета нестационарных температурных полей в аппарате (геометрические параметры теплообменника, теплофизические характеристики рабочей среды и материала конструкционного элемента аппарата, начальные температуры ленты, газовой смеси на первом вдохе и на выдохе).
- Задание температурного поля ленты в начальный момент времени (температура ленты по ее длине принимается равной температуре окружающей среды).
- Установка счетчика проходов в начальное состояние (проход среды через теплообменник представляет собой один вдох или один выдох; расчет температурных полей начинается с вдоха).
- Установка счетчика времени в начальное состояние.
- Задание температуры потока на входе в аппарат.
- Установка счетчика локальных областей в начальное состояние.
- Итерационный расчет взаимосвязанных нестационарных температурных полей потока газовой смеси и ленты регенеративного теплообменника для текущей локальной области.
- Проверка условия расчета последней локальной области. В случае невыполнения условия осуществляется переход к следующей локальной области, иначе выводятся в виде таблицы в файл поля температур потока среды и ленты для текущего момента времени текущего прохода.
- Проверка условия расчета полей температур для последнего момента времени. В случае невыполнения условия осуществляется переход к следующему моменту времени, иначе выполняется следующий пункт.

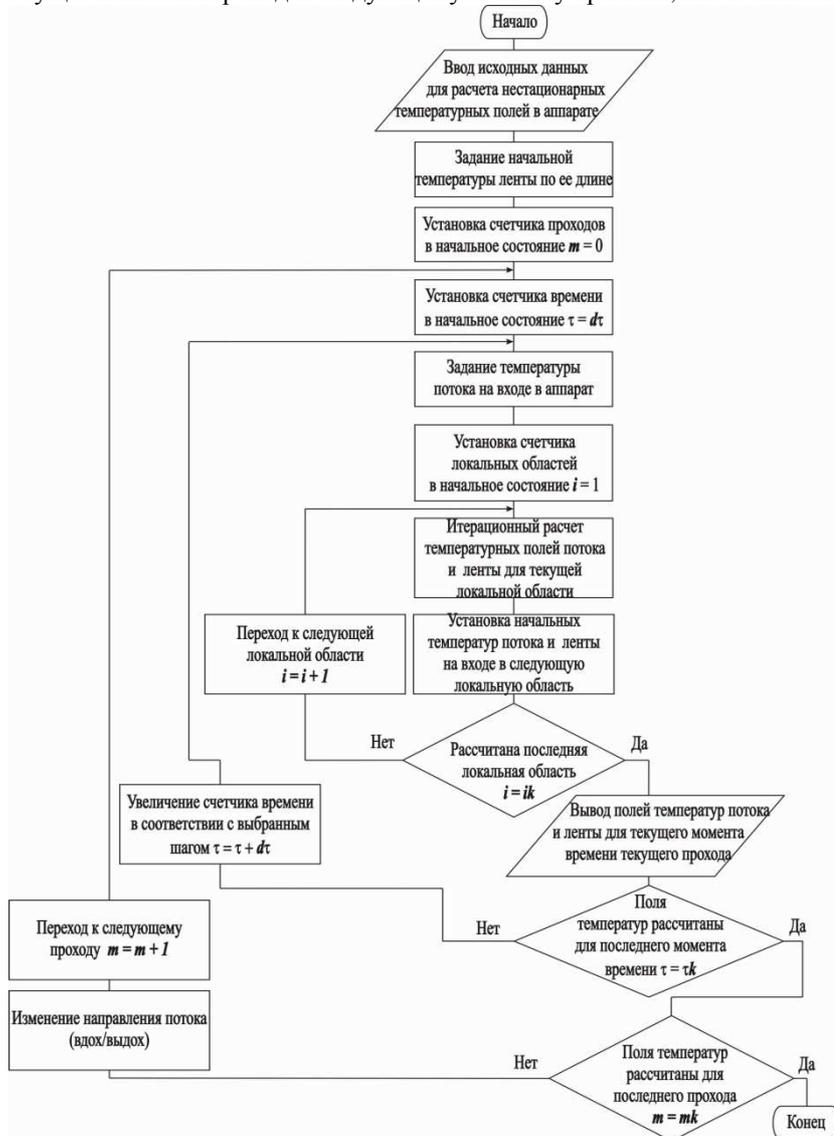


Рис. 3.15. Алгоритм расчета нестационарных температурных полей РТ ИИДА

– Проверка условия расчета полей температур для последнего рассматриваемого прохода газовой смеси через аппарат. В случае невыполнения условия осуществляется изменение направления потока (вдох/выдох) и переход к следующему проходу, иначе работа алгоритма завершается.

### 3.6.4. Результаты математического моделирования РТ ИИДА

На основе представленного выше алгоритма расчета взаимосвязанных нестационарных температурных полей РТ ИИДА разработана и реализована расчетная программа.

В результате работы программа формирует таблицы, в которых содержатся значения температуры по длине пластины в различные моменты времени (временной интервал 1 секунда) для каждого цикла дыхания (период времени, включающий один вдох и один выдох).

По расчетным данным построены кривые, описывающие распределение температуры по длине РТ ИИДА для первых пяти циклов дыхания (на момент окончания цикла) (рис. 3.16 – 3.18), а также кривые, показывающие изменение температуры по длине регенеративного теплообменника индивидуального дыхательного аппарата в различные моменты времени в течение одного цикла дыхания (рис. 3.19).

Совокупность вышеописанных кривых образует две поверхности, одна из которых описывает нестационарное поле температур потока газовой смеси, проходящего через РТ ИИДА, а другая – нестационарное температурное поле гофрированной ленты. На рис. 3.20 – 3.21 показаны фрагменты нестационарных полей температур соответственно газовой смеси, проходящего через РТ ИИДА за первые два цикла дыхания, и гофрированной ленты.

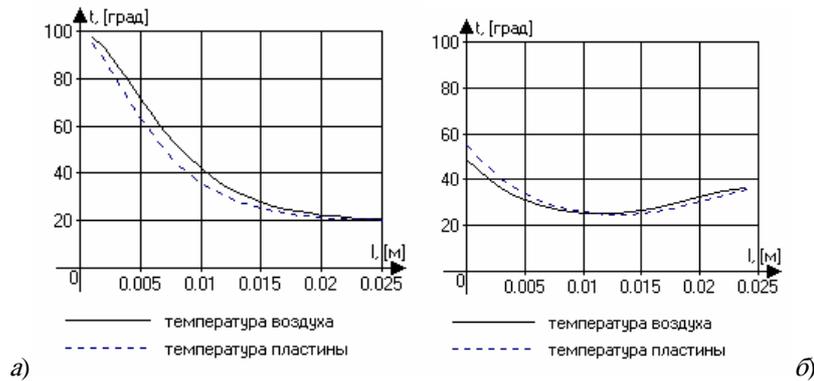


Рис. 3.16. Поле температур РТ ИИДА в первом цикле дыхания:  
а – при вдохе; б – при выдохе

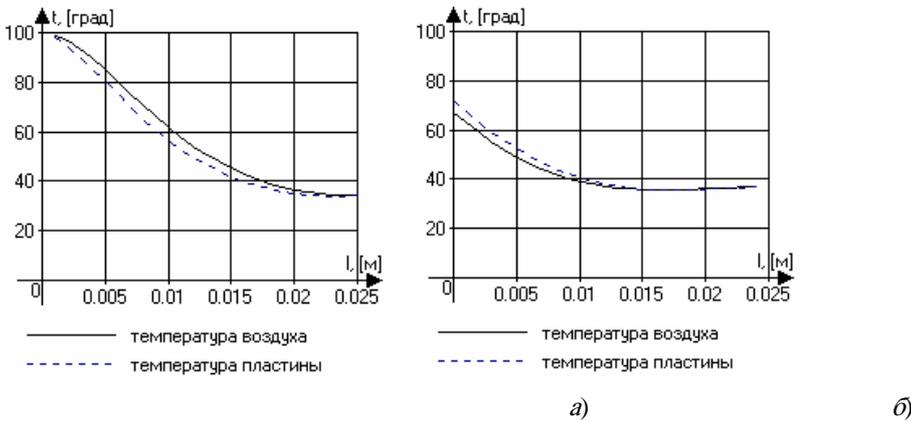


Рис. 3.17. Поле температур РТ ИИДА в третьем цикле дыхания:  
а – при вдохе; б – при выдохе

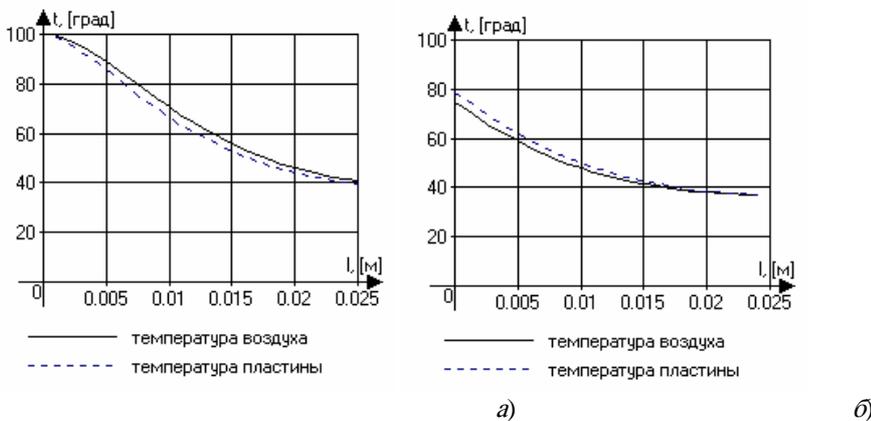
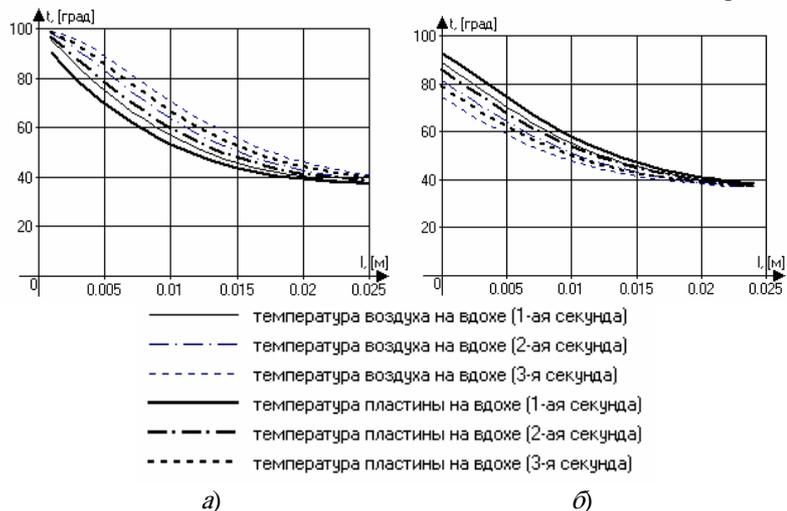


Рис. 3.18. Поле температур РТ ИИДА в пятом цикле дыхания:  
а – при вдохе; б – при выдохе

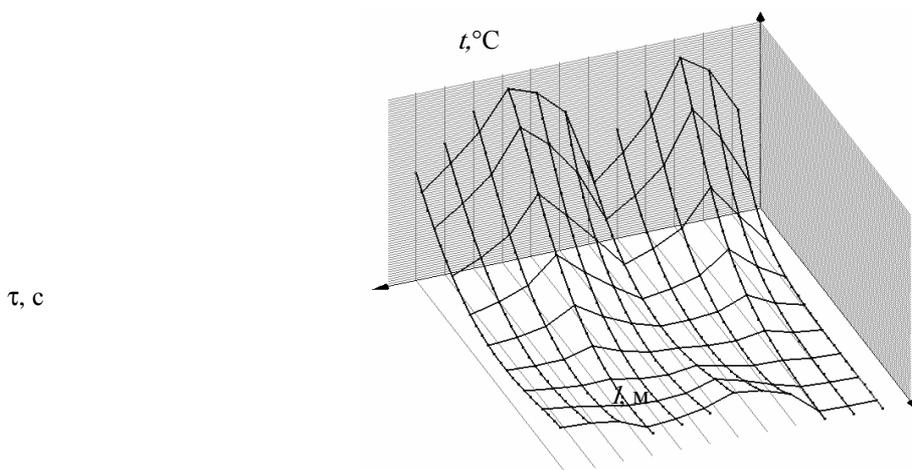
Анализ семейства кривых позволяет оценить количество циклов дыхания, необходимых для выхода системы на стабилизированный циклический режим работы.

Разработанная математическая модель, описывающая тепловые процессы, протекающие в потоке рабочей среды и в конструктивных элементах РТ ИИДА, использована для поиска рациональной конструкции РТ ИИДА.

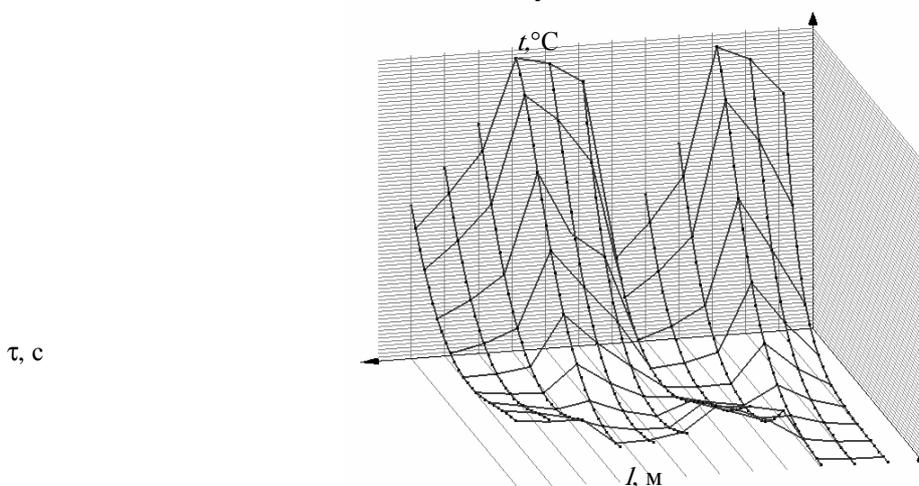
Предложенная программа позволяет исследовать влияние массы алюминиевой ленты на режим работы РТ ИИДА.



**Рис. 3.19. Нестационарное поле температур РТ ИИДА в пятом цикле дыхания:**  
*a* – при вдохе; *б* – при выдохе



**Рис. 3.20. Фрагмент нестационарного поля температур газовоздушного потока, проходящего через РТ ИИДА за первые два цикла дыхания**



**Рис. 3.21. Фрагмент нестационарного поля температур гофрированной ленты РТ ИИДА за первые два цикла дыхания**  
 $L, \text{м}$

В качестве исходных данных для расчета использовались следующие значения геометрических размеров РТ ИИДА и характеристик вдыхаемой газовой смеси:

- длина зоны теплообмена:  $L = 0,025 \text{ м}$ ;

- толщина алюминиевой ленты:  $R = 0,00005$  м;
- начальная температура ленты:  $t_0 = 20,0$  °С;
- начальная температура воздуха:  $tc_1 = 100,0$  °С;
- начальная температура выдоха:  $tc_2 = 37,0$  °С.

Расчеты проводились при условии сохранения одинакового гидравлического сопротивления для различных значений массы конструкционного элемента РТ ИИДА для трех режимов работы.

Результаты работы программы при различных значениях массы алюминиевой ленты представлены в виде сводной таблицы 3.3.

Сравнительный анализ приведенных в табл. 3.3 результатов показывает, что максимальная температура газовой смеси на выдохе снижается с ростом массы алюминиевой ленты. Кроме того, с увеличением массы алюминиевой ленты РТ ИИДА требуется большее число циклов для выхода на стабилизированный циклический режим, т.е. значительно увеличивается время установления стабилизированного циклического режима.

### 3.3. Результаты сравнительных расчетов РТ ИИДА

Сравнительные характеристики	Режим работы	Легкий $V = 20; n = 12;$ $\tau = 5$	Средний $V = 40; n = 20;$ $\tau = 3$	Тяжелый $V = 70; n = 30;$ $\tau = 2$
$m = 0,00960; d = 0,03$				
$K$		13	19	23
$t_{\text{выд}}$		71,8	71,1	71,3
$m = 0,01152 (+ 20\% m); d = 0,031$				
$K$		16	22	31
$t_{\text{выд}}$		67,9	67,3	67,5
$m = 0,0144 (+ 50\% m); d = 0,031$				
$K$		20	27	33
$t_{\text{выд}}$		63,2	62,7	62,9
$m = 0,0192 (+ 100\% m); d = 0,033$				
$K$		29	36	42
$t_{\text{выд}}$		57,8	57,4	57,6

Пр и м е ч а н и я :  $V$  – объемный расход вдыхаемой газовой смеси, л/мин;  $n$  – количество циклов дыхания (цикл дыхания включает в себя вдох и выдох), 1/мин;  $\tau$  – время цикла дыхания, с;  $m$  – масса алюминиевой ленты, г;  $d$  – диаметр регенеративного теплообменника, м;  $t_{\text{выд}}$  – температура газовой смеси на выдохе, °С;  $k$  – число циклов до установления стабилизированного циклического режима, ед.

### Контрольные вопросы

1. В чем различие физической адсорбции и хемосорбции?
2. Что означает термин «изотерма сорбции»?
3. Охарактеризуйте термин «коэффициент конденсации».
4. Какие вещества используются в качестве адсорбентов?
5. Каково влияние температуры на кинетику процесса адсорбции?
6. Что означает термин «теплота адсорбции»?
7. Приведите примеры различных типов фронта сорбции.
8. Приведите пример математического описания процесса массопереноса.
9. Чем отличаются процессы очистки воздуха от процессов регенерации воздуха?
10. Что означают термины «предельно допустимая концентрация» и «максимально допустимая концентрация»?
11. Каков принцип работы фильтров-поглотителей?
12. Чем отличаются статическая и динамическая активности сорбента?
13. Охарактеризуйте термин «время защитного действия».
14. Назовите вещества, наиболее часто применяющиеся в качестве регенеративных продуктов.
15. Приведите пример математического описания процесса теплопроводности.

## 4. МОДЕЛИ ПРОИЗВОДСТВЕННЫХ СИСТЕМ

Производственная система (ПС) или предприятие представляет собой самостоятельную организацию, обладающую материальными, энергетическими, финансовыми и другими ресурсами, которая осуществляет деятельность по производству продукции с целью ее реализации потребителям и извлечения прибыли. Организационно ПС состоит из основных (линейных) и вспомогательных подразделений.

### 4.1. ПРЕДПРИЯТИЕ КАК ОБЪЕКТ МОДЕЛИРОВАНИЯ

Модель ПС представляет собой сложное комплексное многоуровневое описание целей, структуры, материальных, энергетических, информационных потоков, бизнес-процессов и всех других аспектов, необходимых для ее управления. Модель включает словесные описания (миссия, цели), схемы (структура), математические модели (функции затрат, издержек и т.п.) и данные (параметры деталей, характеристики продукции и т.п.). Она должна обеспечивать руководство ПС необходимой

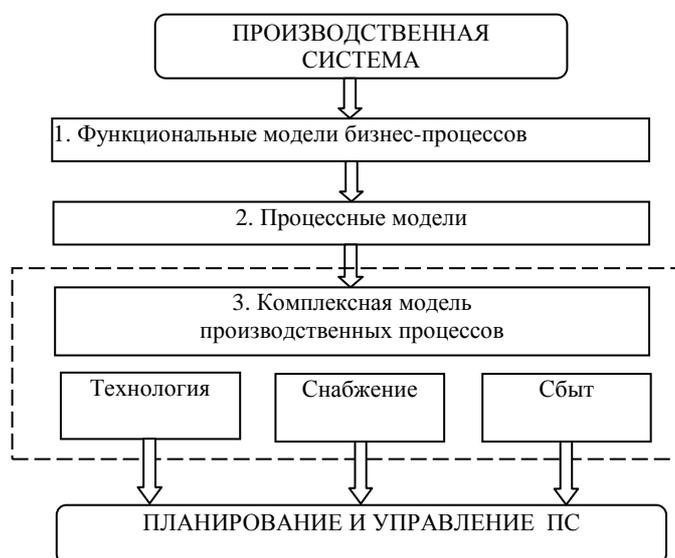
информацией для принятия управленческих решений в достижении поставленных целей. Важнейшими (глобальными) целями любого предприятия являются его устойчивое развитие и конкурентоспособность выпускаемой продукции.

Для описания технологических и организационных процессов или бизнес-процессов в ПС широко используются термины: концептуальная модель, функциональная модель, информационная модель и др. В общем случае бизнес-процесс (БП) представляет собой совокупность последовательно (или параллельно, или последовательно-параллельно) выполняемых операций (действий), в результате которых входные потоки (материальные, информационные) преобразуются в выходные с другими свойствами. В ходе БП потребляются материальные, энергетические, трудовые и финансовые ресурсы, при этом выполняются необходимые ограничения со стороны других БП и внешней среды. Все БП протекают в соответствии с нормативными документами и управляющими директивами, вырабатываемыми на основе целей деятельности предприятия. Для ПС основными БП являются технологические и организационно-деловые.

Для каждого бизнес-процесса разрабатываются качественная и количественная модели, затем на основе интеграции (консолидации) информации о бизнес-процессах получается обобщенное количественное описание предприятия.

Концепция построения модели ПС на основе описания бизнес-процессов представлена на рис. 4.1. На первом этапе строится грубая модель предприятия в виде функциональной модели, которая дает представление о функциях предприятия и о распределении ответственности за их выполнение. На втором этапе строится процессная модель, которая детально описывает потоки ресурсов (материальных, информационных и т.д.), используемых при выполнении функций, а также содержит информацию о взаимосвязях функций. Затем разрабатывается комплексная модель, позволяющая выполнять все расчеты производственного процесса, давать стоимостную оценку всем функциям и решать задачи финансового планирования предприятия. Таким образом, сначала строятся качественные модели БП, затем количественные модели БП, на основе которых строится модель финансового плана ПС.

Основными частями комплексной модели являются описания процессов производства, снабжения материалами и сбыта продукции. В моделях технологических процессов в качестве переменных используются: объемы производства, удельные затраты технологических агрегатов, их мощности, стоимость единицы времени каждого агрегата, платежи за использование агрегатов, в том числе за трудовые ресурсы и др.



**Рис. 4.1. Основные бизнес-процессы производственной системы**

В модели снабжения переменными являются: объемы и стоимость закупаемых материалов, запасы материалов, удельный расход материалов на единицу продукции и т.п.

В качестве переменных модели, описывающей сбыт продукции, используются: объемы реализованной продукции, спрос на каждый вид продукции, запасы готовой продукции, цена реализации продукции и поступления от ее реализации.

В настоящее время модели материальных, энергетических и информационных потоков рассматриваются в рамках логистики, т.е. науки о планировании, управлении и контроле за движением материальных, информационных и финансовых потоков в производственных и других системах.

Модели материальных потоков можно классифицировать по характеру движения потока сырья, продукции и т.п. от поставщика к покупателю. Схемы моделей учитывают разные подходы к формированию запасов.

Схема наиболее распространенной модели приведена на рис. 4.2, а. Эта модель обеспечивает максимальную эффективность производственного процесса, так как поставки сырья на обработку гарантированы и нет необходимости быстро реагировать на изменения покупательского спроса – для этого создаются запасы. Запасы в данном случае страхуют производство от колебаний как поставок, так и спроса. Использование запасов сырья и продукции значительно упрощает управление производственным процессом. Серьезным недостатком модели являются значительные затраты на хранение запасов. Кроме того, данная модель не может применяться в случае изготовления продукции на заказ.

Для сокращения затрат на хранение запасов многие предприятия отказываются от запасов сырья и материалов. При этом поставки материалов организуются так, чтобы они точно соответствовали спросу со стороны производства, схема такой модели приведена на рис. 4.2, б. Предприятие при использовании данной модели находится в большой зависимости от надежной работы поставщика сырья и достаточности у него запасов.

Многие предприятия, изготавливающие продукцию по заказам потребителей, используют модель, схема которой приведена на рис. 4.2, в. Эта модель применима в любых ситуациях, когда покупатели готовы ждать поставки продукции, например, процесс производства очень короткий или готовые изделия нежелательно хранить на складе.

В последние годы стремление сократить затраты на хранение запасов и повысить скорость реакции производства на запросы потребителей привело к появлению стратегии производства под названием «точно – вовремя». Цель данного подхода – нулевые запасы на всех стадиях производственного процесса, схема соответствующей модели приведена на рис. 4.2, г.

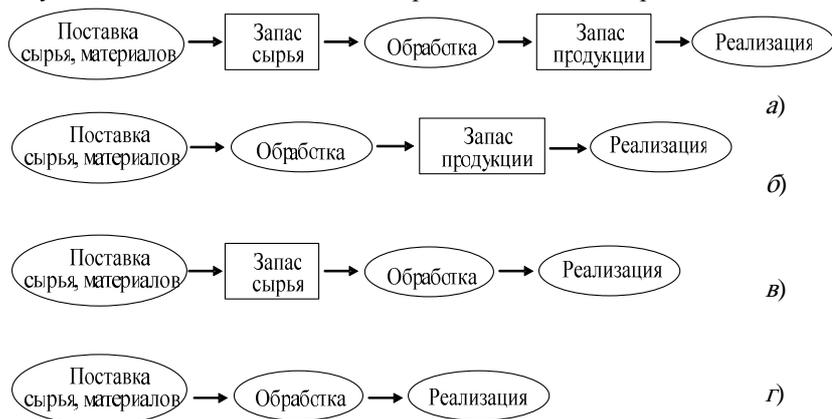


Рис. 4.2

Важное место в комплексной модели производственной системы занимают модели изделий, как конечных в виде заключительной комбинации материалов, предметов, программ и других компонентов, готовых к использованию по назначению, так и компонентов изделия (деталей, сборочных единиц и т.д.). Модель изделия содержит необходимую информацию о его конфигурации, т.е. структуре и составе. Структура представляется в виде описания отношений между компонентами изделия, например, в виде древовидного или сетевого графа, вершинами которого являются компоненты, а ребрами – отношения. Различают функциональные и физические структуры. Функциональная структура формально описывает энергетические и информационные связи между компонентами и обычно представляется блок-схемой. Физическая структура отображает геометрические, электрические, кинематические и другие связи между компонентами.

Состав изделия представляет собой перечень компонентов с указанием их характеристик. Обычно эта часть модели имеет вид иерархически организованного списка, отображающего вхождение деталей в подузлы (модули), подузлы – в узлы (блоки), узлы – в агрегаты (стойки) и т.д.

Все необходимые сведения об изделии для использования в компьютерной среде оформляются в виде электронной модели изделия, т.е. совокупности всех типов электронных моделей (электронная геометрическая модель, электронный чертеж и др.), аннотаций, характеристик, атрибутов, которые вместе определяют и описывают свойства изделия. Понятие «атрибут» в данном случае означает именованную характеристику или параметр изделия (компонента), которая может приобретать конкретное значение в форме числа, вектора, логического значения, символического выражения и т.д. на заданном множестве.

Для успешного решения задач управления предприятием создается интегрированная информационная среда (ИИС) в виде совокупности распределенных баз данных, содержащих сведения об изделиях, производственной среде, ресурсах и процессах предприятия. ИИС должна обеспечивать сохранность, корректность, актуальность и доступность данных персоналу, участвующему в осуществлении жизненного цикла (ЖЦ) изделия. ИИС создается на основе концептуального описания предметной области и объединяет разнородные по природе и форме информации модели, необходимые для управления предприятием.

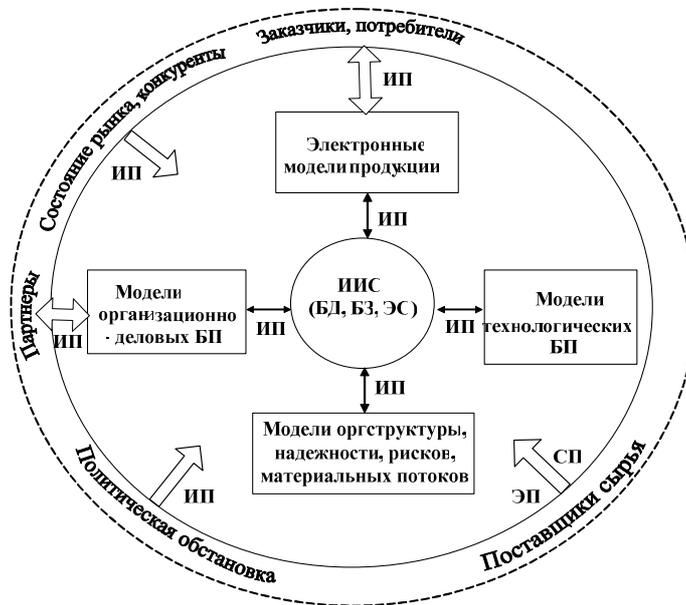
В общем случае под управлением предприятием понимается особый вид бизнес-процесса, в ходе которого определяют цели предприятия, собирается и анализируется информация о ходе производственных процессов, принимаются решения и выполняются действия, необходимые для достижения целей. Множество варьируемых параметров в алгоритмах управления позволяет улучшать динамические свойства предприятия за счет поиска областей эффективных значений в процессе машинных экспериментов с моделями, содержащимися в ИИС.

Модель предметной области, созданная на концептуальном уровне, позволяет глубже проникнуть в процессы, происходящие на каждом из этапов жизненного цикла изделия, и тем самым повысить достоверность использования средств моделирования. Концептуальная модель предметной области создается на основе баз данных (БД), баз знаний (БЗ), экспертных систем (ЭС) и других средств компьютерного моделирования. На рис. 4.3 представлена упрощенная схема модели предметной области производственной системы.

Модели данных, определяющих изделие, представляют собой совокупность информационных объектов и правил их взаимодействия, необходимых для полного его описания, в том числе его геометрии, топологий, свойств и т.д., используемых на всех стадиях ЖЦ. В свою очередь информационный объект рассматривается как совокупность данных и программного кода, которая обладает свойствами и методами, позволяющими определенным образом обрабатывать данные. Понятие «информационная модель» широко используется для описания реальных объектов. Эта модель представляет собой совокупность данных и отношений между ними и предназначена для описания различных свойств реального объекта, которые интересуют разработчика модели и потребителя.

Концептуальная модель должна отражать организационную структуру предприятия. Под структурой обычно понимают совокупность составляющих систему компонентов и устойчивых связей между ними.

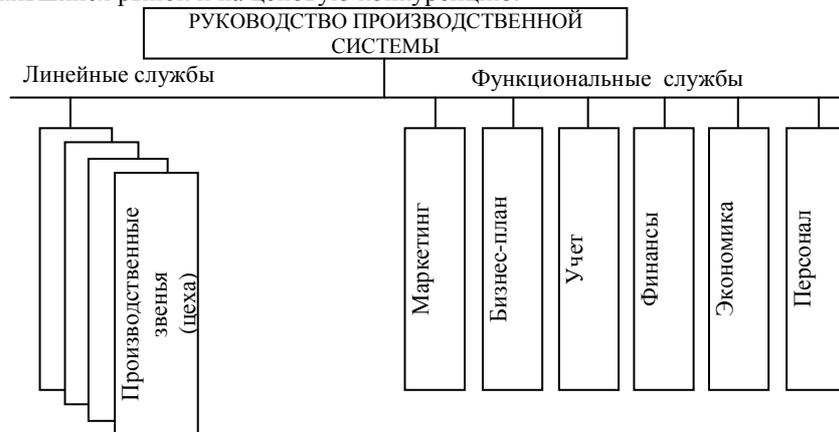
В каждой производственной системе имеется несколько различных структур, в том числе линейная, которая характеризует производственную деятельность; функциональная, объединяющая подразделения с управляющими функциями; структуры центров планирования и учета. Эти структуры в различных вариантах порождают организационные структуры предприятия. Основными видами организационных структур являются: линейно-функциональная, дивизиональная (от division – подразделение) и матричная. Схемы этих структур представлены соответственно на рис. 4.4, 4.5, 4.6.



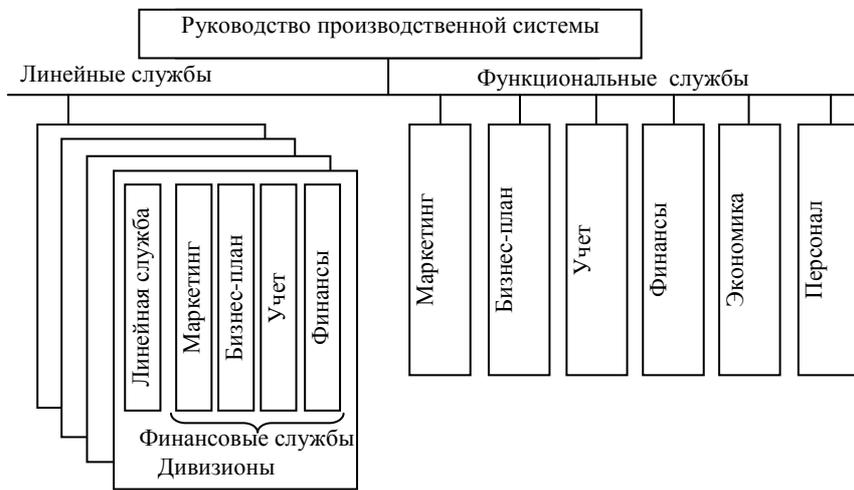
**Рис. 4.3. Концептуальная модель предметной области производственной системы:**

БД – база данных; БЗ – база знаний; БП – бизнес-процессы;  
 ИП – информационные потоки; ИИС – интегрированная информационная среда; СП – сырьевые потоки; ЭП – экономические потоки;  
 ЭС – экстренные системы

Линейно-функциональная и дивизиональная структуры наиболее широко распространены на практике. Основными характеристиками линейно-функциональной структуры являются следующие: эффективность в стабильных условиях рынка; сравнительно низкие затраты на управленческий персонал; быстрое решение простых проблем, находящихся в компетенции одной функциональной службы; специализация и компетентность в решении производственных задач; ориентация на действующие технологии, сложившийся рынок и на ценовую конкуренцию.



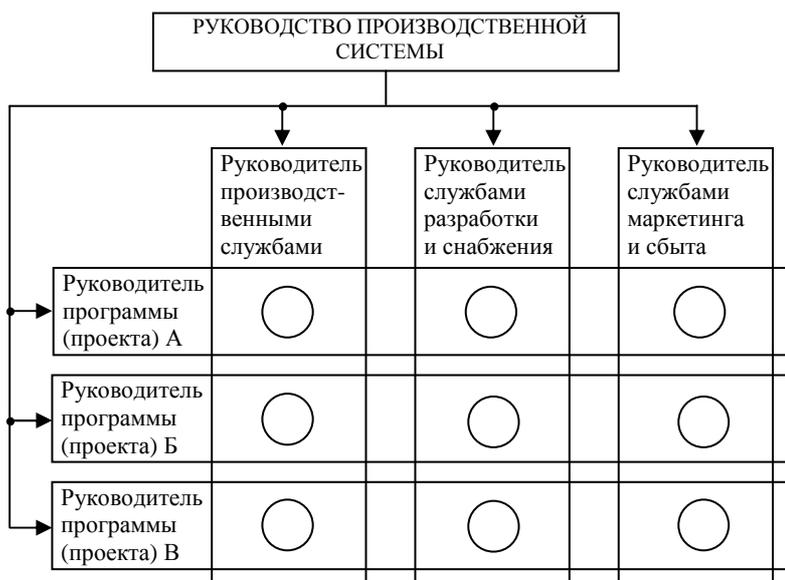
**Рис. 4.4. Схема линейно-функциональной структуры**



**Рис. 4.5. Схема дивизиональной структуры**

Дивизиональная структура более приспособлена к условиям конкурентной среды. Ее характеристиками являются: гибкость, т.е. эффективность в динамичной среде; оперативность в принятии решений по сложным межфункциональным проблемам; ориентация на новые рынки, технологии и на неценовую конкуренцию.

Матричная организационная структура управления представляет собой наиболее современный и эффективный тип структуры (рис. 4.6). Здесь по вертикали осуществляется управление линейными (производственными) и функциональными службами, а по горизонтали – управление программами и проектами (программно-целевой деятельностью). Руководитель соответствующей программы (проекта) работает с виртуальными подразделениями из специалистов различных служб.



**Рис. 4.6. Схема варианта матричной структуры:**  
○ – матричные группы (виртуальные подразделения)

Достоинствами матричной структуры являются: активизация деятельности работников; распределение функций управления между работниками программ и начальниками линейных (функциональных) подразделений; вовлечение руководителей и специалистов в активную творческую деятельность. К недостаткам этой структуры следует отнести следующие: матричные группы не являются устойчивыми образованиями; работники часто перемещаются из одной группы (виртуального подразделения) в другую; имеют место частая смена руководителей и повышенный уровень конфликтности.

Для каждой организационной структуры предприятия имеется своя специфика задач управления и соответствующих моделей. Например, для эффективного функционирования предприятий, имеющих линейно-функциональную или дивизиональную структуру, большое значение имеют задачи планирования и управление запасами. Для предприятий с матричной структурой важную роль играют задачи управления проектами и рисками. Для предприятий с любой структурой, выпускающих наукоемкие изделия, требуется оперативное решение задач управления конфигурацией.

#### 4.2. МОДЕЛИ СИСТЕМ УПРАВЛЕНИЯ ПРЕДПРИЯТИЕМ И ЕГО ЭЛЕМЕНТАМИ

Для эффективного автоматизированного управления предприятием соответствующие модели должны обеспечивать решение следующих задач: планирование сроков выполнения заказов, планирование ресурсов и отслеживание производственных ситуаций; управление на уровне цехов; разработка технологических маршрутов; контроль производства и др.

Схема варианта комплексной модели системы управления предприятием приведена на рис. 4.7. На основе информации о текущем состоянии работ и загрузке производственных мощностей модель позволяет оценить возможность и риски вы-

полнения конкретного заказа в запланированные сроки, определить необходимые для этого ресурсы, получить информацию о загрузке производственных мощностей, разработать сметную калькуляцию и задания участкам.

Компонент модели «Планирование сроков и работ для выполнения заказов» обеспечивает расчет и составление детальных производственных и календарных планов для поступающих заказов. Модель должна



Рис. 4.7. Схема комплексной модели системы управления предприятий

позволять решать задачи оптимизации производственной программы предприятия. Наиболее распространенными являются следующие два класса задач: 1) задачи максимизации целевой функции (критерия) при заданных ограничениях по ресурсам; 2) задачи минимизации ресурсов, необходимых для выполнения производственных заказов.

Все задачи оптимизации, связанные с планированием выполнения заказов, сводятся к тому, чтобы выбрать план из числа альтернативных, который наилучшим образом соответствовал внутренним возможностям предприятия и внешним условиям, при этом прогноз этих возможностей и условий осуществляется с помощью математических моделей.

При решении оптимизационных задач выделяют методический (содержательный), информационный и технический аспекты. Методический аспект связан с формализацией задачи, разработкой требований к модели и ее построением. Информационный и технический аспекты обусловлены проблемами получения необходимых данных о состоянии ПС и компьютерной реализацией модели.

С помощью блока «Отслеживание производственных ситуаций» выполняются: оперативный анализ текущих данных о производственных ситуациях, внесение необходимых изменений в рабочие планы и технологические карты; ведение общего календарного плана и специализированных календарных планов цехов и участков; определение производственной мощности и накладных расходов.

В блоке «Планирование ресурсов» производятся расчеты ресурсов для рабочих мест, графиков загрузки, оборудования на основе текущей информации и прогнозирования ситуаций.

Блок «Технологические маршруты» анализирует информацию, необходимую для составления технологических карт и заданий производственным участкам с учетом фактических ограничений. Технологическая карта определяет прохождение изделия через производственные участки и обрабатывающие центры.

В блоке «Управление на уровне цехов» для каждого производственного заказа обеспечивается выполнение условий о наличии необходимых ресурсов, распределяются работы по участкам.

Параметры производственных операций и нормативы времени для каждого заказа со временем изменяются. Информация об их изменениях вводится в блок «Отслеживание производственных ситуаций», в том числе непосредственно с рабочих мест. Это позволяет получать информацию о текущем состоянии выполнения заказа и о критических моментах в процессе его выполнения. Блоком также отслеживается информация о случаях брака, необходимости доработки, фактической длительности выполнения рабочих операций, об отклонениях производительности работы, оборудования.

Блок «Контроль производства» осуществляет контроль производственных заданий и управление ими в ходе всего технологического цикла. В задачах управления используются результаты моделирования на основе построенной модели.

В общем виде задача планирования выполнения производственного заказа на существующих мощностях предусматривает разработку программ снабжения, производства и финансирования.

Многие производственные задачи управления и планирования решаются методами линейного программирования – симплекс-методом (методом последовательного улучшения плана), транспортным методом, методами целочисленного линейного программирования и др.

Например, симплекс-метод используется для решения следующих задач:

- 1) составление производственных планов с минимальными производственными затратами и с учетом ограничений на различного рода ресурсы;
- 2) определение оптимального состава продукции, в которой составляющие имеют разные стоимости и потребляют различные количества ресурсов;
- 3) определение оптимального маршрута последовательного перемещения продукции в ходе ее обработки от одного обрабатывающего центра к другому с учетом конкретных затрат и производительности оборудования;
- 4) минимизация отходов материалов в процессе раскроя листов или рулонов стали, кожи и т.п.;
- 5) определение оптимальной комбинации различных видов продукции для хранения на складе и др.

Широкое распространение получил также транспортный метод, с помощью которого решаются задачи: нахождения оптимального варианта размещения предприятия, при котором минимизируются затраты на транспортировку грузов с учетом расположения поставщиков и потребителей продукции; определения оптимальных маршрутов движения транспортных средств для перемещения грузов между цехами предприятия с минимальными издержками и др.





Заметим, что методы линейного программирования применяются, если поставлена только одна цель: максимизировать, например, прибыль или минимизировать, например, затраты. Если целей несколько, то используется целевое программирование. В случаях, когда задачу следует решать поэтапно или по временным интервалам, то лучше воспользоваться методом динамического программирования. Для решения более сложных задач могут потребоваться другие методы, например, нелинейное или квадратичное программирование.

#### 4.3. МОДЕЛИ УПРАВЛЕНИЯ ЗАПАСАМИ

Запасы играют важную роль для рациональной и эффективной деятельности производственных предприятий (рис. 4.2). Различают следующие типы запасов.

1. Запасы сырья и материалов, т.е. все закупаемые товары, становящиеся частью выпускаемой продукции.
2. Запасы готовой продукции, в виде товара, находящегося в состоянии запаса и подготовленного для отправки покупателю в любое время. Основная причина создания подобных запасов состоит в том, чтобы разъединить производство и спрос с целью обслуживать непредсказуемый или предсказуемо непостоянный рынок без лишнего изменения объема производственных мощностей.
3. Запасы расходных материалов, которые используются в ходе деятельности организации, но не входят в состав готовой продукции, например, чистящие средства, смазочные материалы и т.д. Этими запасами обычно управляют так же, как запасами сырья и материалов.
4. Запасные части для изделий предприятия или для оборудования обычно рассматриваются как запасы готовой продукции.
5. Запасы полуфабрикатов. На практике всегда существует какое-то незавершенное производство, поскольку для превращения сырья и материалов в готовое изделие требуется время. На промежуточных стадиях производства появляются полуфабрикаты, т.е. продукты, обработка которых уже началась, но еще не завершена в общем производственном цикле. Часть полуфабрикатов свыше необходимого минимума идет для создания промежуточных запасов в виде буфера при наличии «узких мест» или для ускорения выпуска готовых изделий при повышенном спросе.
6. Стратегические запасы, которые создаются из-за возможных изменений в среде поставщиков или политической нестабильности в стране. Их отсутствие может негативно сказаться на будущих поставках или на конкуренции.

Предприятие несет существенные затраты на ведение запасов. К этим затратам относятся затраты на хранение, затраты на приобретение и затраты, связанные с отсутствием запасов (от дефицита запасов). Затраты на хранение включают в себя: стоимость капитала, замороженного в запасах; стоимость хранения, включая занимаемые помещения, оборудование, трудозатраты и т.д.; стоимость потерь из запасов, уровень которых зависит от природы хранимых товаров, но никогда не равен нулю. Затраты на хранение обычно выражаются в процентах от стоимости хранящихся материалов и в большинстве случаев находятся в пределах от 15 до 30 % в год.

Кроме затрат на хранение следует учитывать также затраты на приобретение товаров. Так как помимо стоимости самих товаров их приобретение связано еще с некоторыми дополнительными расходами, в том числе административными, процедурами приемки, расходами на доставку и т.п.

Большое значение для предприятия могут иметь затраты от дефицита, которые обусловлены следующим. При отсутствии запасов производство может остановиться, могут быть приняты шаги к организации экстренной доставки необходимых материалов. В любом случае возникают значительные издержки. Например, конвейер по сборке автомобилей, на котором заканчиваются кузовные детали, обязательно встанет. Эта категория затрат трудно поддается определению.

При рассмотрении задач управления запасами обычно считается, что прием поставок происходит мгновенно и периодически. При одном и том же уровне потребления предприятие может делать закупки материалов редко, но большими партиями или часто, но мелкими. В первом случае сокращаются затраты на приобретение и возрастают издержки на хранение, во втором – наоборот. Для нахождения оптимального соотношения этих статей затрат решается задача определения экономически эффективного размера заказа.

Классификация моделей, используемых в задачах управления запасами, приведена на рис. 4.9. В статических моделях объем спроса на хранимые материалы (продукцию) является постоянным, в динамических моделях он является функцией времени.



Рис. 4.9. Классификация моделей управления запасами

В простейшей детерминированной статической модели управления запасами используются следующие переменные и параметры:  $y(t)$  – уровень запаса в момент времени  $t$ ,  $y_0$  – объем заказа (единиц материала, продукции);  $D$  – интенсивность

спроса или потребления (единиц продукции в единицу времени);  $t_{ц}$  – продолжительность цикла заказа (единиц времени);  $h$  – затраты на хранение (затраты на единицу складированной продукции в единицу времени);  $K$  – затраты на оформление и размещение заказа;  $S$  – суммарные затраты в единицу времени (на приобретение и хранение запаса). Функция  $y(t)$  имеет периодический характер, ее вид показан на рис. 4.10.

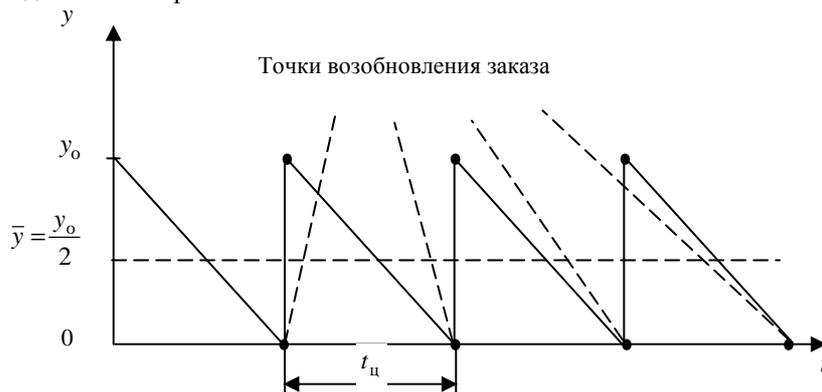


Рис. 4.10. Функции изменения уровня запаса  $y(t)$

Так как средний уровень запаса  $\bar{y} = y_0/2$  и продолжительность цикла заказа  $t_{ц} = y_0/D$ , то суммарные затраты  $S$  рассчитываются по формуле

$$S = \frac{K}{t_{ц}} + h\bar{y} = \frac{KD}{y_0} + \frac{1}{2}hy_0. \quad (4.9)$$

Из выражения (4.9) можно найти оптимальное значение уровня заказа  $y_0^*$ , при котором суммарные затраты минимальны. Дифференцируя функцию  $S(y_0)$  по  $y_0$  (предполагается, что  $y_0$  является непрерывной переменной) и приравняв производную нулю, получаем уравнение для расчета  $y_0^*$ , т.е.

$$\frac{dS(y_0)}{dy_0} = -\frac{KD}{y_0^2} + \frac{1}{2}h = 0,$$

и в результате оптимальные значения  $y_0^*$  и  $t_{ц}^*$  равны

$$y_0^* = \sqrt{\frac{2KD}{h}}, \quad t_{ц}^* = \frac{y_0^*}{D}. \quad (4.10)$$

Таким образом, на основе данной модели оптимальным решением задачи управления запасами будет следующее: заказывать  $y_0^*$  единиц запасаемых материалов через каждые  $t_{ц}^*$  единиц времени. Качественная картина изменения суммарных затрат  $S(y_0)$  и ее составляющих показана на рис. 4.11.

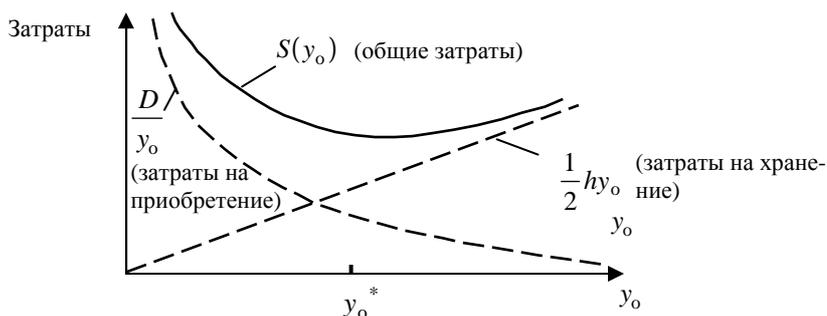


Рис. 4.11. Зависимости изменения затрат на запасы от объема заказа

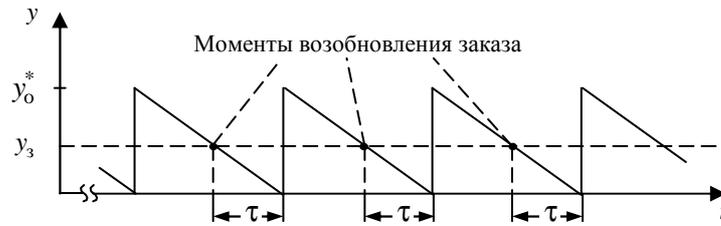
Так как пополнение запаса не может происходить мгновенно в момент оформления заказа, то обычно используют модель управления запасами, учитывающую временное запаздывание от момента заказа до реальной поставки. Обозначим срок выполнения заказа  $\tau$ . В предположении, что  $\tau < t_{ц}^*$ , график изменения уровня запаса и точки возобновления заказа имеют вид, показанный на рис. 4.12.

Если  $\tau > t_{ц}^*$ , то определяется эффективный срок выполнения заказа  $\tau_3$  по формуле

$$\begin{aligned} \tau_3 &= \tau - nt_{ц}^*; \\ n &= E\left[\tau/t_{ц}^*\right], \end{aligned} \quad (4.11)$$

здесь  $E[\tau/t_{ц}^*]$  – целая часть отношения  $\tau/t_{ц}^*$ . В этом случае точки возобновления заказа имеют место при

$$y_3 = \tau_3 D. \quad (4.12)$$



**Рис. 4.12. Изменение уровня запаса и моменты времени возобновления заказа при сроке выполнения заказа  $\tau$**

*Пример 4.2.* Пусть  $D = 100$  ед./день,  $K = 100$  р. за заказ,  $h = 0,02$  р. за хранение единицы в день;  $\tau = 12$  дней. Используя формулы (4.10), получаем

$$y_0^* = \sqrt{\frac{2 \cdot 100 \cdot 100}{0,02}} = 1000 \text{ ед. материала}; \quad t_{ц}^* = \frac{1000}{100} = 10 \text{ дней.}$$

Так как  $\tau > t_{ц}^*$ , то необходимо определить эффективный срок выполнения заказа  $\tau_3$  по формуле (4.11) при

$$n = E[12/10] = 1,$$

т.е.  $\tau_3 = \tau - nt_{ц}^* = 12 - 10 = 2$  дня,

и точка возобновления заказа имеет место при уровне запаса

$$y_3 = \tau_3 D = 2 \cdot 100 = 200 \text{ ед. материала.}$$

Суммарные расходы на содержание запасов при оптимальной стратегии равны

$$S(y_0^*) = \frac{K}{t_{ц}^*} + \frac{1}{2} h y_0^* = \frac{100}{10} + \frac{0,02 \cdot 1000}{2} = 20 \text{ р./день.}$$

Оптимальная стратегия заключается в следующем: заказывать 1000 ед. материала, когда уровень запаса уменьшается до 200 единиц.

Во многих случаях стоимость  $c$  единицы закупаемого материала зависит от величины заказа, обычно эта зависимость имеет вид:

$$c(y_0) = \begin{cases} c_1, & \text{если } y_0 \leq q; \\ c_2, & \text{если } y_0 > q; \quad c_2 < c_1, \end{cases} \quad (4.13)$$

где  $q$  – некоторый пороговый уровень закупки, выше которого материал продается со скидкой (точка «разрыва цены»). Для данного случая модель задачи управления запасами должна учитывать затраты  $S_M$  на приобретение материалов в единицу времени:

$$S_M = \begin{cases} \frac{c_1 y_0}{t_{ц}} = D c_1, & y_0 \leq q; \\ \frac{c_2 y_0}{t_{ц}} = D c_2, & y_0 > q. \end{cases} \quad (4.14)$$

С учетом (4.14) общие затраты в единицу времени определяются соотношением

$$S(y_0) = \begin{cases} S_1(y_0) = D c_1 + \frac{KD}{y_0} + \frac{1}{2} h y_0, & y_0 \leq q; \\ S_2(y_0) = D c_2 + \frac{KD}{y_0} + \frac{1}{2} h y_0, & y_0 > q. \end{cases} \quad (4.15)$$

Как видно из (4.15), функции  $S_1(y_0)$  и  $S_2(y_0)$  отличаются на постоянную величину

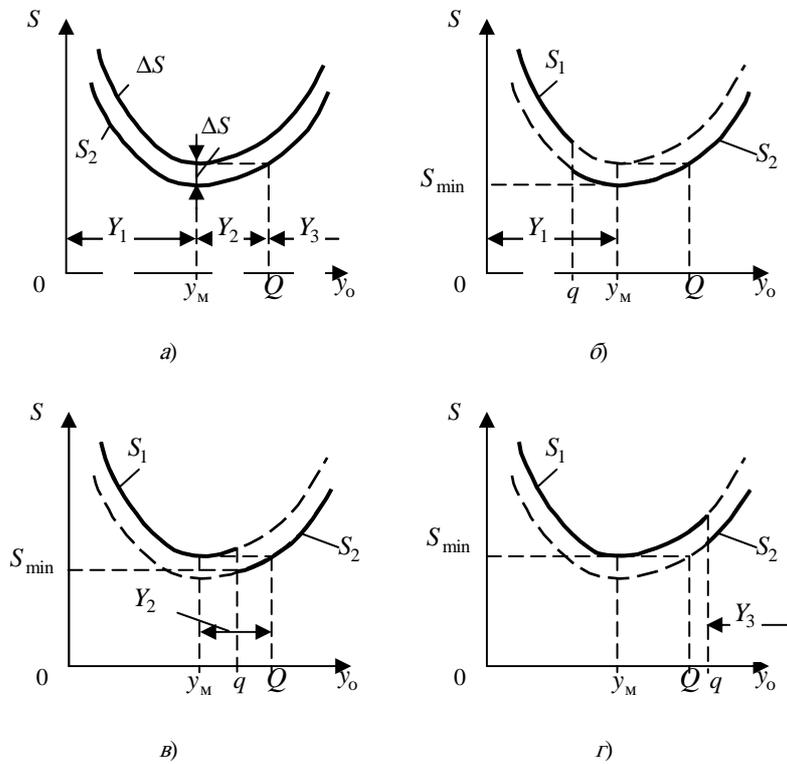
$$\Delta S = S_1(y_0) - S_2(y_0) = D(c_1 - c_2).$$

Поэтому минимумы этих функций совпадают (рис. 4.13, а) и соответствуют значению

$$y_M = \sqrt{\frac{2KD}{h}}. \quad (4.16)$$

Сопоставление функций  $S_1(y_0)$  и  $S_2(y_0)$  показывает, что оптимальное решение  $y_0^*$  зависит от того, в каком интервале значений  $y_0$  находится точка «разрыва цены»  $q$ .

Если  $q \in Y_1 = (0, y_M)$  (рис. 4.13, б), то  $y_0^* = y_M$ .



**Рис. 4.13. Графики функций  $S_1(y_0)$ ,  $S_2(y_0)$  (а) и оптимальные решения для случаев  $q < y_M$ ,  $q \in (y_M, Q)$  (б),  $q > Q$  (г)**

Если  $q \in Y_2 = (y_M, Q)$  (рис. 4.13, в), то  $y_o^* = q$ . Численно значение  $Q$  здесь определяется решением уравнения

$$S_1(y_M) = S_2(Q). \quad (4.17)$$

В случае  $q \in Y_3 = [Q, \infty)$  (рис. 4.13, г) имеет место  $y_o^* = y_M$ , т.е. решение совпадает с решением для первой зоны  $Y_1$ . Таким образом, для оптимального решения можно записать

$$y_o^* = \begin{cases} y_M, & \text{если } q \in (Y_1 \cup Y_3); \\ q, & \text{если } q \in Y_2. \end{cases} \quad (4.18)$$

В рассмотренных моделях предполагалось, что создаются запасы одного вида материалов (продуктов) и дефицит отсутствует. Если требуется хранить  $m$  видов материалов при ограничении на складское помещение, то модель задачи управления запасами существенно усложняется и принимает следующий вид: требуется минимизировать суммарные затраты как функцию  $m$  переменных, т.е.

$$S(y_{o1}, \dots, y_{om}) = \sum_{i=1}^m \left( \frac{K_i D_i}{y_{oi}} + \frac{1}{2} h_i y_{oi} \right) \rightarrow \min_{y_{oi}, i=1, m} \quad (4.19)$$

при ограничениях

$$\sum_{i=1}^m a_i y_{oi} \leq A, \quad (4.20)$$

$$y_{oi} > 0, \quad i = \overline{1, m}, \quad (4.21)$$

здесь  $A$  – максимальная вместимость складских помещений;  $a_i$  – необходимое пространство для хранения  $i$ -го вида материала;  $y_{oi}$ ,  $K_i$ ,  $D_i$ ,  $h_i$  – значения  $y_o$ ,  $K$ ,  $D$ ,  $h$  для материала  $i$ -го вида.

Решение задачи (4.19) – (4.21) с многопродуктовыми запасами решается в следующей последовательности:

1) рассчитываются оптимальные значения  $\tilde{y}_{oi}$ ,  $i = \overline{1, m}$  без учета ограничения (4.20), т.е.

$$\tilde{y}_{oi} = \sqrt{\frac{2K_i D_i}{h_i}}, \quad i = \overline{1, m}; \quad (4.22)$$

2) проверяется выполнение ограничения (4.20) при полученных  $\tilde{y}_{oi}$  по формуле (4.22), если ограничение выполняется, то решение задачи заканчивается и  $y_o^* = \tilde{y}_{oi}$ , если – нет, то переходят к третьему этапу;

3) задача решается с использованием метода множителей Лагранжа, т.е. минимизируется функция

$$L(\lambda, y_{o1}, \dots, y_{om}) = \sum_{i=1}^m \left( \frac{K_i D_i}{y_{oi}} + \frac{1}{2} h_i y_{oi} \right) - \lambda \left( \sum_{i=1}^m a_i y_{oi} - A \right) \rightarrow \min_{\lambda, y_{oi}, i=1, m}, \quad (4.23)$$

где  $\lambda$  – неопределенный множитель Лагранжа,  $\lambda < 0$ .

Значения  $\lambda^*$  и  $y_{oi}^*$ ,  $i = \overline{1, m}$  для (4.23) находятся решением систем уравнений:

$$\frac{\partial L}{\partial \lambda} = - \sum_{i=1}^m a_i y_{oi} + A = 0; \quad (4.24)$$

$$\frac{\partial L}{\partial y_{oi}} = - \frac{K_i D_i}{y_{oi}^2} + \frac{1}{2} h_i - \lambda a_i = 0, \quad i = \overline{1, m}. \quad (4.25)$$

Решая уравнение (4.25), получаем

$$y_{oi}^* = \sqrt{\frac{2K_i D_i}{h_i - 2\lambda^* a_i}}. \quad (4.26)$$

Из формулы (4.26) видно, что последовательно уменьшая множитель  $\lambda$  до тех пор, пока не будет выполняться ограничение (4.20), легко определить  $y_{oi}^*, i = \overline{1, m}$ . Таким образом, задача с  $m$  неизвестными и ограничением решается поиском значения одной переменной  $\lambda$ .

В динамических задачах управления запасами срок на материалы меняется во времени. Эти задачи иногда называются задачами планирования потребностей ресурсов, для их решения обычно используется метод динамического программирования.

В вероятностных моделях управления запасами значение спроса рассматривается как случайная величина с известным законом распределения. В большинстве случаев этот закон считается нормальным. Функция суммарных затрат здесь зависит от плотности распределения спроса в течение срока выполнения заказа и уровня запаса  $R$ , при котором размещается запас. Для определения оптимальных значений  $y_o^*$  и  $R^*$  в данных задачах приходится решать системы нелинейных уравнений.

#### 4.4. МОДЕЛИ ПРИНЯТИЯ РЕШЕНИЙ

Применение современных пакетов, систем и технологий, например, ERP, e-CRM, SCM, XML и других, не снимает полной неопределенности для лица, принимающего окончательное решение, от которого может зависеть успех фирмы или проекта.

Во многих случаях процедура выработки управляющего решения включает следующие этапы: получение информации о состоянии ПС и внешних условиях; идентификация проблемы (задачи) и формулирование цели; формирование множества альтернативных вариантов решения; анализ вариантов с помощью модели выработки решения; определение оптимального (предпочтительного) варианта решения; оценка обоснованности полученного решения и степени соответствия его поставленной цели.

##### 4.4.1. Модель принятия решений с использованием байесовского подхода и экспертных оценок

При оперативном решении ответственных управленческих задач для снижения вероятности ошибок может использоваться модель, представляющая собой комбинацию метода экспертных оценок и байесовского подхода и метода экспертных оценок.

Пусть требуется из множества  $V = \{v_1, v_2, \dots, v_n\}$  вариантов решений, показатели эффективности которых примерно одинаковы, выбрать наиболее целесообразный  $v^*$  для реализации. Обработка результатов работы «узкой» группы экспертов показала, что их мнения не могут быть признаны согласованными (коэффициент конкордации низок) и среди рассматриваемых вариантов нет выделяющегося «лидера».

Идея алгоритма, реализующего данную модель принятия решений, заключается в последовательном привлечении дополнительных экспертов и подсчете для каждого варианта решения  $v \in V$  средней апостериорной вероятности того, что это решение является оптимальным. Работа продолжается до тех пор, пока средняя апостериорная вероятность одного из проектов  $v_a$  множества  $V$  не будет существенно выше, чем для остальных альтернативных вариантов. При соблюдении некоторых условий на возможные исходы последующих экспертиз данный вариант решения  $v_a$  считается оптимальным.

Результат работы каждого дополнительно привлекаемого эксперта рассматривается как исход проведенного опыта, и расчет апостериорной вероятности производится по формуле Байеса, т.е.

$$P(H_i / A_{(j)}) = \frac{P(A_{(j)} / H_i) P(H_i)}{\sum_{i=1}^n P(A_{(j)} / H_i) P(H_i)}, \quad i = \overline{1, n}, \quad (4.27)$$

где  $H_i$  – предположение (гипотеза) о том, что вариант  $v_i$  является оптимальным;  $A_{(j)}$  – результат экспертизы (событие) об оптимальности варианта  $v_j$ ;  $n$  – число рассматриваемых вариантов (мощность множества  $V$ );  $P(H_i), P(H_i / A_{(j)})$  – априорная и апостериорная вероятности гипотезы  $H_i$  соответственно;  $P(A_{(j)} / H_i)$  – вероятность события  $A_{(j)}$ , если имеет место гипотеза  $H_i$  (правдоподобие).

**БУДЕМ ПОЛАГАТЬ, ЧТО СОБЫТИЕ  $A_j$  ПРОИЗОШЛО, ЕСЛИ ВАРИАНТ  $v_j$  ОЧЕРЕДНОЙ ЭКСПЕРТ РАСПОЛОЖИЛ НА 1-Е МЕСТО ПРИ  $n = 2$  ИЛИ 3, И НА 1-Е ИЛИ 2-Е МЕСТО ПРИ  $n > 3$ .**

Если произошло событие  $\bar{A}_{(j)}$ , то апостериорная вероятность  $P(H_i / \bar{A}_{(j)})$  рассчитывается по формуле, аналогичной (4.27), т.е.

$$P(H_i / \bar{A}_{(j)}) = \frac{P(\bar{A}_{(j)} / H_i)P(H_i)}{\sum_{i=1}^n P(\bar{A}_{(j)} / H_i)P(H_i)}, \quad (4.28)$$

где  $P(H_i / \bar{A}_{(j)})$  – апостериорная вероятность гипотезы  $H_i$  при событии  $\bar{A}_{(j)}$ .

По результатам работы очередного  $k$ -го эксперта рассчитываются усредненные апостериорные вероятности по формуле

$$\bar{P}_k(H_i / \mathcal{A}) = \frac{1}{n} \sum_{j=1}^n P(H_i^k / \tilde{A}_{(j)}), \quad i, j = \overline{1, n} \quad (4.29)$$

$$A = \{\tilde{A}_{(j)}, j = \overline{1, n}\},$$

где  $\tilde{A}_{(j)}$  – событие, связанное с проверкой гипотезы  $H_j^k$ , т.е. того, что  $k$ -й эксперт вариант  $v_j$  поставит на первые места, для части слагаемых суммы имеет место  $A_{(j)}$ , для другой –  $\bar{A}_{(j)}$ .

Вероятности  $P(H_i)$ ,  $P(H_i / A_{(j)})$ ,  $P(H_i / \bar{A}_{(j)})$ ,  $\bar{P}_k(H_i / A)$  естественно удовлетворяют условию полноты группы событий, т.е.

$$\sum_{i=1}^n P(H_i) = 1, \sum_{i=1}^n P(H_i / A_{(j)}) = 1, \sum_{i=1}^n P(H_i / \bar{A}_{(j)}) = 1, \sum_{i=1}^n \bar{P}_k(H_i / \mathcal{A}) = 1$$

и

$$P(A_{(j)} / H_i) + P(\bar{A}_{(j)} / H_i) = 1, \quad i, j = \overline{1, n}.$$

В качестве оптимального варианта  $v^*$  после  $k$ -й экспертизы берется тот, для которого вероятность, рассчитанная по формуле (4.29), максимальна и выполняется условие, что некоторое наперед заданное число  $m$  последующих экспертиз не изменяет соотношения

$$\bar{P}_{k+m}(H(v^*) / A) = \max_{v_i \in V} \{\bar{P}_{k+m}(H(v_i) / A)\}, \quad (4.30)$$

где  $H(v^*)$  – гипотеза об оптимальности варианта  $v^*$ ,  $H(v_i) = H_i$ .

При использовании байесовского подхода для решения подобных задач важную роль играет формализация правила «остановки» в процессе проведения экспертиз. С одной стороны, своевременное прекращение итераций экономит средства, затрачиваемые на проведение экспертиз. С другой стороны, необходима уверенность, что дальнейшее привлечение экспертов не приведет к кардинальному изменению усредненной апостериорной вероятности и принятию другого варианта решения для реализации.

Наиболее естественно решение об «остановке» принимать по двум показателям: числе  $m$  дополнительных экспертов, высказывания которых могут изменить выбор оптимального варианта, и вероятности  $P_m$  того, что результаты высказываний этих экспертов приведут к изменению варианта, т.е. гипотезы, для которой усредненная апостериорная вероятность максимальна.

Определение показателей  $m$  и  $P_m$  производится при следующих допущениях:

- 1) в множестве  $V$  можно выделить два лидирующих варианта  $v_a$  и  $v_b$ ;
- 2) проведена обработка мнений  $k$  экспертов, при этом варианту  $v_a$  отдавалось предпочтение (исход  $A$ )  $k_a$  раз ( $k_a \leq k$ ) и варианту  $v_b$  (исход  $B$ ) –  $k_b$  раз ( $k_b < k_a$ ), т.е. по результатам  $k$  итераций вариант  $k_a$  считается предпочтительным (вероятность  $\bar{P}_k(H(v_a) / A)$  – максимальна);
- 3) в качестве вероятностей исходов  $A$  и  $B$  принимаются оценки

$$P_a = \frac{k_a}{k}; \quad P_b = \frac{k_b}{k}, \quad (4.31)$$

причем вероятность  $P_a > 0,5$ ;

- 4) исходы  $A$  и  $B$  при последующих высказываниях экспертов являются независимыми и совместимыми;
- 5) очередность исходов в  $m$  экспертизах не влияет на конечный результат.

При данных допущениях имеют место следующие неравенства. Если

$$\bar{P}_k(H(v_a) / A) > \bar{P}_k(H(v_b) / A) \quad \text{и} \quad K_a > K_b,$$

то соотношение

$$\bar{P}_{k+m}(H(v_a) / A) < \bar{P}_{k+m}(H(v_b) / A) \quad (4.32)$$

становится возможным при

$$m \geq (k_a - k_b) + 1. \quad (4.33)$$

Для определения вероятности  $P_m(b)$ , характеризующей возможность неравенства (4.32), можно использовать модель Бернулли для повторяющихся испытаний. В этом случае если имеет место

$$\bar{P}_k(H(v_a)/A) > \bar{P}_k(H(v_b)/A), \quad k_a > k_b$$

и  $m \geq 2$  (4.33), то вероятность выполнения неравенства (4.32) при минимальном значении  $m$  определяется формулой

$$P_m(b) = (1 - P_a)^m P_b^m. \quad (4.34)$$

Равенство (4.34) означает, что все  $m$  привлекаемых дополнительно экспертов выскажутся отрицательно относительно варианта  $v_a$  (исходы  $\bar{A}$ ) и положительно относительно  $v_b$  (исходы  $B$ ). Формула (4.34) непосредственно следует из распределения вероятностей возможных сложных событий при  $m$  испытаниях, в которых события  $A$  и  $B$  могут принимать по два исхода с разными вероятностями. Такое распределение при использовании моделей Бернулли для событий  $A$  и  $B$  имеет следующий вид:

$$P_m(a, b) = \left( \sum_{v=0}^m C_m^v P_a^v (1 - P_a)^{m-v} \right) \left( \sum_{v=0}^m C_m^v P_b^v (1 - P_b)^{m-v} \right), \quad (4.35)$$

где

$$C_m^v = \frac{m!}{v!(m-v)!}, \quad C_m^m = 1, \quad C_m^0 = 1.$$

Следует заметить, что вероятности  $P_a, P_b$  (4.31) необходимо корректировать после каждой итерации.

Продемонстрируем совместное использование метода экспертных оценок и байесовского подхода на численном примере.

*Пример 4.3.* Пусть из множества проектов  $\mathcal{V} = \{v_1, \dots, v_7\}$  предварительной экспертизой выделено подмножество предпочтительных  $V^n = \{v_5, v_7\}$  решений. Требуется, последовательно привлекая дополнительных экспертов, определить один вариант  $v^*$  для реализации, имеющий максимальную усредненную апостериорную вероятность и удовлетворяющий условию (4.30) при  $m = 2$ . Зададим следующие начальные (априорные) вероятности гипотез:

$$P(H_5^0) = P(H_7^0) = 0,25; \quad P(H_i^0) = 0,1; \quad i = \bar{1}, \bar{4}, \bar{6}. \quad (4.36)$$

Пусть событие  $A_{(5)}$  заключается в том, что очередной эксперт поставил рассматриваемый проект  $v_5$  на 1-е или 2-е места и

$$P(A_{(5)} / H_5^0) = 0,8, \quad P(A_{(5)} / H_i^0, i \neq 5) = 0,6. \quad (4.37)$$

Результаты работы очередного эксперта (эксперт 1) приведены в табл. 4.1. Из таблицы видно, что эксперт 1 поставил вариант  $v_5$  на 3-е место, т.е. произошло событие  $\bar{A}_{(5)}$ , противоположное событию  $A_{(5)}$  и  $P(\bar{A}_{(5)} / H_5) = 1 - P(A_{(5)} / H_5) = 0,2$ .

#### 4.1. Результаты работы первого эксперта

Варианты (проекты)	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$
Ранги эксперта 1	1	3	2	3	3	1	3
События	$A_{(1)}$	$\bar{A}_{(2)}$	$A_{(3)}$	$\bar{A}_{(4)}$	$\bar{A}_{(5)}$	$A_{(6)}$	$\bar{A}_{(7)}$

Расчет апостериорной вероятности гипотезы  $H_5^1$  производится по формуле (4.28), т.е.

$$P(H_5^1 / \bar{A}_{(5)}) = \frac{P(\bar{A}_{(5)} / H_5) P(H_5^0)}{\sum_{i=1}^7 P(\bar{A}_{(5)} / H_i) P(H_i^0)} \approx 0,143.$$

Верхний индекс 1 в  $P(H_5^1 / \bar{A}_{(5)})$  указывает на результат, полученный после высказываний первым экспертом (результат 1-й итерации при использовании формулы Байеса).

Апостериорные вероятности для других гипотез соответственно равны

$$P(H_7^1 / \bar{A}_{(5)}) = \frac{P(\bar{A}_{(5)} / H_7) P(H_7^0)}{\sum_{i=1}^7 P(\bar{A}_{(5)} / H_i) P(H_i^0)} \approx 0,286.$$

$$P(H_i^1 / \bar{A}_{(5)}) = \frac{P(\bar{A}_{(5)} / H_i) P(H_i^0)}{\sum_{i=1}^7 P(\bar{A}_{(5)} / H_i) P(H_i^0)} \approx 0,114, \quad i = 1, 2, 3, 4, 6.$$

Предположим, что событие  $A_{(7)}$  характеризует оптимальность варианта  $v_7$ . В нашем случае имеет место  $\bar{A}_{(7)}$  (табл. 4.1) и при правдоподобиях, аналогичных (4.37), т.е.

$$P(A_{(7)} / H_5) = 0,8, \quad P(A_{(7)} / H_i, i \neq 7) = 0,6,$$

апостериорные вероятности равны:

$$P(H_7 / \bar{A}_{(7)}) = \frac{P(\bar{A}_{(7)} / H_7) P(H_7^0)}{\sum_{i=1}^7 P(\bar{A}_{(7)} / H_i) P(H_i^0)} \approx 0,143;$$

$$P(H_5 / \bar{A}_7) \approx 0,286; \quad P(H_i / \bar{A}_7) \approx 0,114, \quad i = 1, 2, 3, 4, 6.$$

В целях большей достоверности результатов следует рассмотреть и другие гипотезы об оптимальности вариантов. Рассмотрим их схематично.

Событие  $A_{(1)}$  характеризует оптимальность варианта  $v_1$  и при  $P(A_{(1)} / H_1) = 0,8, \quad P(A_{(1)} / H_i, i \neq 1) = 0,6,$

$$P(H_1^1 / A_{(1)}) = \frac{P(A_{(1)} / H_1) P(H_1^0)}{\sum_{i=1}^7 P(A_{(1)} / H_i) P(H_i^0)} \approx 0,129;$$

$$P(H_5^1 / A_{(1)}) = P(H_7^1 / A_{(1)}) = 0,242;$$

$$P(H_2^1 / A_{(1)}) = P(H_3^1 / A_{(1)}) = P(H_4^1 / A_{(1)}) = P(H_6^1 / A_{(1)}) \approx 0,097.$$

Аналогично выполняются расчеты для событий  $A_{(j)}, \quad j = 2, 3, 4$ . Результаты расчетов представлены в табл. 4.2.

В нижней строке таблицы приведены усредненные апостериорные вероятности, рассчитанные по формуле

$$\bar{P}_1(H_i / \mathcal{A}) = \frac{1}{7} \sum_{i=1}^7 P(H_i^1 / A_{(i)}), \quad \mathcal{A} = \{A_{(i)}, i = \bar{1}, \bar{7}\}.$$

Сравнение их с априорными вероятностями гипотез  $P_0(H_i)$  показывает, что средние апостериорные вероятности изменились незначительно, причем вероятности гипотез об оптимальности  $v_5$  и  $v_7$  уменьшились и возросли вероятности для вариантов  $v_1, v_3, v_6$ . Таким образом, высказываний эксперта на первой итерации оказалось недостаточно для принятия решения.

#### 4.2. Апостериорные вероятности на первой итерации

Вероятности	Гипотезы						
	$H_1$	$H_2$	$H_3$	$H_4$	$H_5$	$H_6$	$H_7$
$P(H_i^0)$	0,1	0,1	0,1	0,1	0,25	0,1	0,25
$P(H_i^1 / A_{(1)})$	0,129	0,097	0,097	0,097	0,242	0,097	0,242
$P(H_i^1 / \bar{A}_{(2)})$	0,105	0,053	0,105	0,105	0,263	0,105	0,263
$P(H_i^1 / A_{(3)})$	0,097	0,097	0,129	0,097	0,242	0,097	0,242

$P(H_i^1 / \bar{A}_{(4)})$	0,105	0,105	0,105	0,053	0,263	0,105	0,263
$P(H_i^1 / \bar{A}_{(5)})$	0,114	0,114	0,114	0,114	0,143	0,114	0,286
$P(H_i^1 / A_{(6)})$	0,097	0,097	0,097	0,097	0,242	0,129	0,242
$P(H_i^1 / \bar{A}_{(7)})$	0,114	0,114	0,114	0,114	0,286	0,114	0,143
$\bar{P}_1(H_i / \mathcal{A})$	0,109	0,097	0,109	0,097	0,24	0,109	0,24

Результаты работы эксперта 2 (на второй итерации) представлены в табл. 4.3.

#### 4.3. Результаты работы второго эксперта

Варианты	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$
Ранги эксперта 2	3	3	4	1	1	5	2
События	$\bar{A}_{(1)}$	$\bar{A}_{(2)}$	$\bar{A}_{(3)}$	$A_{(4)}$	$A_{(5)}$	$\bar{A}_{(6)}$	$A_{(7)}$

Используя в качестве априорных вероятностей результаты предыдущего этапа и правдоподобия (4.37) для события  $A_{(5)}$  (вариант  $v_5$  имеет ранг, равный 1), получаем

$$P(H_5^2 / A_{(5)}) = \frac{P(A_{(5)} / H_5) P(H_5^1 / \bar{A}_{(5)})}{\sum_{i=1}^7 P(A_{(5)} / H_i) P(H_i^1 / \bar{A}_{(5)})} = 0,182; \quad (4.38)$$

$$P(H_7^2 / A_{(5)}) = \frac{P(A_{(5)} / H_7) P(H_7^1 / \bar{A}_{(5)})}{\sum_{i=1}^7 P(A_{(5)} / H_i) P(H_i^1 / \bar{A}_{(5)})} \approx 0,273; \quad (4.39)$$

$$P(H_i^2 / A_{(5)}) \approx 0,109, \quad i = 1, 2, 3, 4, 6.$$

Аналогично рассчитываются апостериорные вероятности для событий  $A_{(j)}$ ,  $j = 1, 2, 3, 4, 6, 7$ . Результаты расчетов по высказываниям второго эксперта представлены в табл. 4.4. Из таблицы видно, что вероятности  $\bar{P}_2(H_i / \mathcal{A})$  близки к априорным, поэтому требуется привлечение еще одного эксперта.

Результаты высказываний эксперта 3 представлены в табл. 4.5.

При расчете апостериорных вероятностей здесь в качестве априорных используются значения  $P_2(H_i) = P(H_i^2 / A_{(j)})$ , взятые из табл. 4.4.

#### 4.4. Апостериорные вероятности на второй итерации

Вероятности	Гипотезы						
	$H_1$	$H_2$	$H_3$	$H_4$	$H_5$	$H_6$	$H_7$
$P(H_i^2 / \bar{A}_{(1)})$	0,069	0,104	0,104	0,104	0,259	0,104	0,259
$P(H_i^2 / \bar{A}_{(2)})$	0,108	0,027	0,108	0,108	0,27	0,108	0,27
$P(H_i^2 / \bar{A}_{(3)})$	0,104	0,104	0,069	0,104	0,259	0,104	0,259
$P(H_i^2 / A_{(4)})$	0,103	0,103	0,103	0,069	0,258	0,103	0,258

$P(H_i^2 / A_{(5)})$	0,109	0,109	0,109	0,109	0,182	0,109	0,273
$P(H_i^2 / \bar{A}_{(6)})$	0,104	0,104	0,104	0,104	0,259	0,069	0,259
$P(H_i^2 / A_{(7)})$	0,109	0,109	0,109	0,109	0,273	0,109	0,182
$\sum_j P(H_i^2 / A_{(j)})$	0,706	0,66	0,706	0,706	1,76	0,706	1,76
$\bar{P}_2(H_i / \mathcal{A})$	0,101	0,094	0,101	0,101	0,251	0,101	0,251

#### 4.5. Результаты работы третьего эксперта

Варианты	$\nu_1$	$\nu_2$	$\nu_3$	$\nu_4$	$\nu_5$	$\nu_6$	$\nu_7$
Ранги эксперта 3	5	2	3	4	1	4	3
События	$A_{(1)}$	$\bar{A}_{(2)}$	$A_{(3)}$	$\bar{A}_{(4)}$	$\bar{A}_{(5)}$	$A_{(6)}$	$\bar{A}_{(7)}$

Для правдоподобия (4.37) рассчитанные значения апостериорных вероятностей и усредненные вероятности приведены в табл. 4.6.

#### 4.6. Апостериорные вероятности на третьей итерации

Вероятности	Гипотезы						
	$H_1$	$H_2$	$H_3$	$H_4$	$H_5$	$H_6$	$H_7$
$P(H_i^3 / \bar{A}_{(1)})$	0,36	0,108	0,108	0,108	0,268	0,108	0,268
$P(H_i^3 / A_{(2)})$	0,107	0,036	0,107	0,107	0,268	0,107	0,268
$P(H_i^3 / \bar{A}_{(3)})$	0,108	0,108	0,036	0,108	0,268	0,108	0,268
$P(H_i^3 / \bar{A}_{(4)})$	0,107	0,107	0,107	0,036	0,267	0,107	0,267
$P(H_i^3 / A_{(5)})$	0,103	0,103	0,103	0,103	0,229	0,103	0,257
$P(H_i^3 / \bar{A}_{(6)})$	0,108	0,108	0,108	0,108	0,268	0,036	0,268
$P(H_i^3 / \bar{A}_{(7)})$	0,12	0,12	0,12	0,12	0,3	0,12	0,1
$\bar{P}_3(H_i / \mathcal{A})$	0,098	0,099	0,099	0,099	0,267	0,098	0,242

Таким образом, после высказываний третьего эксперта максимальное значение средней вероятности соответствует гипотезе  $H_5$  ( $\bar{P}_3(H_5 / \mathcal{A}) = 0,267$ ), в качестве оптимального варианта следует принять  $\nu_5$ .

Рассматривая в качестве  $\nu_a$  вариант  $\nu_5$  и в качестве  $\nu_b$  –  $\nu_6$ , при  $k = 3$ ,  $k_a = 2$ ,  $k_b = 1$  на основе формулы (4.33) получаем  $m = 2$ , а согласно (4.31)  $P_a = 2/3$ ,  $P_b = 1/3$ . Для этих значений неравенство

$$\bar{P}_{3+2}(H(\nu_5) / \mathcal{A}) > \bar{P}_{3+2}(H(\nu_7) / \mathcal{A})$$

с вероятностью

$$P_2(a) = 1 - P_2(b) - P_2(a, b),$$

где  $P_2(a, b)$  – вероятность того, что при  $m = 2$  средние апостериорные вероятности для вариантов  $\nu_a$  и  $\nu_b$  примерно сравняются.

Используя формулы (4.34), (4.35), получаем

$$P_2(b) = (1 - P_a)^2 P_b^2 \approx 0,012;$$

$$P_2(a, b) = 2P_a(1 - P_a)P_b^2 \approx 0,036$$

$$P_2(a) \approx 0,95,$$

т.е. дополнительное привлечение двух экспертов с вероятностью 0,95 не изменит «лидерства» проекта  $\cup_5$ , поэтому его можно считать оптимальным, и больше экспертов не привлекать.

Следует заметить, что при обработке табл. 4.1, 4.3, 4.5 обычным способом коэффициент конкордации имеет очень низкое значение (0,094) и естественно мнение экспертов о всех вариантах считаются не согласованными (оценка критерия «хи-квадрат» 1,69, а табличное 12,59). Вместе с тем байесовский подход позволяет сделать достаточно надежные выводы о предпочтительном варианте.

К достоинствам рассмотренного метода можно отнести следующие.

1. Метод позволяет формализовать задачу определения числа привлекаемых экспертов.
2. Расчет средних апостериорных вероятностей дает возможность принимать обоснованные решения, когда мнения экспертов относительно всего множества вариантов считаются несогласованными.
3. Расчет апостериорных вероятностей на каждой итерации и прогнозирование вероятностей  $P_m(b)$  позволяют исключить из рассмотрения заведомо неперспективные варианты.
4. Метод удобен для оперативного принятия решений при работе с экспертами в режиме удаленного доступа (через Internet), когда ответы экспертов поступают не одновременно.

#### 4.4.2. Модель Шортлифа-Бьюкенена

Использование формулы Байеса требует знаний априорных и условных вероятностей, для оценки которых необходимы статистические данные. При этом встречаются следующие трудности: большая трудоемкость получения представительной выборки, особенно в случае многомерных распределений; сложность принятия решений в условиях редко повторяющихся ситуаций, наблюдение за которыми требует длительного времени; возможны изменения характера распределений и взаимосвязей между данными и ситуациями со временем, особенно для экономических показателей развивающихся предприятий и др. В этих случаях предпочтительно использование модель Шортлифа и Бьюкенена (МШБ) или Стендфорской теории фактора уверенности, которая позволяет делать оперативные выводы и принимать решения на основе неполных знаний. Для этого вместо сбора представительной выборки собираются и обрабатываются мнения экспертов и лица, принимающие решения (ЛПР), которые затем интерпретируются в вероятностном смысле.

Преимущества МШБ по сравнению с системой условных вероятностей, применяемых при байесовском подходе, заключаются в следующем:

- используются теоретические знания и накопленный практический опыт экспертов в оценке проблемных ситуаций, для которых не накоплен достаточный статистический материал;
- упрощается модификация алгоритма решения, так как продукционные правила не связаны строгой очередностью одно с другим;
- возможны коррекция и добавление новых правил;
- облегчается поиск потенциальных конфликтов и несовместимостей в базе знаний;
- используются простые механизмы объяснений вычислительного процесса.

Важную роль в МШБ играют понятия мер уверенности и неуверенности.

Мера (measure) уверенности или доверия (belief)  $MB$  в соответствии с равенством  $MB[h, x] = \alpha$  означает, что степень или мера уверенности в некоторой гипотезе  $h$ , основанная на свидетельстве  $x$ , есть  $\alpha$ . Гипотеза  $h$  может заключаться в предпочтительности одного из альтернативных вариантов  $v$  решения.

Мера  $MB$  рассматривается как неформальная оценка, которую эксперт (или ЛПР) добавляет к заключениям типа «вероятно, это так», «почти наверняка, это так» и т.п.

Мера неуверенности или недоверия (distrust)  $MD$  в соответствии с равенством  $MD[h, x] = \beta$  означает, что степень или мера неуверенности в гипотезе  $h$ , основанная на свидетельстве  $x$ , есть  $\beta$ .

Стендфорская теория фактора уверенности основывается на следующих предположениях. Во-первых, в методах, использующих классические положения теории вероятности, при оценке экспертом истинности некоторого отношения (например, значением 0,8) не учитывается, что отношение может быть и ложным. Здесь правило равенства единице суммы вероятностей отношения и его отрицания не распространяется на все ситуации. Во-вторых, во многих случаях при абдуктивном выводе «знание самих правил намного важнее, чем знание алгебры для вычисления их достоверности».

Абдуктивные правила вывода (от частного к частному, по аналогии) не являются строгими, здесь заключения не обязательно истинны для каждой интерпретации, при которой истинны предпосылки.

Значения  $MB$  и  $MD$ , как и для вероятности, всегда должны находиться в интервале  $[0, 1]$ . Свидетельства могут быть не только наблюдаемыми, но и гипотезами. Например,  $MB[h_1, h_2]$  есть мера увеличения уверенности в гипотезе  $h_1$  при условии, что гипотеза  $h_2$  является истинной.

Одно и то же свидетельство  $x$  не может выступать как в пользу, так и против гипотезы, т.е.

$$\text{если } MB[h, x] > 0, \text{ то } MD[h, x] = 0; \quad (4.40)$$

$$\text{если } MD[h, x] > 0, \text{ то } MB[h, x] = 0. \quad (4.41)$$

Если гипотеза  $h$  не зависит от свидетельства  $x$ , т.е. условная вероятность  $P(h/x)$  равна априорной вероятности  $P(h)$ , то

$$MB[h, x] = MD[h, x]. \quad (4.42)$$

Определение  $MB$  и  $MD$  производится с использованием соотношений

$$MB[h, x] = \begin{cases} 1, & \text{если } p(h)=1; \\ \frac{\max\{p(h/x), p(h)\} - p(h)}{1 - p(h)}, & \text{если } p(h) < 1; \end{cases} \quad (4.43)$$

$$MD[h, x] = \begin{cases} 1, & \text{если } p(h)=0; \\ \frac{\min\{p(h/x), p(h)\} - p(h)}{-p(h)}, & \text{если } p(h) > 0, \end{cases} \quad (4.44)$$

где  $p(h)$  – априорная вероятность гипотезы  $h$ ,  $p(h/x)$  – условная вероятность  $h$  при свидетельстве  $x$ .

Вероятность  $p(h)$  отражает уверенность эксперта в  $h$  в любой момент времени, а  $1 - p(h)$  – оценка неуверенности эксперта в истинности  $h$ . Если  $p(h/x) > p(h)$ , то свидетельства  $x$  увеличивает уверенность эксперта в  $h$ . Если  $p(h/x) < p(h)$ , то  $x$  уменьшает уверенность в  $h$  (и увеличивает неуверенность в истинности  $h$ ).

Для расчета  $MB$  и  $MD$  допускается использование упрощенных формул

$$MB[h, x] = \frac{p(h/x) - p(h)}{p(h)}, \text{ если } p(h/x) > p(h); \quad (4.45)$$

$$MD[h, x] = \frac{p(h) - p(h/x)}{p(h)}, \text{ если } p(h/x) < p(h). \quad (4.46)$$

Наряду с  $MB$  и  $MD$  в МШБ используется также коэффициент или фактор уверенности  $CF$  (certainty factor), вычисляемый по формуле

$$CF[h, x] = MB[h, x] - MD[h, x], \quad CF[h, x] \in [-1; 1], \quad (4.47)$$

или

$$CF[h, x] = \begin{cases} \frac{p(h/x) - p(h)}{1 - p(h)}, & \text{если } p(h/x) \geq p(h), \\ \frac{p(h/x) - p(h)}{p(h)}, & \text{если } p(h/x) < p(h), \end{cases} \quad (4.48)$$

$$CF[h, x] = \begin{cases} \frac{p(h/x) - p(h)}{p(h)}, & \text{если } p(h/x) < p(h), \\ \frac{p(h/x) - p(h)}{1 - p(h)}, & \text{если } p(h/x) \geq p(h), \end{cases} \quad (4.49)$$

при этом  $p(h) \notin \{0; 1\}$ .

Например, гипотеза  $h$  – стабильная доходность предприятия региона. Априорная вероятность на основе статистических данных составляет  $p(h) = 0,6$  (для предприятия без указания его профиля). Пусть в качестве свидетельства  $x$  рассматривается, что предприятие производит машиностроительную продукцию и  $p(h/x) = 0,8$ . В этом случае в соответствии с формулами (4.43, 4.44)

$$MB[h, x] = \frac{\max\{0,8; 0,6\} - 0,6}{1 - 0,6} = 0,5; \quad MD[h, x] = \frac{\min\{0,8; 0,6\} - 0,6}{-0,6} = 0; \quad CF[h, x] = 0,5 - 0 = 0,5.$$

Следует заметить, что при данном подходе

$$CF[h, x] + CF[\bar{h}, x] \neq 1, \quad (4.50)$$

здесь  $\bar{h}$  – отрицание  $h$ .

К основным свойствам мер  $MB$  и  $MD$  относятся:

1) если  $h$  – достоверная гипотеза, то

$$p(h/x) = 1, \quad MB[h, x] = 1, \quad MD[h, x] = 0, \quad CF[h, x] = 1; \quad (4.51)$$

2) если достоверна  $\bar{h}$  (отрицание  $h$ ), то

$$p(\bar{h}/x) = 1, \quad MB[h, x] = 0, \quad MD[h, x] = 1, \quad CF[h, x] = -1; \quad (4.52)$$

3) в случае недостатка свидетельств

$$MB[h, x] = 0, \quad MD[h, x] = 0, \quad CF[h, x] = 0, \quad (4.53)$$

т.е. здесь свидетельство  $x$  не подтверждает гипотезу  $h$  и не отвергает ее.

В случае упорядоченного наблюдения двух свидетельств, сначала  $x_1$  и затем  $x_2$ , расчет  $MB$  и  $MD$  производится по формулам:

$$MB[h, x_1 \wedge x_2] = \begin{cases} 0, & \text{если } MD[h, x_1 \wedge x_2] = 1, \\ MB[h, x_1] + MB[h, x_2](1 - MB[h, x_1]), & \text{если } MD[h, x_1 \wedge x_2] \neq 1; \end{cases} \quad (4.54)$$

$$MD[h, x_1 \wedge x_2] = \begin{cases} 0, & \text{если } MB[h, x_1 \wedge x_2] = 1; \\ MD[h, x_1] + MD[h, x_2](1 - MD[h, x_1]), & \text{если } MB[h, x_1 \wedge x_2] \neq 1; \end{cases} \quad (4.55)$$

$$CF[h, x] = \frac{p(h/x) - p(h)}{1 - p(h)}. \quad (4.56)$$

В случае двух гипотез  $h_1, h_2$  для расчетов можно использовать приближенные формулы:

$$MB[h_1 \wedge h_2, x] \approx \min\{MB[h_1, x], MB[h_2, x]\}, \quad (4.57)$$

$$MD[h_1 \vee h_2, x] \approx \min \{ MD[h_1, x], MD[h_2, x] \}, \quad (4.58)$$

$$MB[h_1 \vee h_2, x] \approx \max \{ MB[h_1, x], MB[h_2, x] \}, \quad (4.59)$$

$$MD[h_1 \wedge h_2, x] \approx \max \{ MD[h_1, x], MD[h_2, x] \}. \quad (4.60)$$

Рассмотрим пример использования метода Шортлифа-Бьюкенена.

*Пример 4.4.* Пусть рассматриваются два альтернативных варианта решения:  $v_1$  – модернизировать выпускаемое изделие,  $v_2$  – разработать новое изделие.

Для принятия решения учитываются следующие основные факторы:  $x_1$  – ожидаемое повышение показателей безотказности;  $x_2$  – ожидаемое увеличение срока службы;  $x_3$  – требуемый объем финансовых вложений;  $x_4$  – требуемый объем трудозатрат;  $x_5$  – ожидаемое снижение материалоемкости и энергоемкости;  $x_6$  – ожидаемые сроки завершения работ.

Сформулированы следующие процедурные правила.

П1: «Если  $v_j$  обеспечивает  $x_1$  и  $x_2$ , то вариант  $v_j$  будет принят».

П2: «Если для  $v_j$  выполняются условия  $x_3$  или  $x_4$ , то вариант  $v_j$  будет принят».

П3: «Если  $v_j$  обеспечивает  $x_5$  и  $x_6$ , то вариант  $v_j$  будет принят».

Усредненные доли уверенности  $\tilde{p}(v_j/x_i)$  принятия варианта  $v_j$  на основании данных  $x_i$ , полученных от экспертов, и результаты расчета  $MB[v_j, x_i]$ ,  $MD[v_j, x_i]$ ,  $CF[v_j, x_i]$  для априорной вероятности  $p(v_j) = 0,5$  представлены в табл. 4.7.

#### 4.7. Результаты расчета $MB, MD, CF$ для простых гипотез

$x_i$	$v_1$			
	$\tilde{p}(v_1/x_i)$	$MB[v_1, x_i]$	$MD[v_1, x_i]$	$CF[v_1, x_i]$
$x_1$	0,3	0	0,4	-0,4
$x_2$	0,4	0	0,2	-0,2
$x_3$	0,6	0,2	0	0,2
$x_4$	0,6	0,2	0	0,2
$x_5$	0,4	0	0,2	-0,2
$x_6$	0,6	0,2	0	0,2

Продолжение табл. 4.7

$x_i$	$v_2$			
	$\tilde{p}(v_2/x_i)$	$MB[v_2, x_i]$	$MD[v_2, x_i]$	$CF[v_2, x_i]$
$x_1$	0,8	0,6	0	0,6
$x_2$	0,9	0,8	0	0,8
$x_3$	0,5	0	0	0
$x_4$	0,6	0,2	0	0,2
$x_5$	0,6	0,2	0	0,2
$x_6$	0,6	0,2	0	0,2

При расчете  $MB[v_j, x_i]$ ,  $MD[v_j, x_i]$ ,  $CF[v_j, x_i]$  использованы соотношения (4.43), (4.44), (4.47).

В табл. 4.8 приведены результаты расчета показателей  $MB, MD, CF$  для сложных гипотез, соответствующих сформулированным правилам с использованием соотношений (4.57) – (4.60).

#### 4.8. Результаты расчета $MB, MD, CF$ для сложных гипотез

Правило	$v_1$			$v_2$		
	$MB$	$MD$	$CF$	$MB$	$MD$	$CF$
П1 ( $x_1 \wedge x_2 \sim y_1$ )	0	0,4	-0,4	0,6	0	0,6
П2 ( $x_3 \vee x_4 \sim y_2$ )	0,2	0	0,2	0,2	0	0,2
П1 ( $x_5 \wedge x_6 \sim y_3$ )	0	0,2	-0,2	0,2	0	0,2

Последовательная интеграция значений  $MB$  и  $MD$  по всем трем производственным правилам производится с помощью формул:

$$MB[v_j; y_1, y_2] = MB[v_j; y_1] + MB[v_j; y_2](1 - MB[v_j; y_1]);$$

$$MD[v_j; y_1, y_2] = MD[v_j; y_1] + MD[v_j; y_2](1 - MD[v_j; y_1]);$$

$$MB[v_j; y_1, y_2, y_3] = MB[v_j; y_1, y_2] + MB[v_j; y_3](1 - MB[v_j; y_1, y_2]);$$

$$MD[v_j; y_1, y_2, y_3] = MD[v_j; y_1, y_2] + MD[v_j; y_3](1 - MD[v_j; y_1, y_2]).$$

Итоговые результаты вычислений представлены в табл. 4.9.

Таким образом, при использовании значений  $\tilde{p}(v_j/x_i)$  предпочтительнее вариант решения  $v_2 = \tilde{v}^*$ , так как

$$MB[v_2; y_1, y_2, y_3] = 0,744 > MB[v_1; y_1, y_2, y_3] = 0,2$$

и  $CF[v_2; y_1, y_2, y_3] = 0,744 > CF[v_1; y_1, y_2, y_3] = -0,32.$

#### 4.9. Итоговые результаты вычислений

Сложные гипотезы	$v_1$			$v_2$		
	<i>MB</i>	<i>MD</i>	<i>CF</i>	<i>MB</i>	<i>MD</i>	<i>CF</i>
$[v_j; y_1, y_2]$	0,2	0,4	-0,2	0,68	0	0,68
$[v_j; y_1, y_2, y_3]$	0,2	0,52	-0,32	0,744	0	0,744

#### Контрольные вопросы

1. Что представляет собой модель производственной системы?
2. Какие виды потоков должна учитывать модель производственной системы?
3. Что понимается под концептуальной моделью?
4. Приведите примеры бизнес-процессов в производственной системе.
5. Какие модели используются для решения задач планирования производством?
6. Приведите примеры моделей управления запасами.
7. Какие знаете методы принятия решений в условиях неопределенности?

### 5. МОДЕЛИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ И НАДЕЖНОСТИ

Ряд моделей систем в основном базируются на математическом аппарате теории вероятностей и статистики, так как рассматриваемые в них переменные и процессы имеют случайную природу. К этим моделям, прежде всего, относятся модели систем массового обслуживания и надежности.

#### 5.1. ОБЩИЕ СВЕДЕНИЯ О МОДЕЛЯХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Многие производственные, технические и другие системы при решении задач анализа и синтеза рассматриваются как системы массового обслуживания (СМО). СМО могут быть одноканальные и многоканальные, их функционирование состоит в выполнении поступающего на нее потока требований и заявок. Заявки поступают одна за другой или группами в некоторые обычно случайные моменты времени. Обслуживание поступившей заявки продолжается какое-то время, которое также может рассматриваться как случайное, после чего соответствующий канал освобождается и снова готов для приема следующей заявки. Примерами таких систем являются автоматические системы, автоматические системы управления технологическими процессами (АСУТП), обрабатывающие станки, телекоммуникационные сети, измерительные системы, сервисные системы и т.п.

Любая система массового обслуживания (Queueing System) характеризуется структурой, определяемой составом и функциональными связями. Типовая СМО содержит следующие элементы:

- входящий поток заявок;
- каналы (приборы) обслуживания;
- очередь заявок, ожидающих обслуживания (сценарий обслуживания);
- выходящий поток заявок.

В зависимости от числа каналов и их производительности СМО обладает определенной пропускной способностью, позволяющей с некоторой эффективностью выполнять поток заявок. Предметом теории массового обслуживания является установление зависимости между характером потока заявок, производительностью каналов, числом каналов, сценарием и эффективностью обслуживания. Показателями эффективности обслуживания в зависимости от условий и целей исследуемой задачи могут быть: средний процент заявок, получающих отказ и покидающих систему необслуженными, среднее время ожидания в очереди, вероятность того, что поступившая заявка будет принята к обслуживанию без ожидания в очереди, среднее число заявок в очереди (длина очереди) и т.д. Каждый из этих показателей, с той или другой стороны, характеризует степень приспособленности системы к выполнению потока заявок, т.е. ее пропускную способность.

В узком смысле под пропускной способностью обычно понимают среднее число заявок, которое система может обслужить в единицу времени. Наряду с этим используется термин «относительная пропускная способность», который определяется как среднее отношение количества обслуженных заявок к числу поданных. В общем случае пропускная способность (абсолютная и относительная) зависит как от параметров системы, так и от характера потока заявок.

Если заявки поступают регулярно, т.е. через точно определенные промежутки времени, и обслуживание каждой заявки имеет строго определенную длительность, то расчет пропускной способности системы не представляет особой трудности. На практике заявки обычно поступают в случайные моменты времени, длительность обслуживания заявок тоже носит случайный характер. Поэтому процесс работы системы протекает нерегулярно, в потоках заявок образуются местные сгущения и разрежения. Сгущения входного потока могут приводить к отказам в обслуживании и к образованию очередей, а разрежения – к непроизводительным простоям отдельных каналов или системы в целом. На эти случайности, вызываемые неоднородностью потока заявок, дополнительно накладываются случайности, связанные с задержками обслуживания отдельных заявок. В результате действия этих факторов процесс функционирования системы массового обслуживания представляет

собой случайный процесс. Для решения задач анализа и синтеза СМО, оценки ее пропускной способности необходимо разработать математическую модель случайного процесса, протекающего в системе.

Область применения математических моделей и методов теории массового обслуживания непрерывно расширяется. Многие задачи, связанные с автоматизацией производства, используют модели теории массового обслуживания. Например, потоки деталей, поступающих на технологические машины для выполнения различных операций, могут рассматриваться как потоки заявок, ритмичность поступления которых нарушается за счет случайных причин. Аналогичные задачи возникают в телекоммуникационных и транспортных системах. Многие задачи, относящиеся к надежности технических устройств, например, расчет среднего времени безотказной работы, определение необходимого количества запасных деталей, среднего времени простоя в связи с ремонтом и т.д., решаются методами, заимствованными из теории массового обслуживания.

## 5.2. МОДЕЛИ ПОТОКОВ СОБЫТИЙ

Под потоком событий в теории массового обслуживания и надежности понимается последовательность событий, происходящих одно за другим в моменты времени  $t_1, t_2, \dots$ . Примерами таких потоков могут служить: поток деталей на обработку, поток заявок на обслуживание телекоммуникационной сети, поток отказов в автоматической системе и т.п.

В общем случае события, образующие поток, могут быть различными. Если для моделирования работы СМО рассматривается поток, в котором события различаются лишь моментами появления, то его называют потоком однородных событий. Если события происходят в случайные моменты времени, то соответствующий поток событий называется *случайным*. Поток событий называется *регулярным*, если события следуют одно за другим через строго определенные промежутки времени. Регулярный поток сравнительно редко встречается в реальных системах, однако он представляет интерес как предельный случай для других потоков.

В последующем, если не оговорено особо, поток событий будет считаться однородным и случайным. Важными свойствами этого потока для моделирования СМО являются следующие.

1. Стационарность. Поток событий называется *стационарным*, если вероятность попадания определенного числа событий на участок (интервал) времени длиной  $\tau$  зависит только от длины этого участка и не зависит от того, где на оси времени расположен этот участок.

2. Отсутствие последствия. Поток событий называется *потоком без последствия*, если для любых неперекрывающихся интервалов времени число событий, попадающих на один участок, не зависит от числа событий, попадающих на другие участки времени.

3. Ординарность. Поток событий называется *ординарным*, если вероятность попадания двух или более событий на элементарный (малый) временной участок  $\Delta t$  пренебрежимо мала по сравнению с вероятностью попадания одного события.

4. Поток событий, обладающий перечисленными тремя свойствами (стационарность, ординарность и отсутствием последствия) называется *простейшим*, или стационарным пуассоновским. Для простейшего потока число событий, попадающих на любой фиксированный интервал времени, распределено по закону Пуассона.

Условию стационарности удовлетворяют потоки, вероятностные характеристики которых не зависят от времени. Для стационарного потока характерна постоянная плотность, т.е. среднее число заявок в единицу времени. На практике часто встречаются потоки, которые могут рассматриваться как стационарные на ограниченном интервале времени. Например, поток вызовов на телефонной станции в определенное время суток (рабочее время или ночное) может считаться стационарным, а в течение целых суток поток уже нельзя считать стационарным.

Условие отсутствия последствия означает, что заявки поступают на обслуживание независимо друг от друга. Это условие во многих случаях приемлемо для входных потоков. Однако выходной поток обслуженных заявок обычно имеет последствие, даже если входной поток без последствия.

Условие ординарности потока означает, что заявки на обслуживание приходят поодиночке, а не группами (парами, тройками и т.п.). Если в неординарном потоке заявки поступают одинаковыми по составу группами, то его можно свести к ординарному. Для этого достаточно вместо потока отдельных заявок рассмотреть поток групп заявок (пар, троек и т.д.). Если отдельные заявки случайным образом могут оказаться двойными, тройными и т.д., то приходится рассматривать поток разнородных событий.

Простейший поток заявок играет в теории массового обслуживания особую роль, аналогичную роли нормального закона среди законов распределения случайных величин. Известно, что при суммировании большого числа независимых случайных величин, имеющих различные законы распределения, результирующая величина имеет распределение, близкое к нормальному закону. Аналогично, при суммировании большого числа ординарных стационарных потоков, имеющих последствие, получается поток, близкий к простейшему. Это имеет место при условии, что складываемые потоки должны оказывать на суммарный поток равномерно малое влияние. Обычно достаточно сложить не менее 4–5 потоков, чтобы получить поток, который можно рассматривать как простейший.

Представим отдельный независимый поток событий  $\Pi_i$  как последовательность моментов времени наступления событий (для входного потока – поступления заявок)  $t_1^{(i)}, t_2^{(i)}, \dots$ . Тогда суммирование нескольких потоков  $\Pi_i, i=1, n$  заключается в том, что все моменты времени появления событий сносятся на одну временную ось и для суммарного потока  $\Pi_\Sigma$  можно записать:

$$\Pi_\Sigma = \sum_{i=1}^n \Pi_i. \quad (5.1)$$

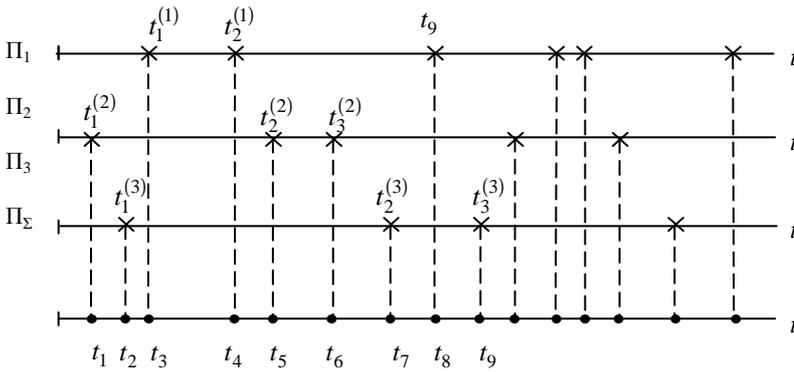
На рис. 5.1 показано суммирование потоков при  $n=3$ .

Широкое использование простейшего потока в теории массового обслуживания объясняется следующим. Во-первых, простейшие и близкие к ним потоки часто встречаются на практике. Во-вторых, при потоках заявок, отличающихся от простейших, во многих случаях можно получить приемлемые по точности результаты, если поток любой структуры заменить простейшим с той же плотностью.

Для простейшего потока  $\Pi$ , как уже отмечалось, число точек, попадающих на временной участок  $\tau$ , распределено по закону Пуассона с математическим ожиданием

$$a = \lambda\tau, \quad (5.2)$$

где  $\lambda$  – плотность потока (среднее число событий, приходящееся на единицу времени).



**Рис. 5.1. Суммирование потоков событий**

Вероятность того, что за время  $\tau$  произойдет ровно  $m$  событий, определяется по формуле

$$p_m(\tau) = \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau}. \quad (5.3)$$

В частности, вероятность того, что участок  $\tau$  окажется пустым, т.е.  $m=0$ , равна

$$p_0(\tau) = e^{-\lambda\tau}. \quad (5.4)$$

Случайное время  $T$  между соседними событиями в простейшем потоке подчиняется показательному (экспоненциальному) распределению с плотностью вероятности:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0; \\ 0, & t < 0, \end{cases} \quad (5.5)$$

где  $\lambda$  – параметр распределения.

Функция распределения времени  $T$  для этого закона имеет вид

$$F(t) = P[T < t] = 1 - e^{-\lambda t}, \quad t \geq 0, \quad (5.6)$$

а математическое ожидание и дисперсия соответственно равны:

$$m_T = M[T] = \int_0^{\infty} t f(t) dt = \frac{1}{\lambda}; \quad (5.7)$$

$$D_T = \lambda \int_0^{\infty} t^2 f(t) dt - m_T^2 = \frac{1}{\lambda^2}, \quad (5.8)$$

здесь  $M[T]$  – знак математического ожидания случайной величины  $T$ .

В случае показательного распределения времени  $T$  любая информация о том, сколько времени уже протекал этот промежуток, не влияет на закон распределения оставшегося времени. Это свойство показательного закона непосредственно связано со свойством отсутствия последствия простейшего потока.

Поток однородных событий, ординарный и без последствия, но нестационарный, называется нестационарным пуассоновским потоком. Такой поток характеризуется мгновенной плотностью потока  $\lambda(t)$  в момент времени  $t$ , т.е.

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{M[0; t + \Delta t] - M[0; t]}{\Delta t}, \quad (5.9)$$

где  $M[0; t]$  – математическое ожидание числа событий на участке  $[0, t]$ .

Для нестационарного пуассоновского потока число событий  $m$ , попадающих на временной интервал  $\tau$ , начинающийся в момент  $t_0$ , подчиняется закону Пуассона:

$$P_m(\tau, t_0) = \frac{a^m}{m!} e^{-a}, \quad m=0,1,2,\dots, \quad (5.10)$$

где  $a$  – математическое ожидание числа событий на временном интервале  $[t_0; t_0 + \tau]$ , т.е.

$$a = \int_{t_0}^{t_0+\tau} \lambda(t) dt. \quad (5.11)$$

Плотность распределения вероятности времени  $T$ , функция распределения и вероятность того, что на интервале  $[t_0; t_0 + \tau]$  не появится ни одного события, для данного потока соответственно равны

$$f_{t_0}(t) = \lambda(t_0 + t) e^{-\int_{t_0}^{t_0+t} \lambda(s) ds}, \quad (t > t_0); \quad (5.12)$$

$$F_{t_0}(t) = 1 - e^{-\int_{t_0}^{t_0+t} \lambda(s) ds} = 1 - e^{-a}; \quad (5.13)$$

$$P(t) = P(T \geq t) = e^{-a}. \quad (5.14)$$

Если в простейшем потоке снять ограничение на отсутствие последствия, при этом промежутки времени между последовательными событиями  $T_1, T_2, \dots$  для потока представляют собой независимые случайные величины, то такой поток называется потоком с ограниченным последствием или потоком Пальма. Модель данного потока широко используется при анализе надежности систем с резервными элементами, а также в виде выходных потоков СМО. Например, если входной поток заявок простейший, то поток необслуженных заявок (в результате выбывания вследствие занятости всех каналов СМО) будет потоком с ограниченным последствием.

Большое применение на практике находят модели потоков с ограниченным последствием в виде потоков Эрланга различного порядка, которые образуются «просеиванием» простейшего потока. Если в простейшем потоке удалить каждую вторую точку, то оставшиеся точки образуют поток Эрланга первого порядка, если же в простейшем потоке сохранять каждую третью точку, то получим поток Эрланга второго порядка и т.д. На рис. 5.2, а, б показаны примеры образования таких потоков, удаленные точки здесь обведены кружками.

Для потока Эрланга  $k$ -го порядка время  $T$  между соседними событиями равно сумме  $k+1$  независимых случайных величин, т.е.

$$T = \sum_{i=1}^{k+1} T_i^n, \quad (5.15)$$

здесь  $T_1^n, T_2^n, \dots, T_{k+1}^n$  – независимые случайные величины, подчиненные одному и тому же показательному закону с параметром  $\lambda$  (см. (4.5)).

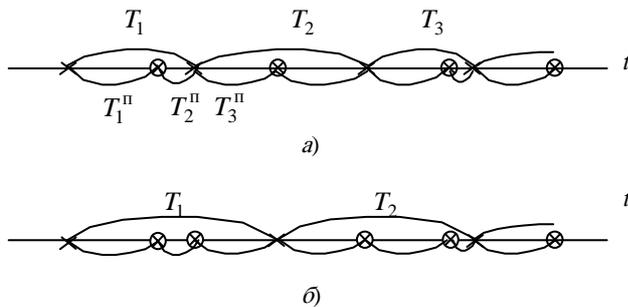


Рис. 5.2. Образование потоков Эрланга первого (а) и второго (б) порядков

Закон распределения времени  $T$  в этом случае называется законом Эрланга  $k$ -го порядка, его плотность вероятности имеет вид

$$f_k(t) = \begin{cases} \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t}, & t \geq 0; \\ 0, & t < 0. \end{cases} \quad (5.16)$$

Математическое ожидание и дисперсия времени  $T$  соответственно равны:

$$m_k = \frac{k+1}{\lambda}, \quad D_k = \frac{k+1}{\lambda^2}, \quad (5.17)$$

а плотность потока

$$\Lambda_k = \frac{\lambda}{k+1}. \quad (5.18)$$

Следует заметить, что при неограниченном увеличении  $k$  поток Эрланга приближается к регулярному потоку.

Показательное распределение и распределение Эрланга широко используются в теории массового обслуживания и в качестве законов распределения случайной величины времени обслуживания одной заявки  $T_{об}$ . В случае показательного закона для времени  $T_{об}$  его характеристики записываются в виде:

$$g(t) = \begin{cases} \mu e^{-\mu t}, & t \geq 0, \\ 0, & t < 0; \end{cases} \quad (5.19)$$

$$G(t) = \text{Вер}[T_{об} < t] = 1 - e^{-\mu t};$$

$$m_{T_{об}} = \frac{1}{\mu}, \quad D_{T_{об}} = \frac{1}{\mu^2},$$

здесь  $\mu$  – параметр распределения времени (обслуживания), т.е. величина, обратная среднему времени обслуживания одной заявки.



$$p_0(0) = 1; \quad p_1(0) = \dots = p_n(0) = 0,$$

т.е. считается, что в начальный момент времени ( $t_0 = 0$ ) все каналы свободны.

Систему линейных дифференциальных уравнений (5.21) можно представить в векторно-матричной форме:

$$\dot{P} = AP(t), \quad (5.22)$$

где  $\dot{P}, P(t)$  соответствующие  $(n+1)$ - векторы-столбцы

$$P = \left( \frac{dp_0(t)}{dt}, \frac{dp_1(t)}{dt}, \dots, \frac{dp_n(t)}{dt} \right)^T;$$

$$P(t) = (p_0(t), p_1(t), \dots, p_n(t))^T$$

и  $A - (n+1) \times (n+1)$  - матрица коэффициентов

$$A = \begin{pmatrix} -\lambda & \mu & 0 & \dots & 0 & 0 \\ \lambda & -(\lambda + \mu) & 2\mu & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & -n\mu \end{pmatrix}.$$

Вероятности  $p_i(t)$  характеризуют изменение средней загрузки системы с течением времени. В частности,  $p_n(t)$  есть вероятность того, что заявка, пришедшая в момент  $t$ , застанет все каналы занятыми, такое состояние может рассматриваться как состояние отказа системы, т.е.

$$P_{\text{отк}}(t) = p_n(t).$$

Величину  $q(t) = 1 - p_n(t)$  называют *относительной пропускной способностью* системы. Она равна отношению среднего числа обслуженных за единицу времени заявок к среднему числу поданных.

Система уравнений (5.21) достаточно легко решается при любом числе каналов  $n$ .

Следует заметить, что уравнения (5.21) справедливы и для зависящих от времени  $\lambda(t), \mu(t)$ , если потоки событий, переводящие систему из одного состояния в другое, остаются пуассоновскими.

При  $t \rightarrow \infty$  вероятности  $p_i(t), i = 0, n$  стремятся к своим предельным (стационарным) значениям  $p_i, i = 0, n$ . Для определения стационарных значений вероятностей состояний системы решается система линейных алгебраических уравнений:

$$AP = 0, \quad P = (p_0, p_1, \dots, p_n)^T. \quad (5.23)$$

При этом одно уравнение системы (5.23) заменяется условием нормировки (см. (5.20)), т.е.

$$\sum_{i=0}^n p_i = 1. \quad (5.24)$$

Например, для  $n = 2$  система алгебраических уравнений принимает вид:

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0; \\ \lambda p_0 - (\lambda + \mu) p_1 + 2\mu p_2 &= 0; \\ p_0 + p_1 + p_2 &= 1. \end{aligned}$$

Для расчета стационарных вероятностей без решения системы (5.23) можно использовать формулы Эрланга, т.е.

$$P_i = \frac{\alpha^i}{\sum_{i=0}^n \alpha^i}, \quad i = \overline{0, n}, \quad \alpha = \frac{\lambda}{\mu}, \quad (5.25)$$

в которых  $\alpha$  называют приведенной плотностью потока заявок, численно эта плотность равна среднему числу заявок, приходящихся на среднее время обслуживания одной заявки, т.е.

$$\alpha = \frac{\lambda}{\mu} = \lambda m_{\text{об}}.$$

Качественная картина изменения вероятностей  $p_i(t), i = \overline{0, n}$ , до предельных значений  $p_i, i = \overline{0, n}$ , при  $n = 2$  показана на рис. 5.4.

Используя формулу (5.25) при  $i = n$ , получаем выражение для расчета вероятностей того, что поступившая заявка найдет все каналы занятыми (вероятность отказа):

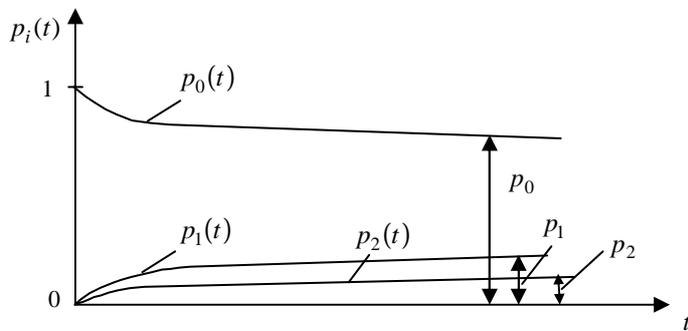


Рис. 5.4. Вероятности состояний  $p_i(t)$  и  $p_i$ ,  $i = 0, 1, 2$  системы массового обслуживания с двумя каналами

$$P_{\text{отк}} = p_n = \frac{\alpha^n}{n!} \cdot \sum_{i=0}^n \frac{\alpha^i}{i!} \quad (5.26)$$

Вероятность  $Q$  того, что заявка будет обрабатываться системой (она характеризует пропускную способность системы) соответственно равна

$$Q = 1 - P_{\text{отк}} \quad (5.27)$$

В случае  $i = 0$  вероятность

$$p_0 = \left( \sum_{i=0}^n \frac{\alpha^i}{i!} \right)^{-1} \quad (5.28)$$

характеризует состояние системы, при котором все каналы свободны.

Важными характеристиками СМО являются среднее число занятых каналов  $\bar{K}$  и среднее число обслуживаемых заявок  $A$  в единицу времени, которые рассчитываются по формулам:

$$\bar{K} = \alpha Q, \quad A = \lambda Q.$$

*Пример 5.2.* Система контроля годности продукции состоит из трех параллельно работающих устройств (каналов). В систему для контроля поступают изделия в среднем через 0,4 часа ( $\lambda = 2,5 \text{ ч}^{-1}$ ). Среднее время проверки одного изделия составляет 0,5 часа ( $\mu = 2 \text{ ч}^{-1}$ ). Требуется рассчитать стационарные вероятности состояний системы контроля  $P_0, P_1, P_2, P_3$ .

В предположении, что процесс изменения состояний в системе марковский и изделия, заставшие все каналы занятыми, не проверяются (система с отказами), для расчета вероятностей можно использовать формулы Эрланга (5.25), т.е. для нашего случая

$$p_i = \frac{\alpha^i}{i!} \left( \sum_{i=0}^3 \frac{\alpha^i}{i!} \right)^{-1}, \quad i = \overline{0,3}, \quad \alpha = \frac{\lambda}{\mu},$$

или

$$p_0 = \left( \sum_{i=0}^3 \frac{\alpha^i}{i!} \right)^{-1}; \quad p_1 = \alpha p_0; \quad p_2 = \frac{\alpha^2}{2} p_0; \quad p_3 = \frac{\alpha^3}{6} p_0.$$

В результате расчетов для  $\alpha = 1,25$  получаем (следует напомнить, что  $0! = 1$ ):

$$p_1 = 0,37238; \quad p_2 = 0,23274; \quad p_3 = 0,09697.$$

Таким образом, вероятность того, что изделие не пройдет контроль (все три устройства заняты) составляет  $p_3 \approx 0,097$ , а вероятность  $Q$  того, что изделие будет проконтролировано, равна  $Q = 1 - p_3 = 0,903$ . Среднее число занятых каналов  $\bar{K}$  (среднее число одновременно контролируемых изделий) и среднее число контролируемых изделий в час  $A$  соответственно равны

$$\bar{K} = \alpha Q = 1,129, \quad A = \lambda Q = 2,2575.$$

Следует заметить, что формулы Эрланга при простейшем входном потоке остаются справедливыми при любом законе распределения времени обслуживания. Кроме того, формулы дают достаточно хорошие результаты, если входной поток имеет незначительное последствие.

Наряду с СМО с отказами на практике широко используются системы с ожиданием. В этой системе, если заявка застала все каналы занятыми, то она становится в очередь и ждет, пока не освободится один из каналов. Если время ожидания в очереди не ограничено, то такую систему называют чистой системой с ожиданием. Если ожидание ограничено какими-либо условиями, то систему называют системой смешанного типа. Последний тип систем наиболее часто встречается на практике.

Ограничения, накладываемые на ожидание, могут быть разных типов. Наиболее часто ограничение накладывается на время ожидания заявки в очереди, это время может быть строго определенным или случайным. При этом начатое обслуживание заявки доводится до конца, независимо от того, сколько времени продолжалось ожидание. В ряде случаев ограничение

накладывается не на время ожидания в очереди, а на общее время пребывания заявки в системе. Встречаются смешанные системы, в которых заявка становится в очередь только в случае, если длина очереди не слишком велика, т.е. число заявок в очереди не превышает допустимого значения.

В системах с ожиданием ожидающие заявки вызываются на обслуживание в соответствии с правилами, называемыми дисциплиной очереди. Заявки могут вызываться в порядке очереди или в случайном порядке. Дисциплиной очереди может быть предусмотрено обслуживание с преимуществами, когда некоторые заявки имеют предпочтения перед другими.

В качестве примера рассмотрим модель смешанной СМО с  $n$  каналами. На вход поступает простейший поток заявок с плотностью  $\lambda$ , время обслуживания одной заявки  $T_{об}$  показательное с параметром  $\mu = \frac{1}{m_{об}}$ . Заявка, заставшая все каналы занятыми, становится в очередь, время ожидания ограничено сроком  $T_{ож}$ . Если до истечения этого срока заявка не поступает на обслуживание, то она покидает систему необслуженной. Время ожидания  $T_{ож}$  считается случайным и распределенным по показательному закону:

$$h(t) = \nu e^{-\nu t}, \quad t > 0, \quad (5.29)$$

где  $\nu$  – параметр, равный обратному значению среднего срока ожидания, т.е.

$$\nu = \frac{1}{m_{ож}}, \quad m_{ож} = M[T_{ож}]. \quad (5.30)$$

Параметр  $\nu$  можно рассматривать как плотность «потока уходов» заявки, стоящей в очереди. При  $\nu \rightarrow \infty$  смешанная система становится системой с отказами, а при  $\nu \rightarrow 0$  – чистой системой с ожиданием.

Для составления системы дифференциальных уравнений (5.21) состояния вида СМО с ожиданием нумеруются с учетом связанных с системой заявок. Заявка называется связанной с системой, если она находится в состоянии обслуживания или ожидает в очереди. При такой нумерации первые  $n+1$  состояния остаются теми же, что в системе с отказами, т.е.  $x_0$  – все каналы свободны (очереди нет);  $x_1$  – занят только один канал (очереди нет); ...;  $x_n$  – заняты все  $n$  каналов (очереди нет), а следующие состояния соответствуют числу состояний, находящихся в очереди:  $x_{n+1}$  – заняты все  $n$  каналов и одна заявка стоит в очереди,  $x_{n+2}$  – заняты  $n$  каналов и две заявки стоят в очереди и т.д.

Так как число заявок, стоящих в очереди, может быть очень большим, то СМО с ожиданием имеет, в общем случае, бесконечное, хотя и счетное число состояний, а соответственно и бесконечное число дифференциальных уравнений для расчета вероятностей  $p_0(t), p_1(t), \dots, p_n(t), p_{n+1}(t), p_{n+2}(t), \dots, p_{n+j}(t), \dots$ . Эта система дифференциальных уравнений записывается следующим образом:

$$\begin{aligned} \frac{dp_0(t)}{dt} &= -\lambda p_0(t) + \mu p_1(t); \\ &\dots \\ \frac{dp_i(t)}{dt} &= \lambda p_{i-1}(t) - (\lambda + i\mu) p_i(t) + (i+1)\mu p_{i+1}(t), \quad (1 \leq i \leq n-1); \\ &\dots \\ \frac{dp_n(t)}{dt} &= \lambda p_{n-1}(t) - (\lambda + n\mu) p_n(t) + (n\mu + \nu) p_{n+1}(t); \\ &\dots \\ \frac{dp_{n+j}(t)}{dt} &= \lambda p_{n+j-1}(t) - (\lambda + n\mu + j\nu) p_{n+j}(t) + [n\mu + (j+1)\nu] p_{n+j+1}(t), \quad (j \geq 1); \\ &\dots \end{aligned} \quad (5.31)$$

Следует заметить, что первые  $n$  уравнений в систе-

ме (5.31) точно совпадают с первыми  $n$  уравнениями для СМО с отказами (см. (5.21)). При  $t \rightarrow \infty$ , т.е. в установившемся режиме обслуживания, система алгебраических уравнений для расчета стационарных вероятностей  $P_0, P_1, \dots, P_n, P_{n+1}, \dots, P_{n+j}, \dots$  принимает вид:

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0; \\ &\dots \\ \lambda p_{i-1} - (\lambda + i\mu) p_i + (i+1)\mu p_{i+1} &= 0, \quad (1 \leq i \leq n-1); \\ &\dots \\ \lambda p_{n-1} - (\lambda + n\mu) p_n + (n\mu + \nu) p_{n+1} &= 0; \\ &\dots \\ \lambda p_{n+j-1} - (\lambda + n\mu + j\nu) p_{n+j} + [n\mu + (j+1)\nu] p_{n+j+1} &= 0; \\ &\dots \\ \sum_{i=1}^n p_i + \sum_{j=1}^{\infty} p_{n+j} &= 1. \end{aligned} \quad (5.32)$$

Расчет стационарных вероятностей удобно производить по следующим конечным формулам:

$$p_i = \frac{\alpha^i}{i!} z^{-1}, \quad i \in [0; n]; \quad (5.33)$$

$$p_{n+j} = \left( \frac{\alpha^n}{n!} \frac{\alpha^j}{\prod_{s=1}^j (n+s\beta)} \right) z^{-1}, \quad j = [1; \infty]; \quad (5.34)$$

здесь

$$z = \sum_{i=0}^n \frac{\alpha^i}{i!} + \frac{\alpha^n}{n!} \sum_{j=1}^{\infty} \frac{\alpha^j}{\prod_{s=1}^j (n+s\beta)}; \quad (5.35)$$

$$\alpha = \frac{\lambda}{\mu} = \lambda m_{об}, \quad \beta = \frac{\nu}{\mu} = \nu m_{об}. \quad (5.36)$$

Следует заметить, что  $p_0 = z^{-1}$  (см. (5.33)).

Используя формулы (5.33) – (5.36), можно подсчитать математическое ожидание  $m_{оч}$  числа заявок, находящихся в очереди,

$$m_{оч} = \sum_{j=1}^{\infty} j p_{n+j} = \left( \frac{\alpha^n}{n!} \sum_{j=1}^{\infty} \frac{j \alpha^j}{\prod_{s=1}^j (n+s\beta)} \right) z^{-1}, \quad (5.37)$$

вероятность  $P_{оч}$  наличия очереди

$$P_{оч} = 1 - \sum_{i=0}^n p_i, \quad (5.38)$$

вероятность  $P_n$  того, что заявка покинет систему необслуженной:

$$P_n = \left( \frac{\beta \alpha^n}{\alpha n!} \sum_{j=1}^{\infty} \frac{j \alpha^j}{\prod_{s=1}^j (n+s\beta)} \right) z^{-1} \quad (5.39)$$

и пропускную способность  $Q$  системы, т.е. вероятность того, что заявка, попавшая в систему, будет обслужена:

$$Q = 1 - P_n. \quad (5.40)$$

Некоторые трудности использования формул (5.33), (5.34), (5.37), (5.38) вызывает наличие в них бесконечных сумм (см. (5.35)). Однако с ростом  $j$  слагаемые в суммах быстро убывают и для оценки ошибки, вызываемой отбрасыванием всех членов сумм, начиная с  $r$ -го, можно использовать формулу:

$$\sum_{j=r}^{\infty} \frac{\alpha^j}{\prod_{s=1}^j (n+s\beta)} < \frac{\left(\frac{\alpha}{\beta}\right)^r}{r!} e^{\frac{\alpha}{\beta}}.$$

Необходимо отметить, что в СМО с ожиданием стационарный режим существует, если выполняется условие  $\alpha < n$ , т.е. когда среднее число заявок, приходящееся на время обслуживания одной заявки, не выходит за пределы возможностей системы с  $n$  каналами. В случае  $\alpha \geq n$  число заявок, находящихся в очереди, с увеличением времени  $t$  неограниченно возрастает.

Для СМО с  $\alpha < n$  в предположении  $\beta \rightarrow 0$  для приближенной оценки вероятностей состояний можно воспользоваться следующими формулами:

$$p_0 = \left( \sum_{i=0}^n \frac{\alpha^i}{i!} + \frac{\alpha^{n+1}}{n!(n-\alpha)} \right)^{-1},$$

$$p_i = \frac{\alpha^i}{i!} p_0, \quad i \in [1; n], \quad (5.41)$$

$$p_{n+j} = \frac{\alpha^{n+j}}{n! n^j} p_0, \quad j \in [n+1; \infty].$$

Среднее число заявок, находящихся в очереди, при  $\beta \rightarrow 0$  равно

$$m_{\text{оч}} = \frac{\alpha^{n+1}}{nn! \left(1 - \frac{\alpha}{n}\right)^2} P_0. \quad (5.42)$$

В завершение данного параграфа рассмотрим СМО с ограничением ожидания по числу заявок, стоящих в очереди. Пусть заявка, заставшая все каналы занятыми, становится в очередь, если в ней находится менее  $m$  заявок, в противном случае заявка в очередь не становится и покидает систему необслуженной. Допущения о простейшем потоке заявок и о показательном распределении времени обслуживания сохраняются. Число состояний такой системы конечно и равно  $n + m + 1$ , т.е.  $x_0$  – все каналы свободны;  $x_i$ ,  $i \in [1, n]$  – занято  $i$  каналов, очереди нет;  $x_{n+j}$ ,  $j \in [1; m]$  – занято  $n$  каналов и  $j$  заявок стоит в очереди. Соответствующая система дифференциальных уравнений записывается в виде

$$\begin{aligned} \frac{dp_0(t)}{dt} &= -\lambda p_0(t) + \mu p_1(t); \\ &\dots \\ \frac{dp_i(t)}{dt} &= \lambda p_{i-1}(t) - (\lambda + i\mu) p_i(t) + (i+1)\mu p_{i+1}(t), \quad i \in [1; n-1]; \\ &\dots \\ \frac{dp_n(t)}{dt} &= \lambda p_{n-1}(t) - (\lambda + n\mu) p_n(t) + n\mu p_{n+1}(t); \\ &\dots \\ \frac{dp_{n+j}(t)}{dt} &= \lambda p_{n+j-1}(t) - (\lambda + n\mu) p_{n+j}(t) + n\mu p_{n+j+1}(t), \quad j \in [1; m-1]; \\ &\dots \\ \frac{dp_{n+m}(t)}{dt} &= \lambda p_{n+m-1}(t) - n\mu p_{n+m}(t). \end{aligned} \quad (5.43)$$

Для определения стационарных вероятностей состояний решается система алгебраических уравнений:

$$\begin{aligned} -\lambda p_0 + \mu p_1 &= 0; \\ &\dots \\ \lambda p_{i-1} - (\lambda + i\mu) p_i + (i+1)\mu p_{i+1} &= 0, \quad i \in [1; n-1]; \\ &\dots \\ \lambda p_{n-1} - (\lambda + n\mu) p_n + n\mu p_{n+1} &= 0; \\ &\dots \\ \lambda p_{n+j-1} - (\lambda + n\mu) p_{n+j} + n\mu p_{n+j+1} &= 0, \quad j \in [1; m-1]; \\ \sum_{i=0}^{n+m} p_i &= 1. \end{aligned} \quad (5.44)$$

Конечные формулы для расчета стационарных вероятностей имеют вид

$$\begin{aligned} p_i &= \frac{\alpha^i}{i!} z_0^{-1}, \quad i \in [0; n]; \\ p_{n+j} &= \frac{\alpha^n}{n!} \left(\frac{\alpha}{n}\right)^j z_0^{-1}, \quad j \in [1; m]; \\ z &= \sum_{i=0}^n \frac{\alpha^i}{i!} + \frac{\alpha^n}{n!} \sum_{j=1}^m \left(\frac{\alpha}{n}\right)^j; \quad \alpha = \frac{\lambda}{\mu}. \end{aligned} \quad (5.45)$$

Вероятность  $P_n$  того, что заявка не будет обслужена, так как в очереди уже стоит  $m$  заявок, равна  $p_{n+m}$ .

## 5.4. МОДЕЛИ НАДЕЖНОСТИ СИСТЕМ

### 5.4.1. Основные понятия

Математические модели надежности необходимы при проектировании систем для выбора варианта, который будет обеспечивать ее эффективное функционирование в процессе реальной эксплуатации. Кроме того, модели надежности требуются при решении задач диагностики неисправностей, планирования ремонтно-профилактических работ и комплектования запасными частями.

При построении модели надежности системы в первую очередь требуется определить:

1) состояния, которые являются отказами системы (отказ – событие, заключающееся в нарушении работоспособности объекта, т.е. когда не обеспечивается нормальное функционирование при выполнении поставленных задач), факт возникновения отказа устанавливается критериями отказа согласно нормативно-технической документации;

2) характер процессов возникновения отказов компонентов исследуемой системы, эти процессы описывают с помощью вероятностных законов и дифференциальных уравнений;

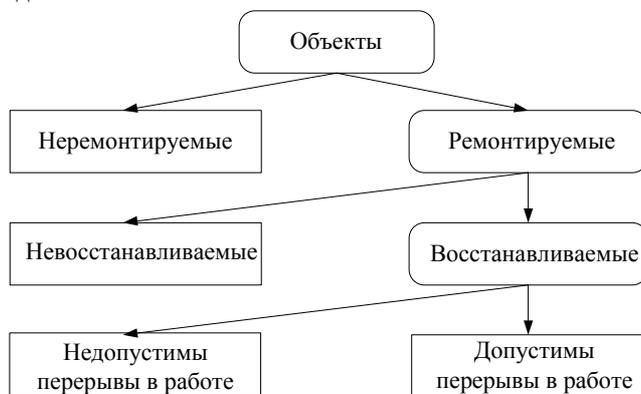
3) конфигурацию (структуру) системы, которая определяет характер соединения компонентов, правила работы, наличие резервирования, схему обслуживания и т.п.

Важную роль для построения модели надежности играет выделение класса системы по отношению к ремонту и восстановлению. Восстановление – это процесс обнаружения и устранения отказа (повреждения) с целью восстановления его работоспособности (исправности). Объект называется восстанавливаемым, если работоспособность его в случае возникновения отказа подлежит восстановлению в рассматриваемых условиях, и невосстанавливаемым – если не подлежит. Один и тот же объект в зависимости от ситуации может быть восстанавливаемым или невосстанавливаемым. Например, аппаратура спутника на этапе хранения и подготовки к старту – восстанавливаемая, а во время полета – невосстанавливаемая.

Ремонт представляет собой комплекс операций по восстановлению исправности или работоспособности объекта и восстановлению ресурса объекта или его составных частей. Заметим, что ресурс (технический ресурс) – это наработка объекта от начала его эксплуатации или ее возобновления после среднего или капитального ремонта до наступления предельного состояния. Под наработкой понимается продолжительность или объем работы объекта. Предельным называется состояние объекта, при котором его дальнейшее применение по назначению должно быть прекращено из-за неустранимого нарушения требований безопасности или неустранимого отклонения заданных параметров за установленные пределы, недопустимого увеличения эксплуатационных расходов или необходимости проведения капитального ремонта. Признаки (критерии) предельного состояния устанавливаются в соответствии с требованиями нормативно-технической документации (НТД) на систему.

Ремонт объекта может выполняться заменой или восстановлением отдельных элементов и сборочных единиц. Объект называется ремонтируемым, если исправность его в случае возникновения отказа или повреждения подлежит восстановлению, проведение ремонта объекта предусматривается в НТД. Объект, исправность и работоспособность которого в случае возникновения отказа (повреждения) не подлежит восстановлению, – неремонтируемый. Таким образом, понятие «ремонт» предусматривает возможность замены отказавших частей и не увязывается с рассматриваемыми условиями (ситуацией). Классификация объектов по отношению к ремонту и восстановлению приведена на рис. 5.5.

Основным компонентом модели надежности объекта, т.е. технического изделия (элемента, системы) определенного целевого назначения, рассматриваемого в период проектирования, является закон распределения случайного времени  $T$  работы до отказа. Существует два основных пути определения распределения времени  $T$ . Первый состоит в утверждении некоторых исходных принципов, которые особенно полезны при недостаточности наших познаний «a priori». При этом подходе нужно полагаться на физическое толкование для постулирования формы кривых интенсивностей отказов. Второй метод оперирует с эмпирическими данными, полученными в результате проведения и обработки результатов испытаний, а также при наблюдении за работой различных образцов оборудования в процессе эксплуатации систем, в которых применяются аналогичные элементы, работающие в подобных



**Рис. 5.5. Классификация объектов по признакам ремонтируемости и восстановления**

условиях. При таком подходе по полученным экспериментальным данным строят гистограмму для времени  $T$  и определяют соответствующую функцию распределения отказов. Наиболее предпочтительна комбинация указанных методов, когда для определения распределения времени  $T$  используются достаточные статистические данные и представление о механизме возникновения отказов на основе физико-химических и других соображений.

Знание законов распределения времени работы до отказа и времени ремонта необходимо при расчете показателей надежности (ПН) систем, т.е. технических характеристик, количественно определяющих одно или несколько свойств, составляющих надежность системы. К ним относятся:

- безотказность – свойство непрерывно сохранять работоспособность в течение некоторого времени или некоторой наработки);
- ремонтпригодность – свойство, заключающееся в приспособлении к предупреждению и обнаружению причины возникновения его отказов, повреждений и устранению их последствий путем проведения ремонтов и технического обслуживания);
- долговечность – свойство объекта сохранять при установленной системе технического обслуживания и ремонтов работоспособность до наступления предельного состояния, при котором прекращается дальнейшая эксплуатация из-за неустраняемых нарушений значений параметров;

- сохраняемость – свойство объекта непрерывно сохранять значения установленных показателей его качества в заданных пределах в течение и после хранения, а также транспортирования.

#### 5.4.2. Неремонтируемые объекты

Неремонтируемые объекты работают до первого отказа, основные показатели безотказности для них и формулы оценки показателей по результатам испытаний приведены в табл. 5.1. Здесь  $N(t)$  – число объектов, работоспособных к моменту времени  $t$ ;  $N$  – число испытываемых объектов;  $T_i$  – наработка до отказа  $i$ -го объекта. Отдельные показатели надежности имеют размерность, например, если в качестве наработки рассматривается время, то средняя наработка до отказа  $m_t$  имеет размерность (ч), а интенсивность отказов  $\lambda(t)$  и плотность распределения времени до отказа  $f(t) - [1/\text{ч}]$ .

#### 5.1. Показатели безотказности

Показатели надежности (безотказности)	Формулы оценки
Вероятность безотказной работы – вероятность того, что в пределах заданной наработки $[0; t]$ отказ не возникает	$\hat{R}(t) = N(t) / N$
Средняя наработка до отказа – математическое ожидание случайной наработки $T$ до первого отказа	$\hat{m}_t = \frac{1}{N} \sum_{i=1}^N T_i$
Интенсивность отказов – условная плотность вероятности возникновения отказа объекта, определяемая для рассматриваемого момента $t$ наработки при условии, что до этого момента отказ не возник	$\hat{\lambda}(t) = \frac{N(t) - N(t + \Delta t)}{N(t)\Delta t}$

Вероятность безотказной работы на временном интервале  $[0; t]$   $R(t)$  и вероятность отказа  $Q(t)$  (в зарубежной литературе иногда их называют соответственно функциями надежности и ненадежности) определяются непосредственно по функции распределения  $F(t)$  случайного времени наработки до отказа  $T$ , т.е.

$$R(t) \stackrel{\Delta}{=} P[T > t] = 1 - F(t), \quad Q(t) \stackrel{\Delta}{=} P[T \leq t] = F(t),$$

здесь  $P[A]$  – вероятность события  $A$ .

Функциональные связи между показателями  $R(t), Q(t), f(t), \lambda(t)$  и  $m_t$  приведены в табл. 5.2. Вероятность безотказной работы в течение интервала  $[t_1; t_2]$  определяется по формуле:

$$R(t_1, t_2) = \frac{R(t_2)}{R(t_1)} = \exp \left[ - \int_{t_1}^{t_2} \lambda(t) dt \right].$$

#### 5.2. Функциональные связи между показателями безотказности

Определяемый ПН	Известный ПН			
	$R(t)$	$Q(t)$	$f(t)$	$\lambda(t)$
$R(t)$	$R(t)$	$1 - Q(t)$	$\int_t^{\infty} f(x) dx$	$\exp \left[ - \int_0^t \lambda(x) dx \right]$
$Q(t)$	$1 - R(t)$	$Q(t)$	$\int_t^{\infty} f(x) dx$	$1 - \exp \left[ - \int_0^t \lambda(x) dx \right]$
$f(t)$	$-\frac{d}{dt} R(t)$	$\frac{d}{dt} Q(t)$	$f(t)$	$\lambda(t) \exp \left[ - \int_0^t \lambda(x) dx \right]$
$\lambda(t)$	$-\frac{d}{dt} R(t) / R(t)$	$\frac{d}{dt} Q(t) / (1 - Q(t))$	$f(t) / \int_t^{\infty} f(x) dx$	$\lambda(t)$

$m_t$	$\int_0^{\infty} R(t)\lambda(t)dt$	$\int_0^{\infty} (1-Q(t))\lambda(t)dt$	$\int_0^{\infty} tf(t)dt$	$\int_0^{\infty} \exp\left[-\int_0^t \lambda(x)dx\right] dt$
-------	------------------------------------	--	---------------------------	--

Конкретный вид функций  $R(t)$ ,  $\lambda(t)$  определяется законом распределения случайной наработки до отказа  $T$ . Во многих случаях на практике функция  $\lambda(t)$  имеет вид, показанный на рис. 5.6.

Здесь можно выделить три участка: приработки (I), нормальной работы (II) и «старения» (III). Иногда в качестве показателей используют среднюю интенсивность отказов  $\bar{\lambda}$  за время, соответствующее техническому ресурсу  $t_p$ , т.е.

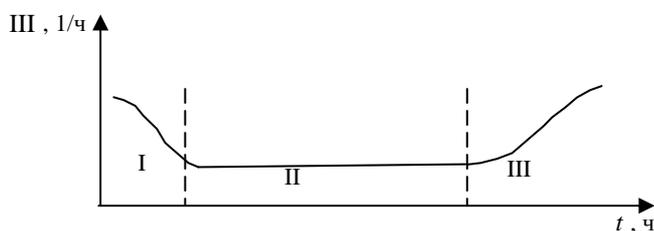


Рис. 5.6. Вид функции  $\lambda(t)$

$$\bar{\lambda} = \frac{1}{t_p} \int_0^{t_p} \lambda(t) dt, \quad (5.46)$$

а также суммарную наработку  $t_n$  до начала массовых параметрических отказов, которые характеризуются отклонением значения хотя бы одного рабочего параметра за пределы допуска.

Если система состоит из  $n$  элементов и отказ любого из элементов приводит к отказу всей системы, то такое соединение элементов в смысле надежности называется последовательным. ПН машин, технологических аппаратов и других систем по известным характеристикам элементов в предположении, что они соединены последовательно, рассчитываются по следующим формулам.

Вероятность безотказной работы  $R(t)$  системы на временном интервале  $[0, t]$  равна произведению вероятностей безотказной работы  $R_i(t), i = \overline{1, n}$  элементов, т.е.

$$R(t) = \prod_{i=1}^n R_i(t), \quad (5.47)$$

или для отрезка времени  $[t_0, t_0 + \tau]$

$$R(t_0, t_0 + \tau) = \prod_{i=1}^n R_i(t_0, t_0 + \tau). \quad (5.48)$$

Интенсивность отказов  $\Lambda(t)$  системы равна сумме интенсивностей отказов  $\lambda_i(t)$  элементов, т.е.

$$\Lambda(t) = \sum_{i=1}^n \lambda_i(t). \quad (5.49)$$

Плотность вероятности  $f(t)$  времени работы до отказа системы определяется по формуле

$$f(t) = R(t) \sum_{i=1}^n \frac{f_i(t)}{R_i(t)}, \quad (5.50)$$

где  $f_i(t)$  — плотность вероятности отказов  $i$ -го элемента.

Среднее время  $m_t$  работы системы до отказа рассчитывается следующим образом

$$m_t = \left( \sum_{i=1}^n \bar{\lambda}_i \right)^{-1}, \quad \bar{\lambda}_i = \frac{1}{t_{\text{сл}}} \int_0^{t_{\text{сл}}} \lambda_i(t) dt, \quad (5.51)$$

здесь  $t_{\text{сл}}$  — срок службы.

Если для элементов имеет место показательный закон распределения времени работы до отказа, т.е. выполняется условие экспоненциального закона надежности, то формулы (5.69) – (5.71) принимают следующий вид

$$R(t) = e^{-\Lambda t}, \quad R(t_0, t_0 + \tau) = e^{-\Lambda \tau},$$

$$\Lambda = \sum_{i=1}^n \lambda_i,$$

$$f(t) = \Lambda e^{-\Lambda t}, \quad m_t = \frac{1}{\Lambda}.$$

Значительная доля отказов систем связана с выходом ее параметров за допустимые пределы, т.е. за пределы допуска. Допуск есть установленные расчетом или экспериментально границы для значений параметров объекта, при которых последний способен выполнять заданные функции, сохраняя свои эксплуатационные показатели в течение требуемого времени

при определенных условиях. Задача установления допусков с учетом требований надежности и проверка их обеспечения является одной из наиболее сложных и ответственных при проектировании систем и анализе метрологических характеристик.

### 5.4.3. Ремонтируемые объекты

Для ремонтируемых объектов важным показателем надежности является параметр потока отказов  $\omega(t)$ . Для ординарных потоков без последствия применительно к ремонтируемым невосстанавливаемым системам параметр потока отказов (рис. 5.7, а) совпадает с интенсивностью потока и представляет собой отношение среднего числа отказов за произвольную малую его наработку к значению этой наработки, т.е.

$$\omega(t) = \lim_{\Delta t \rightarrow 0} \frac{M[r(t, t + \Delta t)]}{\Delta t},$$

где  $r(t, t + \Delta t)$  – число отказов за интервал  $(t, t + \Delta t)$ ;  $M[X]$  – математическое ожидание случайной величины  $X$ .

Среднее значение параметра потока отказов  $\bar{\omega}$  определяется по формуле, аналогичной формуле (5.46). Если при  $t \rightarrow \infty$  плотность

$f(t) \rightarrow 0$ , то существует установившееся значение параметра потока отказов

$$\omega = \lim_{t \rightarrow \infty} \omega(t) = \frac{1}{\tilde{m}_t}. \quad (5.52)$$

В случае экспоненциального распределения наработки между отказами с параметром  $\lambda$  имеет место

$$\omega(t) = \lambda, \quad R(t, t + \tau) = R(\tau) = e^{-\lambda\tau}, \quad \tilde{m}_t = \frac{1}{\lambda}, \quad (5.53)$$

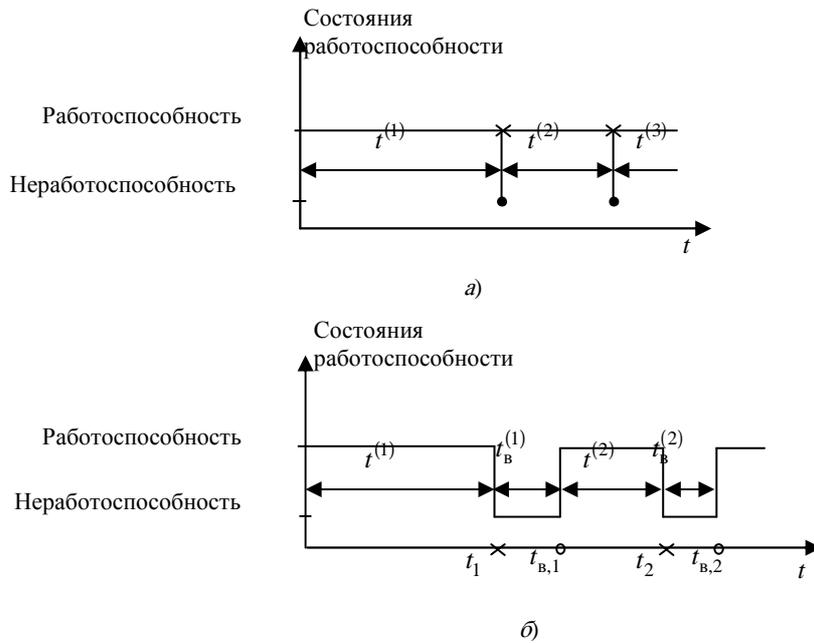


Рис. 5.7. Поток отказов (а) и реализация случайного процесса эксплуатации восстанавливаемого объекта (б)

здесь  $R(t, t + \tau)$  – вероятность безотказной работы на интервале  $[t; t + \tau]$ ;  $\tilde{m}_t$  – средняя наработка между отказами.

Для ремонтируемых восстанавливаемых (обслуживаемых) объектов при их использовании в течение заданного времени работы  $[t_1, t_k]$  допускаются отказы и вызванные ими кратковременные перерывы в работе. В процессе эксплуатации такого объекта чередуются случайные периоды времени безотказной работы  $T^{(i)}$  и времени восстановления (ремонта)  $T_{в}^{(i)}$  (рис. 5.7, б), т.е. имеет место альтернирующий процесс функционирования объекта.

Надежность объектов данного класса характеризуется рядом комплексных (коэффициенты готовности, оперативной готовности, технического использования, характеризующих безотказность и ремонтпригодность) и единичных ( $m_t, R(t_1, t_2)$ , характеризующих безотказность, среднее время восстановления, характеризующее ремонтпригодность) показателей надежности.

Нестационарный коэффициент или функция готовности  $K_r(t)$  есть вероятность того, что в момент времени  $t$  объект находится в состоянии работоспособности (при известных начальных условиях в момент  $t = 0$ ). На основе испытаний с  $N$  объектами  $K_r(t)$  оценивается по формуле

$$\hat{K}_r(t) = N(t) / N,$$

где  $N(t)$  – число объектов, находящихся в момент времени  $t$  в состоянии работоспособности.

Коэффициент готовности (стационарный)  $K_r$  – есть вероятность того, что объект окажется работоспособным в произвольный момент времени, кроме планируемых периодов простоя, и представляет собой предельное значение  $K_r(t)$ , он равен

$$K_r = \lim_{t \rightarrow \infty} K_r(t) = \frac{\tilde{m}_t}{\tilde{m}_t + m_B}, \quad (5.54)$$

где  $\tilde{m}_t$  – средняя наработка между отказами (наработка на отказ);  $m_B$  – среднее время восстановления объекта.

Для ремонтируемых систем коэффициент готовности  $K_r(t)$  в предположении, что ремонт отказавшего узла производится независимо от состояний работоспособности других узлов, это имеет место при неограниченном обслуживании системы, производится по формуле

$$K_r(t) = \prod_{i=1}^n K_{r_i}(t),$$

где  $K_{r_i}(t)$  – коэффициент готовности  $i$ -го узла.

С показателями  $K_r(t)$  и  $K_r$  жестко связаны функция  $K_n(t)$  и коэффициент  $K_n$  простоя (рис. 5.8, а):

$$K_n(t) = 1 - K_r(t), K_n = 1 - K_r.$$

Нестационарный коэффициент оперативной готовности  $K_{ор}(t, t + \tau)$  есть вероятность того, что объект окажется работоспособным в момент  $t$  и проработает безотказно в течение заданного времени  $\tau$ , начиная с этого момента, или другими словами, вероятность того, что интервал времени  $[t, t + \tau]$  целиком попадает внутрь одного из интервалов работоспособности  $t^{(i)}$  (рис. 5.7, б), коэффициент  $R(t, t + \tau)$  оценивается по формуле

$$\hat{K}_{ор}(t, t + \tau) = N(t, t + \tau) / N,$$

где  $N(t, t + \tau)$  – число объектов, работоспособных в момент времени  $t$  и проработавших безотказно до момента времени  $t + \tau$ .

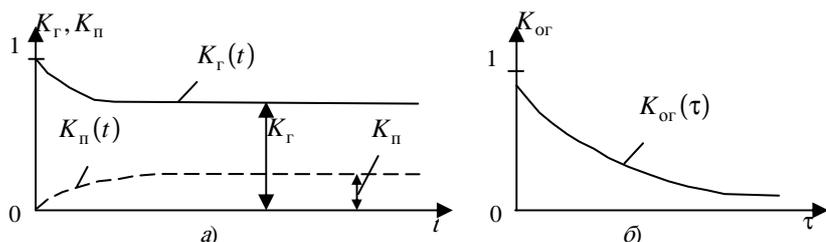


Рис. 5.8. Показатели надежности  $K_r(t)$ ,  $K_r$ ,  $K_n(t)$ ,  $K_n$  (а) и  $K_{ор}(t)$  (б)

Стационарный коэффициент оперативной готовности или просто коэффициент оперативной готовности  $K_{ор}(\tau)$  есть вероятность того, что объект проработает безотказно в течение заданного времени  $\tau$ , начиная с произвольного «достаточно удаленного» момента времени

$$K_{ор}(\tau) = \lim_{t \rightarrow \infty} K_{ор}(t, t + \tau) \text{ или } K_{ор}(\tau) = K_r R(\tau), \quad (5.55)$$

если вероятность безотказной работы в течение времени  $\tau$  не зависит от момента начала работы.

Вид функции  $K_{ор}(\tau)$  показан на рис. 5.8, б.

Для учета всех простоев объекта, в том числе и не связанных с отказами, используется коэффициент технического использования  $K_{ти}$  – отношение математического ожидания времени пребывания объекта в работоспособном состоянии за некоторый период эксплуатации к сумме математических ожиданий времени пребывания объекта в работоспособном состоянии, времени простоев, обусловленных техническим обслуживанием, и времени ремонтов за тот же период эксплуатации, т.е.

$$K_{ти} = \frac{t_{раб.}}{t_{раб.} + t_{рем.} + t_{обсл.}}, \quad (5.56)$$

где  $t_{раб.}$  – суммарное время нахождения объекта в работоспособном состоянии;  $t_{рем.}$  – суммарное время ремонта;  $t_{обсл.}$  – суммарное время профилактического обслуживания.

Показатели ремонтпригодности рассматриваемых объектов представляют собой характеристики случайной величины  $\theta$  – длительности восстановления, она включает и поиск неисправности. Широко используются следующие показатели ремонтпригодности:

– вероятность восстановления за временный интервал  $[0, t]$

$$P_B(t) = P[\theta \leq t] = F_B(t);$$

– среднее время восстановления

$$m_B = M[\theta] = \int_0^{\infty} t f_B(t) dt,$$

где  $F_B(t), f_B(t)$  – функция и плотность распределения времени  $\theta$ .

Если время  $\theta$  имеет показательное распределение с параметром  $\mu$ , т.е.  $\varphi_B(t) = \mu \cdot \exp[-\mu t]$ , то

$$P_B(t) = 1 - e^{-\mu t}, \quad m_B = 1/\mu. \quad (5.57)$$

Стационарный коэффициент готовности при показательных распределениях времени работы между отказами  $T$  и времени восстановления  $\theta$  с параметрами  $\lambda$  и  $\mu$  соответственно рассчитывается по формуле

$$K_r = \frac{\mu}{\mu + \lambda}. \quad (5.58)$$

Для обеспечения надежности ремонтируемых восстанавливаемых объектов, вследствие тяжелых последствий отказов, перерывы в работе которых не допускаются, вводится резервирование. Ремонт отказавшего элемента здесь производится во время выполнения задания. Для таких объектов в основном используются те же показатели надежности, что и для ремонтируемых невосстанавливаемых объектов.

#### 5.4.4. Системы с резервированием

Основным методом повышения надежности систем является введение резервирования (избыточности). Различают пять видов резервирования: структурное (аппаратное), временное, нагрузочное, функциональное и информационное. Наиболее часто используются структурное и временное виды резервирования.

Под структурным резервированием понимается метод повышения надежности системы, предусматривающий использование избыточных структурных элементов. Классификация способов (вариантов) реализации структурного резервирования приведена в табл. 5.3.

Элемент (узел), минимально необходимый для выполнения системой заданных функций, называется основным или рабочим, элементы, обеспечивающие работоспособность системы в случае отказа основного элемента, – резервными. На структурной схеме расчета надежности резервные элементы размещаются параллельно основному. Резервные элементы могут иметь соединение с основным элементом без переключателя и с переключателем (рис. 5.9).

Для невосстанавливаемой системы вероятность  $R_{i(m)}(t)$  безотказной работы  $i$ -го узла на интервале времени  $[0, t]$  при  $m$ -кратном нагруженном резерве (без переключателей), определяется по формуле

$$R_{i(m)}(t) = 1 - \prod_{j=0}^{m-1} [1 - R_{ij}(t)], \quad (5.59)$$

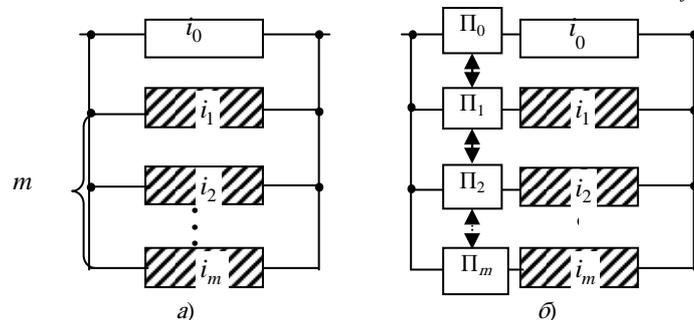


Рис. 5.9. Параллельное соединение элементов в структурной схеме надежности:

*a* – нагруженный резерв; *б* – ненагруженный резерв с переключателями

$\Pi_0, \Pi_1, \dots, \Pi_m$

#### 5.3. Классификация способов резервирования

Признак классификации	Способ резервирования и резерва	Определение
Способ включения резерва	Постоянное резервирование	Резервные элементы участвуют в функционировании наравне с основным.
	Резервирование замещением	Функции основного элемента передаются резервными элементами после его отказа
Схема включения резерва	Общее резервирование	Резерв предусматривается на случай отказа объекта в целом.
	Раздельное резервирование	Резерв предусматривается на случай отказов отдельных элементов или их групп
Состояние резерва	Ненагруженный резерв	Резервные элементы практически не несут нагрузок.
	Облегченный резерв	Резервные элементы находятся

	Нагруженный резерв	в менее нагруженном режиме, чем основной. Резервные элементы работают в том же режиме, что и основной
Однородность резервирования	Однородное резервирование	Резервирование имеет один вид по соответствующему признаку классификации.
	Смешанное резервирование	В объекте различные виды резервирования совмещаются
Кратность резервирования	Однократное резервирование	Кратность резервирования равна единице (дублирование).
	Многократное резервирование	Кратность резервирования выражается числом более единицы

где  $R_{0j}(t)$ ,  $R_{ij}(t)$ ,  $j=1, m$  – вероятности безотказной работы соответственно основного и  $j$ -го резервного элементов.

В данном случае отказ  $i$ -го узла наступает, когда выйдут из строя все  $m+1$  элементов.

Если основной и резервные элементы, работающие в одной режиме, одинаковы, то

$$R_{i(m)}(\tau) = 1 - [1 - R_{ij}(\tau)]^{m+1}. \quad (5.60)$$

Средняя наработка до отказа  $i$ -го резервируемого узла равна

$$m_{ii(m)} = \int_0^{\infty} R_{i(m)}(\tau) d\tau. \quad (5.61)$$

При экспоненциальном распределении времени работы основного и  $m$  резервных элементов с параметром  $\lambda$  для вероятности  $R_{i(m)}(t)$  и среднего времени до отказа имеет место

$$R_{i(m)}(t) \geq 1 - (\lambda t)^{m+1}, \quad (5.62)$$

$$m_{ii(m)} = \frac{1}{\lambda} \left( 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{m+1} \right). \quad (5.63)$$

Недостатком нагруженного резерва является одновременная работа, а, следовательно, повышенный износ всех  $m+1$  элементов. В случаях, когда допустимы перерывы в работе системы, необходимые для замены основного элемента резервными, более экономичными являются облегченный резерв и ненагруженный резерв (резерв замещением).

При ненагруженном резерве без восстановления средняя наработка до отказа группы равна сумме средних наработок всех элементов, т.е.

$$m_{ii(m)} = \sum_{j=0}^m m_{ij}. \quad (5.64)$$

В случае экспоненциального распределения времени работы элементов с параметром  $\lambda_i$ , вероятность  $R_{i(m)}(\tau)$  при ненагруженном резерве определяется по формуле

$$R_{i(m)}(\tau) = e^{-\lambda_i \tau} \sum_{j=0}^m \frac{(\lambda_i \tau)^j}{j!}. \quad (5.65)$$

Следует заметить, что при резерве замещением может потребоваться использование переключающих устройств (рис. 5.9, б).

Для восстанавливаемых систем, когда в процессе функционирования могут восстанавливаться основные и резервные элементы при построении модели надежности необходимо учитывать число  $k$  рабочих элементов в группе, которая резервируется, состав резервных элементов и показатель обслуживания  $r$ , т.е. сколько элементов может одновременно восстанавливаться. Схема такой системы представлена на рис. 5.10, а сокращенно ее модель обозначается кортежем состава  $\langle k, m_1, m_2, m_3, r \rangle$ , здесь  $m_1, m_2, m_3$  – числа резервных элементов соответственно в нагруженном, облегченном (характеризуется показателем  $0 < v < 1$ ) и ненагруженном состояниях.

Данная система с  $m = m_1 + m_2 + m_3$  резервными элементами работает следующим образом. Если при функционировании отказывает один из  $k$  рабочих элементов, то он замещается резервным элементом, находящимся в нагруженном состоянии. Отказавший элемент поступает на восстановление, место восстанавливаемого элемента из числа  $m_1$  занимает элемент из группы  $m_2$ , а группа  $m_2$  пополняется из  $m_3$ . Данное состояние системы обозначается  $x_1$ , т.е. имеется один отказавший элемент. Из состояния  $x_1$  система может вернуться в состояние  $x_0$ , если отказавший элемент будет восстановлен (он пополнит группу  $m_3$ ), или перейти в состояние  $x_2$ , если откажет еще один элемент, и т.д. Следует заметить, отказывать могут не только рабочие элементы, но и элементы из групп  $m_1$  и  $m_2$ . Отказ системы  $\langle k, m_1, m_2, m_3, r \rangle$  наступает при достижении состояния  $x_{m+1}$ , т.е. когда в состоянии отказа будут находиться  $m_1 + m_2 + m_3 + 1$  элементов.

Если время работы элементов до отказа и время восстановления показательные с параметрами  $\lambda$  и  $\mu$  соответственно, то для решения задач моделирования кортеж состава  $\langle k, m_1, m_2, m_3, r \rangle$  дополняется кортежем интенсивностей отказов и восстановления  $\langle \lambda, \lambda, v\lambda, 0, \mu \rangle$ .

Для расчета вероятности безотказной работы  $R(t)$  системы, представленной на рис. 5.10, составляется ориентированный граф  $G_{\Pi}$  изменения состояний системы с поглощающим состоянием  $x_{m+1}$  (рис. 5.11, а). По

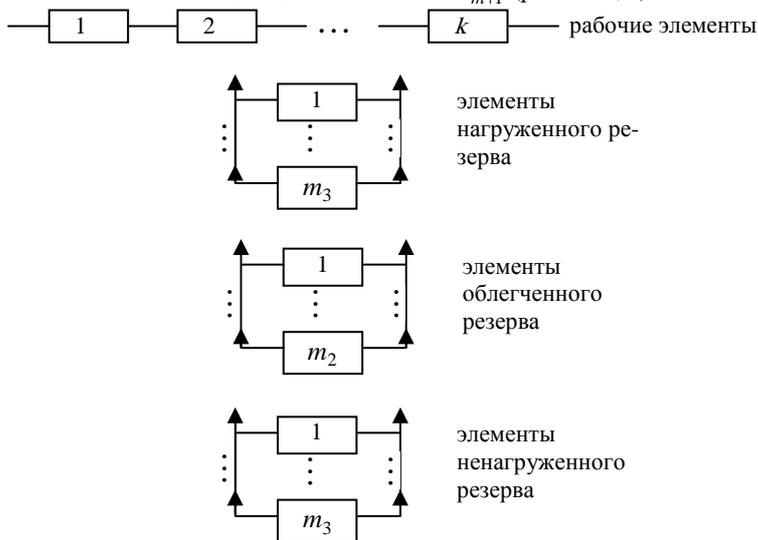


Рис. 5.10. Структурная схема надежности системы  $\langle k, m_1, m_2, m_3, r \rangle$

графу  $G_{\Pi}$  в предположении, что процесс изменения состояний марковский (см. разд. 5.3), записывается система из  $m+2$  дифференциальных уравнений. Интенсивности дуг  $\Lambda_j$  графа, связанные с отказами элементов, определяются из соотношения:

$$\Lambda_j = \begin{cases} k\lambda + m_1\lambda + m_2\nu\lambda, & j = \overline{0, m_3}, \\ k\lambda + m_1\lambda + (m_2 - (j - m_3))\nu\lambda, & j = \overline{m_3 + 1; m_2 + m_3}, \\ k\lambda + (m_1 - (j - m_2 - m_3))\lambda, & j = \overline{m_2 + m_3 + 1; m}, \end{cases}$$

а интенсивности дуг  $M_j$ , связанные с восстановлением,

$$M_j = \begin{cases} j\mu, & j \leq r, \\ r\mu, & j > r. \end{cases}$$

Система дифференциальных уравнений для расчета вероятностей состояний  $p_j(t)$ ,  $j = \overline{0, m+1}$  в компактной форме может быть записана следующим образом

$$\frac{dp_j(t)}{dt} = \Lambda_{j-1} p_{j-1}(t) - (\Lambda_j + M_j) p_j(t) + M_{j+1} p_{j+1}(t), \quad j = \overline{0, m+1}, \quad (5.66)$$

с учетом  $\Lambda_{-1} = \Lambda_{m+1} = M_0 = M_{m+1} = M_{m+2} = 0$ .

На основе решения системы уравнений (а) при начальных условиях

$$p_0(t) = 1, \quad p_j(t) = 0, \quad j = \overline{1, m+1}$$

определяются вероятность безотказной работы

$$R_{(m)}(t) = 1 - R_{m+1}(t)$$

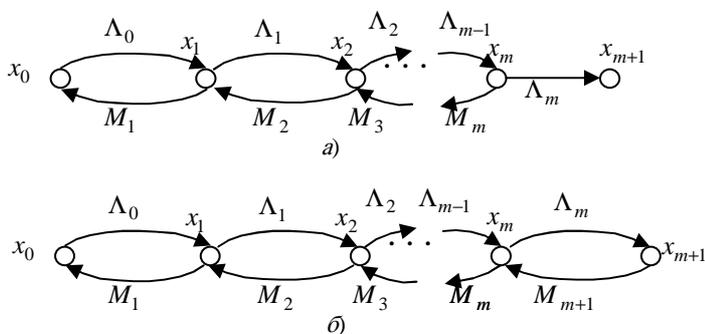


Рис. 5.11. Граф  $G_{\Pi}$  изменения состояний резервированной системы с поглощающим состоянием  $x_{m+1}$  (а) и граф  $G_0$  с отражающим состоянием  $x_{m+1}$  (б)

и средняя наработка системы между отказами

$$m_T = \int_0^{\infty} R_{(m)}(t) dt.$$

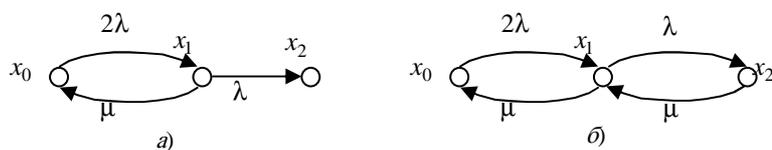
Для определения коэффициента готовности  $K_r(t)$  рассматриваемой системы граф  $G_0$  с отражающим состоянием  $x_{m+1}$  (рис. 5.11, б). Здесь также решается система уравнений (5.66), но при  $M_{m+1} \neq 0$ , т.е. с учетом

$$\Lambda_{-1} = \Lambda_{m+1} = M_0 = M_{m+2} = 0.$$

В результате решения системы уравнений определяются вероятность  $p_j(t)$ ,  $j = \overline{0, m+1}$  и нестационарный коэффициент готовности

$$K_r(t) = 1 - p_{m+1}(t).$$

**Пример 5.3.** Объект состоит из двух элементов – рабочего и резервного, резерв нагруженный. Пусть законы распределения времени работы элементов между отказами и восстановления показательные с параметрами  $\lambda$  и  $\mu$  соответственно. При отказе одного элемента он восстанавливается, объект отказывает, когда при восстановлении отказавшего элемента выходит из строя и другой элемент. Таким образом, имеют место кортежи  $\langle 1, 1, 0, 0, 1 \rangle$  и  $\langle \lambda, \lambda, 0, 0, \mu \rangle$ . Множество  $\mathcal{X}$  включает три состояния:  $x_0$  – работоспособны оба элемента;  $x_1$  – один элемент отказал и восстанавливается;  $x_2$  – отказали оба элемента, это состояние отказа объекта. На рис. 5.12 показаны графы переходов с поглощающим  $G_{\Pi}$  и отражающим  $G_0$  состоянием  $x_2$ . С помощью первого графа определяются вероятности безотказной работы объекта, с помощью второго – функция и коэффициент готовности.



**Рис. 5.12.** Графы переходов с поглощающим (а) и отражающим (б) состоянием системы  $\langle 1, 2, 0, 0, 1 \rangle$

Графу (рис. 5.12, а) соответствует следующая система дифференциальных уравнений

$$\begin{pmatrix} \dot{p}_0 \\ \dot{p}_1 \\ \dot{p}_2 \end{pmatrix} = \begin{pmatrix} -2\lambda & \mu & 0 \\ 2\lambda & -(\mu + \lambda) & 0 \\ 0 & \lambda & 0 \end{pmatrix} \begin{pmatrix} p_0(t) \\ p_1(t) \\ p_2(t) \end{pmatrix},$$

$$\dot{p}_i = \frac{\Delta}{dt} p_i(t), \quad i = 0, 1, 2, \quad p_i(t) = \overset{\Delta}{p}(x_i, t),$$

здесь  $p(x_i, t)$  – вероятность состояния  $x_i$  в момент времени  $t$ .

Решив уравнения при начальных условиях

$$p_0(0) = 1, \quad p_1(0) = p_2(0) = 0,$$

определяем вероятность безотказной работы

$$R^{(0)}(t) = p_0(t) + p_1(t) = 1 - p_2(t).$$

При начальных условиях  $p_1(0) = 1, p_0(0) = p_2(0) = 0$  получаем  $R^{(1)}(t)$ .

Для восстанавливаемых объектов с недопустимыми перерывами в работе могут использоваться также коэффициент оперативной готовности, средняя наработка системы до отказа при различных начальных условиях и др.

### Контрольные вопросы

1. Что понимается под системой массового обслуживания?
2. Какой поток событий считается простейшим?
3. В чем заключается свойство отсутствия последствия?
4. Какой поток событий называется пуассоновским?
5. Какие модели используются для описания марковских систем массового обслуживания?
6. Назовите показатели надежности ремонтируемых систем.
7. Какие способы применяются для повышения надежности систем?

## ЗАКЛЮЧЕНИЕ

Особое внимание в учебном пособии уделяется тому факту, что моделирование стало эффективным инструментом решения задач анализа, оптимизации и синтеза систем. Особенно эффективно применение математического моделирования на этапах проектирования систем жизнеобеспечения и защиты человека в экстремальных условиях, когда особенно высока цена ошибочных решений. В пособии сформулированы математические постановки задач исследования взаимосвязанных процессов адсорбции, тепло- и массопереноса при очистке и регенерации воздуха, приводятся аналитические и численные методы решения задач нестационарной адсорбции, теплопроводности и диффузии, алгоритмы и примеры расчета полей температур и концентраций в изолирующих средствах защиты органов дыхания людей.

Моделирование стало составной частью общих подходов, характерных для современных информационных технологий. Практическая реализация возможностей математического моделирования существенно повышает эффективность инженерных разработок особенно при создании принципиально новых, не имеющих прототипов изделий, технологических машин и приборов, материалов и технологий, что позволяет сократить затраты времени и средств на использование в технике и технологиях передовых достижений фундаментальных наук. Отмеченные возможности математического моделирования еще далеко не исчерпаны, представляются достаточно перспективными и заслуживают детального рассмотрения.

Принимая во внимание тенденции развития технического прогресса и возрастающие возможности вычислительной техники, дальнейшее развитие теории моделирования систем следует ожидать в направлении создания интеллектуальных самосовершенствующихся моделей, способных без участия разработчика изменять свою структуру и подстраиваться под требования новых задач.

## СПИСОК ЛИТЕРАТУРЫ

1. Дорохов, И.Н. Системный анализ процессов химической технологии / И.Н. Дорохов, В.В. Меньшиков. – М. : Изд. «Наука», 2005. – 584 с.
2. Советов, Б.Я. Моделирование систем / Б.Я. Советов, С.А. Яковлев. – М. : Высшая школа, 2005. – 343 с.
3. Самарский, А.А. Математическое моделирование / А.А. Самарский, А.П. Михайлов. – М. : Физматлит, 2001. – 320 с.
4. Норенков, И.П. Основы автоматизированного проектирования : учебник для вузов / И.П. Норенков. – 2-е изд., перераб. и доп. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2002. – 336 с.
5. Леоненков, А.В. Нечеткое моделирование в среде MATLAB и fuzzy TECH / А.В. Леоненков. – СПб. : БХВ – Петербург, 2003. – 736 с.
6. Островский, Г.М. Технические системы в условиях неопределенности : учебное пособие / Г.М. Островский, Ю.М. Волин. – М. : БИНОМ, Лаборатория знаний, 2008. – 319 с.
7. Halemane, K.R. Optimal Process Design under Uncertainty / K.R. Halemane, I.E. Grossmann // AIChE Journal, 1983. V.29. P. 425-433.
8. Swaney, R.E. An index for operational flexibility in chemical process design / R.E. Swaney, I.E. Grossmann // AIChE Journal, 1985. – V. 31. – P. 621 – 644.
9. Shapiro, A. A simulation-based Approach to Two-Step Stochastic Programming with Recourse / A. Shapiro, T.H. De-Mello. – Math. Program., 1998. – V. 81. – P. 301 – 305.
10. Бахвалов, Н.С. Численные методы / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. – Ч. – М. : БИНОМ. Лаборатория знаний, 2007.
11. Островский, Г.М. Новые проблемы теории гибкости химико-технологических процессов / Г.М. Островский, Ю.М. Волин. Доклады РАН, 2000. – Т. 370. – С. 773 – 776.
12. Дудников, Е.Г. Построение математических моделей химико-технологических объектов / Е.Г. Дудников., В.С. Балакирев, В.Н. Кривсунов, А.М. Цирлин. – М. : Химия, 1970. – 312 с.
13. Тихонов, А.Н. Уравнения математической физики : учебное пособие / А.Н. Тихонов, А.А. Самарский. – 4 изд. – М. : Наука, 1972. – 736 с.
14. Кельцев, Н.В. Основы адсорбционной техники / Н.В. Кельцев. – 2-е изд. перераб. и доп. – М: Химия, 1984. – 592 с.
15. Рачинский, В.В. Введение в общую теорию динамики сорбции и хроматографии / В.В. Рачинский. – М. : Наука, 1964. – 135 с.
16. Заде, Л. Понятие лингвистической переменной и ее применение к принятию приближенных решений / Л. Заде. – М. : Мир, 1976. – 167 с.

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ .....	3
1. ОСНОВЫ ТЕОРИИ МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ И СИСТЕМ .....	4

1.1. Принципы системного подхода .....	4
1.2. Общая характеристика проблемы моделирования и оптимизации систем .....	7
1.3. Основные понятия моделирования и оптимизации ХТС в условиях неопределенности .....	19
1.4. Методология компьютерного моделирования ХТС .....	41
1.5. Постановки задач моделирования и оптимизации ХТС в условиях неопределенности .....	48
<b>2. МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И ОПТИМИЗАЦИИ</b> .....	<b>68</b>
2.1. Классификация методов построения моделей систем .....	68
2.2. Экспериментальное определение моделей статики .....	71
2.3. Экспериментальное определение моделей динамики .....	81
2.4. Получение неформальных моделей систем .....	85
2.5. Математическое моделирование нестационарных процессов переноса .....	119
2.6. Методы решения задач оптимизации в условиях интервальной неопределенности .....	139
2.7. Логико-лингвистическое моделирование .....	150
<b>3. АДСОРБЦИОННЫЙ МЕТОД ОЧИСТКИ И РАЗДЕЛЕНИЯ СМЕСЕЙ</b> .....	<b>173</b>
3.1. Адсорбционное равновесие и структура промышленных адсорбентов .....	173
3.2. Механизмы адсорбции .....	185
3.3. Кинетика адсорбции .....	197
3.4. Процессы очистки и регенерации воздуха .....	204
3.5. Моделирование процессов очистки и регенерации воздуха в замкнутом объеме .....	217
3.6. Расчет температурных полей в потоке рабочей среды и в конструктивных элементах регенеративного теплообменника .....	227
<b>4. МОДЕЛИ ПРОИЗВОДСТВЕННЫХ СИСТЕМ</b> .....	<b>247</b>
4.1. Предприятие как объект моделирования .....	247
4.2. Модели систем управления запасами .....	255
4.3. Модели управления запасами .....	261
4.4. Модели принятия решений .....	269
<b>5. МОДЕЛИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ И НАДЕЖНОСТИ</b> .....	<b>285</b>
5.1. Общие сведения о моделях систем массового обслуживания .....	285
5.2. Модели потоков событий .....	287
5.3. Марковские системы массового обслуживания .....	293
5.4. Модели надежности систем .....	304
<b>ЗАКЛЮЧЕНИЕ</b> .....	<b>322</b>
<b>СПИСОК ЛИТЕРАТУРЫ</b> .....	<b>323</b>