

Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное
учреждение высшего образования

**«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ»**

Д.В.Озеркин, Е.М.Покровская

Основы научно-исследовательской деятельности

Учебное пособие
по дисциплине «Научно-исследовательская деятельность»
для обучающихся в аспирантуре

Томск 2018

СОДЕРЖАНИЕ

1. МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ НАУЧНОГО ПОЗНАНИЯ.....	3
1.1 Роль науки в современном обществе	3
1.2 Основные понятия научного познания	5
1.3 Методы научных исследований	8
2. ЭКСПЕРИМЕНТАЛЬНЫЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ.....	22
2.1 Методология эксперимента.....	22
2.2 Основы теории случайных ошибок и методов оценки случайных погрешностей в измерениях	26
2.2.1 Интервальная оценка с помощью доверительной вероятности	26
2.2.2 Определение минимального количества измерений	28
2.3 РЕГРЕССИОННЫЙ АНАЛИЗ	36
2.4 Методы графической обработки результатов эксперимента	40
2.5 Методы подбора эмпирических формул	44
2.6 Оценка адекватности результатов эксперимента	51
2.7 Метрологическое обеспечение эксперимента	55
3. МЕТОД ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА В НАУЧНЫХ ИССЛЕДОВАНИЯХ....	60
3.1 Основные понятия планирования эксперимента	60
3.2 Планирование эксперимента с целью описания исследуемого объекта	63
3.3 Оптимизация технологических процессов с использованием планирования эксперимента	75
4. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В НАУЧНЫХ ИССЛЕДОВАНИЯХ.....	77
4.1 Модели и моделирование	77
4.2 Этапы моделирования. Требования к моделям.....	82
4.3 Структуризация математических моделей сложных систем.....	86
4.4 Классификация математических моделей	100
5. ТЕОРЕТИЧЕСКИЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ	116
5.1 Задачи и виды теоретических исследований	116
5.2 Использование математических методов.....	120
5.3 Аналитические методы	135
5.4 Метод аналогии. Модели подобия.....	148
5.5 Вероятностно-статистические методы	159
6. ИНФОРМАТИКА В НАУЧНЫХ ИССЛЕДОВАНИЯХ	171
6.1 Основные направления информатики	171
6.2 Понятие информации.....	174
6.3 Цели информатизации	178
6.4 Основные проблемы информатизации	180
6.5 Новые информационные технологии.....	183
СПИСОК ЛИТЕРАТУРЫ	187

1. МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ НАУЧНОГО ПОЗНАНИЯ

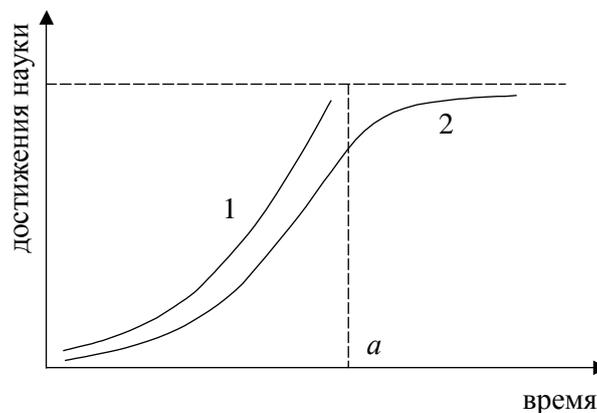
1.1 Роль науки в современном обществе

Термин «*наука*» Советский энциклопедический словарь определяет как «сферу человеческой деятельности, функция которой – выработка и теоретическая систематизация объективных знаний о действительности». При таком определении науки ее следует рассматривать с двух точек зрения. Процесс получения и систематизации научных знаний, т.е. процесс научных исследований, выступает как форма существования науки. Результат научной деятельности – теоретическое осмысленное систематизированное знание о действительности, в том числе и знание методов самих научных исследований, является содержанием самой науки. Таким образом, наука – это непрерывно развивающаяся система знаний об объективных законах природы, общества и мышления, которая создается и превращается в непосредственную практическую силу общества в результате специальной деятельности людей и учреждений. Наука немислима без единства ее формы – научных исследований, и содержания – объективного систематизированного знания. Изучением закономерностей функционирования науки, структуры и динамики научной деятельности, положения науки в жизни общества занимается специальная дисциплина, называемая *науковедением*.

Одна из главных *функций науки* и ее *целей* – познание объективного мира. Наука создана для непосредственного выявления существенных сторон всех явлений природы, общества и мышления. Наряду с этим ускорение прогресса общества в значительной мере зависит от развития науки. Современная наука является двигателем научно-технического прогресса. Она определяет его контуры и темпы развития. Внедрение науки в производство выражается в росте производительности труда, создании новых машин и материалов, улучшении эксплуатационных показателей, надежности и долговечности продукции, снижении ее себестоимости.

Современная наука имеет ряд характерных особенностей. Прежде всего, это бурное лавинообразное развитие.

Количество научных знаний о природе, обществе и мышлении все время возрастает. При этом количество вновь добываемых знаний прямо пропорционально уже известным. Анализ показывает, что основные характеристики научной деятельности за последние 250 лет возрастают по экспоненциальному закону [1]. Через каждые 10-15 лет все показатели удваиваются. Поэтому считают, что основным законом анализа науки является экспоненциальный (рисунок 1.1).



1 – экспонента; 2 – вероятная кривая

Рисунок 1.1 – Закономерности развития результатов научных исследований во времени

Многие ученые полагают, что экспоненциальный закон развития науки со временем должен измениться. Темпы привлечения ресурсов (люди, ассигнования и т.п.) будут замедляться и подчиняться кривой 2 (рисунок 1.1). Однако интенсивность использования ресурсов будет возрастать. Поэтому объем научной продукции, получение новой информации, по-видимому, будет приближаться к экспоненциальному закону.

Лавинообразность развития науки заключается и в систематическом создании новых ее видов, направлений, проблем. Возникает дерево науки. Каждое новое направление (ветвь) рождает новые проблемы. Разветвление науки во многих случаях сопровождается слиянием отдельных ее ветвей. Рождаются пограничные науки на стыке двух, трех и более. Так, в последнее время возникли новые науки – математическая кибернетика, вычислительная техника, криогенная техника, физико-химическая механика, биофизика, биогеохимия, математическая экономика, квалиметрия, прикладная информатика и др.

Важная особенность науки – ее рентабельность. Став непосредственно производительной силой, базой технического прогресса, наука является самой эффективной отраслью, обеспечивающей благодаря внедрению законченных разработок наибольший экономический эффект.

В настоящее время наблюдается интенсивное внедрение научных достижений в производство. Важнейшая задача научных организаций – систематическое внедрение законченных разработок в практическую деятельность. Для этого разрабатываются целевые комплексные программы во всех отраслях промышленности, объединяющие исследовательские институты, вузы, конструкторские бюро, заводы, строительные предприятия в научно-производственные комплексы. В научно-производственных объединениях возникновение новых научных идей определяется актуальной потребностью производства, а законченные научные разработки становятся базой коренной модификации производственных процессов, приводящей к созданию принципиально новых конструктивных решений и технологий, повышению эффективности и качества выпускаемой продукции.

Современная наука превращается в сложный и непрерывно растущий социальный организм, в наиболее подвижную практическую (производительную) силу общества. Развитие науки – исходный пункт для революционирования практики, создания новых отраслей производства и совершенствования общественных отношений.

Наука становится производительной силой общества, что проявляется в глубоких изменениях во взаимоотношениях науки и производства.

Во-первых, многие виды производства и технологические процессы первоначально зарождаются в недрах науки, научно-исследовательских институтах. Развитие атомной энергии, химической технологии, получение сверхтвердых материалов – все это подтверждает сказанное.

Во-вторых, сокращаются сроки между научным открытием и его внедрением в производство, а раньше для этого нужны были десятилетия. Например, от открытия лазера до его применения прошло всего несколько лет. Это же можно сказать об атомной энергии, полупроводниках и др.

В-третьих, в самом производстве успешно развиваются научные исследования, растет сеть научных учреждений в промышленности и сельском хозяйстве. Предприятия перерастают в научно-промышленные комплексы, включающие в свою технологию (наряду с конструированием и производством продукции) также и научные разработки.

1.2 Основные понятия научного познания

Знание – идеальное воспроизведение в различной языковой форме (математической, алгоритмической, лингвистической, экономической и др.) обобщенных представлений и закономерных связей объективного мира [2].

Функциями знания являются обобщение разрозненных представлений о закономерностях природы общества и мышления; хранение в обобщенных представлениях всего того, что может быть передано в качестве устойчивой основы практических действий. Знание является продуктом общественной деятельности людей, направленной на преобразование действительности.

Процесс движения человеческой мысли от незнания к знанию называют **познанием**, в основе которого лежит отражение объективной действительности в сознании человека в процессе его общественной, производственной и научной деятельности, именуемой **практикой**. Потребности практики выступают основной и движущей силой развития познания, его целью. Человек познает законы природы, чтобы овладеть силами природы и поставить их себе на службу; он познает законы общества, чтобы в соответствии с ними воздействовать на ход исторических событий.

Познание вырастает из практики, но затем само направляется на практическое овладение действительностью. От практики к теории и от теории к практике, от действия к мысли и от мысли к действительности – такова общая закономерность отношений человека в окружающей действительности. Практика является началом, исходным пунктом и одновременно естественным завершением всякого процесса познания. Завершение познания всегда относительно, так как в процессе познания, как правило, возникают новые проблемы и новые задачи, которые были подготовлены и поставлены предшествующим развитием научной мысли. Решая эти задачи и проблемы, наука должна опережать практику и таким образом сознательно направлять ее развитие.

В процессе практической деятельности человек разрешает противоречие между наличным положением вещей и потребностями общества. Результатом этой деятельности является удовлетворение общественных потребностей. Указанное противоречие является источником развития познания и, естественно, находит отражение в его диалектике.

Диалектика процесса познания выражается в противоречии между ограниченностью наших знаний и безграничной сложностью объективной действительности, между субъективной формой и объективным содержанием человеческого познания, в необходимости борьбы мнений, позволяющей путем логических доказательств и практической проверки устанавливать истину.

Вся наука, все человеческое познание направлены к достижению истинных знаний, верно отражающих действительность. Только истинное научное знание служит человеку инструментом преобразования действительности, позволяет прогнозировать ее дальнейшее развитие.

В противоположность истинному знанию **заблуждение** представляет собой неверное, иллюзорное отражение мира.

Истинные знания существуют в виде законов науки, теоретических положений и выводов, учений, подтвержденных практикой и существующих объективно, независимо от трудов и открытий ученых. Поэтому истинное научное знание объективно. Вместе с тем **научное знание** может быть относительным и абсолютным. **Относительное знание** – знание, которое, будучи в основном верным отражением действительности, отличается некоторой неполнотой совпадения образа с объектом. **Абсолютное знание** – это полное, исчерпывающее воспроизведение обобщенных представлений об объекте, обеспечивающее абсолютное совпадение образа с объектом. Абсолютное знание не может

быть опровергнуто или изменено в будущем, т.к. является, с системных позиций, религиозной моделью.

Следует отметить, что непрерывное развитие практики исключает возможность превращения знания в абсолютное, но абсолютность практики позволяет отличать объективно истинные знания от заблуждений.

Познание включает в себя два уровня: чувственный и рациональный. **Чувственное познание** формирует эмпирическое знание, а **рациональное** – теоретическое. Чувственное познание обеспечивает непосредственную связь человека с окружающей действительностью. Элементами чувственного познания являются ощущение, восприятие, представление и воображение.

Ощущение – это отражение мозгом человека свойств предметов или явлений объективного мира, которые действуют на его органы чувств. **Восприятие** – отражения мозгом человека предметов или явлений в целом, причем таких, которые действуют на органы чувств в данный момент времени. Восприятие – это первичный чувственный образ предмета или явления. **Представление** – вторичный образ предмета или явления, которые в данный момент времени не действуют на органы чувств человека, но обязательно действовали в прошлом. Представления – это образы, которые восстанавливаются по сохранившимся в мозге следам прошлых воздействий предметов или явлений. **Воображение** – это соединение и преобразование различных представлений в целую картину новых образов.

Рациональное познание дополняет и опережает чувственное, способствует осознанию сущности процессов, вскрывает закономерности развития. Формой рационального познания является абстрактное мышление.

Мышление – это опосредованное и обобщенное отражение в мозгу человека существенных свойств, причинных отношений и закономерных связей между объектами или явлениями. Опосредованный характер мышления заключается в том, что человек через доступные органам чувств свойства, связи и отношения предметов проникает в скрытые свойства, связи, отношения; человек познает действительность не только в результате своего личного опыта, но и косвенным путем, усваивая ее в процессе общения с другими людьми. Мышление неразрывно связано с языком и не может осуществляться вне его. Основной инструмент мышления – логические рассуждения человека, структурными элементами которых (и формами логического отражения действительности) являются понятия, суждения, умозаключения.

Понятие – это мысль, отражающая существенные и необходимые признаки предмета или явления. Понятия могут быть общими, единичными, собирательными, абстрактными и конкретными, абсолютными и относительными. **Общие понятия** связаны не с одним, а с множеством предметов. Наиболее широкие понятия называются **категориями** и к ним относят некоторые философские понятия (о форме и содержании явлений), политэкономии (товар, стоимость) и т.д. **Единичные понятия** относятся всегда только к одному определенному предмету. Под **собирательными** подразумеваются понятия, обозначающие целые группы однородных предметов, представляющих собой известное единство, законченную совокупность (лес, транспортный поток и т.п.).

Понятия конкретные относятся к конкретным предметам, а **абстрактные понятия** – к отдельно взятым признакам этих предметов, например «белые предметы». Особенностью относительных понятий является то, что они всегда мыслятся попарно, например: «правый» и «левый», «начальник» и «подчиненный». Абсолютными называют такие понятия, которые не имеют парных отношений, например «планета», «дом», «дерево».

По признаку отношений между понятиями их делят на тождественные, равнозначные, подчиненные, соподчиненные, частично согласные, противоречащие и противоположные.

Тождественными называют такие понятия, которые имеют одинаковое содержание. Это одни и те же понятия, только выраженные в различной словесной форме. **Равнозначные** понятия имеют один и тот же объем, но отличаются по содержанию.

Понятия характеризуются их объемом и содержанием. Объем понятия – это круг тех предметов, на которые данное понятие распространено. Содержанием называют совокупность признаков, которые объединены в данном понятии.

Отношения тождества и равнозначности понятий имеют чрезвычайно важное значение в науке, так как делают возможным замещение одного понятия другим. Этой операцией широко пользуются в математике при преобразовании и упрощении алгебраических соотношений.

Подчиненными называют понятия, которые по содержанию входят в понятия более высокого ранга или более общие. **Соподчиненными** являются понятия, связанные по объему (объем двух или более понятий входит в объем какого-либо высшего понятия). Например, понятия «многоугольник» и «окружность» являются подчиненными понятию «геометрическая фигура» и соподчиненными между собой. Если отдельные части объема понятий оказываются совпадающими, общими, то их называют **частично согласными**. В подобном отношении находятся, например, такие понятия, как «студент» и «спортсмен».

Понятие, которое отрицает положительное понятие, называют **противоречащим**. Например, понятие «нечеловек» отрицает положительное понятие «человек». Противоречащие понятия не допускают ничего промежуточного; одно понятие полностью исключает другое. Если понятие указывает не только на то, что отрицает, но и на то, что взамен отрицаемого утверждается, то такое понятие называют **противоположным**. У противоположных понятий имеются средние и промежуточные понятия. Так, между понятиями «белый» и «черный» мыслимо понятие «серый».

Для описания процесса формирования новых сложных понятий из более простых используется способ вывода сложных соотношений из элементарных. Формализация процесса часто осуществляется на языке теории множеств.

Раскрытие содержания понятия называют его **определением**. Последнее должно отвечать двум важнейшим признакам: 1) определение должно указывать на ближайшее родовое понятие; 2) определение должно указывать на то, чем данное понятие отличается от других понятий. Так, определяя понятие «квадрат», нужно указать на то, что квадрат относится к роду прямоугольников и выделяется среди прямоугольников признаком равенства своих сторон. Определение понятия не должно быть ни слишком широким, ни слишком узким, т.е. соразмерным и не должно определяться самим собой, т.е. определение понятия не должно делать круга.

Развитие научных знаний заставляет уточнять определение понятий, вносить новые признаки в его содержание. При этом понятие обобщается или ограничивается. В научном исследовании определения обычно завершают процесс исследования, закрепляют те результаты, к которым ученый пришел в своем исследовании. Без определения понятий возможно ложное толкование мыслей автора исследования. Определение понятия оказывается возможным в том случае, когда мы знаем, к какому роду оно относится и какие у него видовые признаки. Установление видовых признаков осуществляется при помощи деления понятия. **Делением понятия** называется раскрытие всех видов, входящих в состав данного понятия. Если определение имеет дело с содержанием изучаемого понятия, то деление – с объемом понятия.

Деление подчиняется следующим правилам: 1) члены деления должны исчерпывать объем делимого понятия; 2) деление должно производиться с точки зрения одного определенного основания; 3) члены деления должны исключать друг друга.

Основанием деления называется тот признак, который является общим всем видам, входящим в объем данного понятия. Особым видом деления понятий является дихотомия,

или двучленное деление, при котором членами деления бывают только два понятия, из которых одно является противоречащим в отношении другого.

Суждение – это мысль, в которой посредством связи понятий утверждается или отрицается что-либо. В речи суждение выражается в виде предложения. Суждение – это сопоставление понятий, устанавливающих объективную связь между мыслимыми предметами и их признаками или между предметом и классом предметов.

Суждения делятся по следующим признакам: качеству, количеству, отношению, модальности. В свою очередь, по качеству суждения делятся на утвердительные и отрицательные, по количеству – на общие, частные и единичные, по отношению – на категорические, условные и разделительные, по модальности – на проблематические, аподиктические и ассерторические. В проблематических суждениях наличие связи понятий отмечается лишь с известной степенью вероятности. В аподиктических суждениях указывается, что связь понятий является безусловно необходимой. Ассерторические суждения указывают только на действительно существующую связь понятий.

Соединение суждений по количеству и качеству приводит к четырем новым видам суждений: общеутвердительному, общеотрицательному, частноутвердительному и частноотрицательному.

К суждению о предмете или явлении человек может прийти или путем непосредственного наблюдения какого-либо факта, или опосредованным путем – с помощью умозаключения. **Умозаключение** – процесс мышления, составляющий последовательность двух или нескольких суждений, в результате которых выводится новое суждение. Часто умозаключение называют выводом, через который становится возможным переход от мышления к действию, практике. Вместе с тем следует подчеркнуть, что не всякая последовательность суждений может быть названа умозаключением или выводом. В умозаключении связь двух суждений иногда обнаруживает подчинение, в силу которого одно (основание) обуславливает другое (следствие).

Умозаключения подразделяются на непосредственные и опосредованные. В непосредственных умозаключениях от одного суждения приходят к другому. В опосредованных суждениях переход от одного суждения к другому осуществляется через посредство третьего. Если в процессе умозаключения изменяется форма суждения, то говорят об ее превращении, например, утвердительное суждение становится отрицательным, и наоборот. При этом смысл и количество суждения сохраняются. Понятия, суждения и умозаключения выражаются в словесной форме.

1.3 Методы научных исследований

Формой осуществления и развития науки является **научное исследование**, т.е. изучение явлений и процессов, анализ влияния на них различных факторов, а также изучение взаимодействия между явлениями с целью получить убедительно доказанные и полезные для науки и практики решения с максимальным эффектом [1]. Научное исследование имеет объект, предмет, на познание которого оно направлено. Объектом (предметом) исследования может быть предмет материального мира, явление, свойства, а также связь между явлениями и свойствами.

Цель научного исследования – определение конкретного объекта и всестороннее, достоверное изучение его структуры, характеристик, связей на основе разработанных в науке принципов и методов познания, а также получение полезных для деятельности человека результатов, внедрение в производство и получение эффекта.

Материалистический принцип науки обуславливает ее предмет – законы природы, общества и мышления.

В каждом научном исследовании можно выделить два уровня:

- 1) эмпирический, на котором идет процесс накопления фактов;
- 2) теоретический, на котором достигается синтез знания (в форме научной теории).

В процессе научного исследования можно отметить следующие этапы: возникновение идей; формирование понятий, суждений; выдвижение гипотез; обобщение научных фактов; доказательство правильности гипотез и суждений.

Научная идея – интуитивное объяснение явления без промежуточной аргументации, без осознания всей совокупности связей, на основании которой делается вывод. Она базируется на уже имеющемся знании, но вскрывает ранее не замечаемые закономерности. Свою специфическую материализацию идея находит в гипотезе.

Научная гипотеза – это предположение о причине, которая вызывает данное следствие. Если гипотеза согласуется с наблюдаемыми фактами, то в науке ее называют теорией или законом. В процессе познания каждая гипотеза подвергается проверке, в результате которой устанавливается, что следствия, вытекающие из гипотезы, действительно совпадают с наблюдаемыми явлениями, что данная гипотеза не противоречит никаким другим гипотезам, которые считаются уже доказанными. Следует, однако, подчеркнуть, что для подтверждения правильности гипотезы необходимо убедиться не только в том, что она не противоречит действительности, но и в том, что она является единственно возможной и с ее помощью вся совокупность наблюдаемых явлений находит себе вполне достаточное объяснение.

С накоплением новых фактов одна гипотеза может быть заменена другой лишь в том случае, если эти новые факты не могут быть объяснены старой гипотезой или ей противоречат. При этом часто старая гипотеза не отбрасывается целиком, а только исправляется и уточняется. По мере уточнения и исправления гипотеза превращается в закон.

Закон – внутренняя существенная связь явлений, обуславливающая их необходимое закономерное развитие. Закон выражает определенную устойчивую связь между явлениями или свойствами материальных объектов.

Закон, найденный путем догадки, должен быть затем логически доказан, только тогда он признается наукой. Для доказательства закона наука использует суждения, которые были ранее признаны истинными и из которых логически следует доказываемое суждение. В редких случаях в равной мере оказываются доказуемыми противоречивые суждения. В таких случаях говорят о возникновении парадокса в науке, что всегда свидетельствует о наличии ошибок в логике доказательства или несостоятельности исходных суждений в данной системе знаний.

Парадокс в широком смысле – это утверждение, резко расходящееся с общепринятым, установившемся мнением, отрицание того, что представляется «безусловно правильным».

Парадокс в узком смысле – это два противоположных утверждения, для каждого из которых имеются представляющиеся убедительными аргументы.

Парадоксальность является характерной чертой современного научного познания мира. Наличие парадоксов становится свидетельством несостоятельности существующих теорий, требованием дальнейшего их совершенствования.

Выявление и разрешение парадоксов стало в современной науке обычным делом. Основные пути их разрешения: устранение ошибок в логике доказательств; совершенствование исходных суждений в данной системе знаний.

Для избежания ошибок логика доказательства должна быть подчинена **законам формальной логики**: закону тождества; закону противоречия; закону исключения третьего и закону достаточного основания.

Закон тождества: объем и содержание мысли о каком-либо предмете должны быть строго определены и оставаться постоянными в процессе рассуждения о нем. **Закон противоречия:** в процессе рассуждения о каком-либо определенном предмете нельзя одновременно утверждать и отрицать что-либо в одном и том же отношении, в противном случае оба суждения не могут быть вместе истинными.

Закон исключения третьего: в процессе рассуждения необходимо доводить дело до определенного утверждения или отрицания, в этом случае истинным оказывается одно из двух отрицающих друг друга суждений. Этот закон имеет силу лишь при условии соблюдения законов тождества и противоречия. **Закон достаточного основания:** в процессе рассуждения достоверными следует считать лишь те суждения, относительно истинности которых могут быть приведены достаточные основания.

В результате проработки и сопоставления с действительностью научная гипотеза может стать теорией.

Теория (от лат. *theoreo* – рассматриваю) – система обобщенного знания, объяснения тех или иных сторон действительности. Теория является духовным, мысленным отражением и воспроизведением реальной действительности. Она возникает в результате обобщения познавательной деятельности и практики. Это обобщенный опыт в сознании людей.

К теории предъявляются следующие требования [3]:

1) научная теория должна быть адекватна описываемому объекту, что позволяет в определенных пределах заменять экспериментальные исследования теоретическими;

2) теория должна удовлетворять требованию полноты описания некоторой области действительности;

3) должны быть объяснены взаимосвязи между различными компонентами в рамках самой теории. Должны существовать связи между различными положениями теории, обеспечивающие переход от одних утверждений к другим;

4) должно выполняться требование внутренней непротиворечивости теории и соответствия ее опытным данным.

Теория должна быть эвристичной, конструктивной и простой. Эвристичность теории отражает ее предсказывательные и объяснительные возможности. Математический аппарат теории должен позволять не только делать точные количественные предсказания, но и открывать новые явления. Конструктивность теории состоит в простой, совершаемой по определенным правилам проверяемости основных ее положений, принципов и законов. Простота теории достигается путем введения обобщенных законов сокращения и уплотнения информации при помощи специальных символов.

Структуру теории формируют принципы, аксиомы, законы, суждения, положения, понятия, категории и факты. Под **принципом** в научной теории понимается самое абстрактное определение идеи (начальная форма систематизации знаний). Принцип – это правило, возникшее в результате субъективно осмысленного опыта людей.

Исходные положения научной теории называются постулатами или аксиомами.

Аксиома (постулат) – это положение, которое берется в качестве исходного, недоказуемого в данной теории, и из которого выводятся все остальные предложения и выводы теории по заранее фиксированным правилам. Аксиомы очевидны без доказательства. В современной логике и методологии науки постулат и аксиома обычно используются как эквивалентные.

Теория строится из относительно жесткого ядра и его защитного пояса. В ядро входят основные принципы. Защитный пояс теории содержит вспомогательные гипотезы, конкретизирующие ее ядро. Этот пояс определяет проблемы, подлежащие дальнейшему исследованию, предвидит факты, не согласующиеся с теорией, и истолковывает их так, что они превращаются в примеры, подтверждающие ее.

Теория является наиболее развитой формой обобщенного научного познания. Она включает в себе не только знания основных законов, но и объяснение фактов на их основе. Теория позволяет открывать новые законы и предсказывать будущее.

Отличительная черта науки заключается в необходимости подвергать выдвинутые гипотезы экспериментальной проверке. Именно экспериментальная проверка, признание того, что можно изменить свое мнение, если к этому принуждают факты, является особенностью науки. Чтобы факты могли заставить изменить мнение исследователя, прежде всего, должны существовать сами факты: заинтересованные наблюдатели должны иметь возможность прийти к согласию в том, что является фактом, а что нет. **Факты** должны влиять на степень уверенности в правильности теории [4]. Экспериментальные факты, которые согласуются с теорией, на самом деле не «доказывают» ее правильности и, наоборот, даже если факты не согласуются с теорией, они не всегда «доказывают», что теория неверна. Проверка теорий часто оказывается тонким и деликатным делом, и в научных исследованиях никогда не достигается абсолютная уверенность в их правильности или ложности. Однако эксперимент оправдан, если в зависимости от исхода он меняет степень уверенности в правильности теории. Если уверенность в правильности теорий не может быть изменена в результате эксперимента, такие концепции не являются частью науки.

Факты окружают человека – это вещи вокруг, которые можно видеть, ощущать, слышать, обонять. Обыденный взгляд на факты, как на неизбежные базисные данные о существующем, не учитывает того, что и в самом простом акте восприятия есть большой компонент научения и опыта. При объяснении научного метода исследователь старается по мере возможности основываться на здравом смысле и повседневном опыте. Обращение к здравому смыслу позволяет лучше передать идеи, объяснить нечто новое и незнакомое с помощью хорошо известного.

Имеется существенное различие между общепринятым взглядом на «факты» как на жесткие, неизбежные, неизменные вещи и реальным положением в науке, где факты менее определены. Факты имеют обусловленную культурой компоненту и до некоторой степени создаются имеющимися теориями и поэтому подвержены изменениям, если меняются сами теории. Если проследить историю развития естественных наук, то можно заметить, что вся она – от Аристотеля до Эйнштейна и Бора – насыщена разрешением противоречий между прямым опытом как совокупностью наблюдаемых фактов и формально-логическими схемами, призванными объяснять эти факты. Естественно стремление при этом из множества формально-логических схем выбрать, в конечном счете, ту, которая наиболее просто и совершенно интерпретирует эти факты.

Согласно общепринятому взгляду наука оперирует набором экспериментально проверяемых фактов, определенным образом упорядоченных. В науке есть общие утверждения, обладающие объяснительной силой, из которых можно вывести множество проверяемых фактов.

Факты не являются действительно независимыми от наблюдателя, его теорий и предпочтений. Тем не менее, в любую конкретную эпоху в любой конкретной культуре большая часть наблюдателей достигает согласия в их трактовке. Факты – это то, в чем согласны все наблюдатели. Это утверждение подразумевает важную особенность научных фактов: у фактов должно быть более одного наблюдателя, должна существовать группа наблюдателей, к которой может присоединиться любой человек. Нередко нельзя оценить истинность наблюдения, не изучив множества вещей, которые большая часть людей не знает автоматически. Нужно быть не просто наблюдателем, а подготовленным и заинтересованным наблюдателем.

Из различий исследуемых объектов вытекают важные практические следствия: наблюдения и выводы в отношении этих объектов очень субъективны и неточны, поэтому требуется их **классификация**. Исходным пунктом научной деятельности является

вычленение проблемы, наблюдение фактов, адекватных ей, классификация этих фактов в связную систему.

Широко распространены неверные оценки роли фактов и их классификаций в науке. Одна из них – убеждение в том, что начальным этапом науки является собирание фактов; лишь после того как они собраны и приведены в определенную систему начинается научное исследование.

На самом деле в мире существует бесчисленное множество фактов, достойных исследования, и необходимо начинать с определенного представления о том, какие из них важны, а какие – нет. Необходимо отбирать факты в соответствии с критериями важности, которые могут основываться на имеющихся у исследователя теориях или чаще на смутном субъективном чувстве, что определенные явления стоит принимать во внимание, а другие – нет. В этом выборе есть очевидный элемент риска.

Проблема классификации фактов аналогична проблеме их отбора. Так же как отбор фактов определяется предубеждениями и теориями, точно так же упорядочение фактов и выявление тех или иных закономерностей зависит от прошлых субъективных критериев.

Движение мысли от незнания к знанию руководствуется методологией. **Методология** – философское учение о методах познания и преобразования действительности, применение принципов мировоззрения к процессу познания, духовному творчеству и практике.

В методологии выявляются две взаимосвязанные функции:

1) обоснование правил применения мировоззрения к процессу познания и преобразования мира;

2) определение подхода к явлениям действительности. Первая функция общая, вторая – частная.

Общая функция базируется на обобщении системы взглядов человека на мир в целом, на место отдельных явлений в мире и на свое собственное место в нем. В зависимости от совокупности научных, политических, правовых, нравственных, религиозных, эстетических убеждений обобщенная система взглядов может носить идеалистический или материалистический характер. Процесс познания органически связан с предметами материального мира, с их движением и развитием. Только такой подход к изучению окружающей нас действительности дает возможность человеку правильно познать материальный мир. Познание есть вечное, бесконечное приближение мышления к объекту. Наука постепенно, диалектически развертывает естественно-научную картину мира, глубже познает ее законы.

Одной из основных задач познания является задача выявления причин изменения и развития конкретных явлений и процессов. Диалектический подход к познанию указывает, что источниками, причинами развития являются внутренние противоречия и борьба противоположностей, которые составляют основу процессов объективной действительности.

В этих процессах единство всегда относительно, временно, преходяще, а борьба взаимоисключающих противоположностей абсолютна, как абсолютно развитие каждого явления, его движения.

Противоположности в науке проявляются в различных формах, вытекающих из конкретно поставленных задач. Это новое и старое, положительное и отрицательное, консервативное и революционное. Новое, положительное и революционное, как более совершенное, пробивает себе дорогу в борьбе со старым, отжившим. Не понимать этого и не изучать с позиций этого закона факты и явления – значит, никогда не подойти к истине.

Не менее важным в процессе познания является вопрос о том, как на основе внешнего воздействия идет процесс усложнения структуры изучаемого объекта или явления, как появляются новые качества?

Поступательный характер, преемственность и тенденции развития объекта позволяют вскрыть третий закон диалектики – отрицание отрицания. Отрицание не отбрасывает все старые представления и взгляды, отрицается то, что исчерпало возможности роста (что устарело, отжило), и удерживается то, что растет и развивается. Одним актом отрицания процесс диалектического движения не завершается. После первого отрицания в силу действия других законов диалектики, в частности закона единства и борьбы противоположностей, в сознании исследователя возникают новые взгляды. Борьба между ними приводит к следующему отрицанию и т.д. Наступает отрицание отрицания.

С общих позиций теории познания различают два уровня знания: эмпирическое и теоретическое [5]. К *эмпирическому знанию* относят знание такого уровня, которое получается непосредственно из опыта путем некоторой обработки экспериментальных данных и данных прямых наблюдений. Эти данные являются базисом, из которого исходит познание в своем последующем развитии. Именно поэтому достоверности эмпирического знания придается особое значение. Степень достоверности полученных результатов должна допускать количественную оценку, причем часто достаточно высокую. Добиться этого можно, лишь правильно выбрав план эксперимента, его структуру и условия проведения, применив соответствующую аппаратуру и методику обработки экспериментальных данных. Такой выбор желательно проводить не на интуитивной основе, а на основе разработанных к настоящему времени теории планирования эксперимента, методов испытаний, использования статистических методов обработки экспериментальных данных. При выполнении соответствующих условий это обеспечивает наивысшую достоверность получаемых выводов и оптимальные оценки искомых параметров исследуемых процессов.

Теоретический уровень знания отражает объект исследования со стороны взаимосвязей его элементов как между собой, так и с внешней средой. Этот уровень получается не только непосредственно в опыте, но и путем абстрактного мышления. Научные исследования теоретического плана часто выходят за пределы экспериментального эмпирического знания, хотя, в конечном счете, лишь опыт, практика могут дать окончательное заключение о правильности теоретических построений. А.Эйнштейн предъявлял к построению научной теории два основных критерия. Первый из них – критерий «внешнего оправдания». Это, фактически, требование согласия теории с опытом. Второй критерий Эйнштейн определяет как ее «внутреннее совершенство». Теория, обладающая «внутренним совершенством», по Эйнштейну, по сравнению с другими теориями исходит в наименьшей степени из произвольных предположений. Сами по себе наблюдения не определяют теории единственным образом. При выборе научной теории, по Эйнштейну, наше сознание действует активно, исходит из критерия внутреннего совершенства, из изящества теории, опирающейся на малое число независимых посылок.

Очевидно, что эмпирическое и теоретическое знание тесно связаны. Эмпирическое знание чаще всего дает основу построения теории. Теория, если она прямо не опирается на эксперимент, в своем развитии может дать лишь некоторую модель явления, проверку справедливости которой дают только данные эксперимента. Поэтому в большинстве научных исследований эксперимент и теория переплетаются. Эксперимент редко ставится наугад. Как правило, еще до его постановки выдвигается некоторая гипотеза относительно изучаемого процесса, строится его математическая модель, параметры которой, так же как и проверка адекватности самой модели, подлежат определению. Когда опыт дает нам подтверждение гипотезы, модели, теории, мы получаем новое знание, критерием истинности которого является дальнейшая практика. Если опыт вступает в противоречие с исходными посылками, его результаты становятся источником новой теории. Таким образом, научные гипотезы, математические модели процессов, сами эксперименты несут

в себе эвристическое начало и являются элементами системы, формирующей новое знание.

Нельзя, конечно, творческое начало, которое заложено в самом процессе познания, ограничить некоторой раз и навсегда заданной схемой. Как экспериментальные, так и теоретические исследования в науке являются творческим процессом, и рождение нового знания часто происходит вопреки тому, что ожидают исследователи. С другой стороны, есть много примеров, когда новое знание появлялось вначале в теориях, для рождения которых, казалось, не было никаких экспериментальных оснований.

Начиная любое исследование, необходимо выдвинуть свою идею, сформулировать научную гипотезу относительно изучаемых процессов. Хорошо, если на основании общих уже известных закономерностей удастся довести эту гипотезу до некоторой математической модели явления, затрагивающей суть явления, дающей его теорию. Хуже, если относительно сути явления, на данном уровне знания, сказать что-либо определенное трудно. Тогда следует построить хотя бы формальную математическую модель, которая позволила бы грамотно спланировать и провести экспериментальные исследования. Теория в этом случае может возникнуть лишь позже, на основе полученного эмпирического знания. А может случиться и так, что выдвинутая гипотеза окажется просто неверной, но тогда эксперимент дает основу для выдвижения новых идей и гипотез. Таким образом, как эксперимент, так и теоретическое знание одинаково важны для науки, значимость уровня знания, в конечном счете, определяется не тем, как он добыт – теоретически или экспериментально, а тем, как в нем отражен объект исследования во всех своих взаимосвязях и прежде всего отражен с тех сторон, которые определяют возможности его использования на практике. В конечном счете, лишь одна общественно-историческая практика является критерием истинности знания.

Метод – это способ (инструмент) достижения цели. Метод объединяет субъективные и объективные моменты познания. Метод объективен, так как в разрабатываемой теории позволяет отражать действительность и ее взаимосвязи. Таким образом, метод является программой построения и практического применения теории. Одновременно метод субъективен, так как является орудием мышления исследователя и в качестве такового включает в себя его субъективные особенности.

С философской точки зрения методы можно разделить на: всеобщие (материалистическая диалектика), действующие во всех областях науки и на всех этапах исследования; общенаучные (т.е. для всех наук); частные (т.е. для определенных наук); специальные или специфические (для данной науки). Разнообразные **методы научного познания** условно подразделяются на ряд уровней: эмпирический, экспериментально-теоретический, теоретический и метатеоретический уровни [2]. По мере развития познания один научный метод может переходить из одной категории в другую.

Методы эмпирического уровня: наблюдение, сравнение, измерение, счет, анкетный опрос, собеседование, тесты, метод проб и ошибок и т.д. Методы этой группы конкретно связаны с изучаемыми явлениями и используются на этапе формирования научной гипотезы.

Наблюдение – это способ познания объективного мира, основанный на непосредственном восприятии предметов и явлений при помощи органов чувств без вмешательства в процесс со стороны исследователя. Чтобы быть плодотворным, наблюдение должно удовлетворять следующим требованиям:

- а) преднамеренности (наблюдение ведется для определенной, четко поставленной задачи);
- б) планомерности (производится по плану, составленному по задачам наблюдения);
- в) целенаправленности (наблюдаются только интересующие стороны явления);
- г) активности (наблюдатель активно ищет нужные объекты, черты явления);

д) систематичности (наблюдение ведется непрерывно или по определенной системе).

Наблюдение как метод познания позволяет получать первичную информацию в виде совокупности эмпирических утверждений. Эмпирическая совокупность дает первичную схематизацию объектов реальности, что и является исходными объектами научного исследования.

Сравнение – это установление различия между объектами материального мира или нахождение в них общего, осуществляемое как при помощи органов чувств, так и при помощи специальных устройств.

Метод сравнения будет плодотворным, если выполняются следующие требования:

а) могут сравниваться только такие явления, между которыми может существовать определенная объективная общность;

б) сравнение должно осуществляться по наиболее важным, существенным (в плане конкретной задачи) признакам.

С помощью сравнения информацию об объекте можно получить двумя путями:

1) непосредственный результат сравнения (первичная информация);

2) результат обработки первичных данных (вторичная или производная информация).

Измерение – это физический процесс определения численного значения некоторой величины путем сравнения ее с эталоном.

Измерение предполагает наличие следующих основных элементов: объекта измерения, эталона, измерительных приборов, метода измерения.

Измерение развилось из операции сравнения, тем не менее оно является мощным и универсальным познавательным средством.

Счет – это нахождение числа, определяющего количественное соотношение однотипных объектов или их параметров, характеризующих те или иные свойства.

Методы экспериментально-теоретического уровня: эксперимент, анализ и синтез, индукция и дедукция, аналогия, моделирование, гипотетический, исторический и логические методы. Эти методы помогают исследователю обнаружить те или иные достоверные факты, объективные проявления в протекании исследуемых процессов. С помощью этих методов производится накопление фактов, их перекрестная проверка. Следует при этом подчеркнуть, что факты имеют научно-познавательную ценность только в тех случаях, когда они систематизированы, когда между ними вскрыты неслучайные зависимости, определены причины следствия. Таким образом, задача выявления истины требует не только сбора фактов, но и правильной их теоретической обработки. Первоначальная систематизация фактов и их анализ проводятся уже в процессе наблюдений, бесед, экспериментов, ибо эти методы включают не только акты чувственного восприятия предметов и явлений, но и их отбор, классификацию, осмысливание воспринятого материала, его фиксирование.

Эксперимент – это такой метод изучения объекта, когда исследователь активно и целенаправленно воздействует на него путем создания искусственных условий или использования естественных условий, необходимых для выявления соответствующих свойств. Эксперимент – одна из сфер человеческой практики, в которой подвергается проверке истинность выдвигаемых гипотез или выявляются закономерности объективного мира. Преимущества экспериментального изучения объекта по сравнению с наблюдением следующие:

а) в процессе эксперимента можно изучать явление «в чистом виде», устранив побочные факторы, затемняющие основной процесс;

б) в экспериментальных условиях можно исследовать свойства объектов;

в) повторяемость эксперимента: можно проводить испытания столько раз, сколько это необходимо.

Эксперимент проводят в следующих случаях:

- 1) при попытке обнаружения у объекта ранее неизвестных свойств;
- 2) при проверке правильности теоретических построений;
- 3) при демонстрации явления.

В научном исследовании эксперимент и теория теснейшим образом взаимосвязаны. Всякое игнорирование эксперимента неизбежно ведет к ошибкам, поэтому всемерное развертывание экспериментальных исследований представляет собой один из наиболее важных путей развития всей современной науки.

Анализ – метод познания при помощи расчленения или разложения предметов исследования (объектов, свойств и т.д.) на составные части. В связи с этим анализ составляет основу аналитического метода исследования.

Синтез – соединение отдельных сторон предмета в единое целое. Анализ и синтез взаимосвязаны, они представляют собой единство противоположностей.

Анализ и синтез бывают:

- 1) прямой, или эмпирический (используется для выделения отдельных частей объекта, обнаружения его свойств, простейших измерений и т.п.);
- 2) возвратный, или элементарно-теоретический (базируется на некоторых теоретических соображениях причинно-следственной связи различных явлений или действия какой-либо закономерности). При этом выделяются и соединяются явления, представляющиеся существенными, а второстепенные игнорируются;
- 3) структурно-генетический (требует вычленения в сложном явлении таких элементов, которые оказывают решающее влияние на все остальные стороны объекта).

Важными понятиями в теории познания являются: **индукция** – умозаключение от фактов к некоторой гипотезе (общему утверждению) и **дедукция** – умозаключение, в котором вывод о некотором элементе множества делается на основании знания общих свойств всего множества. Таким образом, дедукция и индукция – взаимообратные методы познания, широко использующие частные методы формальной логики, а именно:

1. Метод единственного сходства. Если два и более случаев исследуемого явления имеют общим лишь одно обстоятельство, а все остальные обстоятельства различны, то это единственное сходное обстоятельство и является причиной рассматриваемого явления.

2. Метод единственного различия. Если случай, в котором исследуемое наступает, и случай, в котором оно не наступает, во всем сходны и различны только в одном обстоятельстве, то это обстоятельство, присутствующее в одном случае и отсутствующее во втором, является причиной изучаемого явления.

3. Соединенный метод сходства и различия – комбинация двух первых методов.

4. Метод сопутствующих изменений. Если возникновение или изменение одного явления вызывает определенное изменение другого, то оба эти явления находятся в причинной связи друг с другом.

5. Метод остатков. Если сложное явление вызывается сложной причиной, состоящей из совокупности определенных обстоятельств, и известно, что некоторые из этих обстоятельств являются причиной части явлений, то остаток этого явления вызывается остальными обстоятельствами.

Одним из методов научного познания является **аналогия**, посредством которой достигается знание о предметах и явлениях на основании того, что они имеют сходство с другими. Степень вероятности (достоверности) умозаключений по аналогии зависит от количества сходных признаков у сравниваемых явлений (чем их больше, тем большую вероятность имеет заключение и оно повышается, когда связь выводимого признака с каким-либо другим признаком известна более или менее точно). Аналогия тесно связана с моделированием или модельным экспериментом. Если обычный эксперимент непосредственно взаимодействует с объектом исследования, то в моделировании такого

взаимодействия нет, так как эксперимент производится не с самим объектом, а с его заменителем.

Моделирование – метод, основывающийся на использовании модели в качестве средства исследования явлений и процессов природы. Под моделями понимаются системы, замещающие объекты познания и служащие источником информации о нем. Модели – это такие аналоги, сходство которых с оригиналом существенно, а различие – несущественно. Модели делят на два вида: материальные и идеальные. Материальные модели воплощаются в определенном материале – дереве, металле, стекле и др. Идеальные модели фиксируются в таких наглядных элементах, как чертежи, рисунки, схемы и др.

Метод моделирования имеет следующую структуру:

- а) постановка задачи;
- б) создание или выбор модели;
- в) исследование модели;
- г) перенос знания с модели на оригинал.

Гипотетический метод познания предполагает разработку научной гипотезы на основе изучения физической, химической и т.п. сущности исследуемого явления с помощью описанных выше способов познания и затем формулирование гипотезы, составление расчетной схемы алгоритма (модели), ее изучение, анализ, разработка теоретических положений.

При гипотетическом методе познания исследователь нередко прибегает к идеализации – это мысленное конструирование объектов, которые практически неосуществимы (например, идеальный газ, абсолютно твердое тело). В результате идеализации реальные объекты лишаются некоторых присущих им свойств и наделяются гипотетическими свойствами.

Как в социально-экономических и гуманитарных науках, так и в естественных и технических исследованиях часто используют **исторический метод познания**. Этот метод предполагает исследование возникновения, формирования и развития объектов в хронологической последовательности, в результате чего исследователь получает дополнительные знания об изучаемом объекте (явлении) в процессе их развития.

Методы теоретического уровня: абстрагирование, идеализация, формализация, обобщение, анализ и синтез, индукция и дедукция, аксиоматический метод и т.д. На теоретическом уровне производится логическое исследование собранных фактов, выработка понятий, суждений, делаются умозаключения. В процессе этой работы соотносятся ранние научные представления с возникающими новыми. На теоретическом уровне научное мышление освобождается от эмпирической описательности, создает теоретические обобщения. Таким образом, новое теоретическое содержание знаний надстраивается над эмпирическими знаниями.

На теоретическом уровне познания широко используются логические методы сходства, различия, сопутствующих изменений, разрабатываются новые системы знаний, решаются задачи дальнейшего согласования теоретически разработанных систем с накопленным новым экспериментальным материалом.

Абстрагирование – это мысленное отвлечение от несущественных свойств, связей, отношений предметов и выделение нескольких сторон, интересующих исследователя. Процесс абстрагирования проходит две ступени. Первая ступень: вычленение наиболее важного в явлениях и установление независимости или пренебрежимо слабой зависимости изучаемых явлений от определенных факторов (если объект *A* не зависит непосредственно от фактора *B*, то можно отвлечься от последнего как несущественного). Вторая ступень: реализация возможностей абстрагирования. Суть его заключается в том, что один объект заменяется другим, более простым, который выступает в качестве «модели» первого.

Абстрагирование может применяться к реальным и абстрактным объектам (прошедшим ранее абстрагирование). Многоступенчатое абстрагирование ведет к

абстракциям все возрастающей степени общности. Абстрагирование позволяет заменить в познании сложное простым, но таким простым, которое выражает основное в этом сложном.

Существуют следующие основные виды абстракции:

1) отождествления – образования понятий путем объединения предметов, связанных отношениями типа равенства в особый класс (отвлечение от некоторых индивидуальных свойств предметов);

2) изолированная – выделения свойств и отношений, неразрывно связанных с предметами, и обозначения их определенными «именами», что придает абстракции статус самостоятельных предметов («надежность», «технологичность»).

Различие между этими двумя абстракциями состоит в том, что в первом случае изолируется комплекс свойств объекта, а во втором – единственное его свойство;

3) конструктивизации – отвлечение от неопределенности границ реальных объектов (непрерывное движение останавливаем и т.д.);

4) актуальной бесконечности – отвлечение от незавершенности (и незавершимости) процесса образования бесконечного множества, от невозможности задать его полным списком всех элементов. Такое множество рассматривается как существующее;

5) потенциальной осуществимости – отвлечение от реальных границ человеческих возможностей, обусловленных ограниченностью жизни во времени и пространстве (бесконечность выступает уже как потенциально осуществимая).

Ярким примером абстрактной модели действительности является идеальный газ, который широко используется в физике, термодинамике и других науках. Идеальный газ – это теоретическая модель реального газа, в которой молекулы представляют собой материальные точки, не имеющие объема и сил межмолекулярного сцепления.

Идеализация – это мысленное конструирование объектов, несуществующих в действительности или практически неосуществимых (например, абсолютное твердое тело, абсолютное черное тело, линия, плоскость).

Цель идеализации: лишить реальные объекты некоторых присущих им свойств и наделить (мысленно) эти объекты определенными нереальными и гипотетическими свойствами. При этом достижение цели осуществляется:

1) многоступенчатым абстрагированием (например, абстрагирование от толщины приводит к понятию «плоскость»);

2) мысленным переходом к предельному случаю в развитие какого-либо свойства (абсолютное твердое тело);

3) простым абстрагированием (несжимаемость жидкости).

Любая идеализация правомерна лишь в определенных пределах.

Формализация – отображение объекта или явления в знакомой форме какого-либо искусственного языка (математики, химии и т.д.) и обеспечение возможности исследования реальных объектов и их свойств через формальное исследование соответствующих знаков.

Достоинства формализации:

1) она обеспечивает обобщенность подхода к решению проблем;

2) символика придает краткость и четкость фиксации значений;

3) однозначность символики (нет двусмысленности обычного языка);

4) позволяет формировать знаковые модели объектов и заменять изучение реальных вещей и процессов изучением этих объектов.

Обобщение – определение общего понятия, в котором находит отражение главное, основное, характеризующее объекты данного класса. Это средство для образования новых научных понятий, формулирования законов и теорий.

Аксиоматический метод познания – способ построения научной теории, при котором некоторые утверждения (аксиомы) принимаются без доказательств и затем

используются для получения остальных знаний по определенным логическим правилам. Общеизвестной, например, является аксиома о параллельных линиях (не пересекаются), которая принята в геометрии без доказательств.

К методам метатеоретического уровня относят диалектический метод и метод системного анализа. С помощью этих методов исследуются сами теории и разрабатываются пути их построения, изучается система положений и понятий данной теории, устанавливаются границы ее применения, способы введения новых понятий, обосновываются пути синтезирования нескольких теорий. Центральной задачей данного уровня исследований является познание условий формализации научных теорий и выработка формализованных языков, именуемых метаязыками.

При изучении сложных, взаимосвязанных друг с другом проблем используется **системный анализ**, получивший широкое применение в различных сферах научной деятельности человека, и в частности, в логике, математике, общей теории систем, в результате чего сформировались такие науки, как металогика и метаматематика. Металогика исследует системы положений и понятий формальной логики, разрабатывает вопросы теории доказательств, определенности понятий, истины в формализованных языках. Метаматематика занимается изучением различных свойств формальных систем и исчислений.

В основе системного анализа лежит понятие системы, под которой понимается множество объектов (компонентов), обладающих заранее определенными свойствами с фиксированными между ними отношениями. На базе этого понятия производится учет связей, используются количественные сравнения всех альтернатив для того, чтобы сознательно выбрать наилучшее решение, оцениваемое каким-либо критерием, например измеримостью, эффективностью, надежностью и т.п.

Так как системный анализ носит общий, междисциплинарный характер, т.е. касается образования, развития, функционирования, синтеза любых систем, то некоторые идеологи считают, что системный анализ заменяет философию, является новой всеобщей методологией науки. Такое восприятие системного анализа неверно, так как сводит функцию философского знания лишь к методологии научного исследования. Во всех науках существуют философские основания, используются философские категории, но это не повод принятия основания теории за саму теорию. Системный анализ, с одной стороны, позволяет применять ряд общефилософских положений к решению частных задач, а с другой – обогащает саму философию развитием конкретных наук. Чем дальше развивается системный анализ, тем совершеннее развивается его язык, тем он дальше удаляется от своей первоначальной философской основы. Таким образом, отождествление системного анализа с диалектическим методом, с философией неправомерно и может привести к мировоззренческим и методологическим ошибкам.

Системный анализ используется для исследования таких сложных систем, как экономика отдельной отрасли, промышленного предприятия, объединения, при планировании и организации технологии комплексных строительных процессов, выполняемых несколькими строительными организациями, и др.

Системный анализ складывается из основных четырех этапов: **первый** заключается в постановке задачи – определяет объект, цели и задачи исследования, а также критерии для изучения и управления объектом. Неправильная или неполная постановка целей может свести на нет результаты всего последующего анализа. Во время **второго этапа** очерчиваются границы изучаемой системы и определяется ее структура: объекты и процессы, имеющие отношение к поставленной цели, разбиваются на собственно изучаемую систему и внешнюю среду. При этом различают замкнутые и открытые системы. При исследовании замкнутых систем влиянием внешней среды на их поведение пренебрегают. Затем выделяют отдельные составные части системы – ее элементы,

устанавливают взаимодействие между ними и внешней средой. Именно так строится, например, такая фундаментальная наука, как термодинамика.

В последнее время все большее внимание в технике уделяется изучению замкнутых систем, имеющих закрытые технологические циклы, так называемую «безотходную технологию». Такие технологические процессы перспективны как с позиций экономики, так и экологии: «чем меньше отходов, тем выше уровень производства».

Третий, важнейший этап системного анализа заключается в составлении математической модели исследуемой системы. Вначале производят параметризацию системы, описывают выделенные элементы системы и их взаимодействие. В зависимости от особенностей процессов используют тот или иной математический аппарат для анализа системы в целом.

Следует при этом отметить, что аналитические методы используются для описания лишь небольших систем вследствие их громоздкости или невозможности составления и решения сложной системы уравнений. Для описания больших систем, их характеристик не только качественных, но и количественных используются дискретные параметры (баллы), принимающие целые значения. Например, твердость материалов оценивают баллами по шкале Мооса, энергию сейсмических волн при землетрясениях – баллами по И.Рихтеру и др. Методы операций с дискретными параметрами излагаются в теории множеств и, прежде всего, в таких ее разделах, как в алгебре множеств и в алгебре высказываний (математической логике), составляющих основу математического обеспечения современных ЭВМ.

Наряду с аппаратом алгебры множеств и алгебры высказываний при исследовании сложных систем широко используют вероятностные методы, поскольку в них преобладают стохастические процессы. Поэтому наиболее часто исследуют развитие процессов с некоторой вероятностью или же определяют вероятность протекания изучаемых процессов.

Если исследуются сложные системы, именуемые как обобщенные динамические системы, характеризуемые большим количеством параметров различной природы, то в целях упрощения математического описания их расчлняют на подсистемы, выделяют типовые подсистемы, производят стандартизацию связей для различных уровней иерархии однотипных систем. Примерами такого подхода к изучению сложных систем, например управления, являются типовые возмущения, типовые звенья системы с определенными статическими и динамическими свойствами. В результате третьего этапа системного анализа формируются законченные математические модели системы, описанные на формальном, например алгоритмическом языке.

Важным этапом системного анализа является **четвертый**. Это анализ полученной математической модели, определение ее экстремальных условий с целью оптимизации и формулирование выводов.

Оптимизация заключается в нахождении оптимума рассматриваемой функции (математической модели исследуемой системы, процесса) и соответственно нахождения оптимальных условий поведения данной системы или протекания данного процесса. Оценку оптимизации производят по критериям, принимающим в таких случаях экстремальные значения (выражающие, например, максимальный съем продукции с единицы объема аппарата, минимальную стоимость продукции при определенной производительности, минимальный расход топлива и т.д.). На практике выбрать надлежащий критерий достаточно сложно, так как в задачах оптимизации может выявляться необходимость во многих критериях, которые иногда оказываются взаимно противоречивыми. Поэтому наиболее часто выбирают какой-либо один основной критерий, а для других устанавливают пороговые предельно допустимые значения. На основании выбора составляется зависимость критерия оптимизации от параметров модели исследуемого объекта (процесса). Такой результат исследования чрезвычайно важен для

практических целей, дает определенную последующую опытно-конструкторскую проработку задачи.

* * *

Подводя итог в рассмотрении методологических основ научного познания, отметим следующее. Видение изучаемой реальности формируется у исследователя в результате его личного практического опыта и в громадной мере – в результате воспитания и образования, полученного им в процессе обучения, т.е. на основе опыта и знаний, накопленных в практической деятельности за всю предшествующую историю науки. Речь идет не о формальной сумме знаний, а о тех методологических приемах и правилах (или принципах), которыми необходимо владеть в научно-исследовательской работе. Овладеть этими приемами и правилами можно, лишь непосредственно познакомившись с наиболее законченными и совершенными результатами научного творчества.

Для научных исследований нужны и гений, и терпеливый труд, но ни то, ни другое не может в какой бы то ни было степени гарантировать успех. В научных исследованиях есть нечто не программируемое заранее...

2. ЭКСПЕРИМЕНТАЛЬНЫЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ

2.1 Методология эксперимента

Наиболее важной составной частью научных исследований являются эксперименты.

Экспериментальное исследование – один из основных способов получить новые научные знания [1]. В его основе лежит эксперимент, представляющий собой научно поставленный опыт или наблюдение явления в точно учитываемых условиях, позволяющих следить за его ходом, управлять им, воссоздать его каждый раз при повторении этих условий. От обычного, обыденного пассивного наблюдения эксперимент отличается активным воздействием исследователя на изучаемое явление.

Основная **цель эксперимента** – проверка теоретических положений (подтверждение рабочей гипотезы), а также более широкое и глубокое изучение темы научного исследования. Эксперимент должен быть проведен по возможности в кратчайший срок с минимальной затратой материальных и денежных средств при самом высоком качестве полученных результатов.

Различают эксперименты естественные и искусственные. **Естественные эксперименты** характерны для социальных явлений (социальный эксперимент) в обстановке, например, производства, быта и т.п. **Искусственный эксперимент** широко применяется во многих отраслях и в первую очередь в технических науках. В этом случае изучают явление, изолированное до требуемой степени, чтобы оценить его в количественном и качественном отношении.

Иногда возникает необходимость провести **поисковые экспериментальные исследования**. Они необходимы в том случае, если затруднительно классифицировать все факторы, влияющие на изучаемое явление вследствие отсутствия достаточных предварительных данных. На основе предварительного эксперимента строится программа исследований в полном объеме.

Экспериментальные исследования делятся на лабораторные и производственные.

Лабораторные опыты проводят с применением типовых приборов, специальных моделирующих установок, стендов, оборудования и т.д. Эти исследования позволяют наиболее полно и доброкачественно, с требуемой повторностью изучить влияние одних характеристик при варьировании других. Лабораторные опыты при достаточно полном научном обосновании эксперимента (математическое планирование) позволяют получить хорошую научную информацию с минимальными затратами. Однако такие эксперименты не всегда полностью моделируют реальный ход изучаемого процесса, поэтому возникает потребность в проведении производственного эксперимента.

Производственные экспериментальные исследования имеют целью изучить процесс в реальных условиях с учетом воздействия различных случайных факторов производственной среды. Такие эксперименты проводят на строящихся объектах, заводах, эксплуатируемых дорогах, зданиях и сооружениях. Вследствие, как правило, громоздкости опыта требуется особо тщательное продумывание и планирование эксперимента. Важную роль играет обоснование минимального требуемого количества измерений. К производственным исследованиям относят также специальные полевые экспедиции по обследованию эксплуатируемых объектов. Например, для изучения процессов деформаций и разрушений конструкций дорог создают специальные экспедиции, которые обследуют конструкции в осенние и весенние периоды повышенного увлажнения. Для изучения

службы мостов создают специальные мосто-испытательные экспедиции, которые на основе статических и динамических нагрузений исследуют напряженно-деформируемое состояние элементов мостов.

Одной из разновидностей производственных экспериментов является соби́рание материалов в организациях, которые накапливают по стандартным формам те или иные данные. Ценность этих материалов заключается в том, что они систематизированы за многие годы по единой методике. Такие данные хорошо поддаются обработке методами статистики и теории вероятностей.

В ряде случаев производственный эксперимент эффективно проводить методом анкетирования. Для изучаемого процесса составляют тщательно продуманную методику. Основные данные собирают методом опроса производственных организаций по предварительно составленной анкете. Этот метод позволяет собрать очень большое количество данных наблюдений или измерений по изучаемому вопросу. К результатам анкетных данных следует относиться с особой тщательностью, поскольку они не всегда содержат надежные данные. Особую роль здесь играет метод статистической чистки измерений.

Производственные экспериментальные исследования могут быть заменены опытами на специальных полигонах. Полигонные испытания позволяют производить исследования без нарушения технологического производственного ритма, что повышает эффективность использования применяемого в эксперименте оборудования, машин, приборов.

В зависимости от темы научного исследования объем экспериментов может быть различным. В лучшем случае для подтверждения рабочей гипотезы достаточно лабораторного эксперимента, в худшем – приходится проводить серию экспериментальных исследований: предварительные (поисковые), лабораторные, полигонные, на эксплуатируемом объекте.

В ряде случаев на эксперимент затрачивается много средств. Научный работник производит огромное количество наблюдений и измерений, получает множество диаграмм, графиков, выполняет неоправданно большое количество испытаний. На обработку и анализ такого эксперимента затрачивается много времени. Иногда оказывается, что выполнено много лишнего, ненужного. Все это возможно, когда экспериментатор четко не обосновал цель и задачи эксперимента. В других случаях результаты длительного обширного эксперимента не полностью подтверждают рабочую гипотезу научного исследования. Как правило, это также свойственно для эксперимента, четко не обоснованного целью и задачами. Поэтому, прежде чем приступить к экспериментальным исследованиям, необходимо разработать методологию эксперимента.

Методология эксперимента – это общие принципы, структура эксперимента, его постановка и последовательность выполнения экспериментальных исследований. Методология эксперимента включает в себя следующие основные этапы: разработку плана-программы эксперимента; оценку измерений и выбор средств для проведения эксперимента; проведение эксперимента; обработку и анализ экспериментальных данных, установление адекватности.

Приведенное количество этапов справедливо для традиционного эксперимента. Наряду с этим широко применяют математическую теорию эксперимента, позволяющую резко повысить точность и уменьшить объем экспериментальных исследований. В этом случае методология эксперимента включает такие этапы: разработку плана-программы эксперимента, оценку измерений и выбор средств для проведения эксперимента, математическое планирование эксперимента с одновременным проведением экспериментального исследования, обработкой и анализом полученных данных.

План-программа эксперимента включает наименование темы исследования, рабочую гипотезу, методику эксперимента, перечень необходимых материалов, приборов,

установок, список исполнителей эксперимента, календарный план работ и смету на выполнение эксперимента. В ряде случаев включают работы по конструированию и изготовлению приборов, аппаратов, приспособлений, методическое их обследование, а также программы опытных работ на заводах, строительстве и т.д.

Основу плана-программы составляет *методика эксперимента*. Методика представляет собой систему приемов или способов для последовательного наиболее эффективного экспериментального исследования и включает в себя: цель и задачи эксперимента; выбор варьирующих факторов; обоснование средств и требуемого количества измерений; описание проведения эксперимента, обоснование способов обработки и анализа результатов эксперимента.

Определение цели и задачи эксперимента – один из наиболее важных этапов. На основе анализа информации, гипотезы и теоретических разработок обосновывают цель и задачи эксперимента. Вся научная информация позволяет в той или иной степени судить об ожидаемых закономерностях изучаемого процесса, а, следовательно, и определить задачи эксперимента. Четко, конкретно обоснованные задачи – это большой вклад в их решение. Количество задач не должно быть слишком большим (3-4 задачи), в большом исследовании их может быть 8-10.

Выбор варьирующих факторов – это установление основных и второстепенных характеристик, влияющих на исследуемый процесс. Вначале анализируют расчетные (теоретические) схемы процесса. На основе этого классифицируют все факторы и составляют из них убывающий по важности для данного эксперимента ряд. Правильный выбор основных и второстепенных факторов играет важную роль в эффективности эксперимента, поскольку эксперимент сводится к нахождению зависимостей между этими факторами. В отдельных случаях трудно сразу выявить роль основных и второстепенных факторов. При этом необходимо выполнить небольшой по объему предварительный поисковый опыт.

Основным принципом установления степени важности характеристики является ее роль в исследуемом процессе. Для этого изучают процесс в зависимости от какой-то одной переменной при остальных постоянных. Такой принцип проведения эксперимента оправдывает себя только в тех случаях, когда переменных характеристик мало (1-3). Если же переменных величин много, целесообразен принцип многофакторного анализа.

Обоснование средств измерений – это выбор необходимых для наблюдений и измерений приборов, оборудования, машин, аппаратов и др. Экспериментатор должен быть хорошо ознакомлен с выпускаемой измерительной аппаратурой.

В отдельных случаях возникает потребность в создании уникальных приборов, аппаратов, установок, стендов, машин для разработки темы. При этом разработка и конструирование приборов и других средств должны быть тщательно обоснованы теоретическими расчетами и практическими соображениями о возможности изготовления оборудования. Создавая новые приборы, необходимо использовать готовые узлы выпускаемых или реконструировать существующие приборы. Очень ответственной частью является установление точности измерений и погрешностей. Методы измерений должны базироваться на законах специальной науки – метрологии, изучающей средства и методы измерений.

При экспериментальном исследовании одного и того же процесса (наблюдения и измерения) повторные отсчеты на приборах, как правило, не одинаковы. Отклонения объясняются различными причинами – неоднородностью свойств изучаемого тела (грунт, материал, конструкция и т.д.), несовершенностью приборов и классом их точности, субъективными особенностями экспериментатора и др. Чем больше случайных факторов, влияющих на опыт, тем больше отклонения отдельных измерений от среднего значения. Это требует повторных измерений, следовательно, необходимо знать их требуемое минимальное количество. Под требуемым минимальным количеством измерений

понимают такое их количество, которое в данном опыте обеспечивает устойчивое среднее значение измеряемой величины, удовлетворяющее заданной степени точности. Установление потребного минимального количества измерений имеет большое значение, поскольку обеспечивает получение наиболее объективных результатов при минимальных затратах времени и средств.

В методике подробно проектируют процесс проведения эксперимента. Вначале составляют последовательность (очередность) проведения операций измерений и наблюдений. Затем тщательно описывают каждую операцию в отдельности с учетом выбранных средств для проведения эксперимента. Большое внимание уделяют методам контроля качества операций, обеспечивающих при минимальном (ранее установленном) количестве измерений высокую надежность и заданную точность. Разрабатывают формы журналов для записи результатов наблюдений и измерений.

Важным разделом методики является выбор методов обработки и анализа экспериментальных данных. Обработка данных сводится к систематизации всех цифр, классификации, анализу. Результаты экспериментов должны быть сведены в удобочитаемые формы записи – таблицы, графики, формулы, номограммы, позволяющие быстро сопоставлять полученные результаты.

Особое внимание в методике должно быть уделено математическим методам обработки и анализу опытных данных – установлению эмпирических зависимостей, аппроксимации связей между варьируемыми характеристиками, нахождению критериев и доверительных интервалов и др. Далее определяют объем и трудоемкость экспериментальных исследований, которые зависят от глубины теоретических разработок, степени точности принятых средств измерений. Чем четче сформулирована теоретическая часть исследования, тем меньше объем эксперимента. Возможны три случая проведения эксперимента.

1. Теоретически получена аналитическая зависимость, которая однозначно определяет исследуемый процесс. Например, $y = 3e^{-2x}$. В этом случае объем эксперимента для подтверждения данной зависимости минимален, поскольку функция однозначно определяется экспериментальными данными.

2. Теоретическим путем установлен только характер зависимости. Например, $y = ae^{-bx}$. В этом случае задано семейство кривых. Экспериментальным путем необходимо определить a и b . При этом объем эксперимента возрастает.

3. Теоретически не удалось получить каких-либо зависимостей. Разработаны только предположения о качественных закономерностях процесса. Во многих случаях целесообразен поисковый эксперимент. Объем экспериментальных работ возрастает. Здесь уместен метод математического планирования эксперимента.

На объем и трудоемкость существенно влияет вид эксперимента. Полевые эксперименты, как правило, имеют большую трудоемкость. После установления объема экспериментальных работ составляют перечень необходимых средств измерений, объем материалов, список исполнителей, календарный план и смету расходов. План-программу рассматривает научный руководитель, обсуждают в научном коллективе и утверждают в установленном порядке.

2.2 Основы теории случайных ошибок и методов оценки случайных погрешностей в измерениях

2.2.1 Интервальная оценка с помощью доверительной вероятности

Анализ случайных погрешностей основывается на теории случайных ошибок, дающей возможность с определенной гарантией вычислить действительное значение измеренной величины и оценить возможные ошибки.

Основу теории случайных ошибок составляют предположения о том, что при большом числе измерений некоей величины погрешности одинаковые по модулю, но противоположные по знаку встречаются одинаково часто; большие погрешности встречаются реже, чем малые (вероятность появления погрешности уменьшается с ростом ее величины); при бесконечно большом числе измерений истинное значение измеряемой величины равно среднеарифметическому значению всех результатов измерений, а появление того или иного результата измерения как случайного события описывается нормальным законом распределения [2].

Различают генеральную и выборочную совокупность измерений. Под генеральной совокупностью подразумевают все множество возможных значений измерений x_i или возможных значений погрешностей Δx_i . Для выборочной совокупности число измерений n ограничено, и в каждом конкретном случае строго определяется. Обычно считают, если $n > 30$, то среднее значение данной совокупности измерений \bar{x} достаточно приближается к его истинному значению.

Теория случайных ошибок позволяет оценить точность и надежность измерения при данном количестве замеров или определить минимальное количество замеров, гарантирующее требуемую (заданную) точность и надежность измерений. Наряду с этим возникает необходимость исключить грубые ошибки ряда, определить достоверность полученных данных и др.

Для большой выборки и нормального закона распределения общей оценочной характеристикой измерения являются дисперсия D и коэффициент вариации k_B :

$$D = \sigma^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1); \quad k_B = \sigma / \bar{x}. \quad (2.1)$$

Дисперсия характеризует однородность измерения. Чем выше D , тем больше разброс измерений. Коэффициент вариации характеризует изменчивость. Чем выше k_B , тем больше изменчивость измерений относительно средних значений, k_B оценивает также разброс при оценке нескольких выборок.

Доверительным называется интервал значений x_i , в который попадает истинное значение x_d измеряемой величины с заданной вероятностью. Доверительной вероятностью (достоверностью) измерения называется вероятность того, что истинное значение измеряемой величины попадает в данный доверительный интервал, т.е. в зону $a \leq x_d \leq b$. Эта величина определяется в долях единицы или в процентах. Доверительная вероятность p_D описывается выражением

$$p_D = p[a \leq x_d \leq b] = (1/2)[\varphi(b - \bar{x}) / \sigma - \varphi(a - \bar{x}) / \sigma],$$

где $\varphi(t)$ – интегральная функция Лапласа (таблица 2.1), определяемая выражением

$$\varphi(t) = \frac{2}{\sqrt{2\pi}} \int_0^{t_i} e^{-t^2/2} dt.$$

Аргументом этой функции является отношение μ к среднеквадратическому отклонению σ , т.е.

$$t = \mu/\sigma, \quad (2.2)$$

где t – гарантированный коэффициент;

$$\mu = b - \bar{x}; \quad \mu = -(a - \bar{x}).$$

Если же на основе определенных данных установлена доверительная вероятность p_d (часто ее принимают равной 0.90; 0.95; 0.9973), то устанавливается точность измерений (доверительный интервал 2μ) на основе соотношения $p_d = \varphi(\mu/\sigma)$. Половина доверительного интервала равна

$$\mu = \sigma \arg \varphi(p_d) = \sigma t, \quad (2.3)$$

где $\arg \varphi(p_d)$ – аргумент функции Лапласа, а при $n < 30$ – функции Стьюдента (таблица 2.2). Доверительный интервал характеризует точность измерения данной выборки, а доверительная вероятность – достоверность измерения. Пусть, например, выполнено 30 измерений прочности дорожного полотна участка автомобильной дороги при среднем модуле упругости полотна $E = 170$ МПа и вычисленном значении среднеквадратического отклонения $\sigma = 3.1$ МПа.

Таблица 2.1 - Интегральная функция Лапласа

T	p_d	t	p_d	T	p_d
0.00	0.0000	0.75	0.5467	1.50	0.8864
0.05	0.0399	0.80	0.5763	1.55	0.8789
0.10	0.0797	0.85	0.6047	1.60	0.8904
0.15	0.1192	0.90	0.6319	1.65	0.9011
0.20	0.1585	0.95	0.6579	1.70	0.9109
0.25	0.1974	1.00	0.6827	1.75	0.9199
0.30	0.2357	1.05	0.7063	1.80	0.9281
0.35	0.2737	1.10	0.7287	1.85	0.9357
0.40	0.3108	1.15	0.7419	1.90	0.9426
0.45	0.3473	1.20	0.7699	1.95	0.9488
0.50	0.3829	1.25	0.7887	2.00	0.9545
0.55	0.4177	1.30	0.8064	2.25	0.9756
0.60	0.4515	1.35	0.8230	2.50	0.9876
0.65	0.4843	1.40	0.8385	3.00	0.9973
0.70	0.5161	1.45	0.8529	4.00	0.9999

Требуемую точность измерений можно определить для разных уровней доверительной вероятности ($p_d = 0.9; 0.95; 0.9973$), приняв значения t по таблице 2.1. В этом случае соответственно $\mu = \pm 3.1 \cdot 1.65 = 5.1$; $\pm 3.1 \cdot 2.0 = 6.2$; $\pm 3.1 \cdot 3.0 = 9.3$ МПа. Следовательно, для данного средства и метода доверительный интервал возрастает примерно в два раза, если увеличить p_d только на 10%.

Если необходимо определить достоверность измерений для установленного доверительного интервала, например $\mu = \pm 7$ МПа, то по формуле (2.2) $t = \mu/\sigma = 7/3.1 = 2.26$. По таблице 2.1 для $t = 2.26$ определяем $p_d = 0.97$. Это означает, что в заданный доверительный интервал из 100 измерений не попадают только три.

Значение $(1 - p_d)$ называют уровнем значимости. Из него следует, что при нормальном законе распределения погрешность, превышающая доверительный интервал, будет встречаться один раз из $n_{и}$ измерений, где

$$n_{и} = p_d / (1 - p_d), \quad (2.4)$$

или иначе приходится браковать одно из $n_{и}$ измерений.

Таблица 2.2 - Коэффициент Стьюдента $\alpha_{СТ}$

N	p_d					
	0,80	0,90	0,95	0,99	0,995	0,999
2	3,080	6,31	12,71	63,70	127,30	637,20
3	1,886	2,92	4,30	9,92	14,10	31,60
4	1,638	2,35	3,188	5,84	7,50	12,94
5	1,533	2,13	2,77	4,60	5,60	8,61
6	1,476	2,02	2,57	4,03	4,77	6,86
7	1,440	1,94	2,45	3,71	4,32	9,96
8	1,415	1,90	2,36	3,50	4,03	5,40
9	1,397	1,86	2,31	3,36	3,83	5,04
10	1,383	1,83	2,26	3,25	3,69	4,78
12	1,363	1,80	2,20	3,11	3,50	4,49
14	1,350	1,77	2,16	3,01	3,37	4,22
16	1,341	1,75	2,13	2,95	3,29	4,07
18	1,333	1,74	2,11	2,90	3,22	3,96
20	1,328	1,73	2,09	2,86	3,17	3,88
30	1,316	1,70	2,04	2,75	3,20	3,65
40	1,306	1,68	2,02	2,70	3,12	3,55
50	1,298	1,68	2,01	2,68	3,09	3,50
60	1,290	1,67	2,00	2,66	3,06	3,46
∞	1,282	1,64	1,96	2,58	2,81	3,29

По данным приведенного выше примера можно вычислить количество измерений, из которых одно измерение превышает доверительный интервал. По формуле (2.4) при $p_d = 0.9$ определяется $n_{и} = 0.9(1 - 0.9) = 9$ измерений. При p_d равной 0.95 и 0.9973, соответственно 19 и 369 измерений.

2.2.2 Определение минимального количества измерений

Для проведения опытов с заданной точностью и достоверностью необходимо знать то количество измерений, при котором экспериментатор уверен в положительном исходе [2]. В связи с этим одной из первоочередных задач при статистических методах оценки является установление минимального, но достаточного числа измерений для данных условий. Задача сводится к установлению минимального объема выборки (числа измерений) N_{\min} при заданных значениях доверительного интервала 2μ и доверительной вероятности. При выполнении измерений необходимо знать их точность:

$$\Delta = \sigma_0 / \bar{x}, \quad (2.5)$$

где σ_0 – среднеарифметическое значение среднеквадратического отклонения σ , равное $\sigma_0 = \sigma / \sqrt{n}$.

Значение σ_0 часто называют средней ошибкой. Доверительный интервал ошибки измерения Δ определяется аналогично для измерений $\mu = t\sigma_0$. С помощью t легко определить доверительную вероятность ошибки измерений из таблицы 2.1.

В исследованиях часто по заданной точности Δ и доверительной вероятности измерения определяют минимальное количество измерений, гарантирующих требуемые значения Δ и p_d .

Аналогично уравнению (2.3) с учетом (2.5) можно получить

$$\mu = \sigma \arg \varphi(p_d) = \sigma_0 / \sqrt{nt}.$$

При $N_{\min} = n$ получаем

$$N_{\min} = \sigma^2 t^2 / \sigma_0^2 = k_B^2 t^2 / \Delta^2, \quad (2.6)$$

здесь k_B – коэффициент вариации (изменчивости), %; Δ – точность измерений, %.

Для определения N_{\min} может быть принята такая последовательность вычислений: 1) проводится предварительный эксперимент с количеством измерений n , которое составляет в зависимости от трудоемкости опыта от 20 до 50; 2) вычисляется среднеквадратичное отклонение σ по формуле (2.1); 3) в соответствии с поставленными задачами эксперимента устанавливается требуемая точность измерений Δ , которая не должна превышать точности прибора; 4) устанавливается нормированное отклонение t , значение которого обычно задается (зависит также от точности метода); 5) по формуле (2.6) определяют N_{\min} и тогда в дальнейшем, в процессе эксперимента число измерений не должно быть меньше N_{\min} .

Пусть, например, при приемке прецизионных резисторов в качестве одного из параметров измеряется их сопротивление. Согласно инструкции требуется выполнить 25 измерений; допусковое отклонение параметра ± 0.1 Ом. Если предварительно вычисленное значение $\sigma = 0.4$ Ом, то можно определить, с какой достоверностью оценивается данный параметр.

Согласно инструкции $\Delta = 0.1$ Ом. Из формулы (2.6) можно записать $t = \sqrt{n} \cdot \frac{\Delta}{\sigma} = \sqrt{25} \cdot \frac{0.1}{0.4} = 1.25$. В соответствии с таблицей 2.1 доверительная вероятность для $t = 1.25$ $p_d = 0.79$. Это низкая вероятность. Погрешность, превышающая доверительный интервал $2\mu = 0.2$ Ом, согласно выражению (2.4) будет встречаться один раз из $0.79/(1 - 0.79) = 3.76$, т.е. из четырех измерений. Это недопустимо. В связи с этим необходимо вычислить минимальное количество измерений с доверительной вероятностью p_d , равной 0.9 и 0.95. По формуле (2.6) имеем $N_{\min} = 0.4^2 \cdot 1.65^2 / 0.1^2 = 43$ измерения при $p_d = 0.90$ и 64 измерения при $p_d = 0.95$, что значительно превышает установленные 25 измерений.

Оценки измерений с помощью σ и σ_0 по приведенным методам справедливы при $n > 30$. Для нахождения границы доверительного интервала при малых значениях применяют метод, предложенный в 1908 г. английским математиком В.С.Госсетом (псевдоним Стьюдент). Кривые распределения Стьюдента в случае $n \rightarrow \infty$ (практически при $n > 20$) переходят в кривые нормального распределения (рисунок 2.1).

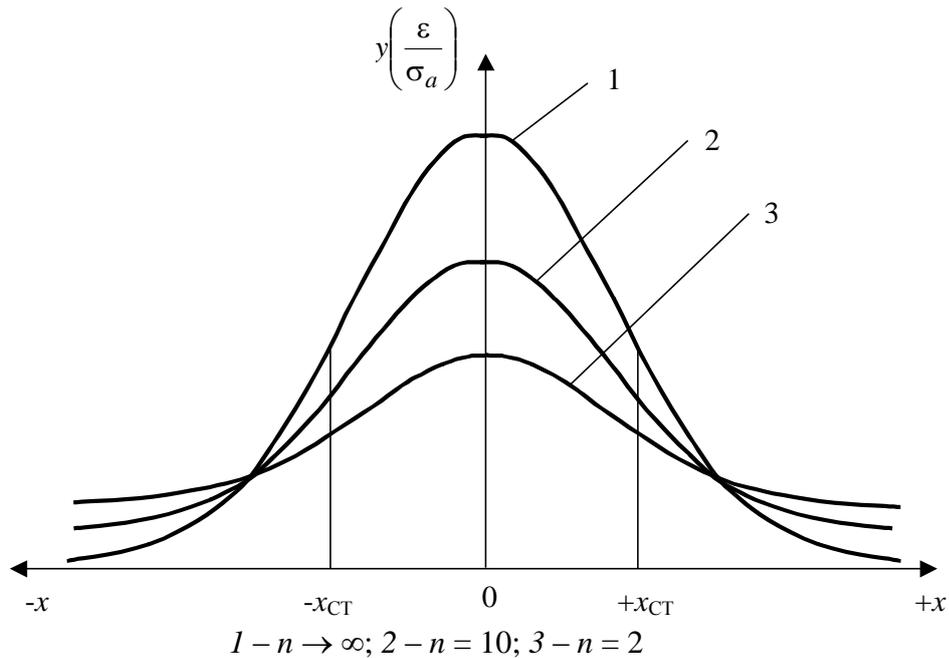


Рисунок 2.1 – Кривые распределения Стьюдента для различных значений объема выборки

Для малой выборки доверительный интервал

$$\mu_{CT} = \sigma_0 \alpha_{CT}, \quad (2.7)$$

где α_{CT} – коэффициент Стьюдента, принимаемый по таблице 2.2 в зависимости от значения доверительной вероятности p_d .

Зная μ_{CT} , можно вычислить действительное значение изучаемой величины для малой выборки

$$x_D = \bar{x} \pm \mu_{CT}. \quad (2.8)$$

Возможна и иная постановка задачи. По n известным измерениям малой выборки необходимо определить доверительную вероятность p_d при условии, что погрешность среднего значения не выйдет за пределы $\pm \mu_{CT}$. Задачу решают в такой последовательности: вначале вычисляется среднее значение \bar{x} , σ_0 и $\alpha_{CT} = \mu_{CT}/\sigma_0$. С помощью величины α_{CT} , известного n и таблицы 2.2 определяют доверительную вероятность.

В процессе обработки экспериментальных данных следует исключать грубые ошибки ряда. Появление этих ошибок вполне вероятно, а наличие их ощутимо влияет на результат измерений. Однако прежде чем исключить то или иное измерение, необходимо убедиться, что это действительно грубая ошибка, а не отклонение вследствие статистического разброса. Известно несколько методов определения грубых ошибок статистического ряда. Наиболее простым способом исключения из ряда резко выделяющегося измерения является правило трех сигм: разброс случайных величин от среднего значения не должен превышать

$$x_{\max, \min} = \bar{x} \pm 3\sigma.$$

Более достоверными являются методы, базируемые на использовании доверительного интервала. Пусть имеется статистический ряд малой выборки, подчиняющийся закону нормального распределения. При наличии грубых ошибок критерии их появления вычисляются по формулам

$$\begin{aligned} \beta_1 &= (x_{\max} - \bar{x}) / \sigma \sqrt{(n-1)/n}; \\ \beta_2 &= (\bar{x} - x_{\min}) / \sigma \sqrt{(n-1)/n}, \end{aligned} \quad (2.9)$$

где x_{\max} , x_{\min} – наибольшее и наименьшее значения из n измерений.

В таблице 2.3 приведены в зависимости от доверительной вероятности максимальные значения β_{\max} , возникающие вследствие статистического разброса. Если $\beta_1 > \beta_{\max}$, то значение x_{\max} необходимо исключить из статистического ряда как грубую погрешность. При $\beta_2 < \beta_{\max}$ исключается величина x_{\min} . После исключения грубых ошибок определяют новые значения x и σ из $(n - 1)$ или $(n - 2)$ измерений.

Второй метод установления грубых ошибок основан на использовании критерия В.И.Романовского и применим также для малой выборки. Методика выявления грубых ошибок сводится к следующему. Задаются доверительной вероятностью p_d и по таблице 2.4 в зависимости от n находят коэффициент q . Вычисляют предельно допустимую абсолютную ошибку отдельного измерения $\varepsilon_{пр} = \sigma q$.

Если $\bar{x} - x_{\max} > \varepsilon_{пр}$, то измерение x_{\max} исключают из ряда наблюдений. Этот метод более требователен к очистке ряда.

При анализе измерений можно применять для приближенной оценки и такую методику: вычислить по (2.1) среднеквадратичное отклонение σ ; определить с помощью (2.5) σ_0 ; принять доверительную вероятность p_d и найти доверительные интервалы $\mu_{ст}$ из (2.7); окончательно установить действительное значение измеряемой величины x_d по формуле (2.8).

Таблица 2.3 - Критерий появления грубых ошибок

N	β_{\max} при p_d			N	β_{\max} при p_d		
	0,90	0,95	0,99		0,90	0,95	0,99
3	1,41	1,41	1,41	15	2,33	2,49	2,80
4	1,64	1,69	1,72	16	2,35	2,52	2,84
5	1,79	1,87	1,96	17	2,38	2,55	2,87
6	1,89	2,00	2,13	18	2,40	2,58	2,90
7	1,97	2,09	2,26	19	2,43	2,60	2,93
8	2,04	2,17	2,37	20	2,45	2,62	2,96
9	2,10	2,24	2,46	25	2,54	2,72	3,07
10	2,15	2,29	2,54	30	2,61	2,79	3,16
11	2,19	2,34	2,61	35	2,67	2,85	3,22
12	2,23	2,39	2,66	40	2,72	2,90	3,28
13	2,26	2,43	2,71	45	2,76	2,95	3,33
14	2,30	2,46	2,76	50	2,80	2,99	3,37

Таблица 2.4 - Коэффициент для вычисления предельно допустимой ошибки измерения

N	Значение q при p _д			
	0,95	0,98	0,99	0,995
2	15,56	38,97	77,96	779,7
3	4,97	8,04	11,46	36,5
4	3,56	5,08	6,58	14,46
5	3,04	4,10	5,04	9,43
6	2,78	3,64	4,36	7,41
7	2,62	3,36	3,96	6,37
8	2,51	3,18	3,71	5,73
9	2,43	3,05	3,54	5,31
10	2,37	2,96	3,41	5,01
12	2,29	2,83	3,23	4,62
14	2,24	2,74	3,12	4,37
16	2,20	2,68	3,04	4,20
18	2,17	2,64	3,00	4,07
20	2,15	2,60	2,93	3,98
∞	1,96	2,33	2,58	3,29

В случае более глубокого анализа экспериментальных данных рекомендуется такая последовательность: 1) после получения экспериментальных данных в виде статистического ряда его анализируют и исключают систематические ошибки; 2) анализируют ряд в целях обнаружения грубых ошибок и промахов: устанавливают подозрительные значения x_{\max} или x_{\min} ; определяют среднеквадратичное отклонение σ ; вычисляют по (2.9) критерии β_1 , β_2 и сопоставляют с β_{\max} , β_{\min} , исключают при необходимости из статистического ряда x_{\max} или x_{\min} и получают новый ряд из новых членов; 3) вычисляют среднеарифметическое \bar{x} , погрешности отдельных измерений $(\bar{x} - x_i)$ и среднеквадратичное очищенного ряда σ ; 4) находят среднеквадратичное σ_0 серии измерений, коэффициент вариации k_v ; 5) при большой выборке задаются доверительной вероятностью $p_d = \Phi(t)$ или уравнением значимости $(1 - p_d)$ и по таблице 2.1 определяют t ; 6) при малой выборке ($n \leq 30$) в зависимости от принятой доверительной вероятности p_d и числа членов ряда n принимают коэффициент Стьюдента α_{CT} ; с помощью формулы (2.2) для большой выборки или (2.7) для малой выборки определяют доверительный интервал; 7) устанавливают по (2.8) действительное значение исследуемой величины; 8) оценивают относительную погрешность (%) результатов серии измерений при заданной доверительной вероятности p_d :

$$\delta = \frac{\delta_0 \alpha_{CT}}{\bar{x}} 100.$$

Если погрешность серии измерений соизмерима с погрешностью прибора $B_{пр}$, то границы доверительного интервала

$$\mu_{CT} = \sqrt{\sigma_0^2 \alpha_{CT}^2 + \left[\frac{\alpha_{CT}(\infty)}{3} \right]^2}. \quad (2.10)$$

Формулой (2.10) следует пользоваться при $\alpha_{CT} \sigma_0 \leq 3B_{пр}$. Если же $\alpha_{CT} \sigma_0 \geq 3B_{пр}$, то доверительный интервал вычисляют с помощью (2.1) или (2.8).

Пусть, например, имеется 18 измерений (таблица 2.5). Если анализ средств и результатов измерений показал, что систематических ошибок в эксперименте не обнаружено, то можно выяснить, не содержат ли измерения грубых ошибок. Если воспользоваться первым методом (критерий β_{\max}), то надо вычислить среднеарифметическое \bar{x} и отклонение σ . При этом удобно пользоваться формулой $\bar{x} = \bar{x}' + \sum (x_i - \bar{x}')/n$, где \bar{x}' – среднее произвольное число. Для вычисления \bar{x} , например, можно принять произвольно $\bar{x}' = 75$. Тогда $\bar{x} = 75 - 3/18 = 74.83$. В формуле (2.1) значение $(\bar{x} - x_i)^2$ можно найти упрощенным методом:

$$\sum (\bar{x} - x_i)^2 = \sum (x_i - \bar{x}') - \frac{(x_i - \bar{x}')^2}{n}.$$

Таблица 2.5 - Результаты измерений и их обработка

x_i	$x_i - \bar{x}'$	$x_i - \bar{x}$	$(x_i - \bar{x}')^2$
67	-8	-7.83	64
67	-8	-7.83	64
68	-7	-6.83	49
68	-7	-6.83	49
69	-6	-5.83	36
70	-5	-4.83	25
71	-4	-3.83	16
73	-2	-1.83	4
74	-1	-0.83	1
75	0	+0.17	0
76	+1	+1.17	1
77	+2	+2.17	4
78	+3	+3.17	9
79	+4	+4.17	16
80	+5	+5.17	25
81	+6	+6.17	36
82	+7	+7.17	49
92	+17	+17.27	289
$\bar{x} = 74.83$	$\Sigma = -3$	Проверка $\Sigma = 0$	$\Sigma = 737$

В данном случае $\sum (\bar{x} - x_i)^2 = 737 - 3^2/18 = 736.5$. По (2.1)
 $\sigma = \sqrt{736.5/(18-1)} = 6.58$, коэффициент вариации $k_B = \frac{6.58}{74.83} \cdot 100 = 8.8\%$. Следовательно,

$$\beta_1 = \frac{(92 - 74.83)}{6.58 \sqrt{\frac{(18-1)}{18}}} = 2.68.$$

Как видно из таблицы 2.3, при доверительной вероятности $p_d = 0.99$ и $n = 18$ $\beta_{\max} = 2.90$. Поскольку $2.68 < \beta_{\max}$, измерение 92 не является грубым промахом. Если $p_d = 0.95$, $\beta_{\max} = 2.58$, то значение 92 следует исключить.

Если применить правило 3σ , то $x_{\max, \min} = 74.83 \pm 3 \times 6.58 = 94.6...55.09$, т.е. измерение 92 следует оставить.

В случае, когда измерение 92 исключается, $\bar{x} = 73.82$, $\sigma = 5.15$. Среднеарифметическое значение среднеквадратичного отклонения для всей серии измерений при $n = 18$ $\sigma_0 = 6.58/\sqrt{18} = 1.55$; при очищенном ряде $\sigma_0 = 5.15/\sqrt{17} = 1.25$.

Поскольку $n < 30$, ряд следует отнести к малой выборке и доверительный интервал вычисляется с применением коэффициент Стьюдента $\alpha_{ст}$. По таблице 2.2 принимается доверительная вероятность 0.95 и тогда $\alpha_{ст} = 2.11$ в случае $n = 18$; $\alpha_{ст} = 2.12$, если $n = 17$. Доверительный интервал при $n = 18$ $\mu_{ст} = \pm 1.55 \cdot 2.11 = 3.3$; при $n = 17$ $\mu_{ст} = \pm 1.25 \cdot 2.12 = 2.7$. Действительное значение изучаемой величины: при $n = 18$ $x_d = 74.8 \pm 3.3$; при $n = 17$ $x_d = 73.8 \pm 2.7$. Относительная погрешность результатов серии измерений: при $n = 18$ $\delta = (3.3 \cdot 100)/74.8 = 4.4\%$; $n = 17$ $\delta = (2.7 \cdot 100)/73.8 = 3.7\%$. Таким образом, если принять $x_i = 92$ за грубый промах, то погрешность измерения уменьшается с 4.3 до 3.7%, т.е. на 14%.

Если необходимо определить минимальное количество измерений при их заданной точности, проводят серию опытов, вычисляют σ , затем с помощью формулы (2.6) определяют N_{min} .

В рассмотренном случае $\sigma = 6.58$; $k_b = 8.8\%$. Если задана точность $\Delta = 5$ и 3% при доверительной вероятности $p_d = 0.95$, $\alpha_{ст} = 2.11$. Следовательно, при $\Delta = 5\%$ $N_{min} = (8.8^2 \cdot 2.11^2)/5^2 = 14$, а при $\Delta = 3\%$ $N_{min} = (8.8^2 \cdot 2.11^2)/3^2 = 40$.

Таким образом, требование повышения точности измерения (но не выше точности прибора) приводит к значительному увеличению повторяемости опытов.

Во многих случаях в процессе экспериментальных исследований приходится иметь дело с косвенными измерениями. При этом неизбежно в расчетах применяют те или иные функциональные зависимости типа

$$y = f(x_1, x_2, \dots, x_n). \quad (2.11)$$

Так как в данную функцию подставляют не истинные, а приближенные значения, то и окончательный результат также будет приближенным. В связи с этим одной из основных задач теории случайных ошибок является определение ошибки функции, если известны ошибки их аргументов.

При исследовании функции одного переменного предельные абсолютные $\varepsilon_{пр}$ и относительные $\delta_{пр}$ ошибки (погрешности) вычисляют так:

$$\begin{aligned} \varepsilon_{пр} &= \pm \varepsilon_x f'(x), \\ \delta_{пр} &= \pm d \ln(x), \end{aligned}$$

где $f'(x)$ – производная функции $f(x)$; $d \ln(x)$ – дифференциал натурального логарифма функции.

Если исследуется функция многих переменных, то

$$\varepsilon_{ПП} = \pm \sum_1^n \left| \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} dx_i \right|, \quad (2.12)$$

$$\delta_{пр} = \pm d \left| \ln(x_1, x_2, \dots, x_n) \right|. \quad (2.13)$$

В (2.12) и (2.13) выражения под знаком суммы и дифференциала принимают абсолютные значения. Методика определения ошибок с помощью этих уравнений следующая: вначале определяют абсолютные и относительные ошибки аргументов (независимых переменных). Обычно величина $x_d \pm \varepsilon$ каждого переменного измерена, следовательно, абсолютные ошибки для аргументов известны, т.е. $\varepsilon_{x1}, \varepsilon_{x2}, \dots, \varepsilon_{xn}$. Затем вычисляют относительные ошибки независимых переменных:

$$\delta_{x1} = \varepsilon_{x1}/x_d; \delta_{x2} = \varepsilon_{x2}/x_d \quad ; \dots; \delta_{xn} = \varepsilon_{xn}/x_d.$$

Находят частные дифференциалы функции и по формуле (2.12) вычисляют $\varepsilon_{пр}$ в размерностях функции $f(y)$ и с помощью (2.13) вычисляют $\delta_{пр}$, %.

Одной из задач теории измерений является установление оптимальных, т.е. наиболее выгодных, условий измерений. Оптимальные условия измерений в данном эксперименте имеют место при $\delta_{\text{пр}} = \delta_{\text{пр min}}$. Методика решения этой задачи сводится к следующему. Если исследуется функция с одним неизвестным переменным, то вначале следует взять первую производную по x , приравнять ее нулю и определить x_1 . Если вторая производная по x_1 положительна, то функция (2.11) в случае $x = x_1$ имеет минимум. При наличии нескольких переменных поступают аналогичным образом, но берут производные по всем переменным x_1, x_2, \dots, x_n . В результате минимизации функций устанавливают оптимальную область измерений (интервал температур, напряжений, силы тока, угла поворота стрелки на приборе и т.д.) каждой функции $f(x_1, x_2, \dots, x_n)$, при которой относительная ошибка измерений минимальна, т.е. $\delta_{xi} = \min$.

В исследованиях часто возникает вопрос о достоверности данных, полученных в опытах. Решение такой задачи можно проиллюстрировать примером.

Пусть установлена прочность контрольных образцов несущих конструкций до закалки $R_1 = \bar{R}_1 \pm \sigma_1 = 20 \pm 0.5$ МПа и прочность этих образцов после закалки $R_2 = \bar{R}_2 \pm \sigma_2 = 23 \pm 0.6$ МПа. Прирост прочности составляет 15%. Это упрочнение относительно небольшое, его можно отнести за счет разброса опытных данных. В этом случае следует провести проверку на достоверность экспериментальных данных по условию

$$\frac{\bar{x}}{\sigma_0} \geq 3.$$

В данном случае проверяется разница $\bar{x} = R_1 - R_2 = 3$ МПа. Ошибка измерения равна $\sigma_0 = \sqrt{\sigma_1^2 + \sigma_2^2}$, поэтому

$$(R_1 - R_2) / \sqrt{\sigma_1^2 + \sigma_2^2} = 3.0 / \sqrt{0.25 + 0.36} = 3.85 > 3.$$

Следовательно, полученный прирост прочности является достоверным.

Выше были рассмотрены общие методы проверки экспериментальных измерений на точность и достоверность. Ответственные эксперименты должны быть проверены также и на воспроизводимость результатов, т.е. на их повторяемость в определенных пределах измерений с заданной доверительной вероятностью. Суть такой проверки сводится к следующему. Имеется несколько параллельных опытов (серий). Для каждой серии вычисляют среднеарифметическое значение \bar{x}_i (n – число измерений в одной серии, принимаемое обычно 3...4). Далее вычисляют дисперсию D_i . Чтобы оценить воспроизводимость, рассчитывают критерий Кохрена (расчетный):

$$k_{KP} = \max D_i / \sum_1^m D_i, \quad (2.14)$$

где $\max D_i$ – наибольшее значение дисперсий из числа рассматриваемых параллельных серий m ; $\sum_1^m D_i$ – сумма дисперсий m серий. Рекомендуется принимать $2 \leq m \leq 4$. Опыты считают воспроизводимыми при

$$k_{KP} \leq k_{KT},$$

где k_{KT} – табличное значение критерия Кохрена (таблица 2.6), принимаемое в зависимости от доверительной вероятности p_d и числа степеней свободы $q = n - 1$. Здесь m – число серий опытов; n – число измерений в серии.

Таблица 2.6 - Критерий Кохрена $k_{кт}$ при $p_d = 0.95$

m	$q = n - 1$									
	1	2	3	4	5	6	8	10	16	36
2	0,99	0,97	0,93	0,90	0,87	0,85	0,81	0,78	0,73	0,66
3	0,97	0,93	0,79	0,74	0,70	0,76	0,63	0,60	0,54	0,47
4	0,90	0,76	0,68	0,62	0,59	0,56	0,51	0,48	0,43	0,36
5	0,84	0,68	0,60	0,54	0,50	0,48	0,44	0,41	0,36	0,26
6	0,78	0,61	0,53	0,48	0,44	0,42	0,38	0,35	0,31	0,25
7	0,72	0,56	0,48	0,43	0,39	0,37	0,34	0,31	0,27	0,23
8	0,68	0,51	0,43	0,39	0,36	0,33	0,30	0,28	0,24	0,20
9	0,64	0,47	0,40	0,35	0,33	0,30	0,28	0,25	0,22	0,18
10	0,60	0,44	0,37	0,33	0,30	0,28	0,25	0,23	0,20	0,16
12	0,57	0,39	0,32	0,29	0,26	0,24	0,22	0,20	0,17	0,14
15	0,47	0,33	0,27	0,24	0,22	0,20	0,18	0,17	0,14	0,11
20	0,39	0,27	0,22	0,19	0,17	0,16	0,14	0,13	0,11	0,08
24	0,34	0,29	0,19	0,16	0,15	0,14	0,12	0,11	0,09	0,07
30	0,29	0,20	0,16	0,14	0,12	0,11	0,10	0,09	0,07	0,06
40	0,24	0,16	0,12	0,10	0,09	0,08	0,07	0,07	0,06	0,04
60	0,17	0,11	0,08	0,07	0,06	0,06	0,05	0,05	0,04	0,02
120	0,09	0,06	0,04	0,04	0,03	0,03	0,02	0,02	0,02	0,01

Пусть, например, проведено три серии опытов по измерению прочности шасси аппарата (таблица 2.7). В каждой серии выполнялось по пять измерений (повторностей). Тогда по формуле (2.14)

$$k_{KP} = \frac{2.96}{2.96 + 2.0 + 0.4} = 0.55.$$

Таблица 2.7 - Результаты измерений прочности шасси аппарата и их обработка

Серия опытов	Измерение величины и повторности					Вычисленные	
	1	2	3	4	5	\bar{x}_i	D_i
1	7	9	6	8	4	6,8	2,96
2	9	7	8	6	5	7,0	2,0
3	8	8	7	9	8	8,0	0,4

Вычислим число степеней свободы $q = n - 1 = 5 - 1 = 4$. Так, например, для $m = 3$ и $q = 4$ согласно таблице 2.6 значение критерия Кохрена $k_{кт} = 0.74$. Так как $0.55 < 0.74$, то измерения в эксперименте следует считать воспроизводимыми. Если бы оказалось наоборот, т.е. $k_{кр} > k_{кт}$, то необходимо было бы увеличить число серий m или число измерений n .

2.3 Регрессионный анализ

Под регрессионным анализом понимают исследование закономерностей связи между явлениями (процессами), которые зависят от многих, иногда неизвестных, факторов. Часто между переменными x и y существует связь, но не вполне определенная, при которой одному значению x соответствует несколько значений (совокупность) y . В таких случаях связь называют регрессионной. Таким образом, функция $y = f(x)$ является

регрессионной (корреляционной), если каждому значению аргумента соответствует статистический ряд распределения y . Следовательно, регрессионные зависимости характеризуются вероятностными или стохастическими связями. Поэтому установление регрессионных зависимостей между величинами y и x возможно лишь тогда, когда выполнимы статистические измерения.

Статистические зависимости описываются математическими моделями процесса, т.е. регрессионными выражениями, связывающими независимые значения x (факторы) с зависимой переменной y (результативный признак, функция цели, отклик). Модель по возможности должна быть простой и адекватной. Например, модуль упругости материала E зависит от его плотности ρ так, что с возрастанием плотности модуль упругости материала увеличивается. Но выявить эту закономерность можно только при наличии большого количества измерений, так как при исследованиях каждой отдельной парной связи в зависимости $E = f(\rho)$ наблюдаются большие отклонения.

Суть регрессионного анализа сводится к установлению уравнения регрессии, т.е. вида кривой между случайными величинами (аргументами x и функцией y), оценке тесноты связей между ними, достоверности и адекватности результатов измерений.

Чтобы предварительно определить наличие такой связи между x и y , наносят точки на график и строят так называемое корреляционное поле (рисунок 2.2). По тесноте группировки точек вокруг прямой или кривой линии, по наклону линии можно визуально судить о наличии корреляционной связи. Так, из рисунка 2.2, *а*, видно, что экспериментальные данные имеют определенную связь между x и y , а измерения, приведенные на рисунке 2.2, *б*, такой связи не показывают.

а)

б)

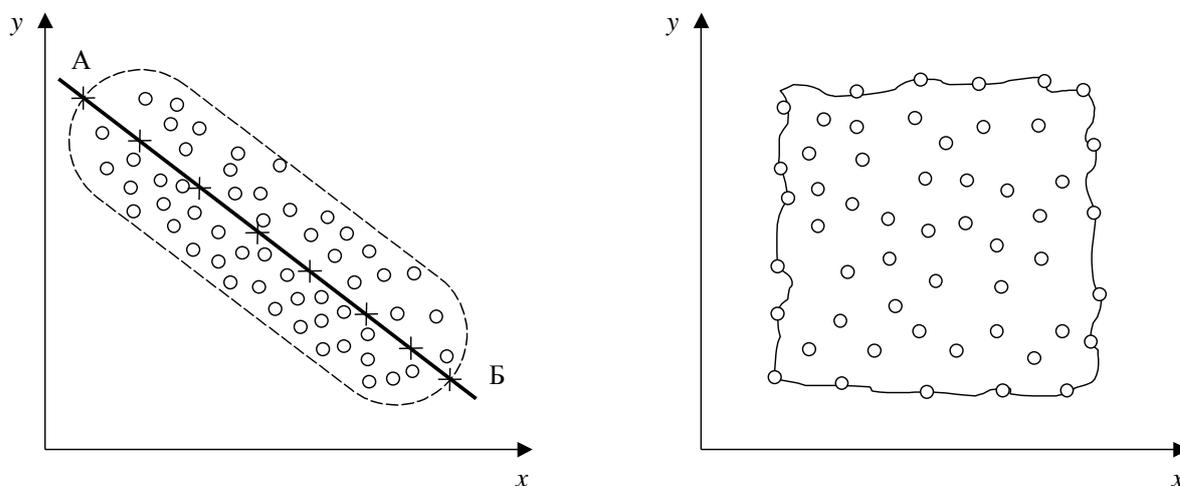


Рисунок 2.2 – Корреляционное поле

Корреляционное поле характеризует вид связи между x и y . По форме поля можно ориентировочно судить о форме графика, характеризующего прямолинейную или криволинейную зависимости. Даже для вполне выраженной формы корреляционного поля вследствие статистического характера связи исследуемого явления одно значение x может иметь несколько значений y . Если на корреляционном поле усреднить точки, т.е. для каждого значения x_i определить \bar{x}_i и соединить точки \bar{y}_i , то можно будет получить ломаную кривую линию, называемую экспериментальной регрессионной зависимостью (линией). Наличие ломаной линии объясняется погрешностями измерений, недостаточным количеством измерений, физической сущностью исследуемого явления и др. Если на корреляционном поле провести плавную линию между \bar{y}_i , которая равноудалена от них,

то получится новая теоретическая регрессионная зависимость – линия AB (см. рисунок 2.2, a).

Различают однофакторные (парные) и многофакторные регрессионные зависимости. Парная регрессия при парной зависимости может быть аппроксимирована прямой линией, параболой, гиперболой, логарифмической, степенной или показательной функцией, полиномом и др. Двухфакторное поле можно аппроксимировать плоскостью, параболоидом второго порядка, гиперболоидом. Для переменных факторов связь может быть установлена с помощью n -мерного пространства уравнениями второго порядка:

$$y = a_0 + \sum_i^n a_i x_i + \sum_i^n \sum_j^n a_{ij} x_i x_j + \sum_i^n a_{ii} x_i^2, \quad (2.15)$$

где y – функция цели (отклика) многофакторных переменных; x_i – независимые факторы; a_i – коэффициенты регрессии, характеризующие влияние фактора x_i на функцию цели; a_{ij} – коэффициенты, характеризующие двойное влияние факторов x_i и x_j на функцию цели.

При построении теоретической регрессионной зависимости оптимальной является такая функция, в которой соблюдаются условия наименьших квадратов $\sum (y_i - \bar{y})^2 = \min$, где y_i – фактические ординаты поля; \bar{y} – среднее значение ординаты с абсциссой x . Поле корреляции аппроксимируется уравнением прямой $y = a + bx$. Линию регрессии рассчитывают из условий наименьших квадратов. При этом кривая AB (см. рисунок 2.2, a) наилучшим образом выравнивает значения постоянных коэффициентов a и b , т.е. коэффициентов уравнения регрессии. Их вычисляют по выражениям

$$b = (n \sum xy - \sum x \sum y) / (n \sum x^2 - (\sum x)^2), \quad (2.16)$$

$$a = y - bx = \frac{\sum y}{n} - b \frac{\sum x}{n}. \quad (2.17)$$

Критерием близости корреляционной зависимости между x и y к линейной функциональной зависимости является коэффициент парной или просто коэффициент корреляции r , показывающий степень тесноты связи x и y и определяемый отношением

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2] \cdot [n \sum y_i^2 - (\sum y_i)^2]}}, \quad (2.18)$$

где n – число измерений. Значение коэффициента корреляции всегда меньше единицы. При $r = 1.0$ x и y связаны функциональной связью (в данном случае линейной), т.е. каждому значению x соответствует только одно значение y . Если $r < 1$, то линейной связи не существует. При $r = 0$ линейная корреляционная связь между x и y отсутствует, но может существовать нелинейная регрессия. Обычно считают тесноту связи удовлетворительной при $r \geq 0.5$; хорошей при $r = 0.8 \dots 0.85$. Для определения процента разброса (изменчивости) искомой функции y относительно ее среднего значения, определяемого изменчивостью фактора x , вычисляют коэффициент детерминации

$$k_d = r^2. \quad (2.19)$$

Уравнение регрессии прямой можно представить выражением

$$y = \bar{y} + r \frac{\sigma_y}{\sigma_x} (x - \bar{x}).$$

Пусть, например, имеется статистический ряд парных измерений:

1	2	3	4	5	6	7	8	9	10
8	11	14	16	21	26	27	32	34	41

по которому нужно найти уравнение прямолинейной регрессии, оценить тесноту связей и оценить степень достоверности. Расчет целесообразно вести в табличной форме (таблица 2.8).

Таблица 2.8 - Расчет уравнения регрессии

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	x^2	y^2	xy	$(x - \bar{x})(y - \bar{y})$
1	8	-4,5	-15	20,25	225	1	64	8	67,5
2	11	-3,5	-12	12,25	144	4	121	22	42,0
3	14	-2,5	-9	6,25	81	9	196	42	22,5
4	16	-1,5	-7	2,25	49	16	256	64	10,5
5	21	-0,5	-2	0,25	4	25	441	105	1,0
6	26	0,5	+3	0,25	9	36	676	156	1,5
7	27	1,5	+4	2,25	16	49	729	189	6,0
8	32	2,5	+9	6,25	81	64	1024	256	22,5
9	34	3,5	+11	12,25	121	81	1156	306	31,5
10	41	4,5	+18	20,25	324	100	1681	410	81,0
55	230	---	---	82,50	1054	385	6344	1558	286,0

В таблице 2.9 приведена сходимость экспериментальной и теоретической регрессии.

Таблица 2.9 - Сходимость экспериментальной и теоретической регрессии

y	8	11	14	16	21	26	27	32	34	41
$y_{\text{э}}$	7,1	10,6	14,2	17,7	21,8	24,8	28,3	31,9	35,4	39,0

$$\bar{x} = \frac{55}{10} = 5.5; \quad \bar{y} = \frac{230}{10} = 23; \quad \sigma_x = \frac{82.50}{10} = 8.25; \quad \sigma_y = \frac{1054}{10} = 105.4.$$

Коэффициент корреляции согласно (2.18)

$$r = \frac{10 \cdot 1558 - 55 \cdot 230}{\sqrt{(10 \cdot 385 - 55^2)(10 \cdot 6344 - 230^2)}} = 0.99.$$

Из (2.16) и (2.17)

$$b = \frac{10 \cdot 1558 - 55 \cdot 230}{10 \cdot 385 - 55^2} = 3.55;$$

$$a = \frac{230}{10} - 3.55 \frac{55}{10} = 3.48.$$

Уравнение регрессии имеет вид

$$y = 3.48 + 3.55x.$$

Как видим из расчетов, сходимость оказалась хорошей.

Коэффициент детерминации, найденный по формуле (2.19), составляет $k_d = 0.99^2 = 0.98$, что означает, что 98% разброса определяется изменчивостью x , а 2% - другими причинами, т.е. изменчивость функции y почти полностью характеризуется разбросом (природой) фактора x .

На практике часто возникает потребность в установлении связи между y и многими параметрами x_1, x_2, \dots, x_n на основе многофакторной регрессии.

Многофакторные теоретические регрессии аппроксимируются полиномами первого или второго (2.15) порядка. Математические модели характеризуют стохастический

процесс изучаемого явления, уравнение регрессии определяет систематическую, а ошибки разброса – случайную составляющие.

Теоретическую модель множественной регрессии можно получить методами математического планирования, т.е. активным экспериментом, а также пассивным, когда точки факторного пространства выбираются в процессе эксперимента произвольно.

2.4 Методы графической обработки результатов эксперимента

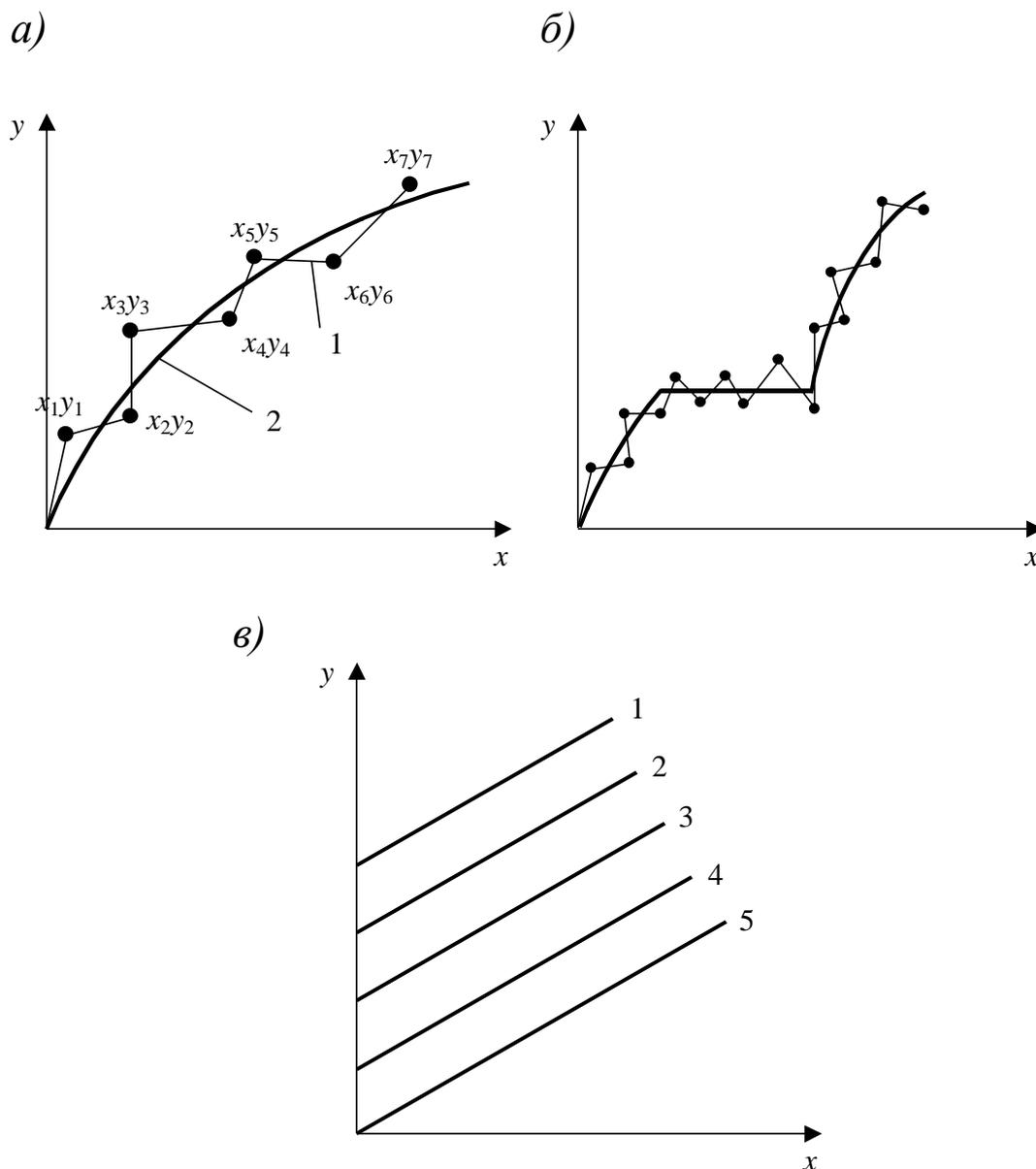
При обработке результатов измерений и наблюдений широко используются методы графического изображения, так как результаты измерений, представленные в табличной форме, иногда не позволяют достаточно наглядно характеризовать закономерности изучаемых процессов. Графическое изображение дает наиболее наглядное представление о результатах эксперимента, позволяет лучше понять физическую сущность исследуемого процесса, выявить общий характер функциональной зависимости изучаемых переменных величин, установить наличие максимума или минимума функции.

Для графического изображения результатов измерений (наблюдений), как правило, применяют систему прямоугольных координат. Если анализируется графическим методом функция $y = f(x)$, то наносят в системе прямоугольных координат значения $x_1y_1, x_2y_2, \dots, x_ny_n$ (рисунок 2.3, а). Прежде чем строить график, необходимо знать ход (течение) исследуемого явления. Как правило, качественные закономерности и форма графика экспериментатору ориентировочно известны из теоретических исследований.

Точки на графике необходимо соединять плавной линией так, чтобы она по возможности проходила ближе ко всем экспериментальным точкам. Если соединить точки прямыми отрезками, то получим ломаную кривую. Она характеризует изменение функции по данным эксперимента. Обычно функции имеют плавный характер. Поэтому при графическом изображении результатов измерений следует проводить между точками плавные кривые. Резкое искривление графика объясняется погрешностью измерений. Если бы эксперимент повторили с применением средств измерений более высокой точности, то получили бы меньше погрешности, а ломаная кривая больше бы соответствовала плавной кривой.

Однако могут быть и исключения, так как иногда исследуются явления, для которых в определенных интервалах наблюдается быстрое скачкообразное изменение одной из координат (рисунок 2.3, б). Это объясняется сущностью физико-химических процессов, например фазовыми превращениями влаги, радиоактивным распадом атомов в процессе исследования радиоактивности и т.д. В таких случаях необходимо особо тщательно соединять точки кривой. Общее «осреднение» всех точек плавной кривой может привести к тому, что скачок функции подменяется погрешностями измерений.

Иногда при построении графика одна-две точки резко удаляются от кривой. В таких случаях вначале следует проанализировать физическую сущность явления, и если нет основания полагать наличие скачка функции, то такое резкое отклонение можно объяснить грубой ошибкой или промахом. Это может возникнуть тогда, когда данные измерений предварительно не исследовались на наличие грубых ошибок измерений. В таких случаях необходимо повторить измерение в диапазоне резкого отклонения данных замера. Если прежнее измерение оказалось ошибочным, то на график наносят новую точку. Если же повторные измерения дадут прежнее значение, необходимо к этому интервалу кривой отнестись особенно внимательно и тщательно проанализировать физическую сущность явления.



a – плавная зависимость: 1 – кривая по результатам непосредственных измерений; 2 – плавная кривая; *б* – при наличии скачка; *в* – при трех переменных: 1 – $z_5 = \text{const}$; 2 – $z_4 = \text{const}$; 3 – $z_3 = \text{const}$; 4 – $z_2 = \text{const}$; 5 – $z_1 = \text{const}$

Рисунок 2.3 – Графическое изображение функции $y = f(x)$

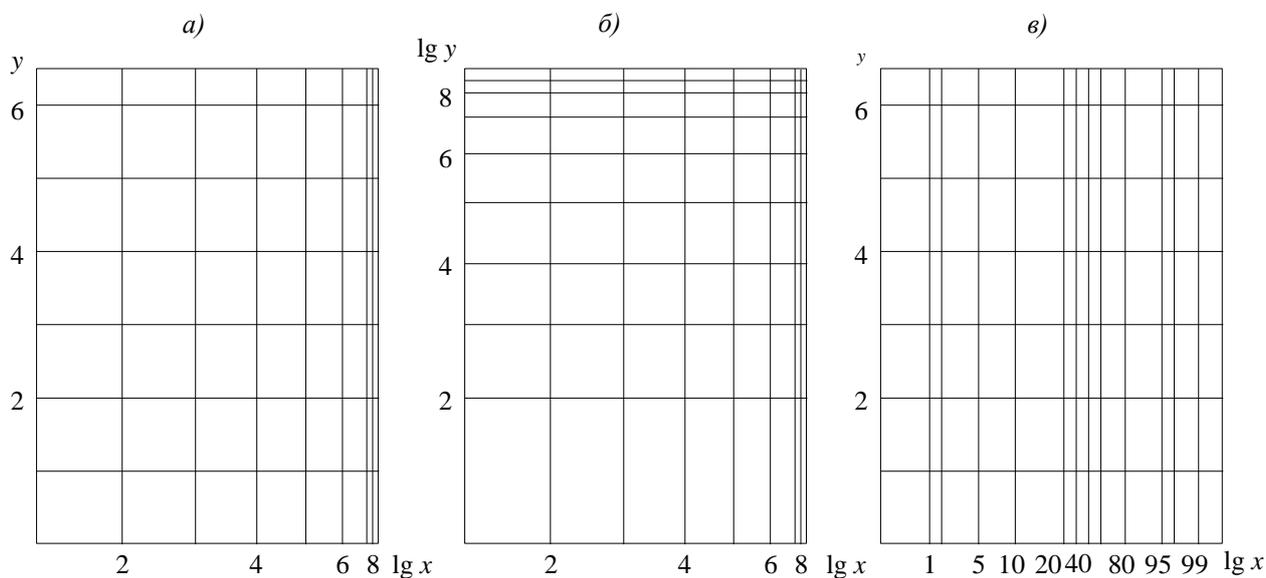
Часто при графическом изображении результатов экспериментов приходится иметь дело с тремя переменными $b = f(x, y, z)$. В этом случае применяют метод разделения переменных. Одной из величин z в пределах интервала измерений $z_1 - z_n$ задают несколько последовательных значений. Для двух остальных переменных x и y строят графики $y = f_1(x)$ при $z_i = \text{const}$. В результате на одном графике получают семейство кривых $y = f_1(x)$ для различных значений z (рисунок 2.3, в). Если необходимо графически изобразить функцию с четырьмя переменными и более $\alpha = f(b, x, y, z)$, то строят серию типа предыдущих, но каждую из них при $b_1, b_2, \dots, b_n = \text{const}$ или принимают из N переменных $(N - 1)$ постоянными и строят графики: вначале $(N - 1) = f_1(x)$, далее $(N - 2) = f_2(x)$, $(N - 3) = f_3(x)$ и т.д. Таким образом, можно проследить изменение любой переменной величины в функции от другой при постоянных значениях остальных. Этот метод графического анализа требует

тщательности, большого внимания к результатам измерений. Однако он в большинстве случаев является наиболее простым и наглядным.

При графическом изображении результатов экспериментов большую роль играет выбор координат или координатной сетки. Координатные сетки бывают равномерными и неравномерными. У равномерных координатных сеток ординаты и абсциссы имеют равномерную шкалу. Например, в системе прямоугольных координат длина откладываемых единичных отрезков на обеих осях одинакова.

Из неравномерных координатных сеток наиболее распространены полулогарифмические, логарифмические и вероятностные. Полулогарифмическая сетка имеет равномерную ординату и логарифмическую абсциссу (рисунок 2.4, а). Логарифмическая координатная сетка имеет обе оси логарифмические (рисунок 2.4, б), вероятностная – ординату обычно равномерную и по абсциссе – вероятностную шкалу (рисунок 2.4, в).

Назначение неравномерных сеток различное. В большинстве случаев их применяют для более наглядного изображения функций. Функция $y = f(x)$ имеет различную форму при различных сетках. Так, многие криволинейные функции спрямляются на логарифмических сетках.



а – полулогарифмическая; б – логарифмическая; в – вероятностная сетка

Рисунок 2.4 – Координатная сетка

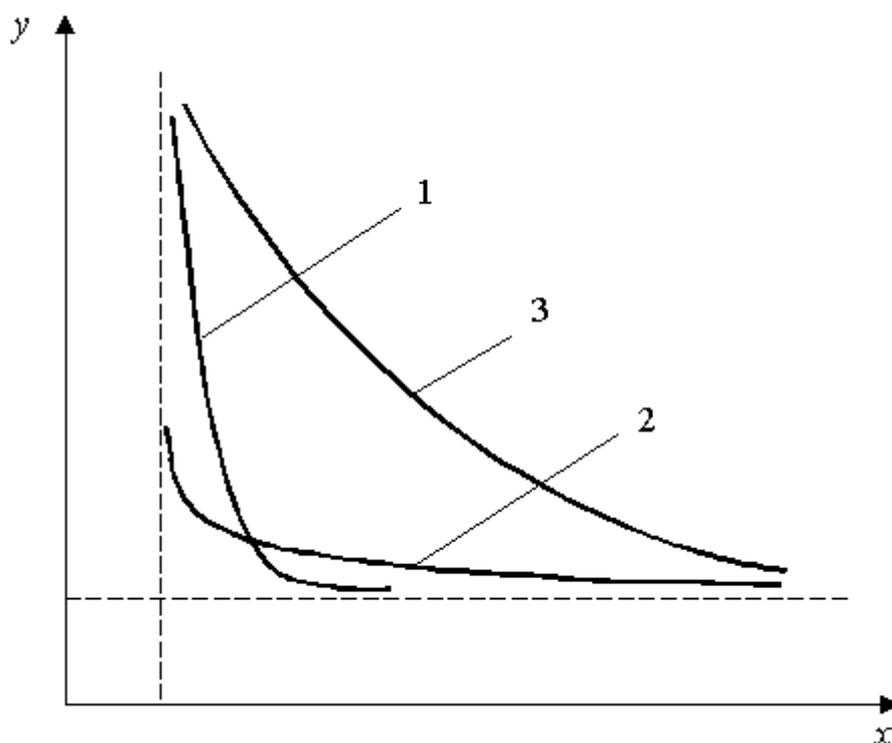
Большое значение в практике графического изображения экспериментальных данных имеет вероятностная сетка, применяемая в различных случаях: при обработке измерений для оценки точности, при определении расчетных характеристик (расчетной влажности, расчетных значений модуля упругости, межремонтных сроков службы и т.д.).

Иногда в процессе обработки экспериментальных данных графическим способом необходимо составить расчетные графики, ускоряющие нахождение по одной переменной других. При этом существенно повышаются требования к точности вычерчивания функции на графике. При вычерчивании расчетных графиков необходимо в зависимости от числа переменных выбрать координатную сетку и определить вид графика – одна кривая, семейство кривых или серия семейств. Большое значение приобретает выбор масштаба графика, что связано с размерами чертежа и соответственно с точностью снимаемых с него величин. Известно, что чем крупнее масштаб, тем выше точность снимаемых значений. Однако, как правило, графики не превышают размеров 20×15 см, что

является удобным при снятии отсчетов. Лишь в отдельных случаях используют графики больших размеров.

Опыт показывает, что применяемая для вычерчивания миллиметровая бумага в пределах размеров 15...20 см дает погрешность, не превышающую $\pm 0.1...0.2$ мм. Это следует иметь в виду при вычерчивании расчетных графиков. Таким образом, абсолютная ошибка снимаемых с графиков величин может достигать $\varepsilon = \pm 0.2M$, где M – принятый масштаб графика. Очевидно, что точность измерений может быть выше точности снимаемых с графика величин.

Масштаб по координатным осям обычно применяют различный. От выбора его зависит форма графика – он может быть плоским (узким) или вытянутым (широким) вдоль оси (рисунок 2.5). Узкие графики дают большую погрешность по оси y ; широкие – по оси x . Из рисунка видно, что правильно подобранный масштаб (нормальный график) позволяет существенно повысить точность отсчетов. Расчетные графики, имеющие минимум или какой-либо сложный вид, особо тщательно необходимо вычерчивать в зонах изгиба. На таких участках количество точек для вычерчивания должно быть значительно больше, чем на плавных участках.



1 – плоская; 2 – уширенная; 3 - нормальная

Рисунок 2.5 – Форма графика в зависимости от масштаба

В некоторых случаях строят номограммы, существенно облегчающие применение для систематических расчетов сложных теоретических или эмпирических формул в определенных пределах измерения величин. Номограммы могут отражать алгебраические выражения и тогда сложные математические выражения можно решать сравнительно просто графическими методами. Построение номограмм – операция трудоемкая. Однако, будучи раз построенной, номограмма может быть использована для нахождения любой из переменных, входящих в номограммированное уравнение. Применение ЭВМ существенно снижает трудоемкость номограммирования. Существует несколько методов построения

номограмм. Для этого применяют равномерные или неравномерные координатные сетки. В системе прямоугольных координат функции в большинстве случаев имеют криволинейную форму. Это увеличивает трудоемкость построения номограмм, поскольку требуется большое количество точек для нанесения одной кривой.

В полу- или логарифмических координатных сетках функции часто имеют прямолинейную форму, и составление номограмм упрощается.

Методика построения номограмм функции одной переменной $y = f(x)$ или многих $f(x_1, x_2, \dots, x_n)$ сводится к построению кривых или их семейств путем принятия постоянными отдельных переменных. Сложные алгебраические выражения целесообразно сводить к простому произведению двух-трех значений, например $d = abc$, где a, b, c – функции двух или трех переменных. В этом случае необходимо вначале, задавшись переменными, вычислить a, b, c . Далее, придавая им постоянные значения, найти d . Величины a, b, c необходимо варьировать в определенных значениях, например, от 0 до 100 через 5 или 10. Наиболее эффективным является такой способ построения номограмм, при котором a, b, c представляются как безразмерные критерии.

2.5 Методы подбора эмпирических формул

В процессе экспериментальных исследований получается статистический ряд измерений двух величин, когда каждому значению функции y_1, y_2, \dots, y_n соответствует определенное значение аргумента x_1, x_2, \dots, x_n .

На основе экспериментальных данных можно подобрать алгебраические выражения функции

$$y = f(x), \quad (2.20)$$

которые называют эмпирическими формулами. Такие формулы подбираются лишь в пределах измеренных значений аргумента $x_1 - x_n$ и имеют тем большую ценность, чем больше соответствуют результатам эксперимента.

Необходимость в подборе эмпирических формул возникает во многих случаях. Так, если аналитическое выражение (2.20) сложное, требует громоздких вычислений, составления программ для ЭВМ или вообще не имеет аналитического решения, то эффективнее пользоваться упрощенной приближенной эмпирической формулой.

Эмпирические формулы должны быть по возможности наиболее простыми и точно соответствовать экспериментальным данным в пределах изменения аргумента. Таким образом, эмпирические формулы являются приближенными выражениями аналитических формул. Замену точных аналитических выражений приближенными, более простыми называют аппроксимацией, а функции – аппроксимирующими.

Процесс подбора эмпирических формул состоит из двух этапов.

I этап. Данные измерений наносят на сетку прямоугольных координат, соединяют экспериментальные точки плавной кривой и выбирают ориентировочно вид формулы.

II этап. Вычисляют параметры формул, которые наилучшим образом соответствовали бы принятой формуле. Подбор эмпирических формул необходимо начинать с самых простых выражений. Так, например, результаты измерений многих явлений и процессов аппроксимируются простейшими уравнениями типа

$$y = a + bx, \quad (2.21)$$

где a, b – постоянные коэффициенты. Поэтому при анализе графического материала необходимо по возможности стремиться к использованию линейной функции. Для этого применяют метод выравнивания, заключающийся в том, что кривую, построенную по экспериментальным точкам, представляют линейной функцией.

Для преобразования некоторой кривой (2.20) в прямую линию вводят новые переменные:

$$X = f_1(x, y), \quad Y = f_2(x, y). \quad (2.22)$$

В искомом уравнении они должны быть связаны линейной зависимостью

$$Y = a + bX. \quad (2.23)$$

Значения X и Y можно вычислить на основе решения системы уравнений (2.22). Далее строят прямую (рисунок 2.6), по которой легко графически вычислить параметры a (ордината точки пересечения прямой с осью Y) и b (тангенс угла наклона прямой с осью X): $b = \operatorname{tg} \alpha = (Y_i - a)/X_i$.

При графическом определении параметров a и b обязательно, чтобы прямая (2.21) строилась на координатной сетке, у которой началом является точка $Y = 0$ и $X = 0$. Для расчета необходимо точки Y_i и X_i принимать на крайних участках прямой.

Для определения параметров прямой можно применить также другой графический метод. В уравнение (2.23) подставляют координаты двух крайних точек, взятых с графика. Получают систему двух уравнений, из которых вычисляют a и b . После установления параметров a и b получают эмпирическую формулу (2.21), которая связывает Y и X , позволяет установить функциональную связь между x и y и эмпирическую зависимость (2.20).

Линеаризацию кривых можно легко осуществить на полу- или логарифмических координатных сетках, которые сравнительно широко применяют при графическом методе подбора эмпирических формул.

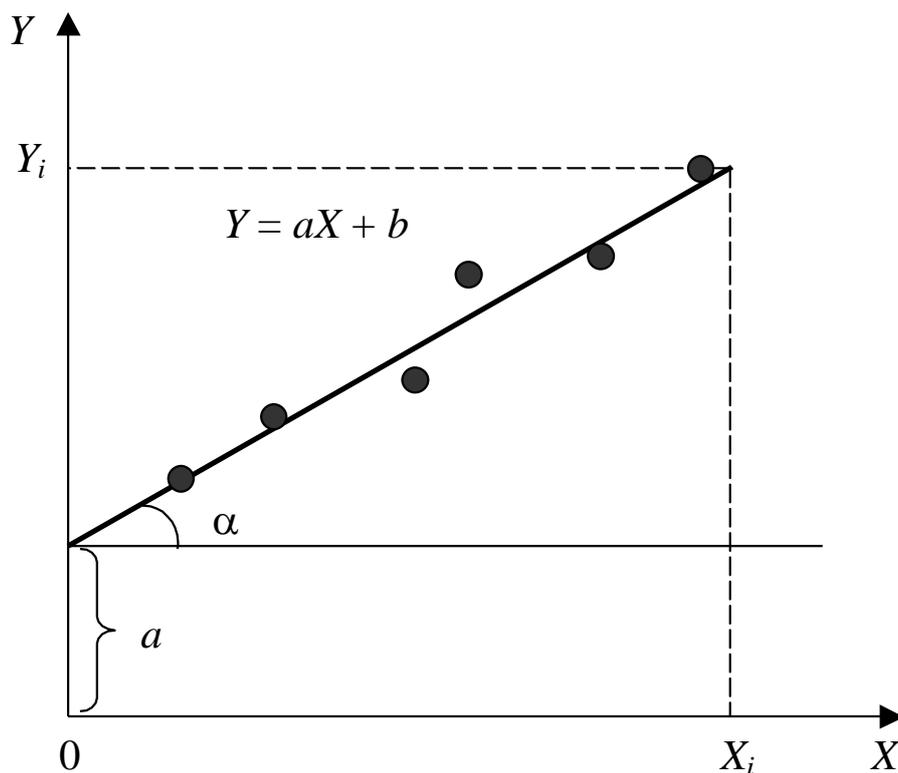


Рисунок 2.6 – Графическое определение параметров x и y

Пример. Подобрать эмпирическую формулу следующих измерений:

12.1	19.2	25.9	33.3	40.5	46.4	54.0
1	2	3	4	5	6	7

Графический анализ этих измерений показывает, что в прямоугольных координатах точки хорошо ложатся на прямую линию и их можно выразить зависимостью (2.21).

Выбираем координаты крайних точек и подставляем в (2.21). Тогда $A_0 + 7A_1 = 54.0$; $A_0 + A_1 = 12.1$, откуда $A_1 = 41.9/6 = 6.98$ и $A_0 = 12.1 - 6.98 = 5.12$. Эмпирическая формула примет вид $y = 5.12 + 6.98A_1$.

Таким образом, аппроксимация экспериментальных данных прямолинейными функциями позволяет просто и быстро установить вид эмпирических формул.

Графический метод выравнивания может быть применен в тех случаях, когда экспериментальная кривая на сетке прямоугольных координат имеет вид плавной кривой. Так, если экспериментальный график имеет вид, показанный на рисунке 2.7, а, то необходимо применить формулу

$$y = ax^b. \quad (2.24)$$

Заменяя $X = \lg x$ и $Y = \lg y$, получим $Y = \lg a + bX$. При этом экспериментальная кривая превращается в прямую линию на логарифмической сетке. Если экспериментальный график имеет вид, показанный на рисунке 2.7, б, то целесообразно использовать выражение

$$y = ae^{bx}. \quad (2.25)$$

При замене $Y = \lg y$ получим $Y = \lg a + bx \lg e$. Здесь экспериментальная кривая превращается в прямую линию на полулогарифмической сетке. Если экспериментальный график имеет вид, представленный на рисунке 2.7, в, то эмпирическая формула принимает вид

$$y = c + ax^b. \quad (2.26)$$

Если b задано, то надо принять $X = x^b$, и тогда получим прямую линию на сетке прямоугольных координат $Y = c + aX$. Если же b неизвестно, то надо принять $X = \lg x$ и $Y = \lg(y - c)$, в этом случае будет прямая линия, но на логарифмической сетке $Y = \lg a + bX$.

В последнем случае необходимо предварительно вычислить c . Для этого по экспериментальной кривой принимают три произвольные точки x_1, y_1 ; x_2, y_2 и $x_3 = \sqrt{x_1 x_2}$, y_3 и вычисляют c в виде отношения

$$c = \frac{y_1 y_2 - y_3^2}{y_1 + y_2 - 2y_3}. \quad (2.27)$$

Если экспериментальный график имеет вид, показанный на рисунке 2.7, г, то нужно пользоваться формулой

$$y = c + ae^{bx}. \quad (2.28)$$

Путем замены $Y = \lg(y - c)$ можно построить прямую на полулогарифмической сетке:

$$Y = \lg a + bx \lg c,$$

где c предварительно определено с помощью формулы (2.27). В этом случае $x_3 = 0.5(x_1 + x_2)$.

Если экспериментальный график имеет вид, представленный на рисунке 2.7, д, то применяется выражение

$$y = a + b/x. \quad (2.29)$$

Путем замены $x = 1/z$ можно получить прямую линию на сетке прямоугольных координат $y = a + bz$.

Если график имеет вид, соответствующий кривым на рисунке 2.7, е, то используют формулу

$$y = 1/(a + bx). \quad (2.30)$$

Если принять $y = 1/z$, то $z = a + bx$, т.е. прямая на сетке прямоугольных координат. Аналогично, уравнению

$$y = \frac{1}{a + bx + cx^2} \quad (2.31)$$

путем замены $y = 1/z$ можно придать вид $z = a + bx + cx^2$.

Сложную степенную функцию

$$y = ae^{nx+mx^2} \quad (2.32)$$

можно преобразовать в более простую. При $\lg y = z$; $\lg a = p$; $n \lg e = q$; $m \lg e = r$ получается зависимость

$$z = p + qx + rx^2.$$

С помощью приведенных на рисунке 2.7 графиков и выражений (2.24) – (2.32) можно практически всегда подобрать уравнение эмпирической формулы.

Пусть, например, необходимо подобрать эмпирическую формулу для следующих измерений:

1	1.5	2.0	2.5	3.0	3.5	4.0	4.5
15.2	20.6	27.4	36.7	49.2	66.0	87.4	117.5

На основе этих данных строится график (рисунок 2.8, а), соответствующий кривым (2.25) (рисунок 2.7, б).

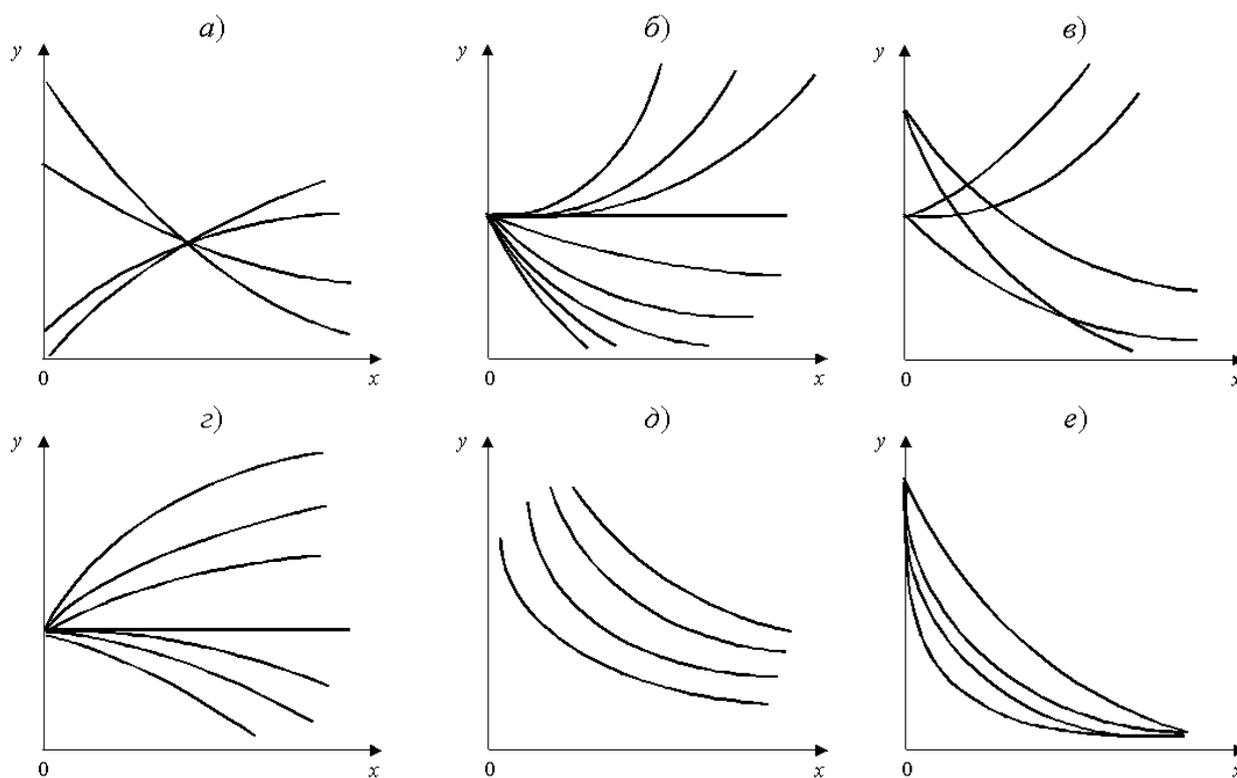


Рисунок 2.7 – Основные виды графиков эмпирических формул

После логарифмирования выражения (2.25) $\lg y = \lg a + bx \lg e$. Если обозначить $\lg y = Y$, то $Y = \lg a + bx \lg e$, т.е. в полулогарифмических координатах выражение для Y представляет собой прямую линию (рисунок 2.8, б). Подстановка в уравнение координат крайних точек дает $\lg 15.2 = \lg a + b \lg e$ и $\lg 117.5 = \lg a + 4.5b \lg e$. Следовательно,

$$\begin{aligned} \lg a + b \lg e &= 1.182; \\ \lg a + 4.5b \lg e &= 2.070, \end{aligned}$$

откуда $b = 0.888/(3.5 \lg e) = 0.584$; $\lg a = 1.182 - 0.254 = 0.928$; $a = 8.47$. Окончательно эмпирическая формула получит вид

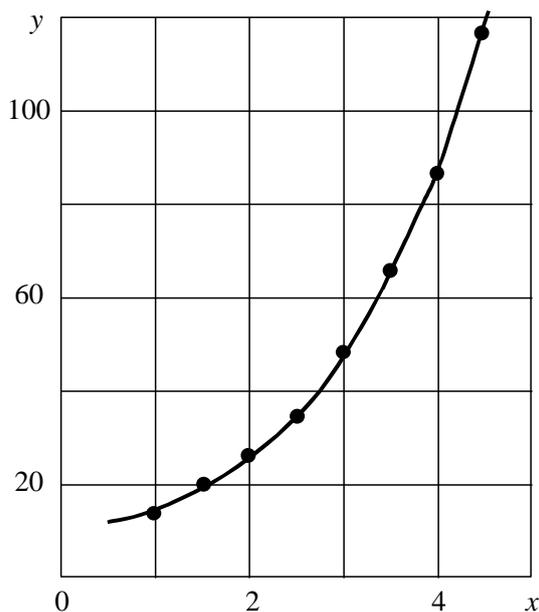
$$y = 8.47 \cdot e^{0.584x}.$$

При подборе эмпирических формул широко используются полиномы

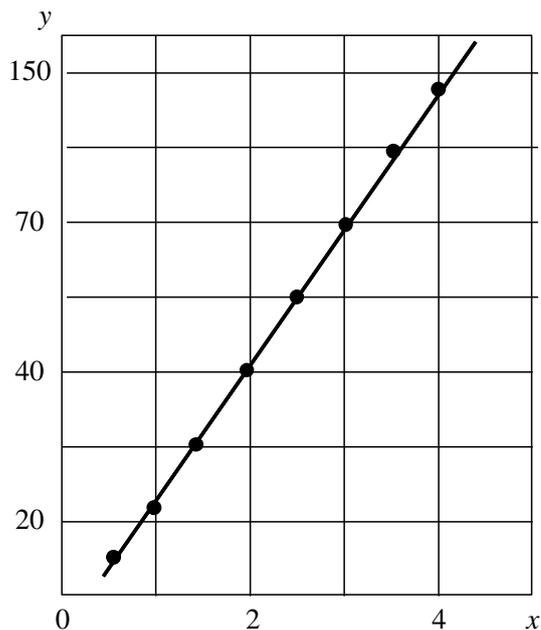
$$y = A_0 + A_1x + A_2x^2 + A_3x^3 + \dots + A_nx^n, \quad (2.33)$$

где $A_0, A_1, A_2, \dots, A_n$ – постоянные коэффициенты. Полиномами можно аппроксимировать любые результаты измерений, если они графически выражаются непрерывными функциями. Особо ценным является то, что даже при неизвестном точном выражении функции (2.33) можно определить значения коэффициентов A . Для определения коэффициентов A кроме графического метода, изложенного выше, применяют методы средних и наименьших квадратов.

a)



б)



a - эмпирическая; *б* - спрямленная

Рисунок 2.8 – Подбор эмпирической характеристики

Метод средних квадратов основан на следующем положении. По экспериментальным точкам можно построить несколько плавных кривых. Наилучшей будет та кривая, у которой разностные отклонения оказываются наименьшими, т.е. $\Sigma \varepsilon \approx 0$. Порядок расчета коэффициентов полинома сводится к следующему. Определяется число членов ряда (2.33), которое обычно принимают не более 3...4. В принятое выражение последовательно подставляют координаты x и y нескольких (m) экспериментальных точек и получают систему из m уравнений. Каждое уравнение приравнивают соответствующему отклонению:

$$\begin{aligned} A_0 + A_1x_1 + A_2x_1^2 + \dots + A_nx_1^n - y_1 &= \varepsilon_1; \\ A_0 + A_1x_2 + A_2x_2^2 + \dots + A_nx_2^n - y_2 &= \varepsilon_2; \\ &\dots\dots\dots \\ A_0 + A_1x_m + A_2x_m^2 + \dots + A_nx_m^n - y_m &= \varepsilon_m. \end{aligned} \quad (2.34)$$

Число точек, т.е. число уравнений, должно быть не меньше числа коэффициентов A , что позволяет их вычислить путем решения системы (2.34).

Разбивают систему начальных уравнений (2.34) последовательно сверху вниз на группы, число которых должно быть равно количеству коэффициентов A_0 . В каждой

группе складывают уравнения и получают новую систему уравнений, равную количеству групп (обычно 2...3). Решая систему, вычисляют коэффициенты A .

Метод средних обладает высокой точностью, если число точек достаточно велико (не менее 3...4). Однако степень точности можно повысить, если начальные условия сгруппировать по 2...3 варианта и вычислить для каждого варианта эмпирическую формулу. Предпочтение следует отдать той формуле, у которой $\sum \varepsilon^2 = \min$. Пусть, например, выполнено семь измерений:

4	5	6	7	8	9	10
10.2	6.7	4.8	3.6	2.7	2.1	1.7

Для подбора эмпирической формулы можно выбрать полином

$$y = A_0 + A_1x + A_2x^2.$$

Путем подстановки в это уравнение значений измерений систему начальных уравнений можно разделить на три группы: 1...2; 3...4; 5...7 в виде

$$A_0 + 4A_1 + 16A_2 - 10.2 = \varepsilon_1;$$

$$A_0 + 5A_1 + 25A_2 - 6.7 = \varepsilon_2;$$

$$A_0 + 6A_1 + 36A_2 - 4.8 = \varepsilon_3;$$

$$A_0 + 7A_1 + 49A_2 - 3.6 = \varepsilon_4;$$

$$A_0 + 8A_1 + 64A_2 - 2.7 = \varepsilon_5;$$

$$A_0 + 9A_1 + 81A_2 - 2.1 = \varepsilon_6;$$

$$A_0 + 10A_1 + 100A_2 - 1.7 = \varepsilon_7.$$

Сложение уравнений в каждой подгруппе дает

$$1\text{-я группа } 2A_0 + 9A_1 + 41A_2 = 16.9;$$

$$2\text{-я группа } 2A_0 + 13A_1 + 85A_2 = 8.4;$$

$$3\text{-я группа } 3A_0 + 27A_1 + 245A_2 = 6.5.$$

Определение из этих выражений коэффициентов A_0 , A_1 и A_2 приводит к эмпирической формуле $y = 26.168 - 5.2168x + 0.2811x^2$.

Метод средних квадратов может быть применен для различных кривых после их выравнивания. Пусть, например, имеется восемь измерений:

3	6	9	12	15	18	21	24
57.6	41.9	31.0	22.7	16.6	12.2	8.9	6.5

Анализ кривой в системе прямоугольных координат дает возможность применить формулу (2.25)

$$y = ae^{-bx}.$$

Произведем выравнивание путем замены переменных $Y = \lg y$, $X = x/2.303$. Тогда $Y = A + BX$, где $A = \lg a$, $B = b$. Так как необходимо определить два параметра, то все измерения делятся на две группы по четыре измерения. Это приводит к уравнениям:

$$1.7604 = A + \frac{3}{2.303} B; \quad 1.2201 = A + \frac{15}{2.303} B;$$

$$1.6222 = A + \frac{6}{2.303} B; \quad 1.0864 = A + \frac{18}{2.303} B;$$

$$1.4914 = A + \frac{9}{2.303} B; \quad 0.9494 = A + \frac{21}{2.303} B;$$

$$1.3560 = A + \frac{12}{2.303} B; \quad 0.8129 = A + \frac{24}{2.303} B;$$

$$6.2300 = 4A + \frac{30}{2.303} B; \quad 4.0688 = 4A + \frac{78}{2.303} B.$$

После суммирования по группам можно получить систему двух уравнений с двумя неизвестными A и B , решение которых дает: $A = 1.8952$; $a = 78.56$; $B = -0.1037$; $b = -0.1037$. Окончательно $y = 78.56e^{-0.1037x}$.

Хорошие результаты при определении параметров заданного уравнения дает использование метода наименьших квадратов. Суть этого метода заключается в том, что если все измерения функций y_1, y_2, \dots, y_n произведены с одинаковой точностью и распределение величины ошибок измерения соответствует нормальному закону, то параметры исследуемого уравнения определяются из условия, при котором сумма квадратов отклонений измеренных значений от расчетных принимает наименьшее значение. Для нахождения неизвестных параметров (a_1, a_2, \dots, a_n) необходимо решить систему линейных уравнений:

$$\begin{aligned} y_1 &= a_1x_1 + a_2u_1 + \dots + a_nz_1; \\ y_2 &= a_1x_2 + a_2u_2 + \dots + a_nz_2; \\ &\dots\dots\dots \\ y_n &= a_1x_m + a_2u_m + \dots + a_nz_m, \end{aligned}$$

где y_1, y_2, \dots, y_n – частные значения измеренных величин функции y ; x, u, z – переменные величины. Эту систему приводят к системе линейных уравнений путем умножения каждого уравнения соответственно на x_1, x_2, \dots, x_m и последующего их сложения, затем умножения соответственно на u_1, u_2, \dots, u_m . Это позволяет получить так называемую систему нормальных уравнений

$$\begin{aligned} \sum_1^m yx &= a_1 \sum_1^m xx + a_2 \sum_1^m xu + \dots + a_n \sum_1^m xz; \\ \sum_1^m yu &= a_1 \sum_1^m ux + a_2 \sum_1^m uu + \dots + a_n \sum_1^m uz; \\ &\dots\dots\dots \\ \sum_1^m yz &= a_1 \sum_1^m zx + a_2 \sum_1^m zu + \dots + a_n \sum_1^m zz, \end{aligned}$$

решение которой и дает искомые коэффициенты.

Пусть, например, необходимо определить коэффициенты a_1 и a_2 в уравнении $k_p = a_1 + a_2u$. Так как требуется определить два параметра, то система уравнений может быть представлена в виде двух уравнений $y = a_1x_1 + a_2x_2$ и $yx_2 = a_1x_1x_2 + a_2x_2^2$, где $y = k_p$; $x_1 = 1$; $x_2 = u$.

Так как уравнения линейные, можно ограничиться четырьмя сериями опытов. Если они сведены в таблицу 2.10, то систему нормальных уравнений можно записать в виде $5.48 = 4a_1 + 1100a_2$; $1519 = 1100a_1 + 307350a_2$, решение которых дает $a_1 = 0.78$; $a_2 = 0.0025$. Следовательно, эмпирическая формула получит вид $k_p = 0.78 + 0.0025u$.

Таблица 2.10 - Результаты опытов

$x_2 = u$	$y = k_p$	x_2^2	yx_2
230	1,26	52900	289,8
255	1,32	65025	336,6
295	1,40	87025	413,0
320	1,50	102400	480,0
1100	5,48	307350	1519,4

Метод наименьших квадратов обеспечивает достаточно надежные результаты. При этом степень точности коэффициентов A в (2.33) должна быть такой, чтобы вычисленные значения y совпадали со значениями в исходных табличных значениях. Это требует вычислять значения A тем точнее, чем выше индекс A , т.е. A_1 должно быть точнее (больше число десятичных знаков), чем A_2 ; A_3 – точнее, чем A_2 , и т.д. Для вычисления коэффициентов $k_p = a_1 + a_2u$ методом наименьших квадратов необходимо пользоваться типовыми программами для ЭВМ.

2.6 Оценка адекватности результатов эксперимента

В результате эксперимента получают статистический ряд обычно парных, однофакторных (x_i, y_i) или многофакторных (a_i, b_i, c_i, \dots) измерений. Статистические измерения подвергают обработке и анализу, подбирают эмпирические формулы и устанавливают их достоверность.

Перед подбором эмпирических формул необходимо еще раз убедиться в достоверности эксперимента, окончательно проверить воспроизводимость результатов по критерию Кохрена. Оценка пригодности гипотезы исследования, а также теоретических данных на адекватность, т.е. соответствие теоретической кривой экспериментальным данным, необходима во всех случаях на стадии анализа теоретико-экспериментальных исследований. Методы оценки адекватности основаны на использовании доверительных интервалов, позволяющих с заданной доверительной вероятностью определить искомые значения оцениваемого параметра. Суть такой проверки состоит в сопоставлении полученной или предполагаемой теоретической функции $y = f(x)$ с результатами измерений. В практике оценки адекватности применяют различные статистические критерии согласия.

Одним из таких критериев является критерий Фишера. Установление адекватности – это определение ошибки аппроксимации опытных данных. Для этого необходимо рассчитать экспериментальное (опытное) значение критерия Фишера – $k_{фэ}$ и сравнить его с теоретическим (табличным) – $k_{фт}$, принимаемым при требуемой доверительной вероятности p_d (обычно $p_d = 0.95$). Если $k_{фэ} < k_{фт}$ – модель адекватна; если $k_{фэ} \geq k_{фт}$ – модель неадекватна. Опытный критерий Фишера вычисляют по формуле

$$k_{фэ} = D_a / D_{ср}, \quad (2.35)$$

где D_a – дисперсия адекватности; $D_{ср}$ – средняя дисперсия всего эксперимента, определяющиеся как

$$D_a = \frac{1}{n-d} \sum_{i=1}^n (y_{iT} - \bar{y}_{iЭ})^2; \quad (2.36)$$

$$D_{cp} = \frac{\sum_{i=1}^m \sum_{j=1}^n (y_{iT} - y_{iЭ})^2}{mn}.$$

Здесь y_{iT} – теоретическое значение функции для каждого измерения; $y_{iЭ}$ – экспериментальное значение функции; $\bar{y}_{iЭ}$ – среднее экспериментальное значение функции из m серий измерений; n – количество измерений в одном опыте (одной серии или количество опытов); d – число коэффициентов уравнения теоретической регрессии.

Значения $k_{фТ}$ принимается по таблице 2.11 для доверительной вероятности 0.95 и числа степеней свободы $q_1 = n - d$, $q_2 = n(m - 1)$. В уравнении (2.36) y_{iT} вычисляются по теоретической регрессии для фактора x_i ; $\bar{y}_{iЭ}$ – как среднее из m серий измерений, т.е.

$$\bar{y}_{iЭ} = \frac{1}{m} (y_{1Э} + y_{2Э} + \dots + y_{mЭ}).$$

Таблица 2.11 - Критерий Фишера

q_1	Значения $k_{фТ}$ при $p_d = 0.95$ для различных q_2								
	1	2	3	4	5	6	12	24	36
1	16	19	21	22	23	23	24	24	25
2	18	19	19	19	19	19	19	19	19
3	10	9.6	9.3	9.1	9.0	8.9	8.7	8.6	8.5
4	7.7	6.9	6.6	6.4	6.3	6.2	5.9	5.8	5.6
5	6.6	5.8	5.4	5.2	5.1	5.0	4.7	4.5	4.4
6	6.0	5.1	4.8	4.5	4.4	4.3	4.0	3.8	3.7
7	5.6	4.7	4.4	4.1	4.0	3.9	3.6	3.4	3.2
8	5.3	4.5	4.1	3.8	3.7	3.6	3.3	3.1	2.9
9	5.1	4.3	3.9	3.6	3.5	3.4	3.1	2.9	2.7
10	5.0	4.1	3.7	3.5	3.3	3.2	2.9	2.7	2.5
11	4.8	4.0	3.6	3.4	3.2	3.1	2.8	2.6	2.4
12	4.8	3.9	3.5	3.3	3.1	3.0	2.7	2.5	2.3
13	4.7	3.8	3.4	3.2	3.0	2.9	2.6	2.4	2.2
14	4.6	3.7	3.3	3.1	3.0	2.9	2.5	2.3	2.1
15	4.5	3.7	3.3	3.1	2.9	2.8	2.5	2.3	2.1
16	4.5	3.6	3.2	3.0	2.9	2.7	2.4	2.2	2.0
17	4.5	3.6	3.2	3.0	2.8	2.7	2.4	2.2	2.0
18	4.4	3.6	3.2	2.9	2.8	2.7	2.3	2.1	1.9
19	4.4	3.5	3.1	2.9	2.7	2.6	2.3	2.1	1.8
20	4.4	3.5	3.1	2.9	2.7	2.6	2.3	2.1	1.8
22	4.3	3.4	3.1	2.8	2.7	2.6	2.2	2.0	1.8
26	4.2	3.4	3.0	2.7	2.7	2.4	2.1	1.9	1.7
30	4.2	3.3	2.9	2.7	2.5	2.4	2.1	1.9	1.7
60	4.0	3.2	2.9	2.5	2.4	2.3	1.9	1.7	1.4
∞	3.8	3.0	2.6	2.4	2.2	2.1	1.8	1.5	1.0

Пусть, например, получено теоретическое выражение $y = 80x$ и для его подтверждения проведен эксперимент. В каждой из пяти серий (повторностей, $m = 5$) выполнено по семь измерений ($n = 7$). Результаты эксперимента приведены в таблице 2.12.

По этим данным можно установить пригодность, т.е. адекватность теоретического выражения.

С этой целью по формуле (2.36) определяется дисперсия адекватности $D_a = 5.36/(7 - 1) = 0.89$. Здесь значение $d = 1$, поскольку в теоретическом выражении один значащий член x . Дисперсия D_{cp} вначале вычисляется построчно для m строк. Для первой строки $D_1 = \Sigma(y_{iT} - y_{iэ})^2/m = 1/5[(12 - 16)^2 + (17 - 16)^2 + (15 - 16)^2 + (14 - 16)^2 + (16 - 16)^2] = 4.4$; для второй строки $D_2 = 1/5[(23 - 24)^2 + (21 - 24)^2 + (24 - 24)^2 + (25 - 24)^2 + (23 - 24)^2] = 2.4$ и т.д. Тогда средняя дисперсия всего эксперимента составит $D_{cp} = \sum_1^n D_i / n = 40.4/7 = 5.77$.

После этого по формуле (2.35) подсчитывается $k_{фэ} = 0.89/5.77 = 0.15$.

Таблица 2.12 - Результаты эксперимента и оценка адекватности его теоретического представления

№ опытов	x_i	Измеренные значения $y_{iэ}$ в серии					Средние значения $\sum_1^m y_{iэ}$ $\bar{y}_{iэ} = \frac{1}{m}$	y_{iT}	$y_{iT} - \bar{y}_{iэ}$	$(y_{iT} - \bar{y}_{iэ})^2$	$\frac{\sum_1^m (y_{iT} - \bar{y}_{iэ})^2}{m}$
		$y_{1э}$	$y_{2э}$	$y_{3э}$	$y_{4э}$	$y_{5э}$					
1	0,2	12	17	15	14	16	14,8	16	1,2	1,44	4,4
2	0,3	23	21	24	25	23	23,2	24	0,8	0,64	2,4
3	0,4	30	34	31	35	35	33,0	32	1,0	1,00	3,8
4	0,5	38	43	40	39	42	40,4	40	0,4	0,16	3,6
5	0,6	52	47	48	49	40	47,2	48	0,8	0,64	16,4
6	0,7	59	58	55	54	53	55,8	56	0,2	0,04	5,4
7	0,8	62	66	62	61	63	62,8	64	1,2	1,44	4,4
Итого										5,36	40,4

Теоретические значения критерия Фишера можно принять по данным таблицы 2.11 при следующих степенях свободы: $q_1 = 7 - 1 = 6$, $q_2 = 7(5 - 1) = 27$, $k_{фт} = 3.75$. Так как $k_{фэ} = 0.15 < k_{фт} = 3.75$, то модель адекватна, т.е. полученная математическая модель с доверительной вероятностью 95% хорошо описывает изучаемый процесс.

Критерий Фишера обычно применяется для определения адекватности малых выборок. В больших выборках целесообразно применять критерии Пирсона, Романовского, Колмогорова. Так, критерий Пирсона наиболее широко применяется при больших статистических измерениях. В соответствии с этим критерием гипотеза о законе распределения подтверждается, если соблюдается условие

$$p(\chi^2, q) > \alpha = 1 - \varphi(x).$$

Здесь $\alpha = 1 - \varphi(x)$ – уровень значимости, обычно принимаемый равным 0.10; χ – критерий согласия Пирсона; q – число степеней свободы, равное

$$q = m - s, \quad (2.37)$$

где m – количество групп (серий, разрядов) большой выборки или число измерений в одной серии при анализе односерийного эксперимента; s – число используемых связей (констант).

Значение χ^2 вычисляют по формуле

$$\chi^2 = \sum_1^m \frac{(y_{Эi} - y_{Ti})^2}{y_{Ti}}, \quad (2.38)$$

где $y_{Эi}$, y_{Ti} – количество измерений (частота) в каждой группе серий соответственно по данным эксперимента и теоретической кривой.

Пусть имеется большая выборка N измерений. Статистические измерения следует разделить на m разрядов: $x_1 \dots x_2$; $x_3 \dots x_4$; $x_5 \dots x_6$ и т.д. По данным измерений в каждом разряде может оказаться $y_{Э}$ измерений. Так, в диапазоне $x_1 \dots x_2$ имеется $y_{Э1}$ измерений (частота); в $x_3 \dots x_4$ имеется $y_{Э2}$ измерений и т.д. Очевидно, $\sum_1^m y_{Эi} = N$. По данным

эксперимента следует построить экспериментальную кривую частот по $y_{Эi} = f(x)$ или $y_{Эi}/N = f(x)$. Эту кривую можно аппроксимировать какой-то теоретической кривой (законом Пуассона, показательным, логарифмическим, нормальным и др.). Для этой теоретической кривой устанавливают соответствующие экспериментальные частоты y_{mi} , производят вычисления критерия Пирсона χ^2 по формуле (2.38) и сравнивают его с данными таблицы 2.13.

Пусть, например, произведено $N = 250$ измерений некоторых величин x_i , и по ним необходимо определить закон распределения. Для этого экспериментальные данные $y_{Эi}$ целесообразно разбить на семь групп, результаты измерений нанести на сетку в прямоугольных координатах и установить, что кривая близка к закону нормального распределения. В таких случаях в качестве аппроксимирующей принимают кривую нормального распределения, по которой устанавливают, соответственно, теоретические частоты:

1	23	50	82	58	28	2
1	27	57	80	57	27	1

и по формуле (2.38) вычисляют критерий согласия

$$\chi_{Э}^2 = \frac{(1-1)^2}{1} + \frac{(23-27)^2}{27} + \frac{(50-57)^2}{57} + \frac{(82-80)^2}{80} + \frac{(58-57)^2}{57} + \frac{(28-27)^2}{27} + \frac{(2-1)^2}{1} = 2.56.$$

По количеству разрядов $m = 7$, констант нормального закона $s = 2$, $q = 7 - 2 = 5$. По данным таблицы 2.13 в соответствии с p (2.56; 5) определяется $\chi_{1}^2 = 0.774$. Это свидетельствует о том, что адекватность удовлетворяется, поскольку $0.774 > 0.10$.

Критерий Романовского определяется отношением

$$k_p = (\chi^2 - q) / \sqrt{2q}.$$

Число степеней свободы q определяется разностью (2.37). Адекватность удовлетворяется при $k_p < 3$. Для рассмотренного примера $k_p = \frac{(2.56 - 5)}{\sqrt{2 \cdot 5}} < 3$. Таким образом, по критериям Пирсона и Романовского гипотеза о принадлежности экспериментальных данных кривой нормального распределения подтверждается.

Критерий Колмогорова k_K применяется для оценки адекватности также при большой статистической выборке N .

Таблица 2.13 - Критерий Пирсона

χ^2	Значения критерия Пирсона $p(\chi^2, q)$ при числе степеней свободы q							
	1	2	3	4	5	6	7	8
1	0,317	0,606	0,801	0,909	0,962	0,985	0,994	0,998
2	0,157	0,367	0,572	0,735	0,849	0,919	0,959	0,981
3	0,083	0,223	0,391	0,557	0,700	0,806	0,885	0,934
4	0,045	0,135	0,261	0,406	0,549	0,767	0,790	0,854
5	0,025	0,083	0,171	0,287	0,415	0,543	0,660	0,757
6	0,014	0,049	0,111	0,199	0,306	0,423	0,539	0,647
7	0,008	0,030	0,071	0,135	0,220	0,320	0,428	0,536
8	0,004	0,018	0,046	0,091	0,156	0,238	0,332	0,433
9	0,002	0,011	0,020	0,061	0,109	0,173	0,252	0,342
10	0,001	0,006	0,018	0,040	0,075	0,124	0,188	0,265
11	0,000	0,004	0,011	0,026	0,051	0,088	0,138	0,201
12	---	0,002	0,007	0,017	0,034	0,062	0,100	0,151
13	---	0,001	0,004	0,011	0,023	0,043	0,072	0,111
14	---	0,000	0,002	0,007	0,014	0,029	0,036	0,059
15	---	---	0,001	0,004	0,010	0,020	0,030	0,042

Чтобы определить этот критерий, статистическую кривую частот преобразовывают в статистическую интегральную функцию, находят наибольшую разность частот между экспериментальной статистической интегральной кривой и соответствующей теоретической интегральной кривой:

$$D_0 = \max (\sum y_{эi} - \sum y_{тi}).$$

Затем вычисляют

$$\lambda = D_0 \sqrt{N}$$

и по значению λ в специальных таблицах находят вероятность $p(\lambda)$. Адекватность удовлетворяется, если $p(\lambda) > 0.05$, т.е. экспериментальные данные подтверждают теоретическое распределение.

2.7 Метрологическое обеспечение эксперимента

Неотъемлемой частью экспериментальных исследований являются средства измерений, т.е. совокупность технических средств (имеющих нормированные погрешности), которые дают необходимую информацию для эксперимента.

Существует основные группы приборов для измерения показателей: физических, механических, химических свойств, а также структуры материала и изделия.

Наряду с этим можно выделить средства измерения, позволяющие непосредственно определить испытуемый показатель (например, процесс для определения прочности материалов), и измерения, которые дают возможность косвенно судить об исследуемом показателе (ультразвуковые дефектоскопы, что позволяет оценить прочность материала по скорости прохождения ультразвука).

К средствам измерений относят **измерительный инструмент, измерительные приборы и установки**. Измерительные средства делят на образцовые и технические.

Образцовые средства являются эталонами. Они предназначены для проверки технических, т.е. рабочих, средств. Образцовые средства не обязательно должны быть точнее рабочих, но они должны иметь большую стабильность и надежность в

воспроизведении. Образцовые средства не применяют для рабочих измерений. С целью повысить точность и чувствительность измерений, а также расширить диапазон измерений дополнительно используют измерительные преобразователи.

Измерительным прибором называют средство измерения, предназначенное для получения определенной информации об изучаемой величине в удобной для экспериментатора форме. В этих приборах измеряемая величина преобразуется в показание или сигнал.

Приборы классифицируют по различным признакам. По способу отсчета значения измеряемой величины их делят на показывающие и регистрирующие.

Измерительная установка (стенд) представляет собой систему, состоящую из основных и вспомогательных средств измерений, предназначенных для измерения одного сложного или нескольких параметров.

Измерительные установки могут вырабатывать также сигналы, удобные не только для снятия наблюдений, но и для автоматической обработки измерений, например, с помощью ЭВМ.

Измерительные приборы (отсчетные устройства) характеризуются погрешностью и точностью, стабильностью измерений и чувствительностью.

Погрешности приборов. Под абсолютной погрешностью измерительного прибора понимают

$$b = \pm(x_n - x_d),$$

где x_n – показание прибора (номинальное значение измеряемой величины); x_d – действительное значение измеренной величины более точным методом.

Погрешность прибора – одна из важнейших его характеристик. Она возникает вследствие ряда причин: недоброкачественных материалов, комплектующих изделий, применяемых для изготовления приборов; плохого качества изготовления приборов; неудовлетворительной эксплуатации его и др. Существенное влияние оказывают градуировка шкалы и периодическая поверка приборов.

Кроме этих систематических погрешностей, возникают случайные, обусловленные сочетанием различных факторов – ошибками отсчета, параллаксом, вариацией и т.д. Таким образом, необходимо рассматривать не какие-либо отдельные, а суммарные погрешности приборов.

Часто для оценки погрешности приборов применяют относительную погрешность (в %):

$$b_{OT} = \pm \frac{x_n - x_d}{x_d} 100.$$

Иногда применяют понятие приведенной погрешности

$$b_{np} = \pm \frac{x_n - x_d}{x_{np}} 100,$$

где x_{np} – какое-либо значение шкалы измерительного устройства (диапазон измерений), длина шкалы и др.

Суммарные погрешности, установленные при определенных условиях ($t_b = 20^\circ\text{C}$, влажность воздуха 80%, $p = 760$ мм рт. ст.), называют основными погрешностями прибора.

Диапазоном измерений называют ту часть диапазона показаний прибора, для которой установлены погрешности прибора. Если известны погрешности прибора, то диапазоны измерений и показаний прибора совпадают.

Диапазон измерений является важной характеристикой прибора. Если шкала измерений изменяется от 0 до N , то в характеристике на прибор диапазон указывают в пределах 0- N . Ряд приборов с нижним пределом измерения 0 имеет большую погрешность

в интервале 0-25% от верхнего предела, т.е. четверть длины шкалы (начало) может давать погрешность, превышающую b .

Приборы нельзя перегружать, т.е. верхний предел измерений не нужно превышать нагрузкой. Некоторые приборы выдерживают перегрузки. Однако со временем погрешности у верхнего предела измерений существенно возрастают. Ряд приборов может выдерживать только ограниченные перегрузки.

Разность между максимальным и минимальным показателями прибора называют размахом. Если эта величина непостоянная, т.е. если при обратном ходе имеется увеличение или уменьшение хода, то эту разницу называют вариацией показаний ω . ω – это простейшая характеристика погрешности прибора. Другой характеристикой прибора является его чувствительность, т.е. способность отсчитывающего устройства реагировать на изменения измеряемой величины. Под порогом чувствительности прибора понимают наименьшее значение измеренной величины, вызывающее изменение показания прибора, которое можно зафиксировать.

Точность прибора – основная его характеристика. Она характеризуется суммарной погрешностью.

Средства измерения делятся на классы точности в зависимости от допускаемых погрешностей. Способы обозначения классов точности приборов различны.

Класс точности прибора (1-й – наивысший, 4-й – наинизший) обозначает допустимую, суммарную, относительную погрешность от верхнего предела измерений. Так, если класс прибора равен 1, то допускаемая относительная погрешность равна $\pm 1\%$. Для прибора с верхним пределом измерения 3 г/см^3 допускаемая погрешность не должна превышать 0.03 г/см^3 всей рабочей шкалы. Если же предел равен ± 300 , или 3000 г/см^3 , допускаемая погрешность соответственно равна ± 3 и $\pm 30 \text{ г/см}^3$.

Стабильность или воспроизводимость прибора – свойство отсчетного устройства обеспечивать постоянство показаний одной и той же величины. В результате старения материалов со временем нарушается стабильность показаний приборов.

Стабильность прибора определяется вариацией показания. Поэтому при установлении стабильности нормируют допускаемую вариацию - ω_1 . Поскольку вариация принимается с одним знаком, а допустимая погрешность имеет плюс и минус, то $\omega_d = 0.5b$.

На все измерительные приборы в той или иной мере действует магнитное поле Земли. Поэтому ряд электроизмерительных приборов должен быть защищен от действия магнитного поля, а также электростатических явлений. В специальной метрологической литературе разработаны схемы защит I (более высокой) и II категорий.

В последние годы при исследовании процессов стали широко применять электрические, электронные, частотные, радиоизотопные и другие приборы. Как правило, такие приборы требуют специальной защиты от пыли, вибрации, газа, света и др. Отсутствие такой защиты может вызвать погрешности, превышающие допустимые.

Проверка средств измерений предусматривает определение и по возможности уменьшение погрешностей приборов. Определение погрешностей позволяет установить, соответствует ли данный прибор регламентированной точности и может ли он быть применен для данных измерений.

Измерительные приборы и установки различных организаций подвергают обязательной государственной проверке раз в 1-2 года.

Под регулировкой прибора понимают операции, направленные на снижение систематических ошибок до величины, меньше допустимой погрешности. Измерительные приборы снабжены двумя регулировочными узлами для регулировки нуля и чувствительности. Регулировка нуля предназначена для устранения систематических ошибок в диапазоне нужного предела измерений.

В ряде случаев возникают систематические погрешности, линейно возрастающие или убывающие с изменением измеряемой величины. Такую погрешность регулировкой нуля устранить невозможно. Ее можно уменьшить с помощью регулировки узла чувствительности. Поскольку погрешность различна на разных участках длины шкалы, то с помощью одновременной регулировки узла нуля и чувствительности достигают существенного снижения систематической ошибки прибора в начале, середине и конце диапазона измерения. Такая регулировка называется градуировкой прибора.

Наиболее распространенным способом поверки приборов и оценки его эксплуатационных характеристик является способ сравнения. Суть его сводится к сопоставлению поверяемого прибора с образцовым. Одна и та же измеряемая величина оценивается поверяемым и образцовым прибором. По отсчетам судят о погрешностях, которые вносят в поверяемый прибор.

Важным моментом в организации эксперимента является выбор средств измерений. Средства измерения должны: максимально соответствовать тематике, цели и задачам НИР; обеспечивать высокую производительность труда с наименьшей затратой времени, выполнение эксперимента в возможно кратчайший срок, требуемое качество экспериментальных работ, т.е. заданную степень точности при минимальном количестве измерений, высокую воспроизводимость и надежность; эргономические требования эксперимента (антропометрические, санитарно-гигиенические, психофизиологические и др.); требования техники безопасности и пожарной профилактики; в наибольшей степени исключать систематические ошибки. Желательно также максимально использовать средства измерений с автоматической записью, иметь высокую экономическую эффективность при минимуме затрат людских, денежных и материальных ресурсов. Следует особо помнить, что стандартные или специально созданные средства измерений перед началом эксперимента должны обязательно пройти метрологическую поверку или аттестацию. В противном случае результаты экспериментов могут вызвать возражение или же быть полностью не признанными.

* * *

Особенность, отличающая науку от других способов понимания и объяснения мира, состоит, в конечном счете, в том, что она полагается на авторитет экспериментальной проверки. Помимо общепринятого способа определения того, какие факты адекватны достоверным теориям, должна существовать готовность подвергнуть их риску проверки и допущение возможности их неправильности.

Экспериментальная проверка – не автоматический и не тривиальный процесс. Эксперименты, казавшиеся решающими на одной стадии понимания, в дальнейшем теряют убедительность, и наоборот. Эксперименты, находившиеся в явном противоречии с теориями, не всегда вели к отказу от них. Теории с полным успехом прошедшие всевозможные проверки, отвергались из-за несоответствия, считавшегося ранее совсем незначительным. Последнее слово всегда остается за согласованным мнением исследователей, использующих критерии истинности и значимости, которые нелегко объективно сформулировать. С течением времени критерии меняются, и решение может быть пересмотрено.

Обращение к экспериментальной проверке не является изобретением современной науки. Древнегреческие философы довольно часто использовали наблюдения природы как свидетельства «за» или «против» определенных теорий.

Любое наблюдение, любое определение факта, даже проведенное компетентными учеными, может быть подвергнуто сомнению. Иногда необходимо повторить наблюдение, чтобы подтвердить или опровергнуть его. Наука, таким образом, ограничивается тем, что

можно назвать «общественными» фактами. Любой исследователь должен иметь возможность проверить их; экспериментальные наблюдения должны быть повторяемыми.

Исследователя не должно смущать нередкое отсутствие достаточно полного экспериментального материала. Зачастую в таких областях исследований, как экономика, социология, военное дело и т.д., невозможно поставить прямой эксперимент, полностью отвечающий на стоящие вопросы. Часто оказывается невозможным сделать это и в физике. Часто не хватает «материала» интеллектуального, а не экспериментального. Нужна концепция – определенное видение изучаемой реальности, которое и придает экспериментальным данным содержательный смысл, превращает экспериментальный материал в объективную информацию о реальности.

Эксперимент дает сведения для теоретических построений и служит той основой, на которой создается здание теории. Для получения фундаментальных экспериментальных фактов необходима громадная поисковая работа, большая часть результатов которой в дальнейшем может оказаться ненужной. «Производственные» отходы в познании нового всегда велики, однако результаты окупают потери. Для построения моделей изучаемой реальности объем экспериментального материала сам по себе не имеет принципиального значения. Более того, одного экспериментального материала, как бы хорош он ни был, недостаточно для построения доброкачественной теории, в желаемой степени адекватно отражающей реальность. Так же, как теория опирается в своей основе на экспериментальные данные, так и эксперимент всегда несет в себе полезную информацию, если он проводится в соответствии с определенной теоретической концепцией. Эксперимент как простая совокупность наблюдаемых фактов при неверной концепции может ввести исследователя в заблуждение, что неоднократно и случалось в истории наук.

Чтобы эксперимент, какова бы ни была его цель, мог считаться проверкой теории, от его исхода должно что-либо зависеть. Если эксперимент имеет ожидаемый результат, это повышает уверенность в правильности теории, если результат другой – это увеличивает сомнения. Если результат никак не влияет на степень доверия к теории, то эксперимент нельзя назвать проверкой, хотя он и мог быть важным и интересным по другим причинам. Чтобы эксперимент повлиял на научные убеждения, исследователь должен быть морально готов к различным результатам, иметь представление об альтернативах и даже предполагать, что выдвинутая гипотеза может оказаться неверной. Иногда альтернатива проста: 1) теория верна или 2) теория ложна. Но нередко существует две или несколько конкурирующих теорий. То, какой эксперимент целесообразнее провести, зависит от выбора альтернатив. Тем не менее, если *не* возникают альтернативные концепции, *не* возникает экспериментальная проверка, так как никакие результаты опыта не могут изменить научные взгляды.

Экспериментальная проверка научной теории – не механический и не автоматический процесс. Не существует заранее предписанного набора процедур, которые можно было бы выполнить и в конце сделать вывод о верности научной теории. Процесс проверки теории, как и процесс ее создания, никогда не заканчивается и требует творчества и воображения. Приходится применять субъективные суждения о том, какие экспериментальные факты важны и действительно окажут влияние на степень достоверности теории.

Когда осуществляется эксперимент нередко вводятся допущения, что он изолирован от любых явлений, не находящихся под контролем исследователя, что используемые для измерения инструменты и проводящий наблюдения экспериментатор каким-то образом оказываются «вне» изучаемых объектов и не влияют на них. На самом деле такая идеализация зачастую неверна. Человек, проводящий эксперимент, сам является его частью, он может повлиять на его исход своим присутствием, своей предвзятостью, самой техникой эксперимента.

3. МЕТОД ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА В НАУЧНЫХ ИССЛЕДОВАНИЯХ

3.1 Основные понятия планирования эксперимента

Основа традиционных методов экспериментального исследования – определение требуемых зависимостей при изменении одного фактора и постоянстве остальных (однофакторный эксперимент). Так, прочность заливочного компаунда зависит от многих факторов, главными из которых являются активность затвердителя, соотношение составных частей, температура смеси и т.д. При таком методе устанавливается степень влияния каждого переменного фактора в отдельности на прочность компаунда.

Эти методы обладают рядом существенных недостатков. Основным из них является не всегда корректное допущение о возможности стабилизации всех переменных исследуемого объекта при последовательном выделении каждой из них в целях изучения объекта с достаточной степенью точности. Кроме того, традиционные методы исследования исключительно трудоемки.

В отличие от традиционных форм выполнения экспериментов в последнее время все чаще применяются методы математического планирования, позволяющие одновременно изучать влияние ряда факторов (многофакторный эксперимент) на исследуемый объект. Они основаны на математической теории эксперимента, которая определяет условия оптимальности проведения эксперимента, в том числе и при неполном знании физической сущности явления. Для этого используют математические методы не только на стадии обработки результатов измерений, как было раньше, но также при подготовке и проведении опытов. Математические методы планирования эксперимента позволяют исследовать и оптимизировать сложные системы и процессы, обеспечивая высокую эффективность эксперимента и точность определения исследуемых факторов.

Видный ученый в области математической теории эксперимента В.В.Налимов считает, что планирование эксперимента – это оптимальное управление экспериментом при неполном знании механизма явлений. Эксперименты обычно ставятся небольшими сериями по заранее согласованному алгоритму, оптимальному в некотором строго сформулированном смысле. После каждой небольшой серии опытов производится обработка результатов наблюдений и принимается строго обоснованное решение о том, что делать дальше. Планирование эксперимента на основе его математической теории можно рассматривать как одно из направлений в автоматизации научных исследований. Во многих случаях, приступая к исследованию какого-либо объекта (процесса), мы не знаем его механизма. Поэтому можно только выделить определяющие условия протекания процесса и требования к его результатам. Тогда представляется целесообразным использовать такой подход, когда объект исследования может быть представлен в виде «черного ящика».

Научиться управлять объектом, информация об элементарных процессах внутри которого чрезвычайно мала, с использованием традиционных методов экспериментальных исследований очень трудно. Это можно сделать путем использования методов планирования эксперимента. При планировании эксперимента возможно решение различных вопросов, связанных с изучением кинетики и механизма явлений, исследованием зависимостей «состав-свойство», выявлением неоднородностей процесса и их влиянием на выходные факторы. Кроме того, планирование эксперимента может осуществляться в целях адаптации технологического процесса к изменяющимся оптимальным условиям его протекания и обеспечения таким образом высокой

эффективности его осуществления и др. В зависимости от поставленных в исследованиях задач применяют различные виды планов.

Теория математического эксперимента содержит ряд концепций, которые обеспечивают успешную реализацию задач исследования. К ним относятся концепции рандомизации, последовательного эксперимента, математического моделирования, оптимального использования факторного пространства и ряд других.

Принцип рандомизации заключается в том, что в план эксперимента вводят элемент случайности. Для этого план эксперимента составляется таким образом, чтобы те систематические факторы, которые трудно поддаются контролю, учитывались статистически и затем исключались в исследованиях как систематические ошибки.

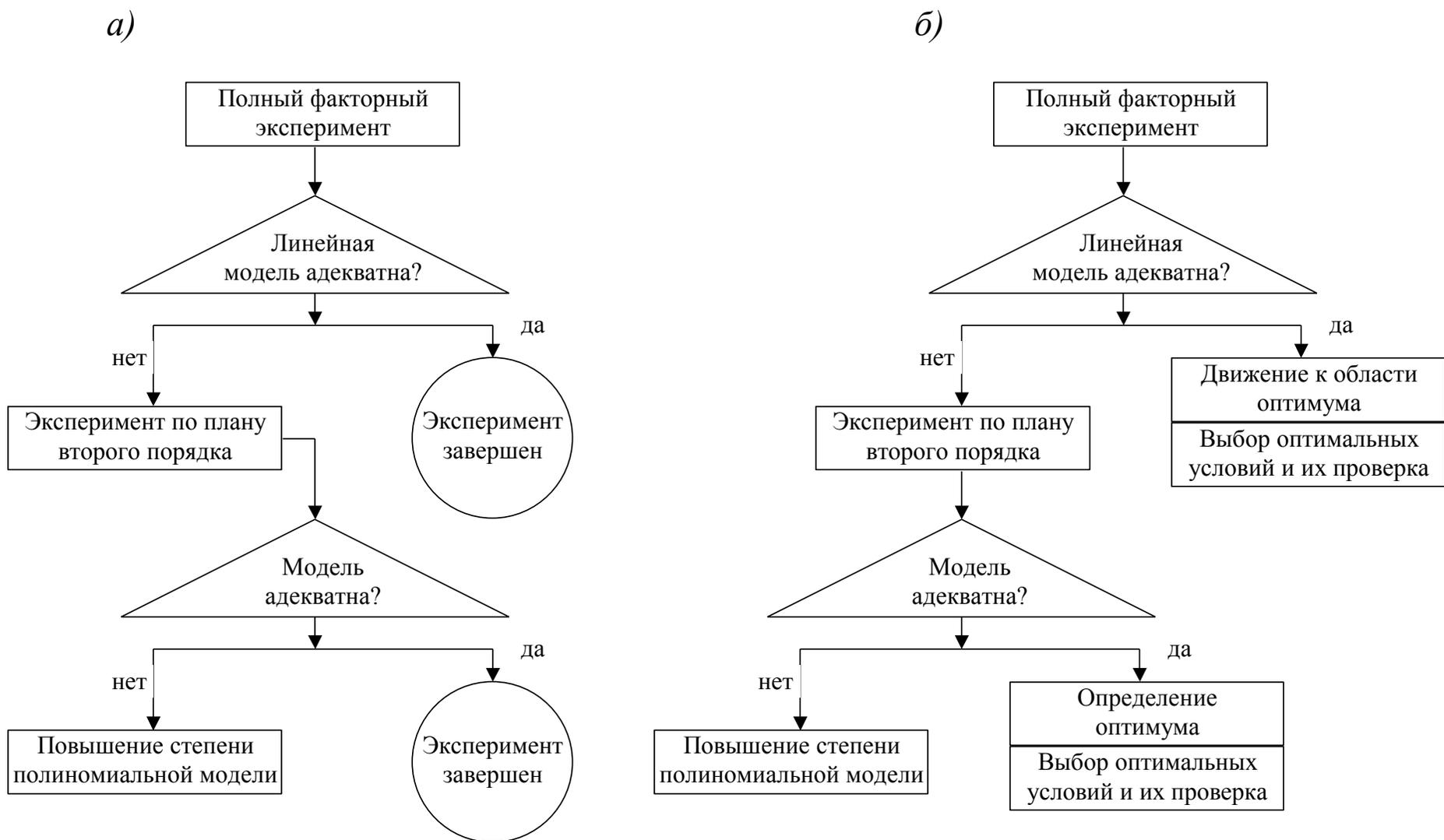
При последовательном проведении эксперимент выполняется не одновременно, а поэтапно, с тем, чтобы результаты каждого этапа анализировать и принимать решение о целесообразности проведения дальнейших исследований (рисунок 3.1).

В результате эксперимента получают уравнение регрессии, которое часто называют моделью процесса. Для конкретных случаев математическая модель создается исходя из целевой направленности процесса и задач исследования, с учетом требуемой точности решения и достоверности исходных данных.

Под моделью понимают не абсолютно точное описание явления (подобно закону), а приближенное выражение неизвестного закона, которое удовлетворительно характеризует явление в некоторой локальной области факторного пространства. Для приближенного описания одного и того же явления можно предложить несколько моделей, оценка которых обычно производится по критерию Фишера. Так как степень полинома, адекватно описывающего процесс, предсказать невозможно, то сначала пытаются описать явление линейной моделью, а затем (если она неадекватна) повышают степень полинома, т.е. проводят эксперимент поэтапно.

Одна из существенных концепций в теории планирования эксперимента – оптимальное использование факторного пространства. Она заключается в том, что состояние объекта в каждом опыте определяется по результату одновременного оптимального варьирования n факторов в n -мерном пространстве. Это позволяет добиться значительного увеличения точности расчета коэффициентов полученной модели и уменьшить трудоемкость эксперимента.

Наиболее изучены, а потому и широко применяются планы экстремального эксперимента, которые позволяют описать исследуемый процесс и выполнить его оптимизацию. Эти планы представляют собой систему опытов, содержащую возможные неповторяющиеся комбинации выбранных факторов при заданных уровнях их варьирования. Данный метод позволяет одновременно изучать влияние многих факторов на исследуемый процесс и дает возможность получить полином k -й степени (функцию отклика) для математического описания исследуемого процесса в некоторой локальной области факторного пространства, лежащего в окрестности выбранной точки с координатами $(z_{01}, z_{02}, \dots, z_{0n})$. Полученную функцию отклика можно использовать также для оптимизации процесса, т.е. определять параметры, при которых явление или процесс будет протекать наиболее эффективно.



а – с целью математического описания исследуемого процесса; *б* – с целью оптимизации исследуемого процесса

Рисунок 3.1 – Структурная схема эксперимента

Экстремальный эксперимент основан на положении о том, что исследуемую непрерывную функцию $y = f(z_1, z_2, \dots, z_n)$, имеющую все производные в заданной точке с координатами $z_{01}, z_{02}, \dots, z_{0n}$, можно разложить в ряд Тейлора:

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_n z_n + \beta_{12} z_1 z_2 + \dots + \beta_{(n-1)n} z_{n-1} z_n + \beta_{11} z_1^2 + \beta_{22} z_2^2 + \dots + \beta_{nn} z_{nn}^2, \quad (3.1)$$

где β_0 – значение функции отклика в начале координат $z_{01}, z_{02}, \dots, z_{0n}$;

$$\beta_i = \frac{\partial y}{\partial z_i}; \quad \beta_{ij} = \frac{\partial^2 y}{\partial z_i \partial z_j}; \quad \beta_{ii} = \frac{1}{2} \frac{\partial^2 y}{\partial z_i^2}.$$

Ряд Тейлора аналогичен уравнению регрессии:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_n x_n + a_{12} x_1 x_2 + \dots + a_{(n-1)n} x_{n-1} x_n + a_{11} x_1^2 + a_{22} x_2^2 + \dots + a_{nn} x_{nn}^2. \quad (3.2)$$

Здесь a_0, a_i, a_{ij}, a_{ii} – коэффициенты регрессии; x_i – кодированная переменная, введенная в целях упрощения арифметических расчетов и равная

$$x_i = \frac{z_i - z_{0i}}{\Delta z_i}; \quad \Delta z_i = \frac{z_{i \max} - z_{i \min}}{2}; \quad z_{0i} = \frac{z_{i \max} + z_{i \min}}{2}.$$

Следовательно, x_i – относительное значение; максимальному значению $z_{i \max}$ соответствует $x_i = +1$, минимальному значению $z_{i \min}$ соответствует $x_i = -1$.

Коэффициенты регрессии в уравнении (3.2) вычисляются методами математической статистики и представляют собой приближенную оценку коэффициентов $\beta_0, \beta_i, \beta_{ii}$ в уравнении (3.1). Следовательно, уравнение (3.2) описывает исследуемый объект (процесс) только с определенной степенью точности.

Уравнение (3.2) широко используют для получения математической модели объектов исследования, хотя оно и не содержит необходимой информации о механизме явления и его физико-химических свойствах. Действительно, зная уравнения, аналогичные (3.2), нельзя восстановить исходную функцию, описывающую объект исследования. В то же время они исключительно полезны для решения экстремальных задач. В математической теории эксперимента разработаны оптимальные планы получения уравнения типа (3.2) и их использования для определения экстремумов.

Следует отметить, что определение коэффициентов регрессии трудоемко и громоздко, поэтому, за исключением простейших случаев, их вычисление и анализ уравнений типа (3.2) производят с помощью ЭВМ по специально составленной программе.

Планы оптимального эксперимента реализуются в такой последовательности: 1) оценка информации и определение n факторов, наиболее существенных для исследуемого процесса; 2) использование математической модели в виде линейной функции отклика; 3) анализ выбранной модели; 4) нахождение экстремума в области n -мерного факторного пространства путем использования полинома k -й степени; 5) если модель неадекватна, то в качестве модели выбирают полиномы более высокого порядка.

3.2 Планирование эксперимента с целью описания исследуемого объекта

Началом планирования эксперимента является сбор, изучение и анализ имеющихся данных об объекте. В результате этого определяют выходной параметр y и входные – z_i .

Выходной параметр (переменная состояния объекта) должен иметь количественную характеристику, т.е. измеряться с достаточной степенью точности и однозначно характеризовать объект исследования, что обеспечивает корректную постановку задачи.

Входные параметры z_i должны иметь границы изменения ($z_{i\max} - z_{i\min}$), причем при различных комбинациях факторов z_i переменная состояния объекта y не должна выходить из области допустимых значений; между факторами z_i и значением y необходимо однозначное соответствие; факторы z_i между собой взаимно независимы. В исследованиях (в целях упрощения вычислений) удобно пользоваться относительными (кодированными) переменными x_i :

$$x_i = \frac{z_i - z_{0i}}{\Delta z_i}; \quad z_{0i} = \frac{z_{i\max} + z_{i\min}}{2}; \quad \Delta z_i = z_{i\max} - z_{0i} = z_{0i} - z_{i\min};$$

$$\Delta z_i = \frac{z_{i\max} - z_{i\min}}{2};$$

$z_{i\max}$, $z_{i\min}$ – наибольшее и наименьшее значения (верхний и нижний уровень) факторов z_i ; z_{0i} – средний (нулевой) уровень факторов z_i ; значениям $z_{i\max}$ и $z_{i\min}$ соответствует $x_{i\max} = 1$, $x_{i\min} = -1$. Очевидно, что если при значениях $x_{i\max}$ и $x_{i\min}$ или кратным им ставятся эксперименты, то вычисления упрощаются. Значение фактора, которое фиксируется при проведении эксперимента, называется его уровнем.

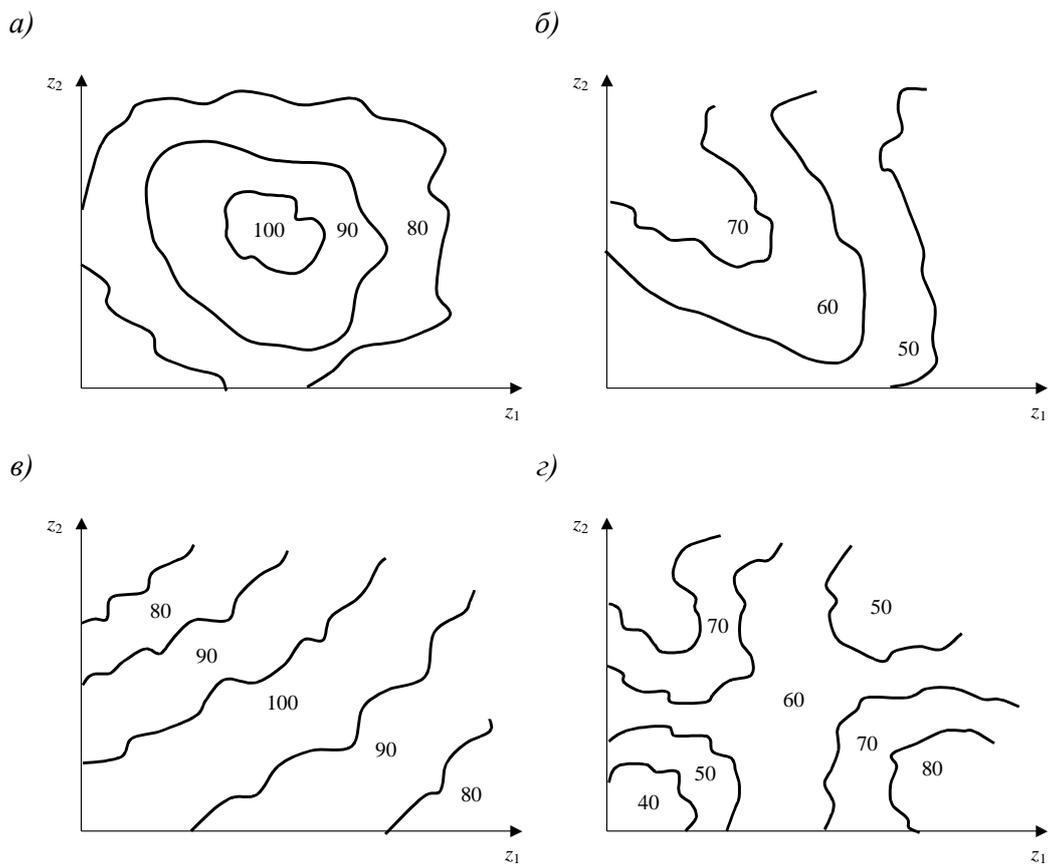
Среди планов экстремального эксперимента наиболее простыми являются планы полного факторного эксперимента (ПФЭ), в случае реализации которых определяется значение параметров состояния объекта y при всех возможных сочетаниях уровней варьирования их факторов z_i . Если мы имеем дело с n факторами, каждый из которых устанавливается на q уровнях, то для того, чтобы осуществить полный факторный эксперимент, необходимо поставить $m = q^n$ опытов. Наибольшее распространение получили планы экспериментов типа 2^n . С увеличением q резко возрастает количество опытов, поэтому если $q > 2$, планы ПФЭ редко используются.

В соответствии с идеей последовательного поиска ПФЭ проводится в несколько этапов. Для упрощения изложения ограничимся полиномами первой и второй степени. В дальнейшем проанализируем только полный двухфакторный эксперимент, составляемый с целью описать поверхность отклика второго порядка (рисунок 3.2).

Допустим, что необходимо исследовать явление в зависимости от изменения двух факторов z_1 и z_2 методом полного двухфакторного эксперимента. Обычно в начале предполагают, что оно описывается линейным полиномом, т.е. поверхность отклика представляет собой плоскость, характеризуемую полиномом

$$y = a_0 + a_1x_1 + a_2x_2.$$

Чтобы построить поверхности отклика в виде плоскости, достаточно провести четыре опыта. Наиболее удобно выбранные факторы варьировать на верхнем и нижнем уровнях, что соответствует $x_i = +1$, $x_i = -1$. Для удобства планирования эксперимента составляют планы (рисунок 3.3) и матрицу (таблица 3.1) двухфакторного эксперимента, в соответствии с которыми и проводят исследования.



а – параболоида; *б* – стационарная возвышенность; *в* – хребет; *г* – седло

Рисунок 3.2 – Виды поверхностей отклика

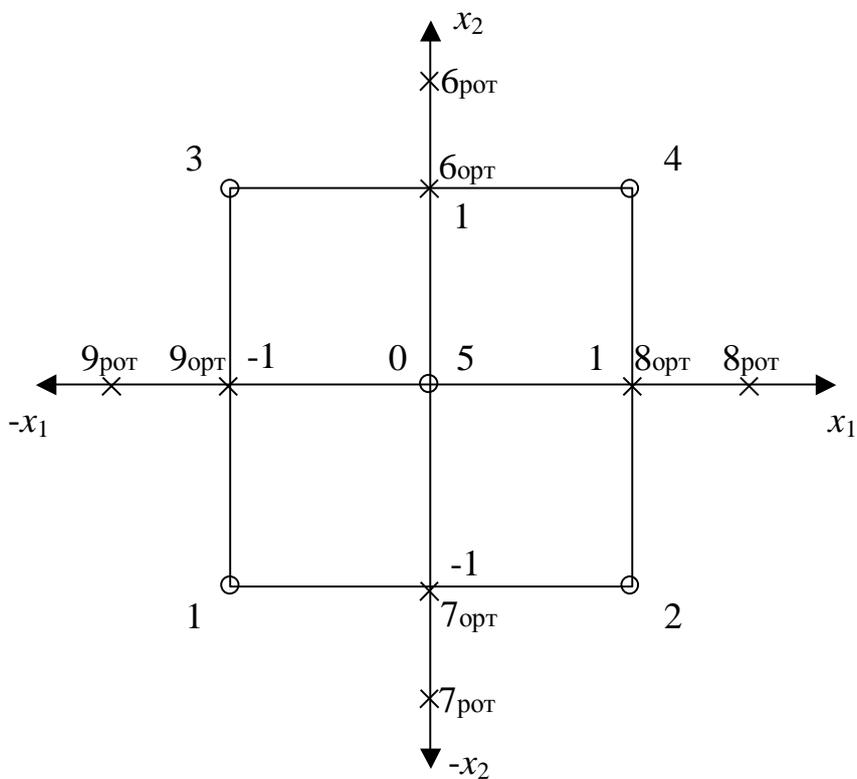


Рисунок 3.3 – Планы для функций $y = f(x_1, x_2)$

Таблица 3.1 - Матрица планирования эксперимента

Номер опыта	Факторы		Функция отклика y
	x_1	x_2	
1	-1	-1	y_1
2	+1	-1	y_2
3	-1	+1	y_3
4	+1	+1	y_4

Как следует из матрицы, первый опыт проводят при минимальных значениях факторов z_1 и z_2 , четвертый – при максимальных значениях z_1 и z_2 , второй – при минимальном значении z_2 и максимальном z_1 , а третий – наоборот.

Аналогично составляют полином, план и матрицу планирования для проведения 3-х, 4-х и большего количества факторов. План трехфакторного эксперимента представлен на рисунке 3.4, матрица – в таблице 3.2.

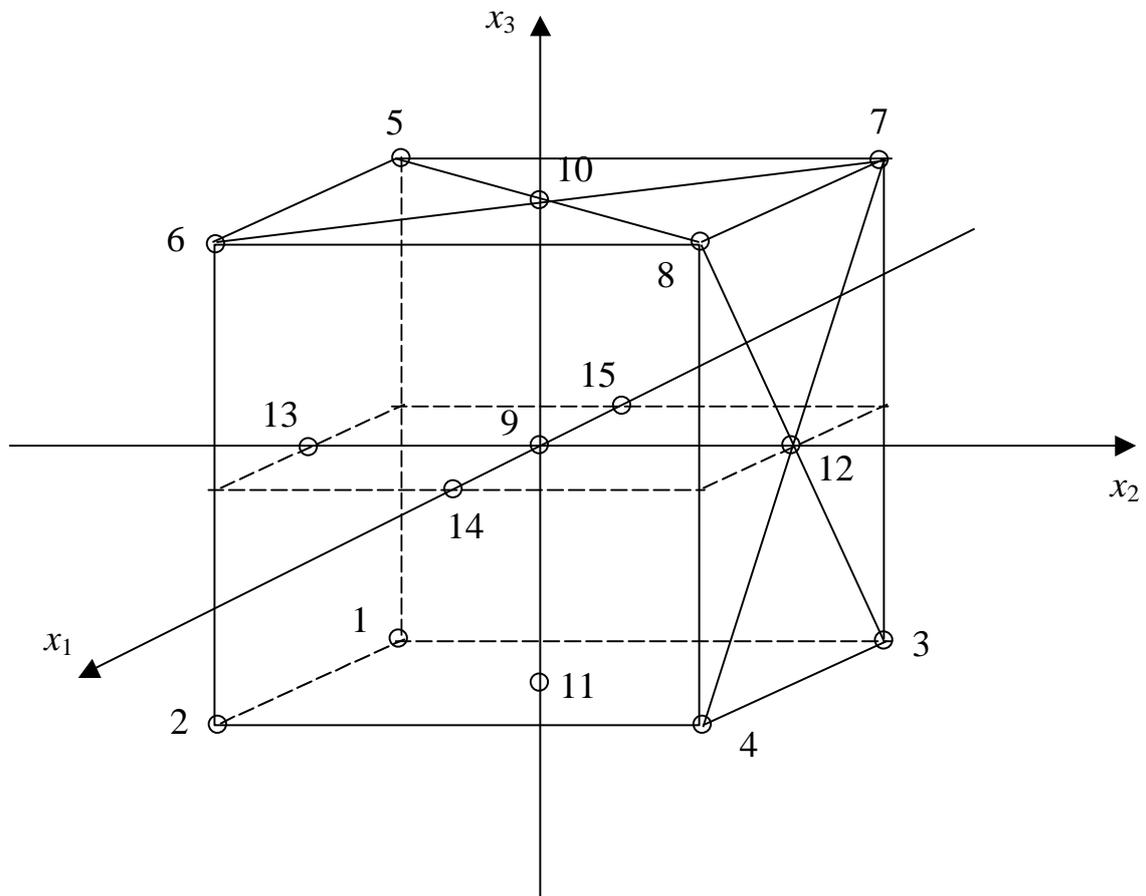


Рисунок 3.4 - Планы для функций $y = f(x_1, x_2, x_3)$

Таблица 3.2 - Матрица планирования эксперимента

Номер опыта	Факторы			Функция отклика y_i
	x_1	x_2	x_3	
1	-1	-1	-1	y_1
2	+1	-1	-1	y_2
3	-1	+1	-1	y_3
4	+1	+1	-1	y_4
5	-1	-1	+1	y_5
6	+1	-1	+1	y_6
7	-1	+1	+1	y_7
8	+1	+1	+1	y_8

Как следует из таблиц 3.1 и 3.2, принцип построения матриц планирования полного факторного эксперимента заключается в том, что уровни варьирования первого фактора чередуются от опыта к опыту, частота смены уровней варьирования каждого последующего фактора вдвое меньше, чем у предыдущего. У последнего фактора уровни изменяются всего два раза.

Матрица планирования полного факторного эксперимента в этом случае обладает следующими свойствами:

$$\sum_{j=1}^m x_{ji} = 0; \quad \sum_{j=1}^m x_{ji}^2 = m; \quad \sum_{j=1}^m x_{ji} x_{ju} = 0. \quad (3.3)$$

Здесь m – число опытов полного факторного эксперимента; j – номер опыта; i, u – номер факторов.

Свойство, выраженное уравнениями (3.3), называется ортогональностью, а матрица – ортогональной (прямоугольной). Это свойство обеспечивает относительную простоту вычисления коэффициентов регрессии, которые определяют по формулам:

$$a_0 = \frac{1}{m} \sum_{j=1}^m y_j; \quad a_i = \frac{1}{m} \sum_{j=1}^m x_{ji} y_j; \quad a_{iu} = \frac{1}{m} \sum_{j=1}^m x_{ji} x_{ju} y_j. \quad (3.4)$$

Формулы (3.4) применимы только для вычисления коэффициентов ортогонального линейного полинома. В самом же общем виде коэффициенты регрессии вычисляют с помощью многочленов Чебышева, обеспечивающих минимум суммы квадратов отклонений.

Некоторые из коэффициентов уравнения регрессии могут оказаться незначительными, т.е. пренебрежимо малыми. Коэффициент регрессии значим и им пренебрегать нельзя, если

$$|a| \geq D_a t, \quad (3.5)$$

где t – критерий Стьюдента; D_a – дисперсия, при определении коэффициента регрессии $D_a = D_{\bar{y}} / \sqrt{m}$, $D_{\bar{y}}$ – дисперсия среднего значения фактора.

Полученные таким образом уравнения линейной регрессии проверяют по условию адекватности (например, по критерию Фишера). Адекватность линейного полинома можно определить и путем вычисления коэффициентов регрессии $a_{12}, \dots, a_{(n-1)n}$, которые при адекватности линейного полинома равны нулю.

В случае линейного полинома для нахождения коэффициентов регрессии, используя метод дробного факторного эксперимента (ДФЭ), можно уменьшить количество

опытов, которые представляют собой часть матрицы полного факторного эксперимента (например, 1/2, 1/4 часть и т.д.). Так, чтобы найти коэффициенты регрессии уравнения

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3,$$

необходимо провести восемь опытов согласно матрице планирования полного трехфакторного эксперимента. Однако их можно найти и при уменьшении количества опытов до четырех, реализовать половину матрицы, поскольку план трехфакторного эксперимента представляется в форме куба, параметры которого полностью определены, если зафиксирована диагональная плоскость и вершины куба, лежащие на ней (рисунок 3.4). В этом случае основу матрицы планирования составляет матрица двухфакторного эксперимента, варьирование третьего фактора соответствует произведению x_1x_2 (таблица 3.3). Это преобразование допустимо, если коэффициент регрессии a_{12} незначим или равен нулю.

Таблица 3.3 - Матрица планирования эксперимента

Номер опыта	x_0	x_1	x_2	$x_3 = x_1x_2$	y_j
1	+1	-1	-1	+1	y_1
2	+1	+1	-1	+1	y_2
3	+1	-1	+1	-1	y_3
4	+1	+1	+1	+1	y_4

Такое планирование эксперимента, когда некоторые факторы приравнивают к произведению нескольких факторов, называется планированием со смешиванием. Его обозначают 2^{n-p} , где n – число факторов; p – число факторов, приравниваемых к произведениям. Например, описанное выше планирование (таблица 3.3) обозначают 2^{3-1} .

Использование метода ДФЭ или дробных реплик дает возможность в несколько раз уменьшить число опытов при исследовании линейных моделей. В таблице 3.4 приведена матрица планирования 2^{7-4} , позволяющая при восьми опытах определить коэффициенты регрессии линейного полинома:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7.$$

Здесь принято $x_4 = x_1x_2$; $x_5 = x_1x_3$; $x_6 = x_2x_3$; $x_7 = x_1x_2x_3$. Такие равенства называют генерирующими соотношениями, их выбор произволен, однако повторение не допускается. Широко используются планы 2^{3-1} , 2^{5-2} , 2^{6-3} , 2^{7-3} и т.д., позволяющие уменьшить количество опытов в два и более раз.

Таблица 3.4 - Матрица планирования 2^{7-4}

Номер опыта	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	+1	-1	-1	-1	+1	+1	+1	-1
2	+1	+1	-1	-1	-1	-1	+1	+1
3	+1	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	+1	-1	+1	-1	-1	-1
5	+1	-1	-1	+1	+1	-1	-1	+1
6	+1	+1	-1	+1	-1	+1	-1	-1
7	+1	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1	+1

Если же экспериментальные данные не согласуются с линейной моделью, то исследуемый процесс стремятся описать поверхностью второго порядка. Использование полинома второго порядка для описания объекта исследования на двух уровнях

варьирования факторов x_i недостаточно. При использовании ПФЭ типа 3^n резко возрастает количество опытов, план становится трудоемким, а формулы для вычисления коэффициентов регрессии – громоздкими.

Используя концепцию факторного пространства, можно дополнить двухуровневый план ПФЭ определенными (звездными) точками (рисунок 3.3). Тогда полное количество опытов можно выразить в виде

$$m = 2^n + 2n + N_0, \quad (3.6)$$

где 2^n – количество опытов ПФЭ при $q = 2$; $2n$ – количество звездных точек; N_0 – нулевые точки.

Такие планы называются центральными, композиционными (ЦКП). Различают ортогональные (почти ортогональные) и ротатабельные ЦКП. При ортогональных ЦКП количество опытов определяют по формуле

$$m = 2^n + 2n + 1.$$

При этом значение звездного плеча α зависит от количества факторов n . Если $n = 2$, то $\alpha = 1.0$; при $n = 3$, то $\alpha = 1.215$; в случае $n = 4$, то $\alpha = 1.414$ и т.д.

Большим преимуществом этих планов является то, что их можно получать из планов 2^n . Для этого к реализованному плану линейного полинома добавляют опыты в промежуточных «звездных» точках и в центре плана. Полученную при этом «композицию» используют для математического описания процесса в виде многочлена второй степени:

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2.$$

С учетом изложенного матрица ортогонального ЦКП для двух факторов представлена в виде таблицы 3.5.

В принятой матрице «0» показывает, что значения z_i приняты в начале координат (рисунок 3.3). Коэффициенты регрессии в этом случае вычисляются с помощью формул:

$$a_0 = \frac{1}{m} \sum_{j=1}^m y_j = \frac{a_{11}}{m} \sum_{j=1}^m x_{j1}^2 - \dots - \frac{a_{uu}}{m} \sum_{j=1}^m x_{ju}^2, \quad (3.7)$$

$$a_i = \frac{\sum_{j=1}^m x_{ij}y_j}{\sum_{j=1}^m (x_{ij})^2} \quad (i \neq 0), \quad (3.8)$$

$$a_{iu} = \frac{\sum_{j=1}^m x_{ij}x_{uj}y_j}{\sum_{j=1}^m (x_{ij}x_{uj})^2} \quad (i \neq u), \quad (3.9)$$

$$a_{ii} = \frac{\sum_{j=1}^m \left(x_{ij} - \frac{1}{m} \sum_{j=1}^m x_{ij}^2 \right) y_j}{\sum_{j=1}^m \left(x_{ij}^2 - \frac{1}{m} \sum_{j=1}^m x_{ij}^2 \right)^2}, \quad (3.10)$$

где j – номер опыта; i, n – номер факторов.

Таблица 3.5 - Матрица планирования эксперимента

Система опытов	Номер опыта	x_1	x_2	X_1x_2	y_j
Полный факторный эксперимент линейной модели	1	-1	-1	+1	y_1
	2	+1	-1	-1	y_2
	3	-1	+1	-1	y_3
	4	+1	+1	+1	y_4
Опыт в центре плана	5	0	0	0	y_5
Опыты в звездных точках	6	0	+1	0	y_6
	7	0	-1	0	y_7
	8	+1	0	0	y_8
	9	-1	0	0	y_9

Для расчета оценки дисперсий при определении коэффициентов регрессии используют следующие выражения:

$$D_{a_0} = \frac{D_y^2}{m} + \frac{mD_{a_{ii}}^2}{m} \sum_{j=1}^m x_{ij}^2,$$

$$D_{a_j} = \frac{D_y^2}{\sum_{j=1}^m (x_{ij})^2} \quad (i \neq 0),$$

$$D_{a_{iu}} = \frac{D_y^2}{\sum_{j=1}^m (x_{ij}x_{iu})^2} \quad (i \neq 0),$$

$$D_{a_{ii}} = \frac{D_y^2}{\sum_{j=1}^m \left(x_{ij}^2 - \frac{1}{m} \sum_{j=1}^m x_{ij}^2 \right)^2}.$$

Коэффициенты a_0 , a_i , a_{iu} значимы, если выполняется условие (3.5). Адекватность полученного уравнения регрессии проверяется с помощью критерия Фишера. Для двух- и трехфакторного эксперимента вычисление коэффициентов регрессии можно производить вручную:

$$a_{iu} = \frac{1}{c} \sum_{j=1}^m \gamma y_j, \quad (3.11)$$

где c и γ - коэффициенты, которые приведены в таблицах 3.6 и 3.7. При проведении опытов с четырьмя и более факторами для вычисления коэффициентов полинома по формулам (3.7-3.10) необходимо использовать ЭВМ.

Таблица 3.6 - Значения коэффициентов c и γ

a_{iu}	a_0	a_1	a_2	a_{11}	a_{22}	a_{12}
c	9	6	6	6	6	4
Номер опыта	Значение γ					
1	-1	-1	-1	1	1	1
2	-1	+1	-1	1	1	-1
3	-1	-1	+1	1	1	-1
4	-1	1	+1	+1	1	1
5	5	0	0	-2	-2	0
6	2	0	1	-2	1	0
7	2	0	-1	-2	1	0
8	2	1	0	1	-2	0
9	2	-1	0	1	-2	0

Таблица 3.7 - Значения коэффициентов c и γ

a_{iu}	a_0	a_1	a_2	a_3	a_{12}	a_{13}	a_{23}	a_{11}	a_{22}	a_{23}
c	45	10	10	10	8	8	8	18	18	18
Номер опыта	Значение γ									
1	-2	-1	-1	-1	1	1	1	1	1	1
2	-2	1	-1	-1	-1	-1	1	1	1	1
3	-2	-1	1	-1	-1	1	-1	1	1	1
4	-2	1	1	-1	1	-1	-1	1	1	1
5	-2	-1	-1	1	1	-1	-1	1	1	1
6	-2	1	-1	1	-1	1	-1	1	1	1
7	-2	-1	1	1	-1	-2	1	1	1	1
8	-2	1	1	1	1	1	1	1	1	1
9	13	0	0	0	0	0	0	-2	-2	-2
10	8	0	0	1	0	0	0	-4	-4	5
11	8	0	0	-1	0	0	0	-4	-4	5
12	8	0	1	0	0	0	0	-4	5	-4
13	8	0	-1	0	0	0	0	-4	5	-4
14	8	1	0	0	0	0	0	5	-4	-4
15	8	-1	0	0	0	0	0	5	-4	-4

С учетом изложенного последовательность оптимального планирования эксперимента заключается в следующем.

1. Выбирают из n действующих в системе факторов наиболее значимые $z_1, z_2, \dots, z_i, \dots, z_n$.
2. Устанавливают пределы изменений выбранных факторов $z_{i\max}$ и $z_{i\min}$, вычисляют основной уровень z_{0i} и интервал варьирования Δz_i и заменяют переменные z_i кодированными x_i .
3. Определяют комбинации факторов z_i , при которых будет изучаться заданная система.

4. Разрабатывают методику измерения выбранных факторов, определяют погрешности и число повторений в каждой из выбранных комбинаций факторов.

5. Проводят эксперимент, одновременно корректируя его с учетом полученных данных. Если линейная модель не согласуется с результатами эксперимента, то проводят опыты в «звездных» точках.

6. Определяют уравнение регрессии, проверяют его адекватность, анализируют и делают соответствующие выводы.

Пример. Необходимо исследовать изменение прочности мелкозернистого бетона в зависимости от состава и водоцементного отношения. Состав бетона изменялся от п/ц = 1.5 до п/ц = 3.5; водоцементное отношение – от в/ц = 0.4 до в/ц = 0.6.

Приводят основные характеристики плана эксперимента (таблица 3.8).

Таблица 3.8 - Основные характеристики плана эксперимента

Характеристика	$\frac{B}{Ц}$	x_1	$\frac{П}{Ц}$	x_2
Основной уровень, z_{0i}	0.5	0	2.5	0
Интервал варьирования, Δz_i	0.1	---	1.0	---
Верхний уровень, z_{imax}	0.6	1	3.5	1
Нижний уровень, z_{imin}	0.4	-1	1.5	-1

$$\text{Кодированные переменные } x_1 = \frac{\frac{B}{Ц} - 0.5}{0.1}; \quad x_2 = \frac{\frac{П}{Ц} - 2.5}{1}.$$

2. Проверяют возможность линейного полинома: $y_p = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2$. Для ведения эксперимента применяют план, приведенный на рисунке 3.3, составляют рабочую таблицу планирования и в соответствии с ней проводят эксперимент, результаты которого записывают в таблицу 3.9.

Таблица 3.9 - Результаты факторного эксперимента

Номер опыта	x_1	x_2	z_1	z_2	$y_э$	y_p	Δy
1	-1	-1	0.4	1.5	410	391	+19
2	+1	-1	0.6	1.5	116	135	-19
3	-1	+1	0.4	3.5	306	323	-17
4	+1	+1	0.6	3.5	88	79	9
5	0	0	0.5	2.0	175	230	-55

3. Определяют коэффициенты линейного полинома: $a_0 = 1/4 \cdot (410 + 116 + 306 + 88) = 920/4 = 230$; $a_1 = 1/4 \cdot (-410 + 116 - 306 + 88) = 512/4 = -128$; $a_2 = 1/4 \cdot (-410 - 116 + 306 + 88) = -132/4 = -33$; $a_{12} = 1/4 \cdot (410 - 116 - 306 + 88) = 76/4 = 19$. Следовательно, $y_p = 230 - 128x_1 - 33x_2 + 19x_1x_2$. Коэффициент a_{12} значим, им пренебрегать нельзя.

4. Контрольный эксперимент в точке $x_1 = 0, x_2 = 0$ показал, что $y_э = 175 \text{ кг/см}^2$, что на 25% меньше $y_p = a_0$. Значит, искомая зависимость не может быть описана линейным полиномом. Продолжают эксперимент, представив исследуемую зависимость в виде квадратного полинома:

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2.$$

5. Проводят опыты в 6, 7, 8, 9 точках (рисунок 3.3). Для этого составляют новую рабочую таблицу, в которую заносят данные измерений в 1-5 точках (таблица 3.10).

Таблица 3.10 - Новая рабочая таблица

Номер опыта	x_1	x_2	$\frac{B}{Ц}$	$\frac{\Pi}{Ц}$	$y_э$	y_p	Δy
1	-1	-1	0,6	1,5	410	399	+11
2	+1	-1	0,4	1,5	116	106	+10
3	-1	+1	0,6	3,5	306	308	-2
4	+1	+1	0,4	3,5	88	92	-4
5	0	0	0,5	2,5	175	173	+3
6	0	+1	0,5	3,5	161	149	+12
7	0	-1	0,5	1,5	187	198	-11
8	+1	0	0,6	2,5	101	99	+2
9	-1	0	0,4	2,5	350	353	-3

6. С помощью формулы (3.11) и таблицы 3.7 определяют коэффициенты регрессии: $a_0 = 1/9 \cdot (-410 - 116 - 306 - 88 + 5 \cdot 175 + 2 \cdot 161 + 2 \cdot 187 + 2 \cdot 101 + 2 \cdot 350) = 1553/9 = 172$; $a_1 = 1/6 \cdot (-410 + 116 - 306 + 88 + 101 - 350) = -761/6 = -127$; $a_2 = 1/6 \cdot (-410 - 116 + 306 + 88 + 161 - 187) = -158/6 = -26$; $a_{11} = 1/6 \cdot (410 + 116 + 306 + 88 - 2 \cdot 175 - 2 \cdot 161 - 2 \cdot 187 + 101 + 350) = 325/6 = 54$; $a_{22} = 1/6 \cdot (410 + 116 + 306 + 88 - 2 \cdot 175 + 161 + 187 - 2 \cdot 101 - 2 \cdot 350) = 16/6 = 2.7$; $a_{12} = 1/4 \cdot (410 - 116 - 306 + 88) = 19$.

Коэффициент a_{22} незначим, им можно пренебречь. Следовательно, имеем полином $y = 172 - 127x_1 - 26x_2 + 19x_1x_2 + 54x_1^2$.

7. По формуле (2.36) вычислим оценку дисперсии адекватности $D_a = (1/(9 - 6)) \cdot (11^2 + 10^2 + 2^2 + 4^2 + 3^2 + 12^2 + 11^2 + 2^2 + 3^2) = 528/3 = 176$. Согласно данным эксперимента $D_{ср} = 256$; критерий Фишера $K_{ф} = 256/176 = 1.44 \leq K_{фт}$. Таким образом, полученный полином адекватно описывает искомую зависимость. Этот полином можно представить в натуральных значениях факторов $\frac{B}{Ц}$ и $\frac{\Pi}{Ц}$.

$$R_0 = 172 - 127 \left(\frac{B}{Ц} - \frac{0.5}{0.1} \right) - 26 \left(\frac{\Pi}{Ц} - 2.5 \right) + 54 \frac{\left(\frac{B}{Ц} - 0.5 \right)^2}{0.1} + 19 \frac{\left(\frac{B}{Ц} - 0.5 \right)^2}{0.1} \cdot 19 \frac{\left(\frac{B}{Ц} - 0.5 \right)}{0.1} \left(\frac{\Pi}{Ц} - 2.5 \right) = 2479 - 7145 \frac{B}{Ц} - 121 \frac{\Pi}{Ц} + 5400 \left(\frac{B}{Ц} \right)^2 + 190 \frac{B}{Ц} \cdot \frac{\Pi}{Ц}.$$

Более точными планами по сравнению с ортогональными являются ротатабельные планы, что достигается благодаря увеличению количества опытов в центре плана и специальному выбору звездного плеча α (таблица 3.11). Однако ротатабельное планирование не только более трудоемкое по сравнению с ортогональным, но и

вычисление коэффициентов регрессии производится по относительно громоздким формулам.

Таблица 3.11 - Значения «звездного» плеча α

Количество факторов	n	2	3	4	5
Звездное плечо	α	1,414	1,680	2,00	2,378
Общее количество опытов	m	13	20	31	52
Число опытов в центре плана	N_0	5	6	7	10

Частое применение ортогональных и ротатабельных планов объясняется тем, что они позволяют относительно просто определять коэффициенты регрессии и дисперсии. Выполнение эксперимента на трех и более уровнях в целях повышения его точности и его обработка представляют значительные трудности, особенно при расчетах, выполняемых вручную. В связи с развитием вычислительной техники появилась возможность считать основным критерием оптимальности неортогональность плана, что позволяет относительно просто выполнять вычисления. В.В.Налимов указывает, что эффективность оценок плана определяется не только оптимальным способом обработки результатов экспериментов, но и оптимальным расположением точек плана в факторном пространстве, что обеспечивает его высокую точность.

Одним из важнейших в современной теории эксперимента является критерий D -оптимальности, который требует, чтобы объем n -мерного эллипсоида рассеивания оценок материалов был минимальный. Очевидно, в этом случае и дисперсия переменного состояния объекта исследования будет также минимальной, т.е. качество эксперимента будет высоким. Планы, удовлетворяющие этому критерию, называются D -оптимальными. Они позволяют проводить эксперименты, в которых факторы варьируются на нескольких уровнях – двух, трех, четырех, пяти и т.д. В этом случае эффективно используется факторное пространство (рисунок 3.5, таблица 3.9). При таком расположении точек в факторном пространстве можно достаточно точно описать исследуемый объект. Количество опытов не превышает необходимого для ортогонального ЦКП (формула (3.6)).

В ряде случаев целесообразно, чтобы уровни факторов изменялись неодинаково, например возможен план $5^1; 3^2; 2^3$. Это означает, что один фактор варьируется на пяти уровнях, два фактора – на трех, три – на двух. Это несимметричный план. Если все факторы имеют одинаковое число уровней, такой план называется симметричный.

Разработаны также A -оптимальные планы (они имеют минимальную длину диагонали прямоугольника, описанного вокруг n -мерного эллипсоида рассеивания параметров) и E -оптимальные планы (достигается минимум максимальной оси эллипсоида рассеивания). Реализация планов эксперимента ограничивается требованиями независимости переменных факторов. Поэтому в ряде случаев, особенно при реализации наиболее универсальных несимметричных планов, на первое место в качестве критерия оптимизации выдвигается коэффициент парной корреляции между коэффициентами, который не должен превышать 0.5 (по Л.П.Рузинову).

Изданы специальные каталоги планов эксперимента (например, каталог, выпущенный Московским государственным университетом), в которых приводится сравнительная оценка планов и даются рекомендации по их выбору применительно к конкретным условиям эксперимента.

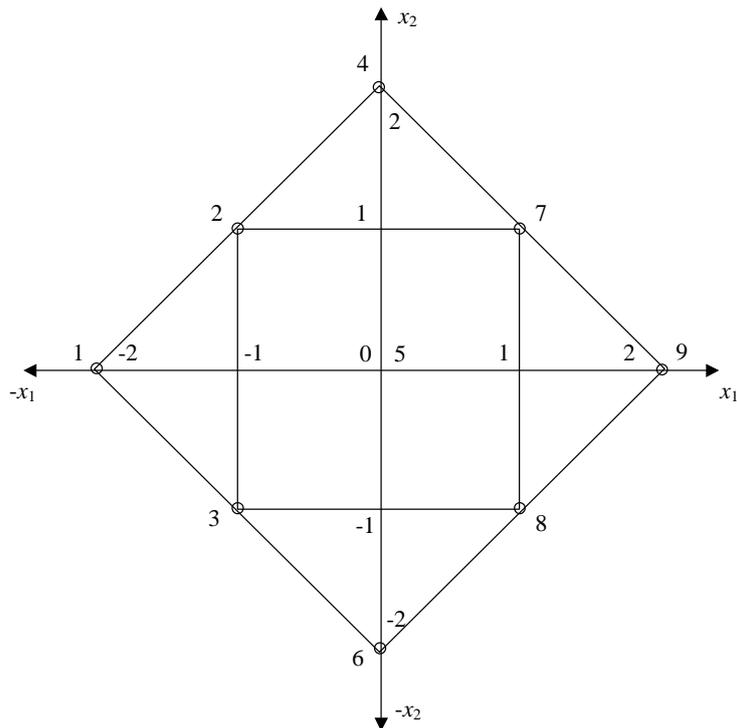


Рисунок 3.5 – D -оптимальный план для функций $f(x_1, x_2)$

Таблица 3.12 - Уровни факторов

№ опыта	Кодированные переменные	
	x_1	x_2
1	-2	0
2	-1	+1
3	-1	-1
4	0	+2
5	0	0
6	0	-2
7	+1	+1
8	+1	-1
9	+2	0

3.3 Оптимизация технологических процессов с использованием планирования эксперимента

Важное место в теории планирования эксперимента занимают вопросы оптимизации исследуемых процессов, свойств многокомпонентных систем или других объектов. Их эффективность наибольшая только в оптимальных условиях, характеризуемых экстремальными значениями y_i – переменными состояниями объекта при определенных значениях параметров x_i . Отметим также, что качество процесса обычно характеризуется несколькими функциями отклика. Как правило, нельзя найти такое сочетание значений влияющих факторов, при котором одновременно достигается

экстремум всех функций отклика. Например, максимальное качество печати лазерного принтера и минимальная стоимость одного отпечатанного листа достигаются при различных режимах работы принтера. Поэтому в большинстве случаев критерием оптимальности выбирают только одну из переменных состояния – функцию отклика, характеризующую процесс, а остальные принимают приемлемыми для данного случая.

Оптимизация процесса представляет собой целенаправленный поиск значений влияющих факторов, при которых достигается экстремум критерия оптимальности. Оптимизацию процессов обычно осуществляют в условиях ограничений на влияющие факторы и исследуемые функции отклика, поскольку как факторы, так и функции могут изменяться только в определенных границах. При этом используют различные виды планов (ПФЭ, ортогональные и ротатабельные ЦКП, D -оптимальные и др.).

Покажем, как можно использовать результаты полного факторного эксперимента для оптимизации процесса методом крутого восхождения или наискорейшего спуска. Допустим, что в некоторой окрестности точки y_i с координатами z_1 и z_2 исследуемая функция отклика, характеризующая процесс, описывается полиномом $y = a_0 + a_1x_1 + a_2x_2$.

Один из факторов, выраженных в натуральных величинах, принимают за базовый, например x_1 . Вычисляют для него произведение $a_1\Delta z_1$, где a_1 – коэффициент регрессии; Δz_1 – интервал варьирования первого фактора. Далее для базового фактора выбирают шаг движения Δz_{01} , с учетом которого производится оптимизация. Обычно $\Delta z_1 > \Delta z_{01}$. После этого определяют

$$v = \frac{\Delta z_{01}}{a_1\Delta z_1}.$$

Затем вычисляют шаги движения к оптимуму для всех остальных факторов, в данном случае $\Delta z_{02} = v a_2\Delta z_2$.

К оптимуму движутся из центра плана. На каждом новом шаге добавляют Δz_{0i} к соответствующим предыдущим значениям факторов z_i . Так осуществляют оптимизацию методом крутого восхождения. Если же определяют минимум функции y , то новые значения факторов находят из предыдущих путем вычитания Δz_{0i} , выполняя наискорейший спуск.

Движение к оптимуму прекращают, если достигнут оптимум функции критерия оптимальности y (в пределах ограничений, наложенных на внешние факторы и функции отклика). Затем в области экстремума функции определяют ее новое выражение в виде полинома.

4. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В НАУЧНЫХ ИССЛЕДОВАНИЯХ

4.1 Модели и моделирование

Любой метод научного познания человеком окружающего мира основывается на **моделировании** – исследовании явлений, процессов, объектов в природе и обществе путем построения и изучения их моделей, то есть таких материально или идеально (мысленно) представляемых объектов, которые в процессе исследования заменяют объекты-оригиналы (явления, процессы, объекты), и изучение которых уточняет или дает новое знание об объекте-оригинале.

Модель представляет объект, систему или понятие (идею) в некоторой форме, отличной от формы их реального существования. Она служит средством, помогающим в объяснении, понимании или совершенствовании объекта. Модель какого-либо объекта может быть или точной копией этого объекта (хотя и выполненной из другого материала и в другом масштабе), или отображать некоторые характерные свойства объекта в абстрактной форме.

Различают материальное и идеальное моделирование.

Материальное, или **физическое моделирование** – это экспериментальный метод. Реальному объекту ставится в соответствие его материальная копия – физическая модель (увеличенная или уменьшенная), допускающая исследование в лабораторных условиях.

Идеальное моделирование – это теоретический метод. Реальный объект заменяется его знаковой, **математической моделью**. Поэтому этот метод обычно называют **математическим моделированием**. Модель формируется в виде математических соотношений и проводится ее исследование – аналитическое или с применением вычислительной техники.

Формальное определение модели, как правило, строится на теоретико-множественном языке: система $S \subseteq X \times Y$ называется математической моделью, если задано семейство задач $D_x, x \in X$ со множеством решений системы Y ; для любого элемента $x \in X$ и $y \in Y$ пара (x, y) принадлежит системе S в том и только в том случае, если существует элемент $y \in Y$, который является решением задачи D_x .

Математическое моделирование напоминает физический эксперимент. В математической модели, как и в лабораторной установке, представлены составные части системы и окружающая ее среда. В ходе испытаний модели через некоторые интервалы времени выдается информация о поведении компонентов и показания приборов. При использовании современных технических средств моделирование на ЭВМ так же наглядно, как и физический опыт (особенно для относительно простых систем). Это позволяет быстро получить сведения о различных вариантах изучаемого объекта или процесса. При этом в относительно короткий срок можно найти оптимальные варианты математической модели, т.е. осуществить ее оптимизацию и, следовательно, оптимизировать сам объект или процесс.

Практически, математическое моделирование как метод **не имеет ограничений**, так как: моделирующая система может одновременно содержать описания элементов непрерывного и дискретного действия и быть подверженной влиянию многочисленных случайных факторов сложной природы; допустимо описание системы соотношений большой размерности; обеспечивается простота перехода от одной задачи к другой введением переменных параметров, возмущений и различных начальных условий.

Последовательным наращиванием элементов модели можно исследовать объекты (системы) любой сложности, для которых достаточно полно известны функционирование

и взаимосвязь относительно несложных исходных элементов. При этом переход на более высокий уровень моделирования звеньев системы связан с увеличением количества участвующих в модели элементов, что приводит к необходимости их упрощения или к представлению в виде обобщенных характеристик, полученных на предыдущем этапе имитации. При этом объем моделей сохраняется в некоторых допустимых пределах.

По сравнению с физическим методом математического моделирования **более универсален**, так как он:

- 1) позволяет с помощью одного «устройства» осуществить решение целого класса задач, имеющих одинаковое математическое описание;
- 2) обеспечивает простоту перехода от одной задачи к другой, введение переменных параметров, возмущений и различных начальных условий;
- 3) дает возможность моделировать по частям (по «элементарным» процессам), что особенно существенно при исследовании сложных объектов;
- 4) использует быстродействующую вычислительную технику, которая непрерывно совершенствуется;
- 5) экономичнее метода физического моделирования как по затратам времени, так и по стоимости.

При использовании математических моделей решается одна из двух **задач** – определение необходимых параметров объекта и выявление желательной его структуры – либо совокупность этих задач.

Модели могут применяться как средства осмысливания действительности, общения, обучения и тренажа, средства постановки экспериментов (в том числе оптимальных), а также в качестве инструмента прогнозирования.

При использовании модели в качестве **средства осмысливания действительности** методами математического моделирования можно проводить эксперименты при полностью контролируемом объеме условий моделирования, что исключает возможность воздействия на результаты моделирования случайных факторов. При этом в рамках используемой модели всегда гарантируется отыскание оптимальных решений, если они требуются.

Как **средство общения** хорошо продуманная модель не имеет себе равных. Все языки, в основе которых лежит слово, в той или иной мере оказываются неточными в случае сложных понятий и описаний. Правильно построенные модели помогают устранить эти неточности, предоставляя в наше распоряжение более действенные способы общения. Преимущество модели перед словесными описаниями – в сжатости и точности представления заданной ситуации. Модель делает более понятной общую структуру исследуемого объекта и вскрывает важные причинно-следственные связи.

Математическую модель можно использовать в качестве **средства обучения и тренажа** в сфере образования и профессиональной подготовки. При разработке и использовании модели экспериментатор видит и «разыгрывает» на ней реальные процессы и ситуации. Это помогает ему понять поставленную задачу, что стимулирует процесс самообучения.

На этапе проектирования важнейшее значение приобретает использование математических моделей в качестве инструмента для анализа, оптимизации и прогнозирования поведения моделируемых объектов. Например, моделирование с помощью ЭВМ аварийных ситуаций в работе системы управления позволяет не только освободить конструктора от утомительных проверок схемы, но и расширить число исследуемых вариантов поведения системы, повысить достоверность выводов в отношении ее безаварийной работы. С помощью ЭВМ можно анализировать на стадии проектирования как внешние, так и внутренние причины возникновения аварийных ситуаций.

Математические модели все шире используются непосредственно в системах управления. Подобные модели необходимы для исследования и совершенствования управления техническими системами и применяются в автоматизированных системах управления (АСУ) как на стадии проектирования, так и эксплуатации. Для решения задач управления моделируются реакции объекта, управляющие воздействия, структура системы управления, контроля и так далее, т.е. имитируются процессы, происходящие в управляющей части системы.

Модель в этом случае входит как структурный элемент в проект АСУ. В такой модели необходимо различать: условия нормального функционирования (например, управление по каналам обратной и прямой связей, статическую и динамическую оптимизацию, адаптивное управление, групповое управление); критические ситуации, когда способ управления зависит от информации о типе и глубине отказа; пусковые и аварийные режимы, когда некоторые элементы программного управления могут зависеть от значений параметров и состояния системы.

Метод математического моделирования давно занимает ведущее место среди методов научного познания, но особую силу, наглядность и широту применения он получил с развитием новых информационных технологий, основанных на быстродействующих компьютерах (*вычислительный эксперимент*, т.е. построение и изучение математических моделей с помощью ЭВМ).

Количественный и качественный *выигрыши* от применения математического моделирования на ЭВМ состоят в следующем:

1. Полностью или частично отпадает необходимость в длительном и трудоемком этапе изготовления лабораторного макета или полупромышленной установки; в затратах на комплектующие изделия, материалы и конструктивные элементы, необходимые для изготовления макетов и установок; в измерительных приборах и оборудовании для испытаний системы.

2. Значительно сокращается время определения характеристик (а следовательно, и доводки объекта) и время испытаний.

3. Появляется возможность разрабатывать системы, содержащие элементы, характеристики которых известны, но самих элементов у разработчика нет в настоящее время; имитировать воздействия, воспроизведение которых при натуральных испытаниях затруднено, требует сложного оборудования, сопряжено с опасностью для установки или экспериментатора, а иногда вообще невозможно; легко получать дополнительные характеристики объекта, которые сложно или невозможно получить с помощью измерительных приборов (характеристики параметрической чувствительности, частотные и пр.).

Метод математического моделирования, как любой численный метод, обладает *существенным недостатком*: решение всегда носит частный характер, соответствуя фиксированным значениям параметров системы и начальных условий. Поэтому для всестороннего анализа объекта (системы) приходится многократно моделировать его процесс функционирования, варьируя исходные данные.

Моделирование тесно связано с такими категориями, как абстракция, аналогия, гипотеза и др. Процесс моделирования обязательно включает и построение абстракций, и умозаключения по аналогии, и конструирование научных гипотез. А модель выступает как своеобразный инструмент познания, который исследователь ставит между собой и объектом и с помощью которого изучает интересующий его объект.

Таким образом, моделированием можно назвать процесс построения, изучения и использования моделей, в котором взаимодействуют три элемента: субъект (исследователь), объект исследования и модель (рисунок 4.1):

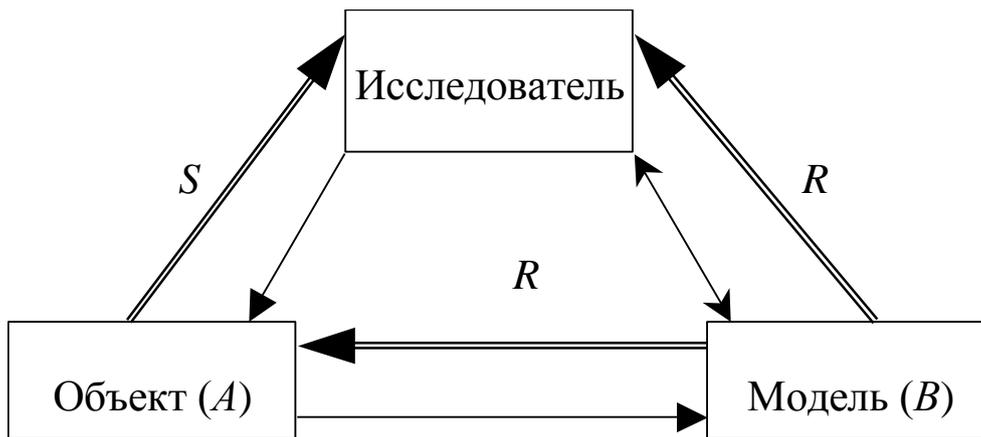


Рисунок 4.1 – Взаимодействие элементов моделирования

Сразу отметим, что исследователь всегда привносит в процесс моделирования индивидуальный субъективный взгляд, как на объект исследования, так и на способы построения модели, на характер интерпретации полученных результатов. Поэтому моделирование всегда представляет собой субъективное отражение объективной действительности.

При использовании математического моделирования разработчик должен, прежде всего, определить, как создать (получить, разработать) модель. Математическое описание является отражением физической сущности процесса со свойственными ему особенностями и ограничениями. Эти особенности и ограничения должны учитываться как при формулировании задачи, так и при составлении описания и выборе численного метода моделирования.

Исходя из реального объекта, формулируются интересующие исследователя свойства на языке той или иной науки, другими словами, строится механическая, либо физическая, либо биологическая, либо социальная и т.п. модель объекта; такая модель называется *содержательной*. При построении содержательной модели формулируются и соответствующие гипотезы (постулаты) модели. При построении содержательной модели отвлекаются от различного рода неидеальностей, неправильностей изучаемого реального объекта, переходят к его упрощенному, схематическому описанию.

Имеется одна важная особенность. Все элементы математической модели (все участвующие величины) являются как бы метками соответствующих реальных элементов. Это дает возможность в процессе решения математической задачи привлекать дополнительные сведения (штриховые стрелки на рисунке 4.2), которые могут упростить этот процесс, либо выделяют из нескольких решений то, которое нужно, и т.д.

Получив решение математической задачи, следует его проанализировать, разобраться в его реальном смысле, сделать выводы. В этом состоит интерпретация (истолкование) результата исследования математической модели. В нее может также входить и контроль правильности (верификация) модели на основе сравнения результата с другими известными фактами, в частности с экспериментальными данными, и т.д.

Математическая модель обычно строится с ориентацией на предполагаемый метод решения математической задачи – в частности, с учетом того, будет ли привлекаться ЭВМ и если будет, то какой мощности. С другой стороны, при проведении математического моделирования или интерпретации решения может понадобиться уточнить или даже существенно изменить математическую модель.

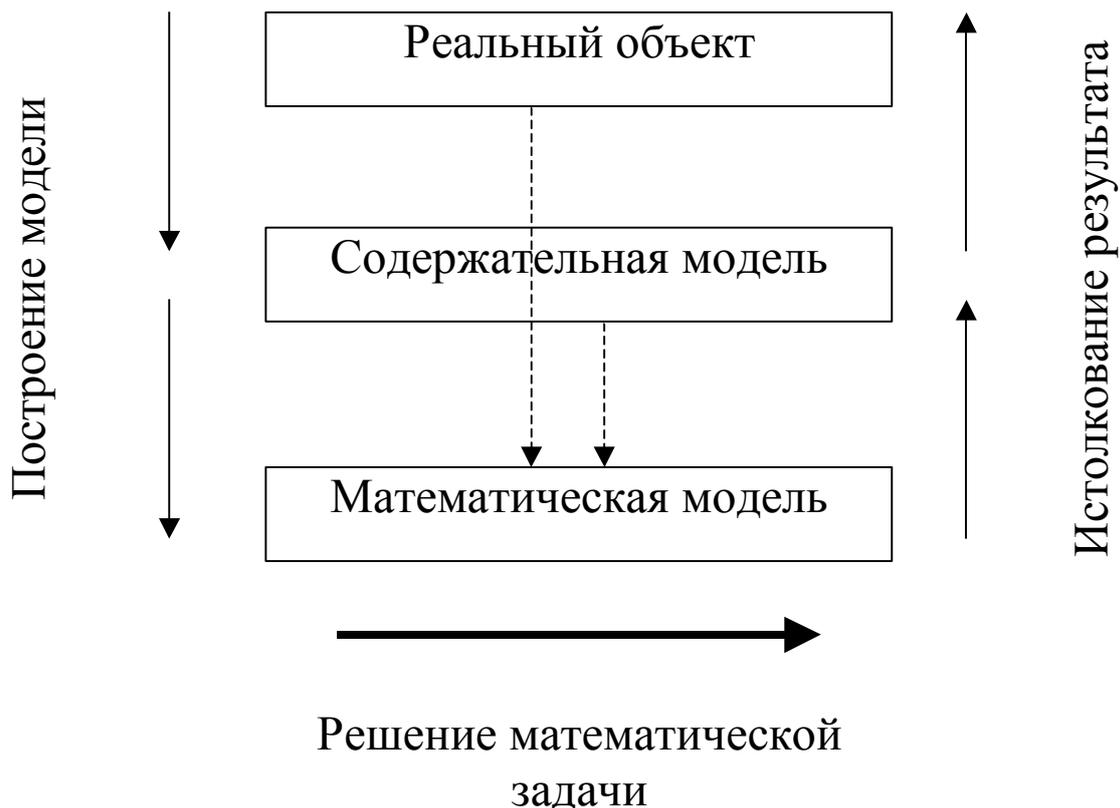


Рисунок 4.2 – Решение математической задачи

Реальный объект может иметь *несколько неравносильных математических моделей*. Это связано, прежде всего, с необходимостью исследования различных систем его свойств. Но даже принципиально разные математические модели рассматриваемого реального объекта могут появиться и при изучении одной и той же системы свойств.

Общие черты математической модели вырисовываются уже при формулировании содержательной модели исследуемого объекта. Однако и после этого обычно бывают возможны различные видоизменения математической модели: в уравнениях можно отбрасывать какие-либо члены или дописывать новые, нелинейные зависимости заменять линейными и наоборот, усложнять или упрощать геометрические формы и т.д.

Возможна и обратная картина: различные реальные объекты или различные содержательные модели могут иметь *одну и ту же математическую модель* – например, описываться одинаковыми дифференциальными уравнениями. Если различные объекты имеют одинаковую математическую модель, то становится возможным моделировать один из этих объектов другим. Например, вместо исследования колебаний сложной линейной механической системы можно производить измерения в соответственно подобранной электрической цепи, имеющей ту же математическую модель. На этом основано действие электромеханических, оптико-механических и других аналоговых устройств. Замечательно, что при применении таких устройств сама математическая модель как бы остается в стороне (значения интересующих механических величин непосредственно получаются по результатам электрических измерений), хотя именно на единстве модели основана возможность этого применения.

В каждой отдельной области знаний цели, задачи и возможности математического моделирования определяются конкретными условиями. Таким образом, математическое моделирование позволяет решать сложные задачи практики, и тем эффективнее, чем сложнее эти задачи.

4.2 Этапы моделирования. Требования к моделям

Пусть имеется некоторый объект A , который необходимо исследовать методом моделирования.

Первый этап построения модели состоит в том, что исследователь конструирует (материально или мысленно) или находит в реальном мире другой объект B – модель объекта A . Этот этап предполагает наличие некоторых знаний об объекте-оригинале. Познавательные возможности модели обуславливаются тем, что модель отражает какие-либо существенные черты объекта-оригинала. Вопрос о необходимости и достаточной мере сходства оригинала и модели требует конкретного анализа. Очевидно, модель утрачивает свой смысл как в случае тождества с оригиналом (тогда она перестает быть оригиналом), так и в случае чрезмерного во всех существенных отношениях отличия от оригинала.

Таким образом, изучение одних сторон моделируемого объекта осуществляется ценой отказа от отражения других сторон. Поэтому любая модель замещает оригинал лишь в строго ограниченном смысле. Из этого следует, что для одного объекта может быть построено несколько «специализированных» моделей, концентрирующих внимание на определенных сторонах исследуемого объекта или же характеризующих объект с разной степенью детализации.

На втором этапе процесса моделирования модель выступает как самостоятельный объект исследования. Одной из форм такого исследования является проведение «модельных» экспериментов, при которых сознательно изменяются условия функционирования модели («входные данные») и систематизируются данные о ее поведении («выходные данные»). Конечным результатом этого этапа является множество знаний о модели R .

На третьем этапе осуществляется перенос знаний с модели на оригинал – формирование множества знаний S об объекте. Этот процесс переноса знаний проводится по определенным правилам. Знания о модели должны быть скорректированы с учетом тех свойств объекта-оригинала, которые не нашли отражения или были изменены при построении модели. Можно с достаточным основанием переносить какой-либо результат с модели на оригинал, если этот результат связан с признаками сходства оригинала и модели. Если же определенный результат модельного исследования связан с отличием модели от оригинала, то этот результат переносить неправомерно.

Четвертый этап – практическая проверка получаемых с помощью моделей знаний и их использование для построения теории поведения объекта, его преобразования или управления им.

От качества подготовки модели зависит качество получаемых результатов, которые заранее предсказать трудно. Однако иногда можно заранее определить, что именно математическое моделирование для решения данной задачи **наиболее приемлемо**. Это возможно в тех случаях, когда:

1) сложно поддерживать одни и те же рабочие режимы при каждом повторении эксперимента на работающем оборудовании или в течение всего времени проведения серии экспериментов;

2) для получения одной и той же величины выборки (и, следовательно, статистической значимости результатов экспериментирования) могут потребоваться чрезмерные затраты времени и средств;

3) при экспериментировании с реальными объектами (системами) невозможно исследование множества альтернативных вариантов, связанных с аварийными или опасными режимами;

4) есть уверенность в успешном создании модели изучаемого объекта, системы или операции. Для этого следует заранее иметь возможность сбора необходимого количества информации об элементах и связях в моделируемом объекте, что обеспечивает достоверность процесса моделирования;

5) имеется возможность (и необходимость) побочного использования процесса построения моделей объекта для их исследования;

б) все другие методы решения непригодны.

Для понимания сущности моделирования важно не упускать из виду, что моделирование – не единственный источник знаний об объекте. Процесс моделирования «погружен» в более общий процесс познания. Это обстоятельство учитывается не только на этапе построения модели, но и на завершающей стадии, когда происходит объединение и обобщение результатов исследования, получаемых на основе многообразных средств познания.

Моделирование – циклический процесс. Это означает, что за первым четырехэтапным циклом может последовать второй, третий и т.д. При этом знания об исследуемом объекте расширяются и уточняются, а исходная модель постепенно совершенствуется. Недостатки, обнаруженные после первого цикла моделирования, обусловленные малым знанием объекта и ошибками в построении модели, можно исправить в последующих циклах. В методологии моделирования, таким образом, заложены большие возможности саморазвития и самосовершенствования.

Рассмотрим и проанализируем последовательность и *содержание этапов моделирования*.

Этап 1. Постановка задачи и ее качественный анализ. На этом этапе формулируется сущность проблемы, принимаются предположения, гипотезы, объясняющие поведение объекта; с помощью методов системного анализа выделяются важнейшие свойства объекта, его структура, изучается взаимосвязь его элементов, цели развития.

В связи со сложностью систем постановка задачи – всегда неизбежный компромисс между учетом всех вероятных факторов, влияющих на функционирование того или иного объекта, и сохранением математической модели достаточно простой, чтобы ее можно было решить с помощью доступных инструментальных и программных средств информационных технологий.

Этап 2. Построение математической модели. Исследуемая проблема формализуется, т.е. выражается в виде конкретных математических зависимостей (уравнений, неравенств и др.). Каждая математическая модель обычно включает в себя три группы элементов:

- совокупность известных внутренних параметров системы $\{X\}$;
- характеристики системы, подлежащие определению $\{Y\}$;
- характеристики внешних изменяющихся условий $\{Z\}$.

Этап 3. Математический анализ модели. С помощью чисто математических приемов исследования выявляются общие свойства модели и ее решения. В частности, доказываются существование решения сформулированной задачи, выясняется, единственно ли оно, определяется область допустимых значений аргументов, область изменения решения, проверяется истинность статистических гипотез о связях между параметрами и переменными модели, сопоставляются размерности величин и т.д.

Этот этап очень важен с точки зрения верификации модели, что позволяет априори выбрать потенциально «правильные» модели. Верификация модели – проверка правильности логической структуры модели.

Этап 4. Подготовка исходной информации. Этот этап, как правило, наиболее трудоемкий.

Во-первых, многие процессы характеризуются закономерностями, которые не обнаруживаются на основании лишь одного или нескольких наблюдений.

Во-вторых, многие процессы характеризуются динамичностью, изменчивостью параметров и структурных отношений. Вследствие этого, такие процессы приходится постоянно держать под наблюдением, необходимо иметь устойчивый поток новых данных. Поскольку наблюдение за динамичными процессами и обработка эмпирических данных обычно занимает довольно много времени, то при построении математических моделей требуется корректировать исходную информацию с учетом ее запаздывания.

В-третьих, используемая при моделировании исходная информация имеет существенно различный характер и происхождение – в зависимости от моделируемых объектов и назначения моделей. Она может быть разделена на две категории: о прошлом развитии и современном состоянии объектов и о будущем развитии объектов, включающая данные об ожидаемых изменениях их внутренних параметров и внешних условий (прогнозы). Вторая категория информации является результатом самостоятельных исследований, которые также могут выполняться посредством моделирования.

Этап 5. Численное решение. Этот этап включает выбор метода решения задачи, разработку алгоритма численного решения, составление и отладку программы расчета на компьютере (или подбор и настройку готовых пакетов прикладных программ) и проведение параметрических исследований, т.е. изучение поведения модели при различных значениях входных параметров.

Этап 6. Анализ полученных результатов и их применение. На этом этапе проверяется адекватность модели, т.е. насколько согласуются полученные знания об объекте-оригинале с практикой. Всесторонний анализ выявляемых расхождений между действительностью и моделью, сопоставление результатов по модели с результатами, полученными иными методами, помогают выработать пути коррекции моделей. Отметим, что адекватность модели – в какой-то степени условное понятие, т.к. полного соответствия модели реальному объекту быть не может. Поэтому при моделировании имеется в виду не просто адекватность, а соответствие по тем свойствам, которые считаются существенными для исследователя.

Если анализ показал неадекватность модели, то, как это отмечалось ранее, происходит переход на первый этап – модель уточняется. Отметим, что такие обратные связи могут существовать на любом этапе процесса моделирования (рисунок 4.2).

При моделировании сложных систем возможны следующие случаи.

1. Моделируемая система достаточно хорошо изучена, что позволяет записать аналитические соотношения, которые и будут служить моделью (законы Кирхгофа, уравнения кинетики, уравнения энергетического и материального балансов и т.п.). Предполагается, что все коэффициенты аналитических соотношений известны.

2. Математическая модель известна с точностью до неизвестных параметров θ , для вычисления которых проводится необходимое число экспериментов.

3. Известно, что моделью может служить одна из функций $\eta_i(x, \theta_i)$ ($i = 1, 2, \dots, l$). Необходимо провести эксперимент для дискриминации моделей и определить неизвестные параметры адекватной модели.

4. Аналитический вид модели не известен вообще.

В трех последних случаях эффективными являются статистические методы моделирования, представляющие собой совокупность методов многомерной статистики (методы регрессионного, дисперсионного, ковариационного, факторного, компонентного и других анализов), которые базируются на наблюдении за функционированием моделируемой системы и обработке результатов наблюдений. При этом наряду с пассивным наблюдением за системой иногда имеется возможность проводить планирование входных возмущений системы. Тогда эффективно применение методов и идей математической теории планирования эксперимента.

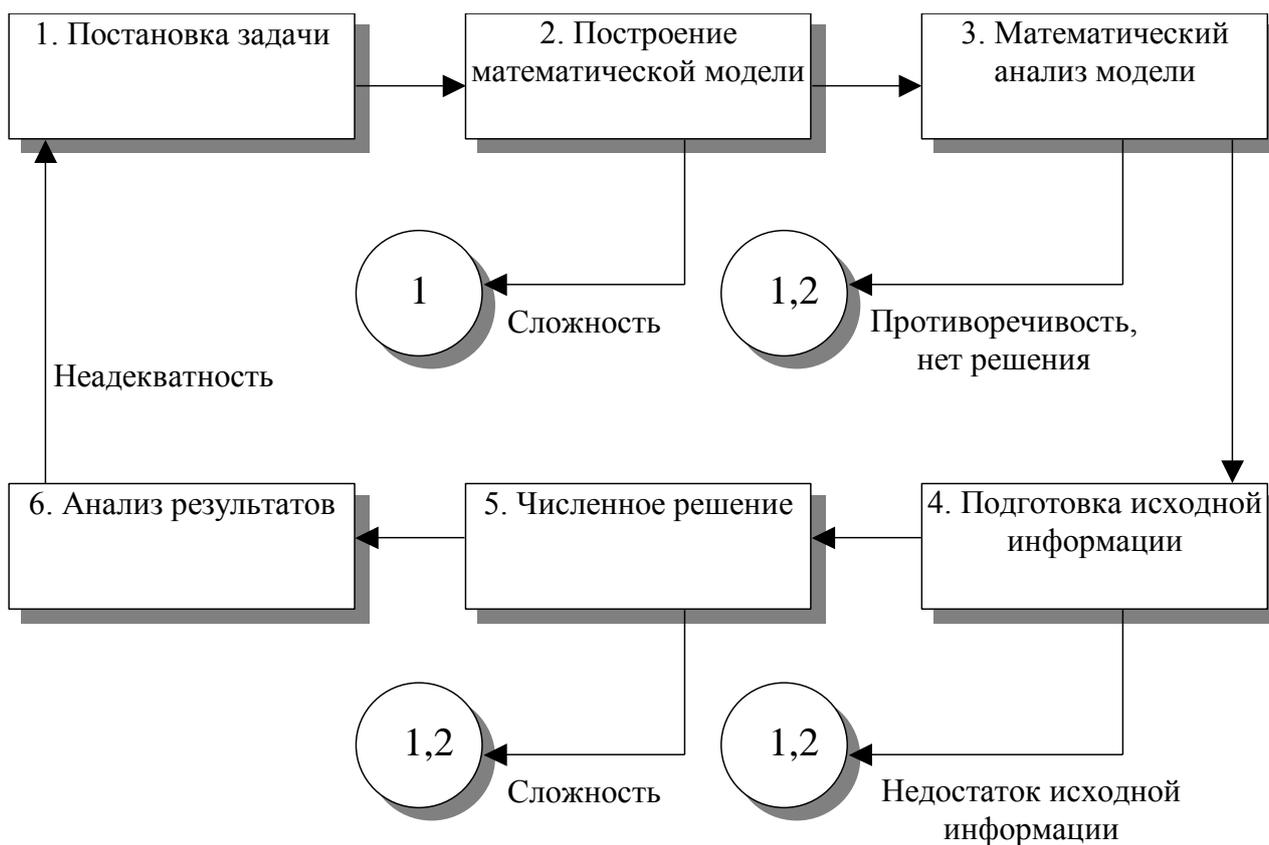


Рисунок 4.2 – Этапы математического моделирования

Важнейшим требованием к математической модели является требование ее **адекватности** (правильного соответствия) изучаемому реальному объекту а относительно выбранной системы S его свойств. Под этим, прежде всего, понимается:

- правильное качественное описание рассматриваемых свойств объекта;
- правильное количественное описание этих свойств с некоторой разумной точностью.

При этом говорят, соответственно, о количественных и качественных моделях. Вместо количественной адекватности говорят также о **точности модели**.

Игнорирование факта, что всякая адекватность математической модели реальному объекту лишь относительна и имеет свои рамки применимости, может привести к грубым ошибкам, основанным на бесконтрольном приписывании реальному объекту свойств его модели. Говоря о математической модели и ее адекватности, часто не упоминают о том, какие именно свойства объекта моделируются. Однако не следует забывать о принципиальной ограниченности области возможного применения любой математической модели.

Если ориентироваться только на требование адекватности, то сложные модели следует предпочитать простым. Усложняя модель, можно учесть большее число факторов, которые могут так или иначе повлиять на изучаемые свойства. Таким образом, приходим к требованию **достаточной простоты модели** по отношению к исследуемой системе его свойств. Именно: модель является достаточно простой, если современные (в частности, вычислительные) средства исследования дают возможность провести экономно по затратам труда и средств, но с разумной точностью качественный или количественный – в зависимости от постановки задачи – анализ исследуемых свойств и осмыслить результат.

Требование простоты модели в каком-то смысле противоположно требованию адекватности: как правило, чем модель более адекватна, тем она менее проста и тем труднее ее анализ. При этом можно упрощать либо содержательную модель объекта, либо ее математическую модель. Опытный исследователь обычно идет по первому пути, так как при этом остаются выполненными наиболее существенные физические соотношения и более ясны постулаты модели.

Существенным является также свойство *полноты математической модели*, состоящее в том, что эта модель дает принципиальную возможность с помощью математических методов получить интересующие утверждения. Еще одно важное требование к математической модели можно назвать ее *продуктивностью*. Оно связано с тем, что изучаемый объект может включать различные параметры – такие как массы, длины и т.п. его компонент, включать функциональные зависимости, которые считаются заданными и описывают связи между рассматриваемыми величинами. Все эти задаваемые параметры и зависимости, называемые исходными данными модели, влияют на значения величин, получаемых в результате решения математической задачи. Требование состоит в том, чтобы в реальных ситуациях исходные данные можно было бы действительно считать заданными, т.е. чтобы их можно было измерить, подсчитать, найти в справочниках и т.п.

Отметим требование *робастности* модели, т.е. ее устойчивости относительно погрешностей в исходных данных. Всегда надо иметь в виду, что эти данные могут быть известны лишь с большей или меньшей точностью и такая неопределенность не должна существенно влиять на результат исследования.

Желательным, хотя и не обязательным является свойство *наглядности* математической модели. Под этим обычно понимают более или менее непосредственный, ясный содержательный смысл ее компонент, который дает возможность не только лишней раз проконтролировать модель, но порой и наметить план решения математической задачи, а также ориентировочно предвидеть результат решения, что может существенно ускорить процедуру.

4.3 Структуризация математических моделей сложных систем

Ставя перед собой задачу наиболее полного учета факторов, влияющих на тот или иной процесс, исследователи приходят к необходимости иметь дело со сложными комплексными явлениями. Это предполагает необходимость совместного изучения различных явлений, в основе которых лежат процессы различной природы, и использование для их анализа различных, но связанных между собой моделей.

Понятие «система» (с греческого: целое, составленное из частей, соединение) в человеческой практике весьма многогранно. Это:

- множество закономерно связанных друг с другом элементов (система предметов, система явлений, система знаний);
- порядок, обусловленный расположением частей в определенном порядке (система работы);
- форма общественного устройства (государственная система);
- совокупность частей, связанных выполнением общей функции (нервная система);
- совокупность хозяйственных единиц, учреждений, организационно объединенных в единое целое (производственно-хозяйственная система) и т.д.

Обобщая все эти понятия, можно дать следующее определение системы. *Система* – это некоторая совокупность элементов произвольного множества, их взаимосвязей,

свойств и взаимоотношений, представляющих целостный комплекс и функционирующих в соответствии с определенными закономерностями, присущими данному комплексу.

Сложность системы определяется количеством входящих в нее элементов, количеством и характером связей между этими элементами, взаимоотношениями между системой и внешней средой.

В наиболее общем виде **системный анализ** обычно определяют двояко:

- 1) как научную дисциплину, разрабатывающую общие принципы исследования сложных объектов с учетом их комплексного характера;
- 2) как методологию анализа объектов путем представления их в качестве систем и исследования этих систем.

Отметим основные идеи, характерные для системного анализа.

1. С позиции системного анализа исследователя, прежде всего, интересует описание места и роли каждого элемента в системе в целом.

2. Системный анализ, как правило, выделяет наличие различных уровней системного объекта и их соподчиненность. Это вызывает необходимость описания взаимосвязи между ними. Наиболее часто встречающаяся форма реализации взаимосвязи – это управление. Поэтому проблема управления возникает практически в любом системном исследовании.

3. Системный анализ ориентируется не только на изучение отдельных элементов, но и ставит задачу выявления и исследования эмерджентности системы, т.е. наличие у системы таких свойств, которые не характерны ни одному элементу системы, взятому в отдельности. Это требует особого выделения в системе синергетических связей, то есть таких связей, которые при кооперированных (совместных) действиях независимых элементов системы обеспечивают больший эффект, чем сумма эффектов каждого из ее элементов, действующих автономно.

Эмерджентность (англ. *emergency*: внезапное появление, возникновение из ничего) – появление в целом нечто качественно нового, такого, чего не было и не могло быть без этого объединения. Синергетика (греч. *synergos*: совместный, согласованно действующий) – наука, изучающая общие закономерности образования, устойчивости и разрушения упорядоченных временных и пространственных структур в сложных неравновесных системах различной природы (физических, химических, экологических и др.). Образно говоря, предмет ее исследования – «возникновение порядка из беспорядка и хаоса», самоорганизованность систем, появление новых качественных свойств.

Возникновение качественно новых свойств при соединении отдельных элементов в систему – это частное проявление всеобщего закона диалектики – закона перехода количества в качество. И чем больше отличаются свойства совокупности от суммы свойств элементов, тем выше организованность системы. Поэтому свойство эмерджентности можно считать проявлением внутренней целостности системы, ее системообразующим фактором.

Приведем пример эмерджентности.

Пусть имеется некий цифровой автомат S , преобразующий любое целое число на его входе в число, на единицу больше входного (рисунок 4.3, *а*).

Если соединить два таких автомата последовательно в кольцо (рисунок 4.3, *б*), то в полученной системе обнаружится новое свойство: она генерирует возрастающие последовательности: одна – последовательность только четных чисел, другая – последовательность только нечетных чисел.

Параллельное же соединение (рисунок 4.3, *в*) ничего не изменяет в смысле проявления новых «арифметических» свойств, но можно отметить появление эмерджентного свойства другого характера – увеличение надежности работы автомата (реализовано дублирование).

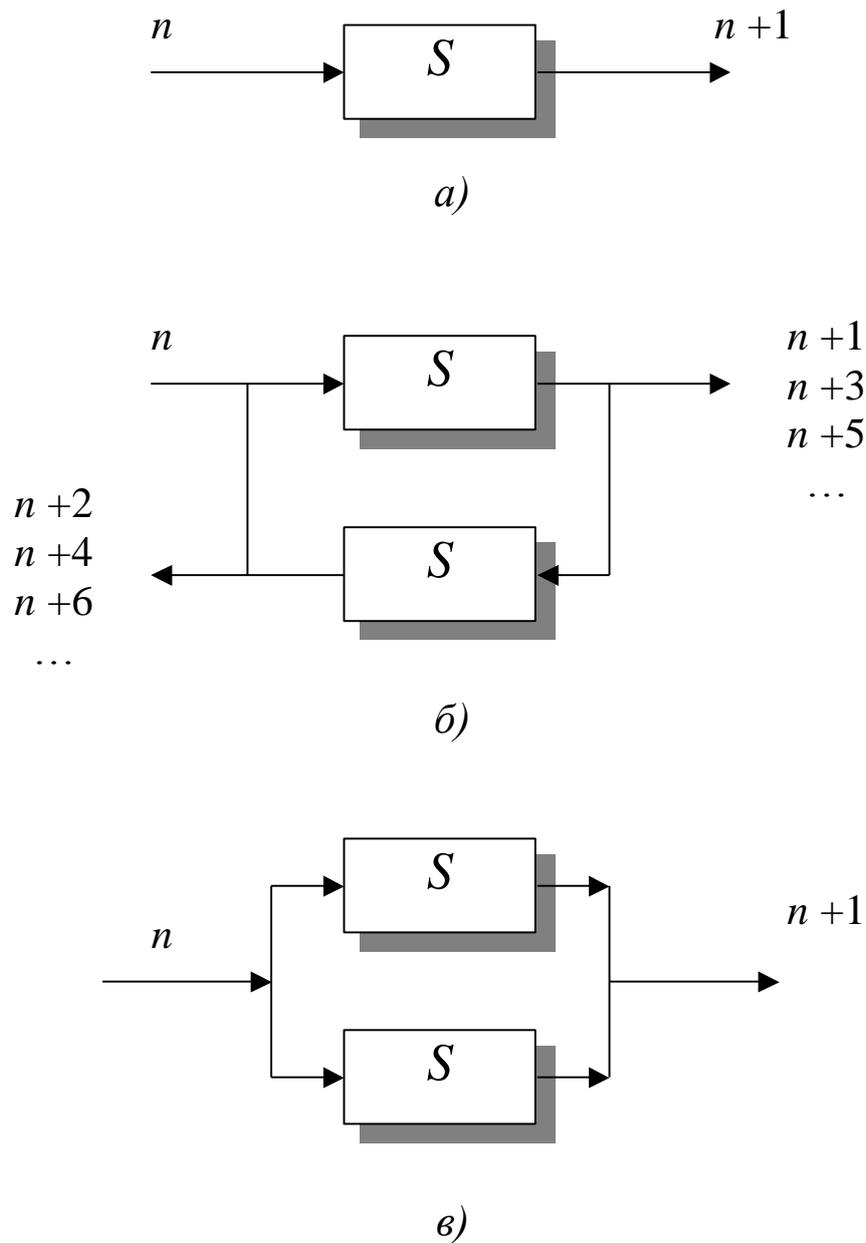


Рисунок 4.3 – Пример проявления эмерджентности

В рамках системного анализа выявляется и исследуется целенаправленность в развитии системного объекта.

При моделировании часто рассматриваются целенаправленные системы, то есть системы, которым небезразлично, в каком состоянии они находятся. Так или иначе, они стремятся к некоторому целесообразному поведению, направленному на достижение наиболее предпочтительных состояний.

Построение любой математической модели для сложной системы начинают с блочного формализованного описания объекта моделирования, т.е. составлению полного математического описания предшествует анализ отдельных «элементарных» процессов, протекающих в объекте моделирования. Полная модель процесса получается как комбинация вариантов моделей отдельных блоков.

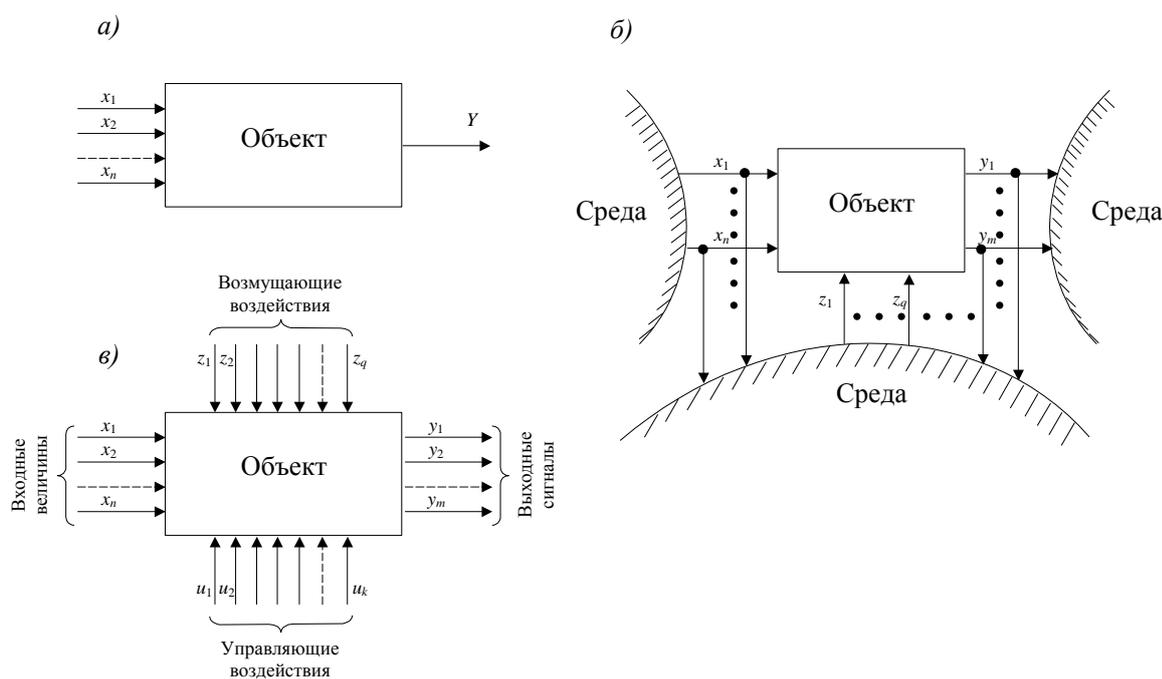
Решаемые с помощью математического моделирования задачи исследования сложных систем подразделяются на три класса:

1) хорошо структурированные или количественно сформулированные задачи, в которых существенные зависимости выяснены настолько точно, что могут быть выражены в числах или символах;

2) неструктурированные или качественно выраженные задачи, содержащие лишь описание важнейших ресурсов, признаков и характеристик, количественные зависимости между которыми не известны;

3) слабо структурированные или смешанные задачи, содержащие как качественные, так и количественные элементы.

Структура модели большой системы показана на рисунке 4.4, а, где Y – вектор выходов. Для определения структуры математической модели в таком виде (т.е. как преобразователь типа «черного ящика» со многими входами и выходами) необходимо выяснить, какие именно входы и выходы объекта будут включены в его модель.



а – модель как обобщенный преобразователь; б – связи объекта (модели) со средой; в – схема объекта с внешними воздействиями

Рисунок 4.4 – Схема модели объекта с n входами и m выходами

Реальные системы функционируют в условиях большого количества случайных факторов, источниками которых являются воздействия внешней среды, а также ошибки, шумы и отклонения различных величин, возникающие внутри системы. Объект связан со средой бесконечным числом связей (рисунок 4.4, б), определяющих его состояние. В общем виде схема объекта изображена на рисунке 4.4, в.

Совокупность параметров среды, которые воздействуют на объект, разделяют на группы в зависимости от характера и доли их участия в процессе. В самом общем случае объект характеризуют следующие параметры: входные величины (входы) – x_1, x_2, \dots, x_n ; управляющие воздействия (управления) – u_1, u_2, \dots, u_k ; возмущающие воздействия (возмущения) – z_1, z_2, \dots, z_s ; выходные величины (выходы) – y_1, y_2, \dots, y_m .

Входными принято называть параметры, значения которых могут быть измерены, но возможность воздействия на них отсутствует. Предполагается также, что эти параметры не зависят от режима работы объекта.

Управляющими называются параметры, на которые можно оказывать прямое воздействие в соответствии с выбором разработчика или предъявляемыми требованиями,

что позволяет управлять процессом. Управляющими параметрами могут быть, например, регулируемое давление в аппарате, температура теплоносителя и т.п.

Возмущающими называют параметры, значения которых случайным образом изменяются с течением времени и которые не доступны для измерения. Ими могут быть, например, содержания различных примесей в исходном сырье, активность катализатора, температура процесса, изменяющаяся за счет реакций, а также другие возмущения.

Выходными называют параметры, значения которых определяются режимом работы объекта. Эти параметры характеризуют его состояние как результат суммарного воздействия входных, управляющих и возмущающих параметров. Поскольку назначение выходных параметров – описывать состояние объекта, их иногда называют параметрами состояния.

Почти каждая модель представляет собой некоторую комбинацию таких составляющих, как компоненты, параметры (переменные), функциональные зависимости, ограничения, целевые функции.

Например, в модели ракеты или космического корабля **компонентами** являются такие объекты, как система тяги, система наведения, система управления, несущая конструкция и т.п. Обобщенная модель города может состоять из таких компонентов, как система образования, система здравоохранения, транспортная система и т.п. В экономической модели производственной системы компонентами могут быть отдельные производственные объединения, отдельные потребители и т.п.

Описание каждого **параметра** (переменной) в модели должно производиться стандартным образом, например: определение и символ; текстовое описание; единицы измерения; диапазон изменения; характеристики (однозначный, многозначный; параметр числовой или с кодированным значением; регулируемая, нерегулируемая или случайная переменная и т.д.); место применения в модели; источник параметра (переменной); примечания.

Функциональные зависимости описывают поведение параметров (переменных) в пределах компонентов системы или выражают соотношения между ними. Эти соотношения, или операционные характеристики, по своей природе являются либо детерминированными, либо стохастическими. Детерминированные соотношения – это тождества или уравнения, которые устанавливают зависимость между переменными или параметрами в тех случаях, когда процесс на выходе системы однозначно определяется заданной информацией на входе. В отличие от этого стохастические соотношения представляют собой такие зависимости, которые при заданной входной информации дают на выходе неопределенный результат.

Ограничения представляют собой устанавливаемые пределы изменения значений переменных или ограничивающие условия распределения и расходования тех или иных средств (энергии, запасов, времени и т.п.). Они могут вводиться либо разработчиком (искусственные ограничения), либо самой системой вследствие присущих ей свойств (естественные ограничения). Например, физической системе такого типа, как ракета, искусственным ограничением может быть заданный минимальный радиус действия или максимально допустимый вес. Большинство технических требований к системам представляет собой набор искусственных ограничений.

Целевая функция, или функция критерия, – это точное отображение целей или задач системы и необходимых правил оценки их выполнения.

Математическую структуру модели объекта в некоторых случаях удобно описывать на языке теории графов. Построение математической модели в виде графа удобно при декомпозиции объекта на элементы. Граф служит источником информации о соподчиненности и связях элементов. Элементы системы – вершины графа, связи – дуги графа.

Формально *элементом* считается объект, не подлежащий дальнейшему расчленению на части (при данном рассмотрении системы). Существенны только свойства элемента, определяющие его взаимодействие с другими элементами системы и влияющие на свойства системы в целом.

Элементный состав может быть *гетерогенным* (содержащим разнотипные элементы), *гомогенным* (содержать однотипные элементы) и *смешанным* (содержать группы однотипных элементов и группы разнотипных элементов).

Связи могут быть нейтральные, прямые (рисунок 4.5, *а*) и обратные (рисунок 4.5, *б*), как положительные, так и отрицательные. Нейтральные связи не связаны с функциональной деятельностью системы, непредсказуемы или случайны.

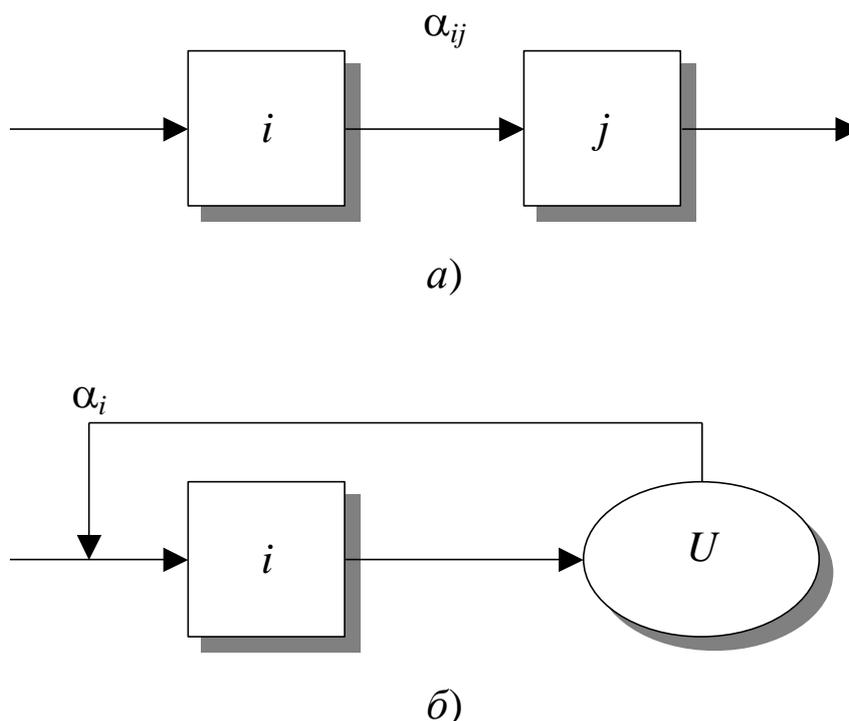


Рисунок 4.5 – Прямая (α_{ij}) и обратная (α_i) связи; i, j – элементы системы; U – блок сравнения выходных значений с каким-либо заданным

Наиболее часто встречающиеся типы структуры приведены на рисунке 4.6.

Структуры с обратной связью реализуются в управляемых системах. Процесс управления можно представить как взаимодействие двух систем – управляющей и управляемой. Назначение управляющей системы – формировать такие воздействия на управляемую систему, которые бы побуждали бы последнюю принять состояние, определяемое целью управления.

В больших системах нельзя установить непроницаемые перегородки, разграничивающие действия переменных различной физической природы. Например, практически во всех химико-технологических процессах нужно одновременно учитывать такие, не поддающиеся в реальных условиях разграничиванию процессы, как теплопередача, аэродинамические и гидравлические процессы, кинетику множества одновременно протекающих реакций. Понятие элемента такой системы и расчленение системы на элементы условны и зависят от целей анализа, так как каждый элемент можно рассматривать как систему.

Элементы могут накапливать, передавать, преобразовывать и рассеивать энергию или информацию. Типичными элементами являются трубы, краны, теплообменники, емкости, компрессоры, двигатели, триоды, конденсаторы и пр.

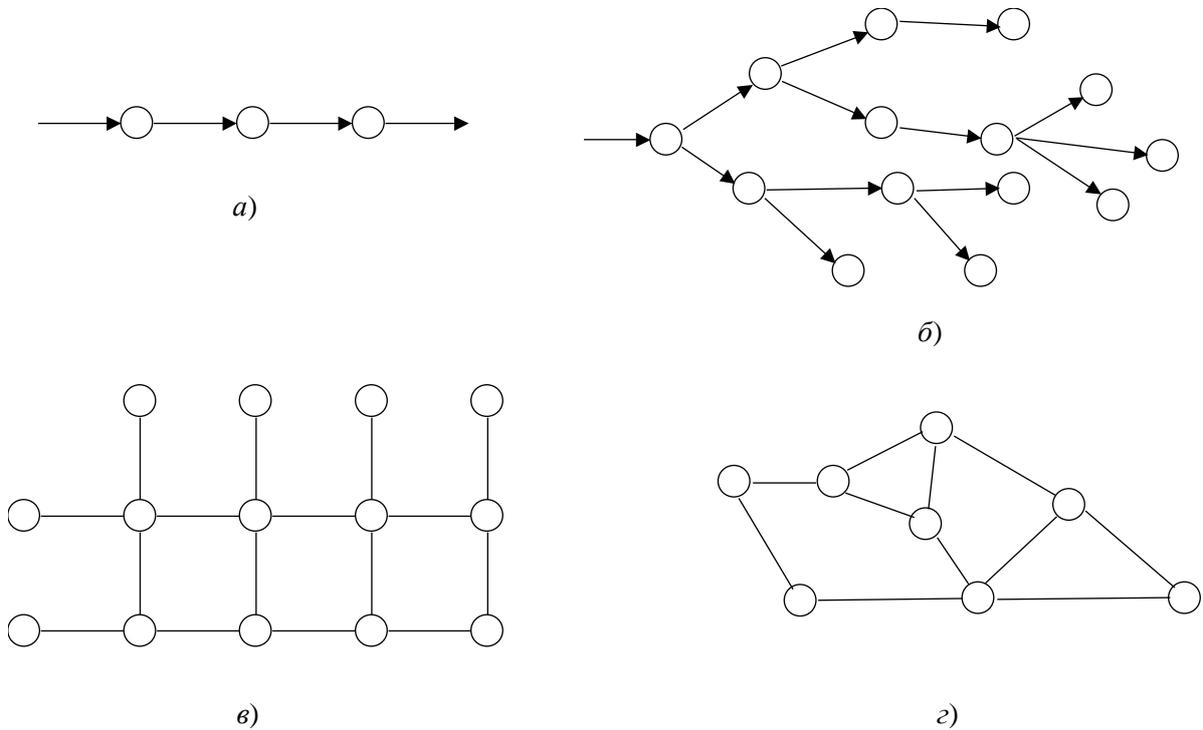


Рисунок 4.6 – Типы структур: а) линейная; б) древовидная; в) матричная; г) сетевая

Точка, в которой соединяются элементы, называется узлом. В узлах не происходит никакого накопления, преобразования или рассеивания энергии; они похожи на абстрактные точки системы координат. Узлами, например, являются электрические шины, патрубки и муфты для присоединения труб. При математическом моделировании практически каждую систему можно представить состоящей из элементов, узлов и подсистем.

Формально любая совокупность элементов данной системы может рассматриваться как ее **подсистема**. Обычно подсистемы являются некоторыми самостоятельно функционирующими частями системы. Например, в производственном комплексе предприятия можно выделить подсистемы, соответствующие отдельным цехам или технологическим линиям. Правильное выделение подсистем сложной системы способствует упрощению расчетов при моделировании и более наглядной интерпретации его результатов. Модель подсистемы составляется в виде структуры из моделей элементов и целиком входит в полную модель управляемой системы. Поскольку подсистема – это самая крупная, функционирующая отдельно от общих связей, структурная единица, важным этапом работы является ее **декомпозиция**, основанная на сборе фактов, выявлении и оценке различных воздействующих факторов. Как правило, в ходе моделирования приходится разделять систему на составные части, т.е. выполнять декомпозицию, а затем обследовать каждую часть в отдельности и объединять полученные сведения в единое целое.

Общая идея модели отображается в виде **логической структурной схемы системы**. Принято строить модель по модульному принципу, т.е. в виде совокупности стандартных блоков-модулей. Такой подход достаточно эффективен, логически оправдан и может быть легко осуществлен и проверен. При этом можно строить и совершенствовать

модель итерационным методом, добавляя к основной схеме блок за блоком. Построение модели из стандартных блоков дает возможность экспериментировать при ее реализации и в процессе машинной имитации.

При построении блочной модели разделяют ее функции на логические подфункции с более высоким уровнем детализации. Каждая модель может быть разделена на **блоки**, а блоки – на **подблоки**. Этот процесс деления блоков на подблоки продолжается до необходимого уровня детализации описания системы. Таким образом, модель функционально подразделяется на подмодели. Используя современные языки программирования, можно получить модель, максимально приближенную к изучаемой системе (как в структурном, так и терминологическом отношении).

Далее выясняются, какие классы объектов должны находиться в модели, и какими параметрами каждый из них характеризуется; выбираются входные и выходные переменные. Обычно выходные переменные модели выбрать нетрудно, так как они определяются уже в процессе формулировки целей моделирования. Чем меньше входных переменных, тем легче процесс моделирования. Однако, если входных переменных слишком мало, модель может стать неадекватной реальности, если слишком много, – из-за недостаточного объема памяти ЭВМ или сложности вычислительных процедур машинная имитация оказывается нереализуемой.

Если некоторые первоначально выбранные подсистемы оказываются чрезмерно сложными, каждую из них расчленяют (с сохранением связей) на конечное число более мелких подсистем нижнего уровня. Процедуру расчленения подсистем продолжают до получения таких подсистем, которые в условиях данной задачи будут признаны достаточно простыми и удобными для непосредственного математического описания. Подсистемы, не подлежащие дальнейшему расчленению, являются, как это сказано выше, элементами сложной системы. Таким образом, в общем случае сложная система является многоуровневой, состоящей из взаимосвязанных элементов, объединяемых в подсистемы различных уровней.

Использование понятия многоуровневой системы существенно расширяет возможности формального описания и моделирования объектов материального мира. При этом объекты большой сложности становятся предметом системного анализа, точного математического расчета. Они могут быть подвергнуты (с помощью ЭВМ) различным количественным исследованиям.

Представление исследуемого объекта в виде многоуровневой конструкции из элементов обычно называют **структуризацией объекта**. Структуризация – первый шаг на пути формального описания сложной системы. Другие необходимые шаги связаны с формализацией элементов системы и взаимодействий между ними. В структурированной системе объектами материального мира являются только элементы.

При декомпозиции сложных систем удобно расчленять их на типовые элементы, в которых протекают сходные между собой процессы. Для выделения типовых элементов (процессов) и определения их природы используют несколько основных критериев:

1) общность математического описания (модели) процессов, т.е. идентичность материальных и энергетических связей. Такая общность модели учитывает физико-химические особенности процессов;

2) общность аппаратно-технологического оформления процессов, отражающая их целевое назначение и условия реализации;

3) общность особенностей автоматического управления, которая связана с природой процессов.

Взаимодействие элементов в процессе функционирования сложной системы рассматривается как результат совокупности воздействий каждого элемента на другие элементы. Воздействие, представленное некоторым набором характеристик, называют **сигналом**. Каждый элемент системы в общем случае может принимать входные сигналы и

выдавать выходные. Сигналы передаются по каналам связи, проложенным между элементами сложной системы.

Совокупность алгоритмов, моделирующих элементы, с учетом алгоритмов их взаимодействия определяет исходный моделирующий алгоритм системы. В большинстве случаев исходный алгоритм нельзя положить в основу модели системы из-за его громоздкости и трудностей реализации на средствах используемой вычислительной техники, поскольку конечные цели моделирования элементов и всей системы различны. Исследователи, занимающиеся оценкой характеристик какого-либо конкретного элемента, разрабатывают моделирующий алгоритм так, чтобы получить оценки характеристик именно этого элемента с максимальной или заданной точностью. Конечные же цели моделирования системы в том, чтобы суммарная ошибка оценки выходных показателей системы не превосходила некоторых наперед заданных величин. В суммарную ошибку входят ошибки случайные (из-за конечного числа реализаций на модели) и детерминированные (обусловленные неточностями структурного описания элементарных процессов).

Обычно стремление точнее описывать процессы в элементах сопровождается усложнением моделирующих алгоритмов, что приводит к увеличению времени счета одной реализации и при ограниченном времени, отведенном на моделирование, – к уменьшению числа реализаций на модели системы. Это, в свою очередь, сопровождается увеличением случайных ошибок в получаемых оценках. Поиск компромиссного соотношения между случайными и детерминированными ошибками с учетом ошибок моделирования, обусловленных ограниченным объемом имеющихся данных, практически всегда связан с анализом допустимых упрощений, как исходных алгоритмов элементов, так и алгоритмов их взаимодействия.

При составлении модели, состоящей из отдельных функциональных блоков, возможны два подхода в зависимости от назначения модели:

Структурный подход – моделирование внутреннего механизма блока. В этом случае математическая модель должна отражать механизм взаимодействия узлов, элементов и деталей рассматриваемого блока; должны моделироваться как внутренняя структура блока, так и функционирование его элементов. Этот подход должен применяться тогда, когда задачей моделирования является, например, проверка структуры блока, правильности взаимодействия его частей и общей логики работы модели. Критерием правильности структуры блока является выполнение блоком заданной в ходе моделирования функции.

Функциональный подход – моделирование функции блока. В этом случае блок рассматривается как «черный ящик», его внутренний механизм может не моделироваться; задается лишь передаточная функция блока в целом. Этот подход применим к тем блокам, внутреннее содержание которых не описывается данной моделью. Такие блоки рассматриваются как неделимые элементы моделируемой системы.

Выбор того или иного подхода к моделированию функциональных блоков зависит от поставленной задачи. В ряде случаев моделирующий алгоритм бывает настолько сложным для реализации с помощью имеющихся в наличии вычислительных средств, что требуется изменить формулировку исходной задачи моделирования для упрощения математического описания. Это упрощение часто достигается за счет снижения точности математической модели путем сокращения полноты математического описания при исключении из модели части параметров или взаимодействий моделируемого объекта.

Основная цель разбиения полной системы на элементы, блоки и подсистемы – это построение ограниченного набора соотношений между характеристиками системы. В общем случае для полной сложной системы эта задача оказывается непосильной. Поэтому обычно приходится расчленять систему на большое количество элементов, математическое описание которых может быть выполнено.

Для определенности процесса декомпозиции введем численную меру сложности модели. В основу этой меры положим трудоемкость синтеза модели, т.е. затраты, необходимые для создания модели. Назовем эту меру *сложностью*. Сложность на стадии анализа (стадии «черного ящика») должна учитывать лишь число входов n и выходов m модели (для простоты не будем различать управляемые и неуправляемые входы). Пусть сложность имеет вид $L = L(n, m)$. Например, $L = n^\gamma m$, где $\gamma > 1$, так как число входов сильнее влияет на сложность, чем число выходов. Величину γ следует определять в зависимости от того, во сколько раз увеличивается трудоемкость синтеза модели при увеличении числа ее входов на единицу.

Функцию сложности в некоторых случаях можно считать аддитивной, т.е. если модель объекта состоит из нескольких (g) подсистем, то общая сложность равна сумме сложностей этих подсистем, т.е. $L = \sum_{i=1}^g L_i$, где L_i – сложность i -й подсистемы исходного объекта.

Учитывая сказанное, процесс декомпозиции модели можно рассматривать как процесс минимизации ее сложности, т.е. решение следующей минимизационной задачи:

$$L_D \rightarrow \min_{D \in \{D\}} \Rightarrow D^*,$$

где D – операция декомпозиции; $\{D\}$ – множество допустимых вариантов декомпозиции; D^* – оптимальная декомпозиция, минимизирующая сложность L декомпозируемой системы.

Пример. Рассмотрим пример декомпозиции объекта на две различные подсистемы ПС1 и ПС2 N различными способами (рисунок 4.7). Это означает, что множество $\{D\}$ состоит из N различных вариантов, для которых числа связей k и q принимают различные значения. Хорошей декомпозицией следует считать ту, в которой эти числа минимальны. Действительно, используя приведенную выше меру сложности, получаем:

$$L = L(k, q) = (n + q)^\gamma k + k^\gamma (m + q).$$

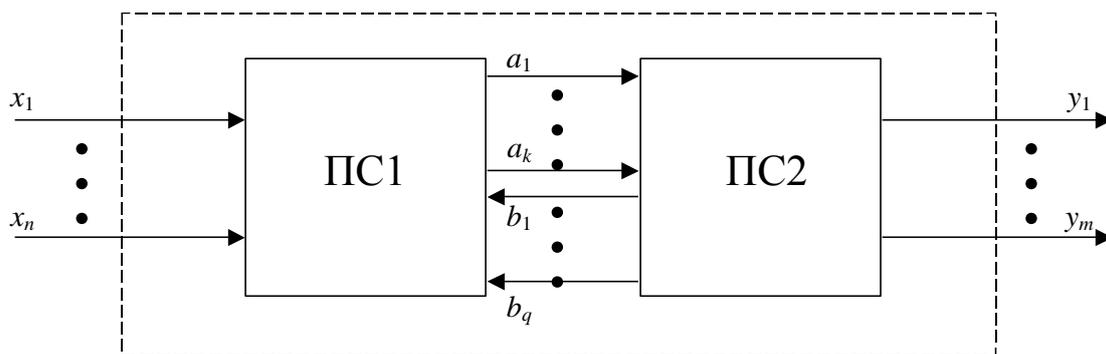


Рисунок 4.7 – Пример декомпозиции сложной системы на два элемента

Оптимальной декомпозицией D^* из $\{D\}$ будет та, которая минимизирует $L(k, q)$. Пусть D_i – i -я декомпозиция, которая определяется двумя числами:

$$D_i = \langle k_i, q_i \rangle, i = 1, \dots, N.$$

Решение получаем в виде $D^* = D_i$, если

$$L(k_i, q_i) = \min_{i=1, \dots, N} \{L(k_i, q_i)\}.$$

Таким образом, цель декомпозиции модели состоит, прежде всего, в том, чтобы упростить последующий синтез модели объекта «расщеплением» ее на более простые

элементы. Этот процесс должен производиться с учетом априорной информации о структурных особенностях объекта.

Рассмотрим множество систем $C = \{C_k\}$, $k = 1, \dots, K$. Для любой отдельной системы $C_{k_0} \in \{C_k\}$ остальные C_k , для которых $k \neq k_0$, есть среда M . C_k можно рассматривать как множество непересекающихся подсистем. Подсистему также можно рассматривать состоящей из множества непересекающихся подсистем низшего уровня.

Для подсистем самого нижнего уровня в иерархии определений, задающих разбиение множества $\{C_k\}$ на непересекающиеся подмножества, описания соответствуют их экспериментальным свойствам. Систему, для которой все свойства определены экспериментально (т.е. описание ее дано не через описание множества подсистем, ее составляющих), назовем *элементом* Э. В иерархии определений элементы находятся на самом нижнем уровне:

$$C_k = \{\mathcal{E}_{c,k,l}\}, l = 1, \dots, L_{c,k}.$$

Описание свойств подсистем и систем более высоких уровней может быть дано через описание свойств подсистем более низкого уровня.

Экспериментируя с некоторой подсистемой C_k как с элементом (т.е. не разделяя на подсистемы), можно получить ее описание и тем самым уменьшить количество подсистем в модели.

Состояние всей совокупности подсистем $\{C_k\}$, а значит, и всей системы C , находящейся на самом высоком уровне в системе определений, обозначим X :

$$X = \{x_1, x_2, \dots, x_j, \dots, x_n\}; \quad X \in \{X\},$$

где $\{X\}$ – пространство состояний совокупности рассматриваемых систем C_k ; $\{X_k\}$ – пространство состояний C_k ; $\{X_M\}$ – пространство состояний среды M (эти пространства имеют общие координаты); x_j – координата вектора пространства состояний – действительное число, обозначающее величину, полученную при физических измерениях или наблюдениях.

Любой подсистеме C_k соответствует совокупность свойств (совокупность закономерностей ее функционирования):

$$\Omega = \{\omega_h\}, \quad h = 1, 2, \dots, H.$$

где ω_h – закономерность функционирования C_k в определенных условиях.

Свойства C_k проявляются в результате взаимодействия со средой; в различных условиях могут проявляться различные свойства – закономерности функционирования системы или элемента.

Закономерность ω_h представляет собой *совокупность описаний* множества частных закономерностей (ω_q), проявляемых системой в условиях, которые соответствуют отдельным конкретным экспериментам. Закономерность ω_h – это функциональное соотношение, которое может быть задано как в виде формулы (некоторой аналитической зависимости), так и в виде таблицы отдельных пар чисел, соответствующих ω_q (значения аргумента и функции, заданные на некотором подмножестве или всем пространстве состояний $\{X_q\}$).

Таким образом, $\omega_h = \{\omega_q\}$, $q = 1, 2, \dots, Q_h$, где ω_q – экспериментальный физический факт, выявляющий элементарное свойство системы и проявляющийся в определенных условиях.

Началу процесса функционирования C_k в соответствии с ω_q , т.е. конкретной реализации физической закономерности, однозначно соответствует начальное состояние

среды – $X_{M,t}^0$ и начальное состояние подсистем $C_k – X_{C,t}^0$, где состояние $M – X_M = (x_{M1}, \dots, x_{Md})$, а состояние $C_k – X_C = (x_{C1}, \dots, x_{Cy})$.

В каждый момент времени t система C и среда M находятся в одном определенном состоянии $(X_{C,t}, X_{M,t})$, которое в общем случае может быть известно лишь с некоторой вероятностью.

Событием в C назовем всякое изменение состояния X_C . Изменение можно зафиксировать только сравнением векторов, соответствующих состояниям C в различные моменты времени t и $t + \varepsilon$ ($\varepsilon > 0$).

Рассмотрим события, происходящие в результате взаимодействия системы и среды. Обозначим событие в C_k как $X_{C,t,t+\varepsilon} = \langle X_{C,t}, X_{C,t+\varepsilon} \rangle$; событие в среде M – как $X_{M,t,t+\varepsilon} = \langle X_{M,t}, X_{M,t+\varepsilon} \rangle$.

Результат взаимодействия – это пара событий (событие в отдельной системе C_k и событие в среде):

$$\langle \omega_{C,M,t,t+\varepsilon} \rangle = \{ X_{C,t,t+\varepsilon}, X_{M,t,t+\varepsilon} \}.$$

При моделировании удобно наблюдать за изменениями в полной системе C через изменения в отдельных подсистемах C_k .

Всякое взаимодействие отдельной подсистемы C_{k0} с другими $C_k \in M$ порождает пары событий, наблюдаемых в C_{k0} и в M , разделенных между собой во времени интервалом $\tau = t' - t$, имеющем, в общем случае, случайную продолжительность.

Взаимодействием элементов системы, или взаимодействием элемента системы со средой, назовем пары событий, отнесенные к различным моментам времени t и t' ($t' \geq t$). Причем событию $\langle \omega_{C,M,t} \rangle$, отнесенному к моменту t (процесс изменений в $\langle X \rangle$, заканчивающийся в момент t), будет всегда с определенной вероятностью соответствовать событие $\langle \omega_{C,M,t'} \rangle$ (процесс заканчивается в момент t').

Любое свойство системы проявляется в определенных условиях в результате взаимодействия со средой. Конкретному эксперименту ω_q всегда соответствуют в определенный момент времени t определенные **условия**, возникающие в результате некоторого предшествующего события в C . Эти условия заключаются в следующем:

1) система C должна находиться в одном из допустимых для данного эксперимента ω_q состояний $X_q \in \{X\}_q$;

2) для каждой из координат вектора $X_q – X_{q,j}$ определяется значение функции $\psi_{q,j}$ в момент времени t :

$$\psi_{q,j} = 1, \text{ если } x_{q,j} \in \{x_j\}^*, \omega, q, j;$$

$$\psi_{q,j} = 0, \text{ если } x_{q,j} \notin \{x_j\}^*, \omega, q, j,$$

где $\{x_j\}^*, \omega, q, j$ – множество допустимых состояний по координате j в момент времени t .

Событие в системе C , соответствующее появлению условий, определяющих проявление свойства ω_q , выявляется функцией

$$*\omega_q = *\omega_q(\psi_{q,1}, \psi_{q,2}, \dots, \psi_{q,j}, \dots, \psi_{q,N}).$$

Здесь $*\omega_q$ принимает значение 1, если $X_q \in \langle X_q \rangle$, и 0 – в противном случае. Если $*\omega_q = 0$, то никаких событий не происходит; если $*\omega_q = 1$, т.е. возникают условия, в которых начинается физический процесс, соответствующий ω_q , то наблюдаемое событие

$$\langle \omega_{C,M,t,t+\varepsilon} \rangle = \{ X_{C,t,t+\varepsilon}, X_{M,t,t+\varepsilon} \}.$$

Событию в C_k $X_{C,t,t+\varepsilon}, t'$, которое рассматривается в описании ω_q , соответствует оператор

$$\Phi_q^*(X_{t'-\varepsilon'}, \varepsilon + \varepsilon' \leq \tau,$$

где Φ_q^* задает отображение $X_{t'-\varepsilon'}$ в $\{X\}_{q, t'}$.

Оператор $\varphi = \varphi^*(\varphi^*(X_t))$ задает отображение X_t в $\{X\}_{q, t'}$.

Конкретный вид функций ω и φ для всех ω_q , т.е. для всех возможных взаимодействий по всем C_k , задает пересечение $\{X_C, K\}$ и $\{X_M\}$ по координатам и тем самым определяет структуру C (множество элементов системы и их взаимодействия).

Изменения структуры системы определяются изменением состава C , т.е. свойств элементов системы в результате их взаимодействий. Описание изменений системы C состоит в описании изменений ω , φ для различных ω_q при всевозможных взаимодействиях. Если считать элементарными терминами составляющие ω_q : ω , φ , φ^* , то изменения структуры могут описываться через изменения состава ω_q , для чего может вводиться функция \mathfrak{Z} , определяющая соответствие состава оператора ω_q событию в системе.

Пример. Рассмотрим некоторый процесс A функционирования сложной системы C , для которого в качестве характеристик состояний выбраны функции $x_1(t), x_1(t), \dots, x_n(t)$, а в качестве параметров – величины a_1, a_2, \dots, a_k . Математическая модель для процесса A – система соотношений вида

$$\begin{aligned} x_1(t) &= f_1(t, a_1, a_2, \dots, a_k); \\ &\dots\dots\dots \\ x_n(t) &= f_n(t, a_1, a_2, \dots, a_k). \end{aligned}$$

Если бы функции f_1, f_2, \dots, f_n были известны (точно или с необходимой степенью приближения), приведенные соотношения оказались бы идеальной в данных условиях математической моделью процесса. На практике модели такого вида, когда характеристики процесса являются явными функциями только его параметров и времени, встречаются редко.

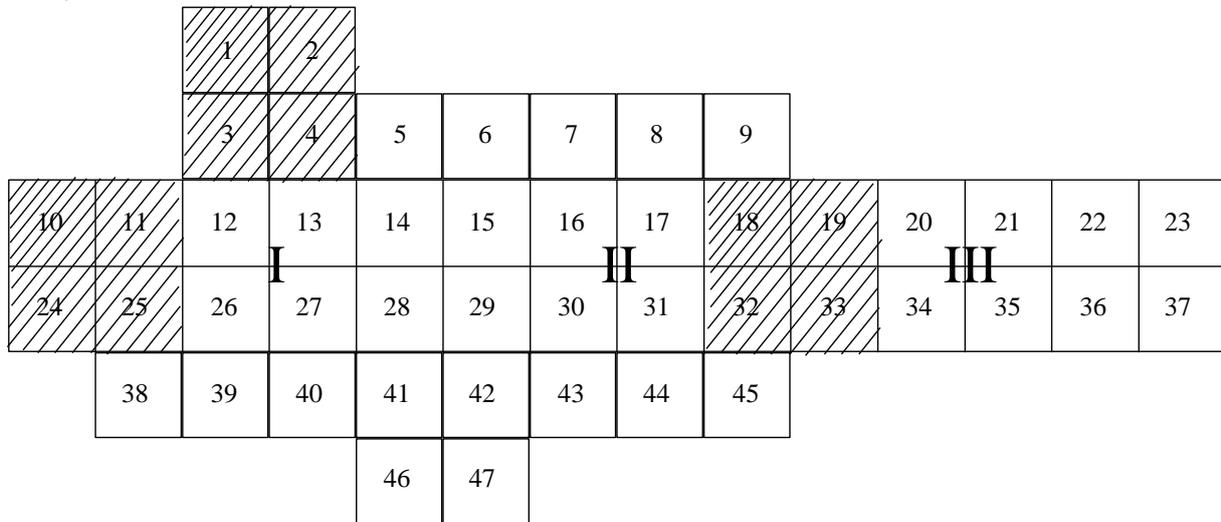
Рассмотрим некоторый процесс A , который расчленим на ряд элементарных актов (блоков, подпроцессов) A_i ($i = 1, 2, \dots, m$) таким образом, чтобы построение математической модели для каждого из них было заведомо возможно (рисунок 4.8). Граница между ними условна и проходит «по телу» системы, рассекая многочисленные связи с циркулирующими по ним потоками информации. При технологически грамотном расчленении системы часть ее элементов из модели исчезает, поэтому такие элементы не оказывают значимого влияния на ход процессов, исследуемых с помощью модели.

Разделение модели процесса A на блоки неоднозначно и зависит от того, какие части системы ранее анализировались автономно, от имеющихся стандартных программ, от традиций исследователя и т.п. Однако при прочих равных условиях обмен информацией между блоками должен быть по возможности минимальным. При решении вопроса о допустимости удаления блока без замены его эквивалентом несущественными и подлежащими удалению считаются блоки модели, маловлияющие на принятый критерий интерпретации результатов моделирования. Правила замены блоков упрощенными эквивалентами различаются в зависимости от характера взаимодействия блоков с оставшейся частью системы. Например, удаляя конечные блоки, составляющие описание взаимодействующего с моделируемой системой «потребителя», следует отразить интересы последнего при формировании критерия интерпретации результатов моделирования.

Рассмотрим способы замены блока, воздействующего на исследуемую часть системы. Это воздействие зависит не только от структуры блока, но и от реакции со стороны исследуемой части. Поэтому характеристики воздействия в общем случае нельзя

однозначно определить при автономном исследовании блока, и его нельзя заменить одним, не зависящим от исследуемой части эквивалентом.

a)



b)

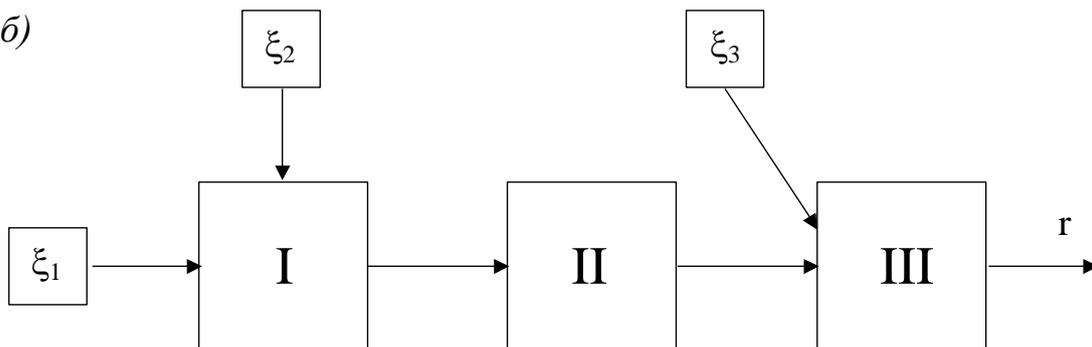


Рисунок 4.8 – Исходный технологический процесс как сумма подпроцессов 1-47 (a) и его декомпозиция (б)

Блок модели, воздействующей на исследуемую часть системы, можно заменить множеством упрощенных эквивалентов, не зависящих от исследуемой части. Каждый эквивалент формирует одно из возможных воздействий в пределах заданного диапазона, а моделирование проводится в нескольких (по числу воздействий) вариантах.

Применим рассмотренные правила к схеме рисунка 4.8. При удалении конечных элементов (22, 23, 36, 37 на рисунке 4.8, a), составляющих описание взаимодействующего с системой «потребителя», часто невозможно наглядно представить результаты моделирования. Поэтому функционирование этих элементов следует отразить при конструировании критерии r интерпретации результатов (рисунок 4.8, б).

Ряд элементов (14, 15, 28, 29) заменяются пассивными связями, транслирующими без искажения информацию, которой обмениваются сохранившиеся элементы. Некоторая часть элементов заменяется внешними воздействиями. Например, элементы 10, 11, 24, 25 заменены воздействием ξ_1 , элементы 1-4 – воздействием ξ_2 . Возможны комбинированные замены: элементы 18, 19, 32, 33 заменены пассивной связью и воздействием ξ_3 . Оставшиеся элементы группируются в блоки I, II, III. Описания этих блоков могут быть полностью воспроизведены в модели или, если это необходимо, заменены упрощенными операторами, характеризующими отдельные аспекты функционирования соответствующей части системы.

Пример. Пусть характеристиками состояний подпроцессов в блоках A_i будут соответственно функции z_{ij} , $j = 1, 2, \dots, r_i$. В качестве параметров для описания подпроцессов A_i выберем соответственно величины β_{il} , $l = 1, 2, \dots, h_i$. Совокупность математических моделей для подпроцессов A_i , рассматриваемых совместно для всех $i = 1, 2, \dots, m$, в общем случае еще не составляет математической модели для процесса A . Эта совокупность характеризует отдельные изолированные подпроцессы A_i :

$$\Phi_{A_i q}(z_{i1}, z_{i2}, \dots, z_{ir_i}, \beta_{i1}, \beta_{i2}, \dots, \beta_{ih_i}, t) = 0.$$

Кроме этих соотношений, необходимо иметь соотношения вида

$$\Psi_p(x_1, x_2, \dots, x_n, a_1, \dots, a_k, z_{11}, \dots, z_{mr}, \beta_{11}, \dots, \beta_{mh}, t) = 0, \\ p = 1, 2, \dots, s,$$

связывающие характеристики z_{ij} подпроцессов A_i с характеристиками x_1, x_2, \dots, x_n процесса A . Совокупность соотношений двух последних видов может служить математической моделью процесса A . Получившаяся таким образом блочная модель (рисунок 4.8, б) предназначена для моделирования взаимодействия блоков через немногочисленные связи, пропускающие ограниченную и обзримую информацию.

Блочные модели являются естественным средством для исследования управляемых систем, конструктивно расчлененных на отдельные аппараты и установки. В связи с этим процесс построения моделей иерархичен: каждый блок (аппарат) в свою очередь допускает блочное представление. Это облегчает управление моделью и организацию (разделение) работ по ее программированию. Каждый блок может быть исследован автономно: аналитически, экспериментально или посредством специального «внутреннего» моделирования. При этом описание некоторых блоков может быть задано стохастически.

4.4 Классификация математических моделей

Общая **классификация** современных моделей показана на рисунке 4.9. Модели подразделяют на две основные группы: вещественные (материальные, приборные) и символические (языковые).

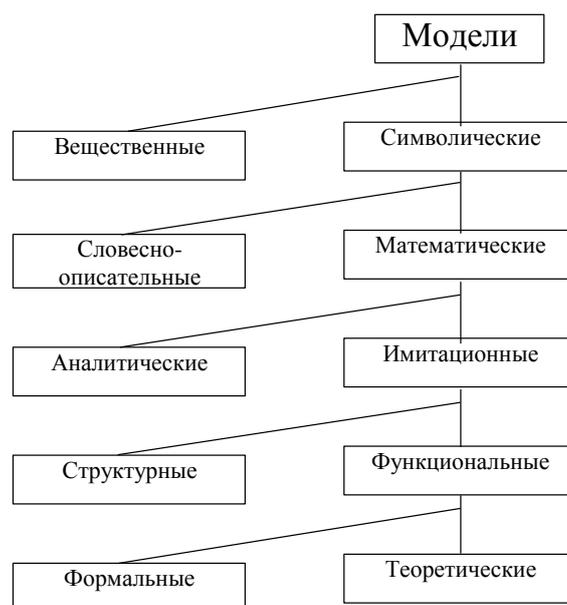


Рисунок 4.9 – Общая классификация моделей

Вещественные модели часто называют просто «модели» (авиамодели, автомодели и пр.). Примерами вещественных моделей являются также пилотные установки (для изучения химических процессов), полигоны с соответствующими макетами для испытаний машин, макеты городов и т.д. Широкое проведение моделирования связано с построением специальных аналоговых (или цифровых) устройств и моделей установок, входящих также в класс вещественных моделей.

В символических моделях фиксация, построение, описание объекта или явления даются на том или ином языке. При этом не имеет значения, на каком конкретном языке описан тот или иной объект, так как переход с одного языка описания объекта на другой не представляет принципиальных трудностей.

Примерами символических моделей являются, например, дифференциальные уравнения второго порядка, описывающие колебания в электрическом контуре или маятнике, чертеж изделия, схема технологической обработки, географические карты, описания, данные на разговорном языке и т.д.

Символические модели делятся на модели словесно-описательные и математические.

К словесно-описательным (дескриптивным) моделям относятся технические задания, пояснительные записки к проектам и отчетам, постановки задач в словесно-описательной форме. Такие модели позволяют достаточно полно описать объект или ситуацию, однако их невозможно использовать непосредственно для анализа процессов формализованным путем с помощью ЭВМ. Поэтому словесно-описательные модели обычно преобразуют в математические для удобства дальнейшего оперирования с ними.

Математическими моделями называются комплексы математических зависимостей и знаковых логических выражений, отображающих существенные характеристики изучаемого явления. Во многих случаях математические модели наиболее полно отображают объект. Примером являются системы алгебраических и дифференциальных уравнений. Поскольку последние представляют собой наиболее абстрактные и, следовательно, наиболее общие модели, математические модели широко применяются в системных исследованиях.

Однако каждое применение математической модели должно быть обоснованным и осторожным: символическая модель всегда является абстрактной идеализацией задачи, поэтому при решении последней необходимы некоторые упрощающие предположения, которые могут привести к тому, что модель не будет служить действительным представлением данной задачи.

Математические модели могут быть аналитическими или имитационными. При использовании **аналитических моделей** процессы функционирования элементов сложной системы записываются в виде некоторых функциональных соотношений (алгебраических, интегро-дифференциальных, конечно-разностных и т.п.) или логических условий. Аналитическая модель может исследоваться одним из следующих способов:

- 1) аналитически, – когда получают в общем виде явные зависимости для искомых величин;
- 2) численно, – когда, не имея решения уравнений в общем виде, применяют средства вычислительной техники, чтобы получить числовые результаты при конкретных начальных данных;
- 3) качественно, – когда, не имея решения в явном виде, можно найти некоторые свойства решения, например, оценить устойчивость решения и т.п.

При использовании **имитационных моделей**, в отличие от аналитических, в ЭВМ воспроизводится текущее функционирование технической системы в некотором масштабе времени. При этом требуется воспроизводить входные воздействия в виде наборов чисел – реализаций процессов (а не числовых характеристик, как при аналитическом моделировании). В зависимости от характера решаемой задачи в процессе имитационного

моделирования с различной степенью точности воспроизводятся и промежуточные преобразования сигнала. Например, если при анализе динамического режима работы блока на его вход подается набор чисел, отображающий процесс с заданной корреляционной функцией, то в ходе моделирования получается реализация выходного процесса, по которой в случае необходимости может быть дана выборочная оценка корреляционной функции выходного сигнала.

Имитационное моделирование напоминает физический эксперимент. Отсюда первое достоинство имитационных моделей – наглядность результатов моделирования (как окончательных, так и промежуточных). Если при аналитическом моделировании обеспечивается подобие характеристик объекта и модели, то при имитационном подобие имеется в самих процессах, протекающих в модели и реальном объекте.

Одно из основных достоинств имитационных моделей – возможность моделирования даже в тех случаях, когда аналитические модели либо отсутствуют, либо (из-за сложности системы) не дают практически удобных результатов. Достаточно просто при имитационном моделировании реализуются алгоритмы обработки результатов измерений для выработки, например, управляющих воздействий в АСУ технологическими процессами (АСУТП), что позволяет оценить точностные характеристики управляющих сигналов. При наличии соответствующих данных можно включить в сферу моделирования объект, управляемый АСУТП, и тем самым оценить качество управления объектом по некоторому показателю эффективности.

Имитационное моделирование позволяет учесть влияние большого числа случайных и детерминированных факторов, а также сложных зависимостей при вводе в модель соответствующих элементов и операций. С точки зрения сбора статистических данных имитационная модель дает возможность проводить активный эксперимент с помощью целенаправленных изменений параметров модели на некотором множестве реализаций. Последнее позволяет исследовать оптимизируемые функции качества (функционалы) системы с помощью ЭВМ.

Для анализа функциональных зависимостей с помощью полученного в результате моделирования ряда числовых результатов могут быть использованы методы поиска: регулярные методы, методы случайного поиска и методы теории статистических решений. Таким образом, в отличие от решения отдельных задач имитационное моделирование на ЭВМ является качественно более высокой степенью изучения сложных систем и применения ЭВМ.

При решении ряда задач могут применяться имитационные математические модели, отображающие только структурные (в частности, геометрические) свойства объекта. Такие **структурные модели** могут иметь форму матриц, графов, списков векторов и выражать возможное расположение элементов в пространстве, непосредственные связи между элементами в виде каналов, проводов, трубопроводов и т.п. Структурные модели используют в случаях, когда задачи структурного синтеза удается ставить и решать, не учитывая особенности физических процессов в объекте.

При моделировании сложных объектов возрастает объем входной информации (описание связей, задание параметров элементов модели и т.д.). Укрупненность же элементов моделей приводит к разрастанию необходимой номенклатуры элементарных моделей, к увеличению объема моделирующей программы. Компромиссным вариантом может быть соответствие разбиения модели делению моделируемой системы на функциональные блоки. **Функциональные модели** отображают как структуру, так и процессы функционирования объекта и чаще всего имеют форму систем уравнений.

По способам получения функциональные математические модели делят на теоретические и формальные. **Теоретические модели** получают на основе изучения физических закономерностей; структура уравнений и параметры моделей имеют определенное физическое толкование. **Формальные модели** получают на основе

проявления свойств моделируемого объекта во внешней среде. Теоретический подход в большинстве случаев позволяет получать математические модели более универсальные, справедливые для широких диапазонов изменения внешних параметров. Формальные модели по сравнению с теоретическими более точны в окрестностях той точки пространства параметров, вблизи которой они определялись, но менее точны вдали от этой точки.

Математические модели классифицируются в соответствии с теми характеристиками моделируемого объекта, которые обуславливают применение того или иного математического аппарата при его моделировании. Рассмотрим следующие виды моделей: стационарные и нестационарные; динамические; линейные и нелинейные; распределенные и сосредоточенные в пространстве; непрерывные и дискретные во времени; непрерывные и дискретные по величине; детерминированные и случайные; информационные.

Стационарные и нестационарные модели. Если свойства преобразования входных сигналов (функций), т.е. структура и свойства оператора $A\{ \}$, не изменяются со временем, то систему и ее модель называют стационарной; в противном случае – нестационарной. Реакция стационарной системы на любой заданный тип возмущения зависит только от интервала времени между моментом начала действия входного возмущения и данным моментом времени, т.е. свойство стационарности означает, что процесс преобразования входных сигналов (функций) инвариантен относительно сдвига как от текущего времени, так и от момента приложения входного сигнала. Реакция нестационарной системы зависит как от текущего времени, так и от момента приложения входного сигнала. В этом случае при сдвиге входного сигнала во времени (без изменения его формы) выходные сигналы не только сдвигаются во времени, но и изменяют свою форму.

К стационарным моделям можно обычно отнести и модели одномоментные, используемые в тех случаях, когда моделируется система, для которой необходимо получить какое-то решение в определенный момент времени. Примером могут служить системы управления запасами материалов, в которых одномоментные модели применяются повсеместно (определение однократного объема заказа на пополнение запасов или времени подачи заказа).

Частным случаем стационарных моделей являются модели *статические*, которые включают описание связей между основными переменными процесса в установившихся режимах (в равновесном состоянии без изменения во времени). Например, математическое описание статики химико-технологического процесса состоит обычно из трех видов уравнений: материального и теплового балансов, термодинамического равновесия системы (характеристика движущей силы) и скоростей протекания процессов (химических реакций, тепло- и массопередачи и т.п.). Для расчетов медленных процессов или процессов, протекающих с небольшими отклонениями от стабильных условий, принимается допущение, позволяющее считать процесс установившимся. Подобное допущение принимается, например, для расчета теплового баланса турбины при половинной, трехчетвертной или полной нагрузке или для решения методами линейного программирования задачи смешения материалов.

Стационарные математические модели (кроме статических) обычно состоят из дифференциальных уравнений, статические – из уравнений алгебраических.

Динамические модели позволяют рассчитать стационарные или нестационарные режимы объектов. Стандартные динамические модели включают переменные и соотношения между ними:

- 1) вектор независимых переменных X ;
- 2) добавочную независимую переменную t , называемую временем, хотя она может не представлять физическую временную размерность;

3) вектор неизвестных параметров θ ;

4) вектор переменных Y состояния системы, который является функцией от t , X и θ .

Эти функции, например, определяются неявно с помощью системы обыкновенных дифференциальных уравнений первого порядка

$$Y = dY/dt = h(t, X, Y, \theta),$$

где X – вектор заданных функций, и системы начальных условий $Y(0) = Y_0(X, \theta)$. Здесь Y_0 – вектор заданных функций;

5) вектор наблюдаемых переменных Z , точными значениями которых Z_T являются заданные функции от переменных состояния и от других переменных: $Z_T = Z_T(t, X, Y, \theta)$.

Общеизвестный специальный случай – когда переменные состояния наблюдаются непосредственно, т.е. $Z = Y$.

Стандартные динамические модели характеризуются множеством переменных состояния системы, которые изменяются со временем (или в зависимости от некоторой другой независимой переменной) в соответствии с определенными дифференциальными уравнениями первого порядка. Начальные условия могут быть известны полностью или частично. Состояние системы наблюдается в различные моменты времени, но иногда переменные состояния не являются непосредственно измеряемыми, и вместо них приходится измерять связанные с ними наблюдаемые переменные. Неизвестные параметры могут появляться в начальных условиях, в дифференциальных уравнениях и в уравнениях наблюдений. В последнем случае они представляют неизвестные характеристики измерительных приборов, например константы калибровки.

Если в модели объекта содержатся дифференциальные уравнения порядка выше первого, сложность их анализа возрастает с ростом порядка уравнения (или с ростом числа дифференциальных уравнений в системе, поскольку уравнение m -го порядка можно преобразовать в систему из m уравнений 1-го порядка). Другая трудность, возникающая иногда при анализе систем дифференциальных уравнений, связана с особенностями задания начальных условий. Чаще всего начальные условия задаются при одном и том же значении независимой переменной. Для протекания химических реакций, например, начальными условиями обычно служат значения концентраций в один и тот же момент $t = 0$; в описаниях реакторов – это концентрации и температура в одной и той же точке – на входе в аппарат. Задачи с начальными условиями, заданными таким образом, называются задачами Коши. Моделировать их сравнительно просто.

Но встречаются задачи, в которых различные начальные условия заданы в разных точках. Например, во многих аппаратах с противотоком часть условий может быть задана со стороны входа одного потока, часть – со стороны входа другого. Это краевые задачи. Если краевую задачу не удастся свести к задаче Коши с помощью дополнительных уравнений (например, уравнений рабочей линии), то решение усложняется. При этом требуется, как правило, применение специальных расчетных приемов – итерации и др.

Состояние системы можно представить как точку с координатами y_1, y_2, \dots, y_m в некотором пространстве m измерений, называемом фазовым пространством или пространством состояний. Эта точка называется изображающей. Изменению состояния системы отвечает некоторое движение изображающей точки в этом пространстве. Путь изображающей точки при этом есть интегральная кривая системы. Эта кривая носит название фазовой траектории.

При построении фазового пространства добиваются взаимно однозначного и непрерывного соответствия между состояниями системы и точками фазового пространства, т.е. каждому состоянию системы должна соответствовать одна и только одна точка фазового пространства, а каждой точке фазового пространства – одно и только одно состояние системы. При этом близким состояниям системы должны соответствовать

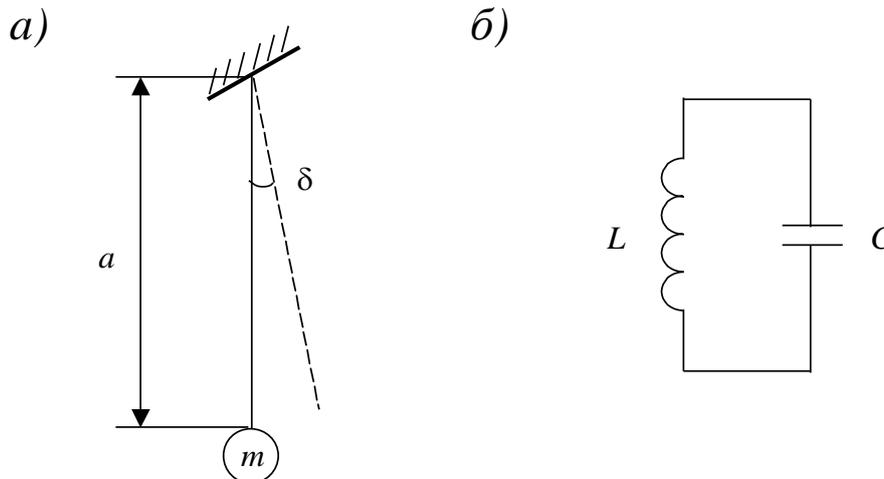
близкие точки фазового пространства. В силу этих требований фазовое пространство не всегда может быть обычным евклидовым m -мерным пространством.

Пример. Физический маятник, состояние которого характеризуется угловой скоростью S и углом δ отклонения от состояния устойчивого равновесия. Уравнение движения маятника при отсутствии сопротивлений $dS/dt = -a \sin \delta$; $d\delta/dt = S$, где $a = \text{const}$ определяется конструкцией маятника. Фазовое пространство в данном случае двумерно. Если принять, что оно является плоскостью, на которой оси δ и S – соответственно оси абсцисс и ординат, то состояние системы с одинаковыми угловыми скоростями S и со значениями $\delta = \alpha + 2\pi - \varepsilon$ при произвольном α и сколь угодно малом ε будут сколь угодно близки. Однако соответствующие им точки плоскости будут отстоять одна от другой на расстоянии $2\pi - \varepsilon$, которое при $\varepsilon \rightarrow 0$ не стремится к нулю. При таком выборе фазового пространства нарушены как требование взаимной однозначности, так и требование непрерывности соответствия состояния системы точкам фазового пространства. Чтобы удовлетворить этим требованиям, необходимо взять за фазовое пространство цилиндрическую поверхность с осью S , направленной по образующей, и с угловой координатой δ .

Приведенные выше уравнения для малых колебаний маятника (рисунок 4.10, а) преобразуются в дифференциальное уравнение второго порядка

$$a \frac{d^2\delta}{dt^2} + \delta = 0,$$

где a – постоянный множитель.



а – маятник; б – электрический контур

Рисунок 4.10 – Простейшие динамические объекты

Аналогично анализ процессов в электрическом колебательном контуре (рисунок 4.10, б) сводится к исследованию дифференциального уравнения

$$L \frac{d^2q}{dt^2} + \frac{1}{C} q = 0,$$

где q – мгновенное значение заряда на обкладках конденсатора. Из последнего уравнения можно получить все сведения об исследуемом процессе, например, период электрических колебаний

$$T = 2\pi\sqrt{LC}.$$

Из сравнения дифференциальных уравнений маятника и контура следует, что они по существу одинаковы. Таким образом, совершенно разные явления могут описываться

моделями в виде дифференциальных уравнений одного и того же вида, а именно – дифференциальными уравнениями второго порядка

$$a_0 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_2 y = 0.$$

Здесь y – обобщенная координата, определяющая состояние движения системы; a_0 , a_1 , a_2 – коэффициенты, зависящие от параметров системы. В случае маятника обобщенной координатой является угол отклонения от вертикали, в случае колебательного контура – заряд конденсатора.

Пример. Пусть сфера радиусом R и массой M свободно падает в несжимаемой Ньютоновской жидкости вязкости μ . Сила тяжести, действующая на сферу, равна $g(M - M_0)$, где M_0 – масса жидкости, вытесняемая сферой. В соответствии с законом Стокса торможение, препятствующее движению сферы (когда это движение медленное), равно $6\pi R\mu V$, где V – скорость падения сферы. Первый закон движения Ньютона приобретает вид

$$M\dot{V} = g(M - M_0) - 6\pi R\mu V,$$

где $V = dV/dt$ – ускорение. Уравнение можно переписать так:

$$\dot{V} + \alpha V = \beta,$$

где $\alpha = 6\pi R\mu/M$, $\beta = g(M - M_0)/M$. Таким образом, получаем стандартную динамическую модель.

Пример. Примером динамической системы может быть гомогенизатор или смесительный силос. Для поддержания постоянного объема цементной сырьевой смеси в гомогенизаторе следует установить равные скорости входного и выходного потоков смеси. Необходимо определить зависимость между их составами. В число материалов, участвующих в цементообразовании, входят окись кальция CaO , двуокись кремния SiO_2 , окись алюминия Al_2O_3 , окись железа Fe_2O_3 и другие соединения. Выведем сначала зависимость между содержанием окиси кальция во входном и выходном потоках материалов, а затем и для других компонентов смеси. Обозначим: $C_{\text{вых}}$, $C_{\text{вх}}$ – количество CaO , выводимое из гомогенизатора и вводимое в него за время Δt в результате прохождения через него определенного количества материала, равного $F\Delta t$; ΔC_{Γ} – изменение количества CaO в гомогенизаторе; V – емкость гомогенизатора; c_{Γ} – относительное содержание CaO в гомогенизаторе (и одновременно в выходном потоке при допущении, что смешение осуществляется идеально); $c_{\text{см}}$ – относительно содержание CaO в выходном потоке сырьевой мельницы и во входном потоке гомогенизатора; F – скорость потока материала, проходящего через гомогенизатор; Δt – интервал времени.

При условии, что значения всех переменных остаются постоянными в интервале времени Δt , количество CaO , выгружаемое из гомогенизатора, составляет $C_{\text{вых}} = c_{\Gamma} F\Delta t$. Аналогично количество CaO , поступающего в гомогенизатор, $C_{\text{вх}} = c_{\text{см}} F\Delta t$. Увеличение количества CaO в гомогенизаторе

$$\Delta C_{\Gamma} = C_{\text{вх}} - C_{\text{вых}}, \text{ т.е. } \Delta C_{\Gamma} = (c_{\text{см}} - c_{\Gamma}) F\Delta t.$$

Увеличение относительного содержания CaO в гомогенизаторе $\Delta c_{\Gamma} = \Delta C_{\Gamma}/V$, откуда

$$\Delta C_{\Gamma} = \Delta c_{\Gamma} V; \quad \frac{\Delta c_{\Gamma}}{\Delta t} = \frac{F}{V} (c_{\text{см}} - c_{\Gamma}).$$

Устремляя Δt к нулю, получаем

$$dc_{\Gamma}/dt = F(c_{\text{см}} - c_{\Gamma})/V.$$

Такая модель технологического процесса показывает, что скорость изменения процентного содержания CaO в гомогенизаторе равна константе F/V (отношение скорости

прохождения материала через гомогенизатор к его емкости), умноженной на разность между относительным содержанием CaO на входе c_{CM} и на выходе c_{Γ} .

Применяя аналогичные обозначения, получаем для SiO_2 , Al_2O_3 и Fe_2O_3 такие же уравнения:

$$\frac{\Delta s_{\Gamma}}{\Delta t} = \frac{F}{V}(s_{CM} - s_{\Gamma}); \quad \frac{\Delta a_{\Gamma}}{\Delta t} = \frac{F}{V}(a_{CM} - a_{\Gamma}); \quad \frac{\Delta f_{\Gamma}}{\Delta t} = \frac{F}{V}(f_{CM} - f_{\Gamma}).$$

Таким образом, модель технологического процесса, выполняемого гомогенизатором, представляет собой ряд дифференциальных уравнений.

Линейные и нелинейные модели. Линейность или нелинейность анализируемого процесса оказывает решающее влияние на вид модели, метод программирования и быстродействие программы при ее выполнении на ЭВМ. Благодаря быстродействию и простоте линейные модели широко применяются разработчиками, хотя большинство природных и промышленных процессов – нелинейно. Примером линейной модели является зависимость между напряжением и силой тока в электрической цепи, хотя это справедливо в ограниченном диапазоне токов и напряжений.

Линейность или нелинейность по отношению к входным сигналам – это не то же самое, что линейность или нелинейность выходных переменных (функций) по параметрам. Оператор $A\{ \}$ и задаваемая им модель называются линейными, если для системы справедлив принцип суперпозиции. Он состоит в том, что линейной комбинации произвольных входных сигналов ставится в соответствие та же линейная комбинация сигналов на выходе из системы:

$$A\left\{\sum_{i=1}^n C_i x_i\right\} = \sum_{i=1}^n C_i A\{x_i\} = \sum_{i=1}^n C_i \eta_i.$$

Математическую модель с использованием линейного оператора можно представить в виде $\eta = AX$.

Нелинейные уравнения, в свою очередь, можно разделить на два подкласса: алгебраические (в которых над переменными производятся только действия сложения, вычитания, умножения, деления и возведения в степень с рациональным показателем) и трансцендентные, в которые входят другие функции от переменных (показательные, тригонометрические и др.). В любом случае сложность модели существенно зависит от числа уравнений и от вида входящих в них функций. Обычно наиболее просто решаются алгебраические уравнения 1-й степени (линейные), наиболее сложно – трансцендентные.

В математических моделях часто используется ряд нелинейных математических структур, в первую очередь, это степенные полиномы. Методы интерполяции дают возможность существенно упростить способы расчета коэффициентов степенного полинома при наличии точных данных о входных и выходных параметрах. Например, при выборе в качестве математической структуры полинома вида

$$P(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_{n+1}(x - x_0)\dots(x - x_n)$$

для данных с равноотстоящими значениями независимой переменной коэффициенты рассчитываются по упрощенному способу:

$$a_0 = y_0; \quad a_i = \frac{\Delta^i y_0}{i! h^i}, \quad i = 1, 2, \dots, n,$$

где h – шаг интерполяции; $\Delta^i y_0$ – разность i -го порядка. В результате полином $P(x)$ записывается в виде

$$P(x) = y_0 + \frac{\Delta y_0}{1! h} (x - x_0) + \frac{\Delta^2 y_0}{2! h^2} (x - x_0)^2 + \dots + \frac{\Delta^n y_0}{n! h^n} (x - x_0)^n$$

и называется интерполяционным полиномом Ньютона.

В интерполяционном полиноме Ньютона используются разности, полученные при движении по одну сторону от значения, взятого в качестве начального. При использовании центральных разностей, т.е. разностей, полученных в результате применения как последующих значений функций (следующего после значения, выбранного в качестве начального), так и предыдущих, коэффициенты полинома рассчитываются следующим образом:

$$a_0 = y_0; \quad a_1 = \frac{\Delta y_0}{1!h}, \quad \dots, \quad a_{2n} = \frac{\Delta^{2n} y_{-n}}{(2n)!h^{2n}}.$$

Введя обозначение $g = (x - x_0)/h$, получим интерполяционный полином Гаусса для равноотстоящих $2n$ точек интерполяции:

$$P(x) = y_0 + g\Delta y_0 + \frac{g(g-1)}{2!}\Delta^2 y_{-1} + \dots + \frac{(g+n-1)(g+n-2)\dots(g-n)}{(2n)!}\Delta^{2n} y_{-n}.$$

Модификациями указанных полиномов являются интерполяционные полиномы Стирлинга

$$P(x) = y_0 + g \frac{\Delta y_{-1} + \Delta y_0}{2} + g^2 \frac{\Delta^2 y_{-1}}{2} + \dots + \frac{g(g^2 - 1^2)}{3!} \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots + \frac{g(g^2 - 1^2)(g^2 - 2^2)\dots[g^2 - (n-1)^2]}{(2n)!} \Delta^{2n} y_{-n}$$

и Бесселя

$$P(x) = \frac{y_0 + y_1}{2} + \left(g - \frac{1}{2}\right)\Delta y_0 + \frac{g(g-1)}{2} \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \dots$$

Для неравноотстоящих узлов интерполяции используется интерполяционный полином Лагранжа

$$P(x) = \sum_{i=0}^n y_i \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)}.$$

Особое место среди известных математических структур занимают ортогональные полиномы. Ортогональным называется полином вида

$$P(x) = k_0\varphi_0(x) + k_1\varphi_1(x) + \dots + k_n\varphi_n(x),$$

если всякая функция $\varphi_i(x)$ системы $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ нормальна и две различные функции $\varphi_i(x)$ и $\varphi_j(x)$ указанной системы функций ортогональны в заданном интервале $a \leq x \leq b$, т.е. имеет место равенство

$$\int_a^b \varphi_i(x)\varphi_j(x)dx = \begin{cases} 1 & \text{при } i = j; \\ 0 & \text{при } i \neq j. \end{cases}$$

Подобный выбор функций $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ дает возможность при квадратичном приближении произвести упрощенный расчет коэффициентов k_0, k_1, \dots, k_n .

Наиболее часто используют те ортогональные полиномы, для которых легко найти выражение в явном виде. Такими являются, например, приведенные ниже полиномы Чебышева, используемые для промежутка $[-1, +1]$. Полином Чебышева первого рода определяется выражением

$$P(x) = \frac{a_0}{2} T_0(x) + a_1 T_1(x) + a_2 T_2(x) + \dots,$$

где значения $T_i(x)$ определяются по выражению

$$T_i(x) = \cos i \arccos x, \quad i = 0, 1, 2, \dots$$

Для целых положительных значений i это выражение является обычным степенным полиномом, так как $T_0(x) = 1$; $T_1(x) = x$; $T_2(x) = 2x^2 - 1$; $T_3(x) = 4x^3 - 3x$.

При изменении x в промежутке $[-\infty, +\infty]$ используется полином Эрмита

$$P(x) = a_0 H_0(x) + a_1 H_1(x) + a_2 H_2(x) + \dots,$$

где

$$H_i(x) = (-1)^i e^{x^2} d^i e^{-x^2} / dx^i; i = 0, 1, 2, \dots$$

В частности,

$$H_0(x) = 1; H_1(x) = 2x; H_2(x) = 4x^2 - 2;$$

$$H_3(x) = 8x^3 - 12x; H_4(x) = 16x^4 - 48x^2 + 12.$$

Для периодических функций используются тригонометрические полиномы

$$P(x) = a_0 + \sum_{i=1}^n (a_i \cos ix + b_i \sin ix).$$

Каждая из математических структур обладает своей спецификой, что и определяет область ее применения при моделировании. Интерполяционные полиномы используются для объектов с известными зависимостями или с точными данными о значениях входных и выходных параметров.

Модели распределенные и сосредоточенные в пространстве. Исследуемый объект (процесс) может быть распределенным или сосредоточенным в пространстве и одновременно изменяться во времени. Модели, описывающие распределенные процессы, называются моделями с распределенными параметрами. Обычно они имеют вид дифференциальных уравнений в частных производных. Если основные переменные процесса не изменяются в пространстве, а только во времени, то математические модели, описывающие такие процессы, называют моделями с сосредоточенными параметрами и представляют их в виде обыкновенных дифференциальных уравнений.

Если процесс развивается одновременно и во времени, и в пространстве (по одной координате l), то оператор A может преобразовывать входную векторную функцию $X(t, l)$ в выходную векторную функцию $Y(t, l)$ и зависеть от обоих аргументов:

$$A = A(t, l) = A_{t, l}.$$

Пример. Рассмотрим твердый брус, нагреваемый с одной стороны и изолированный с другой. Соотношение между температурой, временем и расстоянием от точки нагрева описывается дифференциальным уравнением в частных производных

$$\frac{dT}{dt} = a^2 \frac{d^2 T}{dl^2},$$

где a – коэффициент температуропроводности.

Температура в этом уравнении является функцией двух переменных: времени t и расстояния l , т.е. в любой момент времени t_i температура изменяется с изменением расстояния l_i или, наоборот, в любом месте l_i температура изменяется со временем.

Модели непрерывные и дискретные во времени. Непрерывной во времени модель является в том случае, когда характеризующая ее переменная определена для любого значения времени; дискретной во времени, – если переменная получена только в определенных моменты времени. Непрерывный во времени процесс определяется моделью $Y = A(t)$, где t может принимать любое значение.

Дискретный во времени процесс определяется моделью $Y = f(k\Delta t)$, где $k = 0, \pm 1, \pm 2, \dots$

Так, если сигнал $x(t)$ некоторого компонента системы определен только в моменты времени t_1, t_2, \dots , то такой дискретный сигнал при моделировании записывают в виде последовательности $(x(t_1), x(t_2), \dots)$. Дискретность модели может также возникнуть в том

случае, если она состоит из непрерывных компонентов, но информация переходит от одной компоненты к другой по заданной схеме (такие переходы возможны только по окончании соответствующих операций).

Непрерывные модели применяются при изучении систем, связанных с непрерывными процессами, которые описываются с помощью систем дифференциальных уравнений, задающих скорость изменения переменных системы во времени. Непрерывные модели можно описать с помощью конечно-разностных уравнений, которые в пределе переходят в соответствующие дифференциальные уравнения. При этом программирование непрерывных моделей сводится к дискретной вычислительной задаче, которую можно символически записать в виде уравнения

$$y(t + \Delta t) = g[y(t), y(t - \Delta t), \dots, x(t), \theta],$$

где $y(t)$, $y(t - \Delta t)$ – соответственно векторы состояния системы с компонентами, определяемыми во все предыдущие моменты времени; $x(t)$ – вектор внесистемных переменных; θ – вектор параметров системы; g – функция, определяющая поведение системы.

Конечно-разностные уравнения (линейные и нелинейные, стационарные и нестационарные, уравнения первого и более высокого порядков, одномерные и многомерные) позволяют описать самые разнообразные динамические системы как с дискретным временем, так и, в пределе, с непрерывным, т.е. это достаточно универсальный метод моделирования.

Рассмотрим линейное одномерное конечно-разностное уравнение q -го порядка ($q > 1$) с постоянными коэффициентами. Предположим, что наблюдения производятся в дискретные, равноотстоящие моменты времени. Примем также, что реакция на выходе такой динамической стационарной системы появляется с некоторой временной задержкой (на b интервалов квантования) по отношению к входному сигналу. С учетом указанных предположений запишем конечно-разностное уравнение в виде

$$y(k) = \delta_1 y(k-1) + \delta_2 y(k-2) + \dots + \delta_q y(k-q) + \omega_0 u(k-b) + \omega_1 u(k-b-1) + \dots + \omega_p u(k-b-p),$$

где $\delta_1, \delta_2, \dots, \delta_q, \omega_0, \omega_1, \dots, \omega_p$ – параметры уравнения; k – номер очередного интервала квантования.

Линейное конечно-разностное уравнение высокого порядка подобного типа всегда можно заменить системой конечно-разностных уравнений первого порядка, выраженных в форме Коши, т.е. разрешенных относительно первых разностей. Это обстоятельство привело к тому, что в качестве наиболее общего вида представления математического описания стали широко использоваться конечно-разностные модели в форме Коши. Стационарная линейная система может быть описана моделью в стандартной форме:

$$\begin{cases} \eta(k+1) = \Phi \eta(k) + G u(k); \\ y(k) = H \eta(k), \end{cases}$$

где $y(k)$, $\eta(k)$, $u(k)$ – соответственно векторы откликов, переменных состояния и управляющих входных сигналов.

Матрицы Φ , G и H не зависят от момента времени, т.е. являются постоянными. Они могут включать неизвестные параметры модели, подлежащие оцениванию. Приведенную модель называют также канонической формой модели стационарной линейной системы с дискретным временем. Если система нестационарна, то матрицы Φ , G и H будут зависеть от k .

Нелинейная импульсная система в достаточно общем случае может быть описана моделью

$$\begin{cases} \eta(k+1) = v[\eta(k), u(k), k, \theta]; \\ y(k) = \phi[\eta(k), u(k), k, \theta], \end{cases}$$

где $v[]$, $\varphi[]$ – векторные функции.

Дискретизация возможна не только по аргументу – времени, но и по уровню (величине) сигналов. В этом случае сигналы имеют конечное число значений в некоторой заданной области существования. Подобные сигналы называют дискретными по величине или квантованными. И если непрерывному скалярному сигналу y_n соответствует плавная кривая, то квантованному скалярному сигналу y_k отвечает кусочно-постоянная линия (рисунок 4.11).

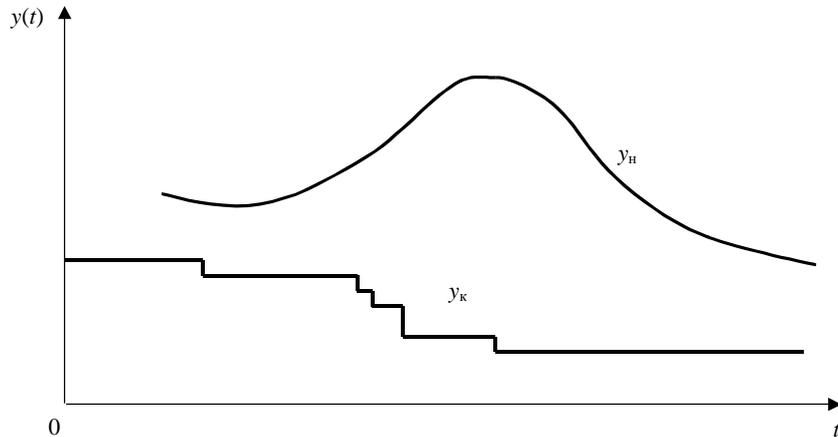


Рисунок 4.11 – Непрерывный (y_n) и дискретный по величине (y_k) сигналы

Системы (модели), у которых входные и выходные сигналы являются непрерывными по времени и по величине, называют непрерывными. Если же входные и выходные сигналы дискретны по времени, то системы называют системами с дискретными временем или импульсными. Системы, у которых входные и выходные сигналы дискретны или по времени, или по величине, называют дискретными. Существуют также и системы промежуточного типа, у которых свойства сигналов как функций от времени различны. Например, часть сигналов может быть непрерывной во времени, тогда как остальные сигналы могут быть дискретными. Такая система называется дискретно-непрерывной по времени.

Детерминированные и случайные модели. По наличию в модели случайных элементов, т.е. в зависимости от способа задания параметров, исходной информации, начальных условий и способа нахождения характеристик системы, математические модели можно подразделить на два больших класса: детерминированные и случайные (вероятностные, стохастические). В детерминированных моделях все исходные данные, ограничения и целевая функция (т.е. некоторое соотношение, количественно характеризующее поставленную перед системой цель) задаются в виде конкретных чисел, векторов или числовых функций.

В детерминированных моделях используются различные классические методы математики: дифференциальные, линейные, разностные и интегральные уравнения, операторы для сведения к алгебраическим моделям и др. При совместном рассмотрении этих соотношений состояния системы в заданный момент времени однозначно определяются через параметры системы, входную информацию и начальные условия.

По степени математической абстракции детерминированные модели можно разделить на сложные, описывающие все причинные связи какой-то реальной системы и позволяющие точно прогнозировать поведение системы в зависимости от изменения переменных (или параметров), и упрощенные, – в которых выбирается ряд основных, существенных зависимостей; устанавливаются и математически описываются связи между отдельными параметрами, соответствующие причинно-следственным закономерностям; другие, несущественные, связи отбрасываются (идеализированные модели).

Между этими двумя моделями существует ряд моделей, отличающихся степенью детализации. Первые модели, являясь наиболее точными и достоверными в чистом виде, из-за сложности не могут широко применяться в моделировании. На практике чаще всего применяются упрощенные идеализированные модели. При этом считается, что имеются существенные и несущественные факторы: существенные учитываются, несущественные отбрасываются. Между принятыми в модели факторами и результирующими показателями устанавливается жесткая детерминированная связь. Широкое распространение идеализированных моделей вызвано их простотой и возможностью логического обоснования.

Любому реальному процессу присущи случайные флуктуации. Однако выбор детерминированной или вероятностной математической модели зависит от того, учитываются ли случайные факторы. Выделение детерминированных моделей в отдельный класс объясняется широким их применением и разнообразием математических методов решения детерминированных задач.

Если хотя бы один параметр модели или ограничительная функция имеет в качестве своих значений случайный вектор или случайную величину, то это случайная (стохастическая) модель. В этом случае под однозначностью определения характеристик моделируемого процесса понимается однозначное определение распределений вероятностей для характеристик процесса при заданных распределениях вероятностей для начальных условий и возмущений.

Стохастический характер модели связан с наличием в объекте и в среде различных неконтролируемых, но существенных факторов, которые можно моделировать статистически. Состояние системы в этом случае $Y = F(X, U, E(t))$, где $E(t)$ – случайный процесс, моделирующий имеющуюся неопределенность объекта и среды. Эта неопределенность может быть связана как с быстрым изменением параметров объекта, так и с помехами, накладывающимися на измеряемые значения сигналов на входе и на выходе объекта.

Стохастический объект и его модель ведут себя неоднозначно в одинаковых ситуациях, что моделируется случайным вектором $E(t)$, статистические свойства которого должны быть заданы. В простейшем случае

$$Y = F(X, U) + E(t).$$

Примером стохастического объекта является любой биологический организм, который в одинаковых условиях ведет себя по-разному. В этом случае Y описывает поведение объекта, которое строго зависит от внешних условий, а все отклонения от этого регулярного поведения образуют «случайную помеху» $E(t)$.

Переход от детерминированной модели к стохастической осуществляется таким образом, чтобы она отражала в себе случайный характер данных и самой модели. Способ перехода выбирается в зависимости от сведений об изучаемой модели: уверенности в правильности и надежности данных и модели. При этом возможно, что эти сведения ошибочны.

Например, в случае детерминированного безынерционного объекта, когда возмущение и реакция могут рассматриваться как случайные величины X и Y соответственно, математическая модель, описывающая объект, дается в виде условного математического ожидания Y относительно X , т.е. объект описывается уравнением вида

$$M\{Y | X\} = f(X),$$

где $M\{Y | X\}$ – условное математическое ожидание Y относительно X , f – неслучайный закон преобразования.

Так, для усилительного элемента, на входе которого действует случайная величина X , выходной сигнал Y имеет вид

$$Y = M\{Y | X\} = KX.$$

В общем случае для стохастических объектов оператор является случайным (например, коэффициенты линейного дифференциального уравнения, весовые функции и т.д.).

Информационные модели. С помощью информационных (процедурных) моделей моделируются сложные устройства и комплексы типа вычислительных машин, радиолокационные станции, системы управления большими промышленными установками, летательными аппаратами и т.д. Функционирование таких систем представляет собой цепь событий, происходящих в дискретные моменты времени и заключающихся в изменении состояний элементов. Дискретное представление пространства и времени обуславливает дискретность фазовых переменных, которыми являются величины, характеризующие состояния элементов. Роль элементов и внутренних параметров выполняют системы и выходные параметры некоторых подсистем. Так, элементами ЭВМ можно считать арифметическое устройство, оперативную память, устройство ввода и вывода и т.п. Фазовые переменные, характеризующие состояния этих элементов, могут принимать только два значения: «занято», если в данный момент устройство работает, или «свободно», если устройство находится в состоянии ожидания. Примерами выходных параметров служат вероятность обслуживания поступивших в систему заявок (сообщений), среднее время простоя в очереди на обслуживание, быстродействие устройства. Для построения математических информационных моделей широко используют математическую логику, теорию массового обслуживания, методы теории автоматического управления.

При анализе сложных промышленных производств особый интерес представляют информационные процедурные модели, а также модели режимов и обеспечения безопасности работы. Информационные процедурные модели определяют содержание, формат и скорость (или частоту) потока информации. Эти модели охватывают также контроль и проверку информации, учет и отчетность по ней, получение разрешений и представление некоторых видов информации, меры предосторожности против потерь информации в аварийных случаях и порядок работы по ее восстановлению при неисправностях или поломках.

Процедурные модели режимов и обеспечения безопасности работы описывают действия, изменяющие состояние комплекса оборудования предприятия, а также совокупность предписаний об ограничениях, налагаемых на ход работы по соображениям безопасности. К типичным режимам относятся пуск, останов оборудования, изменения нагрузки. При разработке этих моделей человек-оператор рассматривается как составная часть комплекса оборудования; ему отведены следующие функции: обеспечение ввода данных в ЭВМ; слежение за выходными данными с помощью измерительной аппаратуры и ЭВМ; привлечение к работе резервного оборудования в случае неисправности основного; отыскание и устранение ошибок в программе и неполадок в ЭВМ. Эти функции отведены оператору, так как он может выполнять их лучше и с большей экономией средств, чем любая автоматизированная система.

Пример. В исходном режиме турбину медленно вращает небольшой электродвигатель с помощью специальной передачи. Задача состоит в том, чтобы освободить турбину от этой передачи и ускорить ее вращение в соответствии с графиком ускорения, рекомендованным заводом – изготовителем турбины. В ходе процесса ускорения производятся проверки на соответствие вибрации и эксцентриситета допустимым пределам. Если пределы, установленные для вибрации и эксцентриситета, превышены, турбину необходимо «придерживать» на определенных безопасных скоростях, которые ниже критических скоростей, вызывающих очень большую вибрацию. После «придерживания» при соблюдении определенных условий возможно возобновление процесса ускорения.

При выполнении этой процедуры методом ручного управления для ускорения вращения турбины необходимо регулировать зазор перепускного вентиля. При полуавтоматическом управлении этой операцией для разгона турбины требуется устанавливать контрольные точки регулятора числа оборотов во всем диапазоне скоростей. На рисунке 4.12 показана процедурная модель машинного управления разгоном турбины с помощью импульсов, подаваемых управляющей ЭВМ на гистерезисный двигатель, приводящий в действие перепускной клапан запорного вентиля.

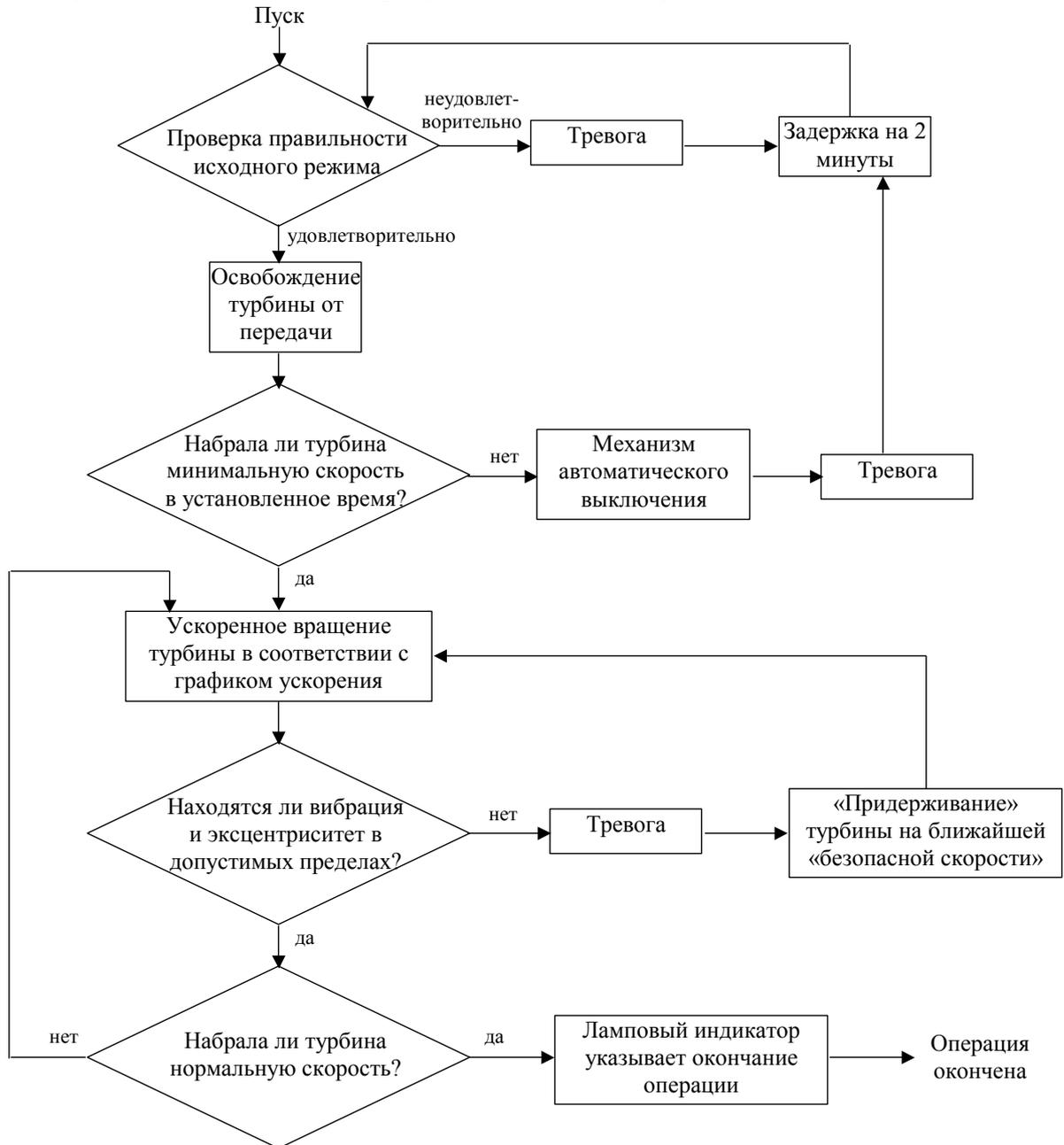


Рисунок 4.12 – Информационная модель пуска турбины

Существует много проблем математического моделирования: как методологического и теоретического, так и организационного и прикладного характера.

Во-первых, это несовершенство существующих моделей, что неизбежно ведет к дискредитации идеи полезности математического моделирования, к медленному внедрению математических моделей в теорию и практику. Моделировать можно объект любой природы и любой сложности. И как раз сложные объекты представляют наибольший интерес для моделирования; именно здесь моделирование может дать результаты, которые нельзя получить другими способами исследования.

Во-вторых, сложность некоторых процессов и явлений затрудняют не только построение математических моделей, но и проверку их адекватности, истинности получаемых результатов. В естественных науках достаточным условием истинности результатов моделирования и любых других методов познания является совпадение результатов исследования с наблюдаемыми фактами. Категория «практика» совпадает здесь с категорией «действительность».

В-третьих, недостаток системного, комплексного подхода к моделированию задач, который бы объединял коллективы математиков, физиков, программистов и т.д. Этому, кстати, способствует и борьба научных школ, научных направлений, порою диаметрально противоположно трактующих одни и те же явления.

Поэтому, потенциальная возможность математического моделирования любых объектов и процессов не означает, разумеется, ее успешной осуществимости при данном уровне математических знаний, имеющейся конкретной информации и вычислительной техники. И хотя нельзя указать абсолютные границы математической формализуемости проблем, всегда будут существовать еще неформализованные проблемы, а также ситуации, где математическое моделирование недостаточно эффективно.

5. ТЕОРЕТИЧЕСКИЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ

5.1 Задачи и виды теоретических исследований

Целью теоретических исследований является выделение в процессе синтеза знаний существенных связей между исследуемым объектом и окружающей средой, объяснение и обобщение результатов эмпирического исследования, выявление общих закономерностей и их формализация.

Теоретическое исследование завершается формированием теории, не обязательно связанной с построением ее математического аппарата. Теория проходит в своем развитии различные стадии от качественного объяснения и количественного измерения процессов до их формализации и в зависимости от стадии может быть представлена как в виде качественных правил, так и в виде математических уравнений (соотношений).

Задачами теоретического исследования являются: обобщение результатов исследования, нахождение общих закономерностей путем обработки и интерпретации опытных данных; расширение результатов исследования на ряд подобных объектов без повторения всего объема исследований; изучение объекта, недоступного для непосредственного исследования; повышение надежности экспериментального исследования объекта (обоснование параметров и условий наблюдения, точности измерений).

При проведении теоретических исследований, основанных на общедоступных методах анализа и синтеза, широко используются декомпозиция и композиция исследуемой системы (объекта, явления).

На всех этапах построения модели объекта производится его упрощение, и вводятся определенные допущения. Последние должны быть осознанными и обоснованными. Неверные допущения могут приводить к серьезным ошибкам при формулировании теоретических вопросов.

При построении моделей объекта исследования должны использоваться наиболее общие принципы и закономерности. Это позволяет учесть все допущения, принятые при получении формализованных теорий, и точно определить область их применения.

Противоположным расчленению является метод объединения и связанный с ним комплексный подход к изучению объекта, которые чаще всего объединяются под названием «общая теория систем» или «системология».

Общая теория систем (ОТС) возникла на основе изучения некоторых биологических объектов и явлений и впервые была сформулирована Л.Берталанфи.

Со временем в структуре общей теории систем выделились два направления. Цель первого направления – развитие ОТС как некоторой философской концепции, включающей в себя такие понятия, как принцип системности, системный подход, системный анализ и т.д. В другом направлении общая теория систем представляет собой некоторый математический аппарат, претендующий на строгое описание закономерностей формирования и развития любых систем.

ОТС базируется на трех постулатах. Первый постулат утверждает, что функционирование систем любой природы может быть описано на основе рассмотрения формальных структурно-функциональных связей между отдельными элементами систем. Влияние материала, из которого состоят элементы систем, проявляется в формальных характеристиках системы (ее структуре, динамике и т.д.). Второй постулат состоит в том, что организация системы может быть определена на основе наблюдений, проведенных

извне посредством фиксирования состояний только тех элементов системы, которые непосредственно взаимодействуют с ее окружением. Третий постулат заключается в том, что организация системы полностью определяет ее функционирование и характер взаимодействия с окружающей средой. Эти постулаты дают возможность определить организацию системы, исходя из характеристик взаимодействия с внешней средой, и характеристики взаимодействия, исходя из организации системы.

Диалектическое требование изучать объект во всех его связях получило в общей теории систем свое дальнейшее развитие в форме ряда принципов: системности (целостное представление объекта); релятивности системы (любое множество предметов можно рассматривать как систему и как несистему); универсальности системы. Этот принцип направлен против абсолютизации отдельных систем и способов их образования, т.е. любое множество можно рассматривать как систему и как несистему в определенных аспектах и фиксированных условиях.

Теоретические исследования включают: анализ физической сущности процессов, явлений; формулирование гипотезы исследования; построение (разработка) физической модели; проведение математического исследования; анализ теоретических решений; формулирование выводов. Если не удастся выполнить математическое исследование, то формулируется рабочая гипотеза в словесной форме с привлечением графиков, таблиц и т.д. В технических науках необходимо стремиться к применению математической формализации выдвинутых гипотез и выводов.

В процессе теоретических исследований приходится непрерывно ставить и решать разнообразные по типам и сложности задачи в форме противоречий теоретических моделей, требующих разрешения.

В логико-психологическом аспекте задача – это несогласованные или противоречивые информационные процессы (системы), соотношение между которыми вызывает потребность в их преобразовании. В процессе решения задачи противоречия между указанными информационными процессами или системами устраняются.

Структурно любая задача включает условия и требования (рисунок 5.1). Условия – это определение информационной системы, из которой следует исходить при решении задачи. Требования – это цель, к которой нужно стремиться в результате решения. Условия и требования могут быть исходными, привлеченными и искомыми. Исходные условия даются в первоначальной формулировке задачи (исходные данные). Если их оказывается недостаточно для решения задачи, то исследователь вынужден привлекать новые данные, называемые привлеченными. Искомые данные или искомые условия – это привлеченные условия, которые требуется отыскать в процессе решения задачи.

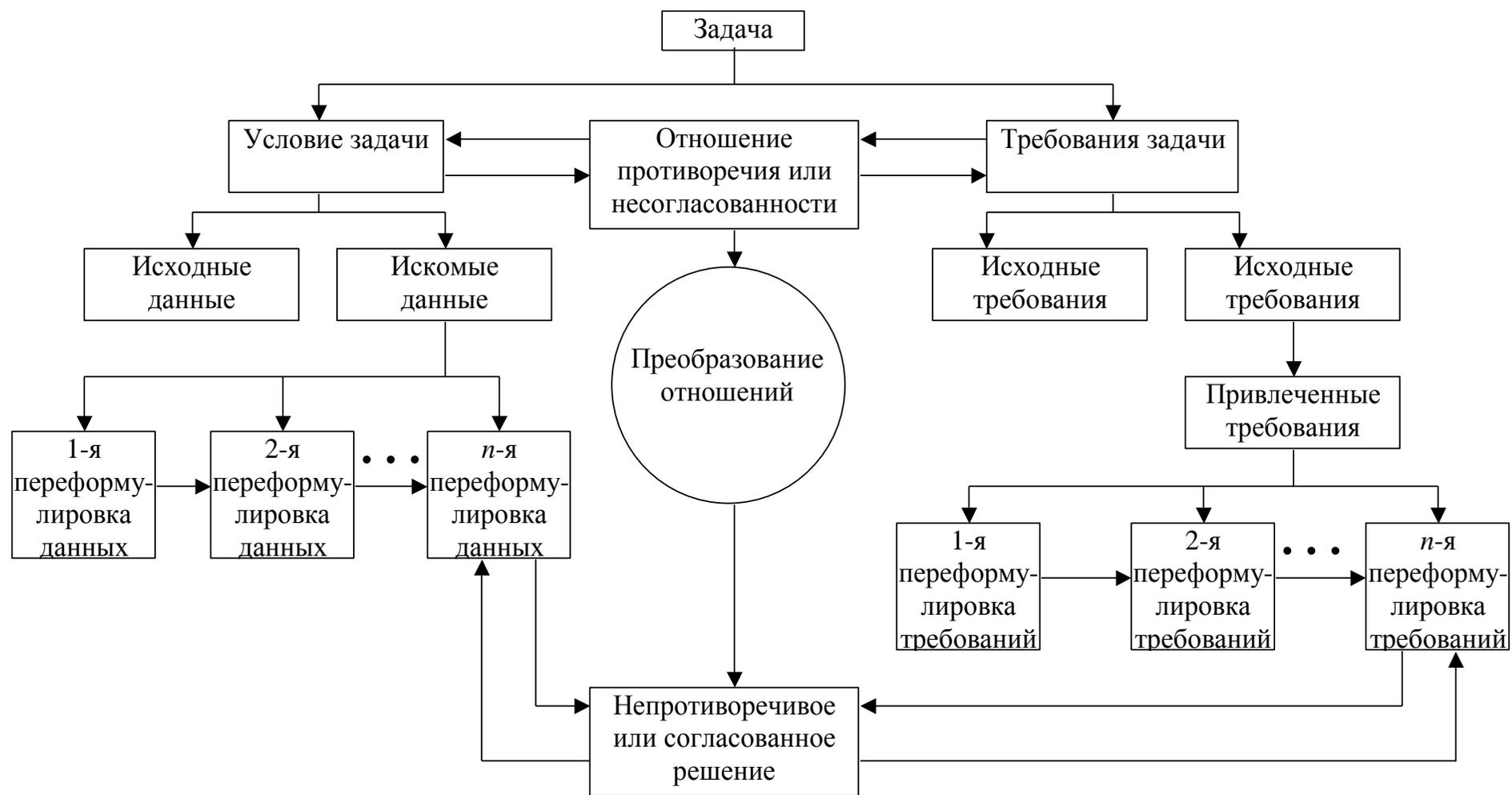


Рисунок 5.1 – Структурные компоненты решения задачи

Условия и требования задачи находятся в противоречии, они неоднократно сталкиваются, сопоставляются, сближаются между собой. Такое преобразование структурных компонентов задачи продолжается до тех пор, пока не будет решена сама задача.

Процесс проведения теоретических исследований состоит обычно из нескольких стадий. Оперативная стадия включает проверку возможности устранения технического противоречия, оценку возможных изменений в среде, окружающей объект, анализ возможности переноса решения задачи из других отраслей знания (ответить на вопрос: «Как решаются в других отраслях знаний задачи, подобные данной?»), применение «обратного» решения (ответить на вопрос: «Как решаются задачи, обратные данной, и нельзя ли использовать эти решения, взяв их со знаком минус?») или использования «прообразов» природы (ответить на вопрос: «Как решаются в природе более или менее сходные задачи?»). Вторая стадия исследования является синтетической, в процессе которой определяется влияние изменения одной части объекта на построение других его частей, определяются необходимые изменения других объектов, работающих совместно с данным, оценивается возможность применения измененного объекта по новому, и найденной технической идеи при решении других задач.

Выполнение названных предварительных стадий дает возможность приступить к стадии постановки задачи, в процессе которой определяется конечная цель решения задачи, проверяется возможность достижения той же цели решения задачи «обходными» (может быть, более простыми) средствами, выбирается наиболее эффективный путь решения задачи и определяются требуемые количественные показатели. В связи с этим при необходимости уточняются требования применительно к конкретным условиям практической реализации полученного решения задачи.

Аналитическая стадия включает определение идеального конечного результата (ответить на вопрос: «Что желательно получить в самом идеальном случае?»), выявляются помехи, мешающие получению идеального результата, и их причины, определяются условия, обеспечивающие получение идеального результата с целью найти, при каких условиях исчезнет «помеха».

Постановка задачи является наиболее трудной частью ее решения. Умение увидеть скрытое основное отношение задачи в самом начале решения, а, следовательно, умение поставить задачу, выделить ее из огромной массы окружающих, приводящих обстоятельств и, наконец, добраться до ее завуалированной сущности – залог успеха в достижении поставленной цели. Чем быстрее задача ставится, тем быстрее она приходит в состояние предрешения. Все это указывает на то, что четкая формулировка основного отношения задачи – важнейший этап ее решения. Следует при этом иметь в виду, что преобразование в начале расплывчатой формулировки задачи в четкую, определенную (переформулировка) часто облегчает решение задач.

Решение теоретических задач должно носить творческий характер. Творческие решения часто не укладываются в заранее намеченные планы. Иногда оригинальные решения появляются «внезапно», после, казалось бы, длительных и бесплодных попыток. Часто удачные решения возникают у специалистов смежных областей знания, на которых не давит груз известных решений. Творческие решения представляют по существу разрыв привычных представлений и взгляд на явления с другой точки зрения. Следует особо подчеркнуть, что собственные творческие мысли (оригинальные решения) возникают тем чаще, чем больше сил, труда, времени затрачивается на постоянное обдумывание путей решения теоретической задачи, чем глубже научный работник увлечен исследовательской работой.

При разработке теорий наряду с вышеизложенными методами используются и другие. Немалую роль при построении любых теорий играют, например, логические методы и правила, носящие нормативный характер. К числу таких правил относятся

правила вывода, образования сложных понятий из простых, установление истинности сложных высказываний и т.д. Специальными принципами построения теорий служат также принципы формирования аксиоматических теорий, критерии непротиворечивости, полноты и независимости систем аксиом и гипотез и др.

Теоретические исследования играют большую роль в процессе познания объективной действительности, поскольку они позволяют глубоко проникнуть в сущность природных явлений, создавать постоянно развивающуюся научную картину мира. Теоретическое исследование является функцией мышления, которая состоит в том, чтобы открыть, проверить, частично освоить различные области природы, создать и развить мировоззрение.

В этом процессе познание природы раскрывается все более полно, но с каждой новой подтвержденной гипотезой выдвигает все больше проблем. Таким образом, с ростом объективных знаний одновременно увеличивается и область открытых вопросов, подлежащих решению, так как каждый найденный ответ лишь приближает к познанию абсолютной истины, но не может достигнуть ее.

5.2 Использование математических методов

Решение практических задач математическими методами последовательно осуществляется путем математической формулировки задачи (разработки математической модели), выбора метода проведения исследования полученной математической модели, анализа полученного математического результата.

Математическая формулировка задачи обычно представляется в виде чисел, геометрических образов, функций, систем уравнений и т.п. Описание объекта (явления) может быть представлено с помощью непрерывной или дискретной, детерминированной или стохастической и другой математической формой.

Первым этапом является постановка задачи, определение объекта и целей исследования, задание критериев (признаков) изучения объектов и управления ими. Неправильная или неполная постановка задачи может свести на нет результаты всех последующих этапов.

Весьма важным на этом этапе является установление границ области влияния изучаемого объекта. Границы области влияния объекта определяются областью значимого взаимодействия с внешними объектами. Данная область может быть определена на основе следующих признаков: границы области охватывают те элементы, воздействие которых на исследуемый объект не равно нулю; за этими границами действие исследуемого объекта на внешние объекты стремится к нулю. Учет области влияния объекта при математическом моделировании позволяет включить в эту модель все существенные факторы и рассматривать моделируемую систему как замкнутую, т.е., с известной степенью приближения, независимую от внешней среды. Последнее значительно упрощает математическое исследование.

Следующим этапом является выбор типа математической модели. Выбор типа математической модели является важнейшим моментом, определяющим направление всего исследования. Обычно последовательно строится несколько моделей. Сравнение результатов их исследования с реальностью позволяет установить наилучшую из них.

На этапе выбора типа математической модели при помощи анализа данных поискового эксперимента устанавливаются: линейность или нелинейность, динамичность или статичность, стационарность или нестационарность, а также степень детерминированности исследуемого объекта или процесса.

Установление общих характеристик объекта позволяет выбрать математический аппарат, на базе которого строится математическая модель. Выбор математического аппарата может быть осуществлен в соответствии со схемой, представленной на рисунке 5.2. Как видно из данной схемы, выбор математического аппарата не является однозначным и жестким.

Так, для детерминированных объектов может использоваться аппарат линейной и нелинейной алгебры, теории дифференциальных и интегральных уравнений, теории автоматического регулирования.

Адекватным математическим аппаратом для моделирования вероятностных объектов являются теория детерминированных и случайных автоматов с детерминированными и случайными средами, теория случайных процессов, теория марковских процессов, эвристическое программирование, методы теории информации, методы теории управления и оптимальные модели.

При описании квазидетерминированных (вероятностно-детерминированных) объектов может использоваться теория дифференциальных уравнений с коэффициентами, подчиняющимися определенным законам.

Цель и задачи, которые ставятся при математическом моделировании, играют немаловажную роль при выборе типа (класса) модели. Практические задачи требуют простого математического аппарата, а фундаментальные – более сложного, допускают прохождение иерархии математических моделей, начиная от чисто функциональных и кончая моделями, использующими твердо установленные закономерности и структурные параметры.

Не меньшее значение на выбор модели оказывает анализ информационного массива, полученного как результат аналитического обзора результатов исследований других авторов или поискового эксперимента. Деление массива на зависимые и независимые факторы, на входные и выходные переменные, предварительный поиск взаимосвязи между различными данными выборки позволяет определить адекватный математический аппарат.

Анализ информационного массива позволяет установить непрерывность или дискретность исследуемого показателя и объекта в целом.

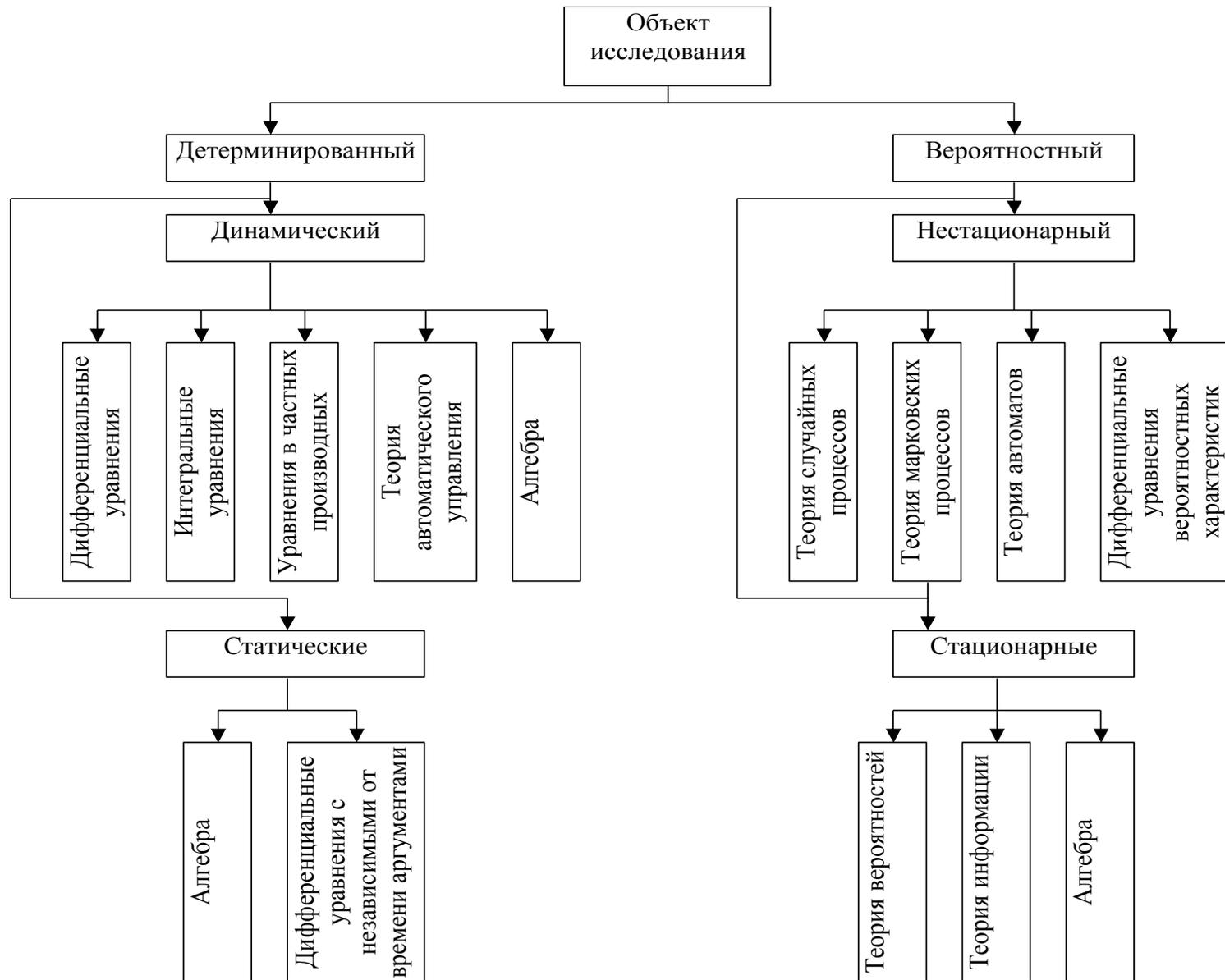


Рисунок 5.2 – Математический аппарат для построения математической модели

Установление непрерывности объекта позволяет использовать для его моделирования дифференциальные уравнения. В свою очередь, дискретность объекта предопределяет использование для математического моделирования аппарата теории автоматов.

Кроме вышеизложенного на установление типа (класса) математической модели может оказать существенное влияние необходимость определенного отображения гипотезы.

Учет целей и задач математического моделирования, характера гипотезы и анализа информационного массива позволяет конкретизировать модель, т.е. в выбранном типе (классе) моделей определить их вид. Вид модели связан с заданием областей определения исследуемых параметров объекта, т.е. значения, которые являются допустимыми, и установлением зависимостей между ними. Для количественных (числовых) параметров зависимости задаются в виде систем уравнений (алгебраических или дифференциальных), для качественных – используются табличные способы задания функций.

Если параметры описываются противоречивыми зависимостями, то определяются их весовые коэффициенты, выраженные в долях единицы, баллах. Тем самым противоречивые зависимости переводятся в вероятностные.

Для описания сложных объектов с большим количеством параметров возможно разбиение объекта на элементы (подсистемы), установление иерархии элементов и описание связей между ними на различных уровнях иерархии.

Особое место на этапе выбора вида математической модели занимает описание преобразования входных сигналов в выходные характеристики объекта.

Если на предыдущем этапе было установлено, что объект является статическим, то построение функциональной модели осуществляется при помощи алгебраических уравнений. При этом кроме простейших алгебраических зависимостей используются регрессионные модели и системы алгебраических уравнений.

Если заранее известен характер изменения исследуемого показателя, то число возможных структур алгебраических моделей резко сокращается и предпочтение отдается той структуре, которая выражает наиболее общую закономерность или общеизвестный закон. Если характер изменения исследуемого показателя заранее неизвестен, то ставится поисковый эксперимент. Предпочтение отдается той математической формуле, которая дает наилучшее совпадение с данными поискового эксперимента.

Результаты поискового эксперимента и априорный информационный массив позволяют установить схему взаимодействия объекта с внешней средой по соотношению входных и выходных величин. В принципе возможно установление четырех схем взаимодействия:

1) одномерно-одномерная схема (рисунок 5.3, *а*) – на объект воздействует только один фактор, а его поведение рассматривается по одному показателю (один выходной сигнал);

2) одномерно-многомерная схема (рисунок 5.3, *б*) – на объект воздействует один фактор, а его поведение оценивается по нескольким показателям;

3) многомерно-одномерная схема (рисунок 5.3, *в*) – на объект воздействует несколько факторов, а его поведение оценивается по одному показателю;

4) многомерно-многомерная схема (рисунок 5.3, *г*) – на объект воздействует множество факторов и его поведение оценивается по множеству показателей.

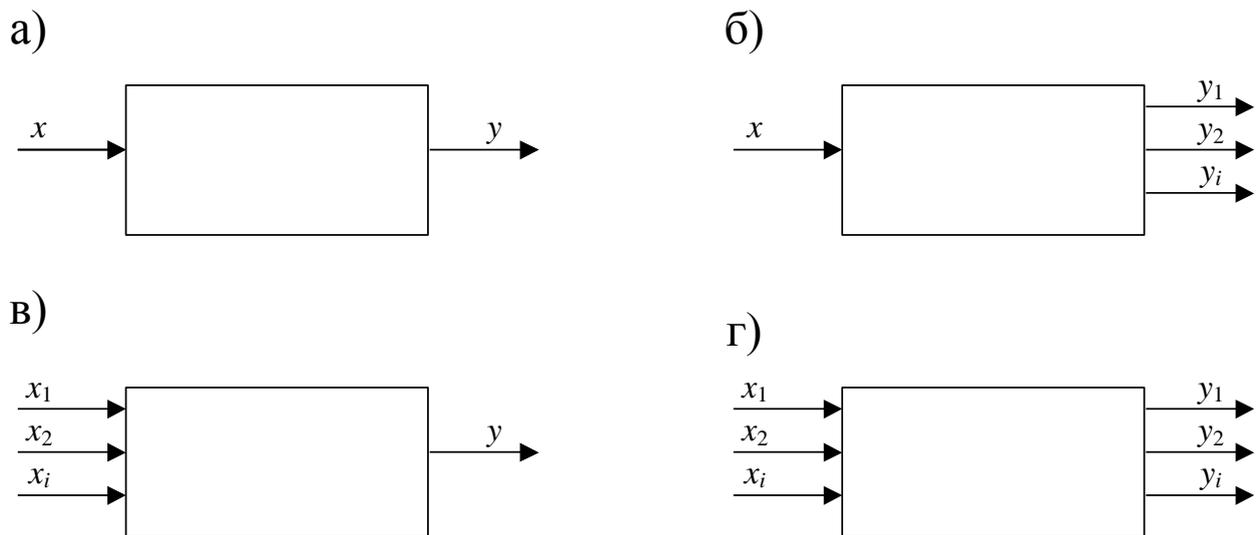


Рисунок 5.3 – Схемы воздействия объекта с внешней средой

При одномерно-одномерном взаимодействии статического стационарного детерминированного объекта с внешней средой постоянное входное воздействие связывается с постоянным выходным сигналом через постоянный коэффициент. Если же этот объект является нестационарным, то указанная связь описывается различными функциями $y = f(x)$. Чаще всего данная функция описывается полиномом.

В случае обнаружения многомерно-одномерной схемы статический стационарный детерминированный объект описывается следующей моделью:

при равнозначности внешних воздействий

$$y = a \sum_{i=1}^m x_i ;$$

при неравнозначности внешних воздействий

$$y = \sum_{i=1}^m a_i x_i ,$$

где a_i – постоянный коэффициент; m – число внешних воздействий (факторов).

Для статического нестационарного объекта (при той же схеме взаимодействия) часто используется модель в виде полного степенного полинома:

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^{m_1} \sum_{j=1}^{m_1} a_{ij} x_i x_j + \sum_{i=1}^{m_2} \sum_{j=1}^{m_2} \sum_{v=1}^{m_2} a_{ijv} x_i x_j x_v + \dots,$$

где m_1, m_2 – число парных и тройных сочетаний факторов ($m_1 = C_m^2; m_2 = C_m^3$).

При одномерно-многомерной схеме статический стационарный и нестационарный объект описывается аналогично одномерно-одномерной схеме взаимодействия статического стационарного объекта с внешней средой. При этом определяются отдельно математические модели входного воздействия с каждым выходным сигналом. Выходные сигналы считаются независимыми.

Многомерно-многомерное взаимодействие сводится к многомерно-одномерному, и математическая модель объекта принимается аналогичной изложенной выше. Для нестационарного одномерно-одномерного (многомерного) взаимодействия алгебраические функции могут представлять собой решение дифференциальных уравнений. При этом необходимо рассматривать производные математического ожидания по переменному фактору. Например, экспоненциальная зависимость может являться решением следующего дифференциального уравнения:

$$\frac{d\bar{y}}{dx} + a\bar{y} = a\bar{y}^m \quad (\bar{y} = 0 \text{ при } x = 0),$$

где \bar{y} – максимальное значение математического ожидания.

Выбор вида модели динамического объекта сводится к составлению дифференциальных уравнений. Модель динамического объекта может быть построена и в классе алгебраических функций. Однако такой подход является ограниченным, так как не позволяет в математическом описании учесть влияния входных воздействий на динамику выхода без существенной перестройки самих алгебраических функций (структуры и коэффициентов).

Поэтому по полноте модели отдается предпочтение математическим моделям, построенным в классе дифференциальных уравнений.

Если интересующие исследователя переменные являются только функциями времени, то для моделирования используются обыкновенные дифференциальные уравнения. Если же эти переменные являются также функциями пространственных координат, то для описания таких объектов недостаточно обыкновенных и следует пользоваться более сложными дифференциальными уравнениями в частных производных.

Методология моделирования динамических систем в классе дифференциальных уравнений существенно зависит от схемы взаимодействия объекта со средой и степени знания входа и выхода объекта.

Рассмотрим случай, когда вход и выход объекта известны.

При одномерно-одномерном и одномерно-многомерном взаимодействии детерминированного объекта со средой структура дифференциального уравнения определяется по виду выходной характеристики объекта для типового входного воздействия (например, ступенчатого).

Одной из наиболее простых выходных характеристик объекта является линейная (рисунок 5.4, а). Такое изменение выхода определяется решением дифференциального уравнения

$$\frac{dy}{dt} = kx, \quad y_0 = 0,$$

где $k > 0$ – коэффициент размерности и пропорциональности; y_0 – начальное значение выходного сигнала; t – время.

Если $y_0 \neq 0$, то выходная характеристика объекта соответствует рисунку 5.4, б. Однако дифференциальное уравнение остается неизменным.

Более сложный вид реакции объекта на ступенчатое входное воздействие (рисунок 5.4, в) может быть описан полным неоднородным дифференциальным уравнением первого порядка:

$$\frac{dy}{dt} + a_0 y = kx, \quad y(0) = y_0,$$

где a_0 – коэффициент дифференциального уравнения.

Реакция объекта, соответствующая рисунку 5.4, г, позволяет использовать в качестве математической модели объекта дифференциальное уравнение второго порядка:

$$\frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = kx, \quad y(0) = y_0.$$

Рассмотренные виды математических моделей соответствуют постоянному входному воздействию ($x = \text{const}$). Если входные воздействия являются некоторыми функциями времени, в частности алгебраическими, то в приведенных дифференциальных уравнениях изменяются правые части $x = f(t)$.

При многомерно-одномерном взаимодействии (в случае детерминированного объекта) динамические модели также ищут в классе дифференциальных уравнений. При этом допускается, что входные факторы являются независимыми по их действию на объект.

Все факторы приводятся к сумме коэффициентами чувствительности в правой части дифференциального уравнения. Дифференциальное уравнение подбирается по виду выходной характеристик объекта.

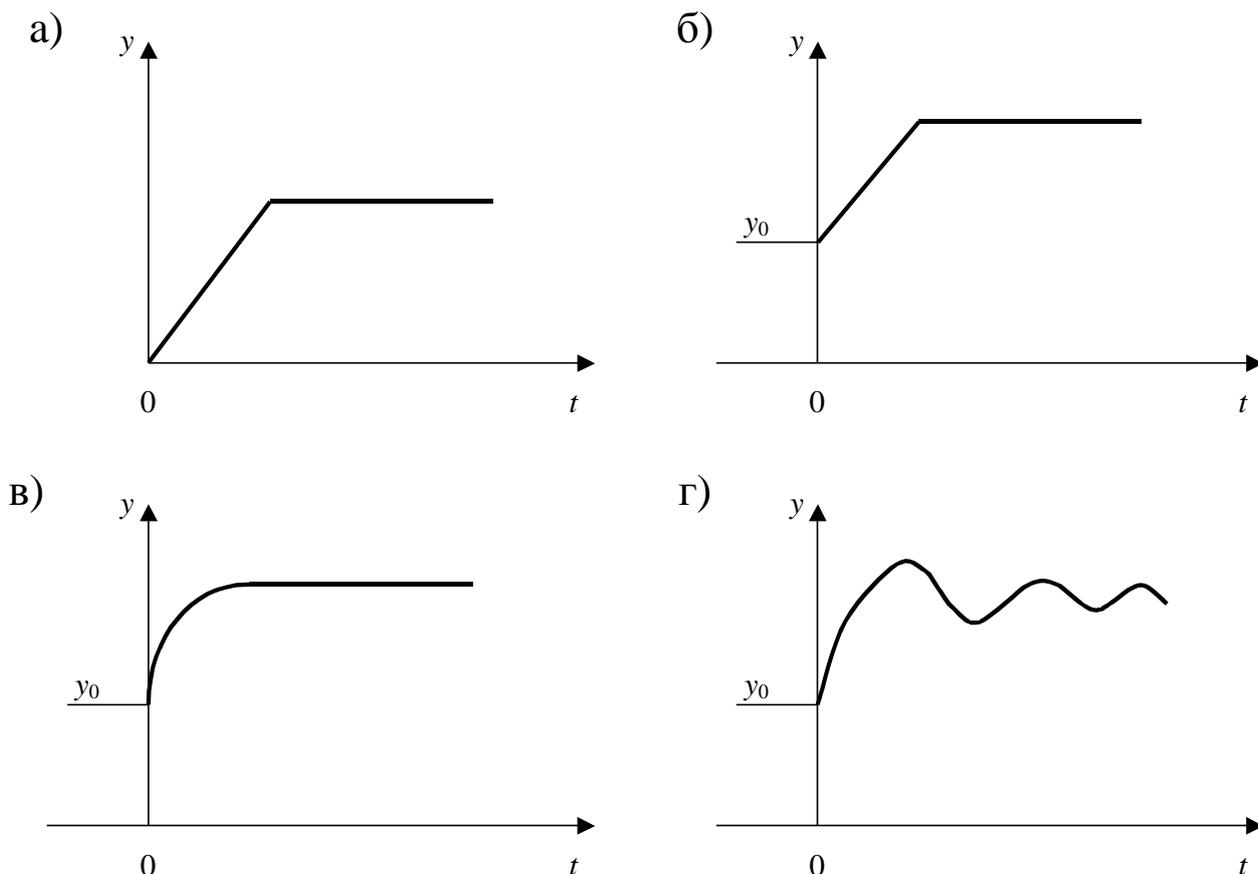


Рисунок 5.4 – Выходные характеристики детерминированного объекта при ступенчатом внешнем воздействии

Если допущение о независимости действия факторов неправомерно, то предварительно устанавливается влияние каждого из факторов на выходной показатель объекта и по виду выходных характеристик подбираются соответствующие дифференциальные уравнения. При одновременном действии факторов полагается, что выходная характеристика объекта представляет собой сумму решений независимых дифференциальных уравнений, соответствующих каждому фактору.

Для многомерно-многомерного взаимодействия построение динамических моделей может быть сведено к многомерно-одномерной схеме взаимодействия.

При отсутствии априорной информации о входе и выходе объекта дифференциальные уравнения, моделирующие динамику объекта, составляются на основе предположений или знаний о свойствах и структуре объекта.

Универсального метода составления дифференциальных уравнений нет, можно лишь использовать некоторые общие подходы к составлению уравнений первого порядка.

Геометрические или физические задачи обычно приводят к одному из следующих трех видов уравнений:

- дифференциальные уравнения в дифференциалах;
- дифференциальные уравнения в производных;
- простейшие интегральные уравнения с последующим преобразованием их в дифференциальные уравнения.

Рассмотрим, как составляются уравнения каждого их приведенных видов в отдельности.

1. **Уравнения в дифференциалах.** При составлении дифференциальных уравнений первого порядка удобно применять «метод дифференциалов». Сущность его заключается в том, что из условия задач составляются приближенные соотношения между дифференциалами. Для этого малые промежутки величин заменяются их дифференциалами, неравномерно протекающие физические процессы в течение малого промежутка времени dt рассматриваются как равномерные.

Эти допущения и замены не отражаются на окончательных результатах вследствие того, что замена приращений дифференциалами сводится к отбрасыванию бесконечно малых высших порядков малости. Так как отношение дифференциалов функции и аргумента является пределом отношения их приращений, то по мере того, как приращения стремятся к нулю, принятые допущения выполняются с большей точностью. Получающиеся при этом дифференциальные уравнения оказываются точными, если они однородны и линейны относительно дифференциалов.

Рассмотрим геометрический пример на применение метода дифференциалов.

Пусть перед исследователем стоит задача определения поверхности вращения, по которой нужно отшлифовать зеркало рефлектора оптического излучателя, чтобы выходящие из одной точки световые лучи после отражения в зеркале пересекались в другой точке.

По сути, данная задача сводится к нахождению уравнения сечения искомой поверхности меридианной плоскостью, проходящей через точку F_2 , в которой помещается источник света, и точку F_1 , в которой пересекаются отраженные лучи (рисунок 5.5).

Пусть MQ – малая дуга этого сечения. Будем считать ее прямолинейным отрезком. Опишем из точек F_1 и F_2 , как из центров, дуги MN и MP окружностей радиусами $F_1M = r_1$ и $F_2M = r_2$. Эти дуги также будем считать прямолинейными отрезками.

Треугольники MQN и MQP – прямоугольные ($\angle MNQ$ и $\angle MPQ$ – прямые) с общей гипотенузой MQ .

Пользуясь известным законом оптики о равенстве углов падения и отражения, а также свойством равенства вертикальных углов, находим, что $\angle MQN = \angle MQP$ и $\Delta MQN = \Delta MQP$. Отсюда следует, что $QN = QP$. Так как $QN = \Delta r_1$, а $QP = \Delta r_2$, то, заменяя приращения радиус-векторов r_1 и r_2 их дифференциалами, имеем

$$dr_1 + dr_2 = 0.$$

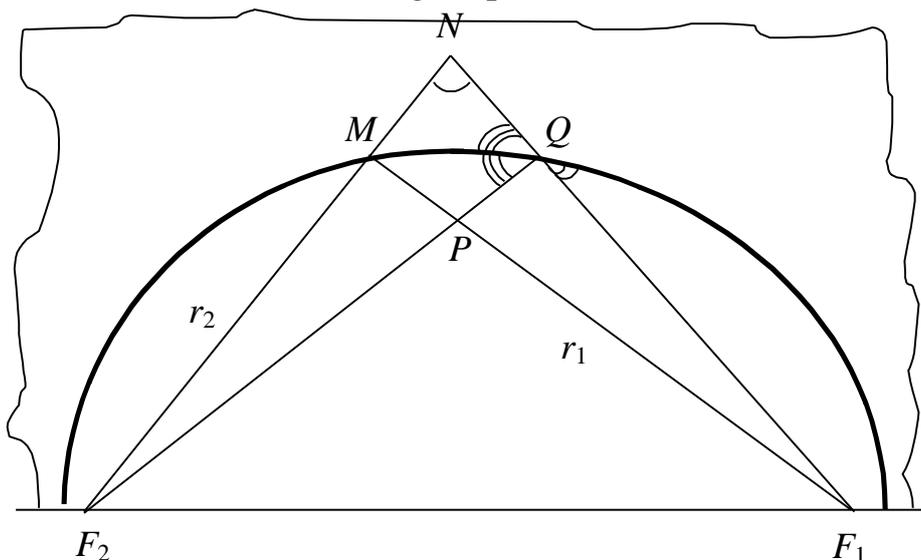


Рисунок 5.5 – Расчетная схема

Дифференциальное уравнение составлено. Оно легко интегрируется. Для этого перепишем его следующим образом:

$$d(r_1 + r_2) = 0,$$

откуда находим общий интеграл

$$r_1 + r_2 = C.$$

Итак, сечение искомой поверхности меридианной плоскостью является эллипсом. Следовательно, зеркало рефлектора надо отшлифовать по поверхности эллипсоида вращения.

2. Уравнения в производных. Для составления дифференциальных уравнений используется видоизмененный метод дифференциалов, который именуют методом производных.

Сущность метода производных заключается в том, что из условия задачи составляются приближенные соотношения между скоростями изменения функции и аргумента. При этом часто используется геометрическая интерпретация скорости – угловой коэффициент касательной. Отсутствие бесконечно малых в методе производных кажущееся, поскольку скорость изменения исследуемой величины сама появилась из рассмотрения бесконечно малых элементов.

При исследовании роста числа публикаций в науке исходят из допущения, что скорость роста $\frac{dy}{dt}$ пропорциональна достигнутому уровню y числа публикаций. Это тождественно утверждению, что относительная скорость роста $\frac{1}{y} \frac{dy}{dt} = const$.

Указанное допущение позволяет составить дифференциальное уравнение в форме

$$\frac{dy}{dt} = ky,$$

или

$$\frac{1}{y} \frac{dy}{dt} = k \quad (k > 0),$$

где $k = const$, характеризующая в среднем отклики на публикации в той или иной области знания.

Решение этого дифференциального уравнения имеет вид

$$y = ae^{kt},$$

где a – постоянная, характеризующая некоторое начальное число публикаций.

3. Простейшие интегральные уравнения. При рассмотрении работы сил, объемов тел, площадей криволинейных поверхностей их можно описать при помощи определенного интеграла или интегральных формул. В случае если при таком описании неизвестные функции попадают под знак интеграла, то получаемая формальная запись называется интегральным уравнением.

Последующее дифференцирование интегрального уравнения преобразует его в дифференциальное.

Для иллюстрации метода построения интегральных уравнений с последующим преобразованием их в дифференциальные рассмотрим следующую задачу.

Пусть нужно найти закон прямолинейного движения материальной тела массы m , если известно, что работа действующей на тело силы пропорциональна времени t , прошедшему от начала движения. Начальный путь и начальная скорость равны соответственно s_0 и v_0 .

Из курса механики известно, что в случае прямолинейного перемещения точки, когда направления силы и скорости совпадают, работа $A = \int_{s_0}^s F(u) du$, где $F(u)$ – действующая на точку сила.

По условию задачи

$$A = kt.$$

Сравнивая оба выражения для A , находим

$$\int_{s_0}^s F(u) du = kt.$$

Дифференцированием по s получаем

$$F(s) = k \frac{dt}{ds},$$

а так как $\frac{ds}{dt} = v$ – скорость движения и

$$\frac{dt}{ds} = \frac{1}{ds/dt} = \frac{1}{v},$$

то

$$F(s) = \frac{k}{v}.$$

С другой стороны, из второго закона Ньютона следует, что

$$F(s) = m \frac{dv}{dt}.$$

Сравнивая оба выражения для $F(s)$, составляем дифференциальное уравнение

$$m \frac{dv}{dt} = \frac{k}{v}.$$

Общее решение этого уравнения представляется в виде

$$\frac{mv^2}{2} = kt + C_1.$$

При начальном условии $v = v_0$ при $t = 0$ находим, что

$$C_1 = mv_0^2 / 2.$$

Следовательно,

$$v = \sqrt{\frac{2k}{m}t + v_0^2}.$$

Заменяя v на $\frac{ds}{dt}$ и интегрируя, находим

$$s = \frac{m}{3k} \left(\frac{2k}{m}t + v_0^2 \right)^{3/2} + C_2.$$

При $t = 0$ $s = s_0$, следовательно,

$$C_2 = s_0 - \frac{mv_0^3}{3k}.$$

Таким образом, закон движения материальной точки принимает вид

$$s = \frac{m}{3k} \left(\frac{2k}{m}t + v_0^2 \right)^{3/2} + s_0 - \frac{mv_0^3}{3k}.$$

При составлении дифференциальных уравнений регулируемых объектов необходимо, прежде всего, определить условия получения равновесного режима работы объекта, т.е. уравнение статического равновесия.

Во многих случаях уравнение статического равновесия оказывается общим для различных объектов исследования. Например, при поступательном движении исследуемый объект (управляемый ракетный снаряд) будет находиться в состоянии статического

равновесия (движение будет равномерным) только в том случае, когда движущие силы F_D равняются силам сопротивления F_C . Уравнение статического равновесия принимает вид

$$F_D - F_C = 0.$$

В том случае, когда объект исследования (коленчатый вал двигателя внутреннего сгорания, роторы электродвигателя и генератора, турбины и т.д.) совершает вращательное движение, то условием статического равновесия является равенство крутящего момента M_D моменту сопротивления M_C , т.е.

$$M_D - M_C = 0.$$

При выполнении этого условия вращение вала будет равномерным.

Для регулируемого резервуара, в котором необходимо поддержать постоянный уровень, условием статического равновесия является уравнение

$$V_{\text{пр}} - V_{\text{расх}} = 0,$$

где $V_{\text{пр}}$ – приход жидкости в резервуар; $V_{\text{расх}}$ – расход жидкости.

Аналогично выглядит уравнение статического равновесия регулятора ресивера с газом или паром

$$G_{\text{пр}} - G_{\text{расх}} = 0,$$

где $G_{\text{пр}}$, $G_{\text{расх}}$ – массы приходящего и уходящего из ресивера газа.

Если регулируемым объектом является объем, в котором должна поддерживаться постоянная температура, то условие статического равновесия получает вид

$$Q_{\text{пр}} - Q_{\text{расх}} = 0,$$

где $Q_{\text{пр}}$ – количество теплоты, поступающей в объем за единицу времени; $Q_{\text{расх}}$ – количество теплоты, уходящей из объема в единицу времени.

Переходный процесс в исследуемом объекте может появиться только в том случае, если будет нарушено статическое равновесие. Появление приращения одного из членов уравнения статического равновесия неминуемо повлечет приращения и второго члена уравнения, причем такие приращения, как правило, не равны между собой. В результате условие статического равновесия нарушается и тогда при поступательном движении

$$F_D + \Delta F_D \neq F_C + \Delta F_C;$$

при вращательном движении

$$M_D + \Delta M_D \neq M_C + \Delta M_C;$$

при заполнении резервуара жидкостью

$$V_{\text{пр}} + \Delta V_{\text{пр}} \neq V_{\text{расх}} + \Delta V_{\text{расх}};$$

при заполнении ресивера газом

$$G_{\text{пр}} + \Delta G_{\text{пр}} \neq G_{\text{расх}} + \Delta G_{\text{расх}};$$

при нарушении теплового режима

$$Q_{\text{пр}} + \Delta Q_{\text{пр}} \neq Q_{\text{расх}} + \Delta Q_{\text{расх}}.$$

Полученные неравенства могут быть упрощены, если учесть в них условие статического равновесия:

$$\Delta F_D \neq \Delta F_C; \Delta M_D \neq \Delta M_C; \Delta V_{\text{пр}} \neq \Delta V_{\text{расх}}; \Delta G_{\text{пр}} \neq \Delta G_{\text{расх}}; \Delta Q_{\text{пр}} \neq \Delta Q_{\text{расх}}.$$

Таким образом, при нарушении условий статического равновесия в исследуемом объекте возникает избыток или недостаток движущих сил (или моментов), поступления жидкости, газа, избыток или недостаток теплоты. Этот избыток или недостаток вызывает изменение характера работы объекта.

Дальнейшее преобразование полученных неравенств осуществляется путем привлечения общеизвестных зависимостей, принципов, законов или анализа выходных характеристик объекта, полученных в поисковом эксперименте. При этом могут использоваться феноменологические законы (такие, как законы Гука, Фурье), полуэмпирические соотношения (например, дополнительное соотношение Ньютона в теории удара) и чисто эмпирические соотношения.

Нарушение статического равновесия исследуемого движущегося объекта, имеющего поступательное движение, приведет к ускоренному или замедленному движению. Так как движущийся объект обладает массой m и ускорением a , то на основании принципа Даламбера уравнение динамического равновесия (уравнение движения) может быть записано в виде

$$ma = \Delta F_{\text{Д}} - \Delta F_{\text{С}},$$

или, так как

$$a = \frac{dv}{dt},$$

где v – скорость движения; t – время, то

$$m \frac{dv}{dt} = \Delta F_{\text{Д}} - \Delta F_{\text{С}}.$$

Для объекта, имеющего вращательное движение, уравнение динамического равновесия записывается также на основании принципа Даламбера. Если J – приведенный к оси вала двигателя момент инерции вращающихся или возвратно-поступательно движущихся деталей и ω – угловая скорость вращения вала, то уравнение движения получает вид

$$J \frac{d\omega}{dt} = \Delta M_{\text{Д}} - \Delta M_{\text{С}}.$$

При нарушении статического равновесия работы резервуара в нем происходит накапливание (в алгебраическом смысле) жидкости и, как следствие, изменение ее уровня.

Если dV – изменение объема жидкости в резервуаре за время dt , то

$$dV = (\Delta V_{\text{пр}} - \Delta V_{\text{расх}})dt,$$

или

$$\frac{dV}{dt} = \Delta V_{\text{пр}} - \Delta V_{\text{расх}}.$$

При нарушении равновесия в работе ресивера количество газа в нем изменяется на величину dG за элементарный интервал времени dt , поэтому

$$dG = (\Delta G_{\text{пр}} - \Delta G_{\text{расх}})dt,$$

откуда

$$\frac{dG}{dt} = \Delta G_{\text{пр}} - \Delta G_{\text{расх}}.$$

Нарушение теплового режима некоторого объема приводит к изменению количества теплоты, сосредоточенной в выбранном объеме, на dQ за интервал времени dt , поэтому

$$dQ = (\Delta Q_{\text{пр}} - \Delta Q_{\text{расх}})dt,$$

или

$$\frac{dQ}{dt} = \Delta Q_{\text{пр}} - \Delta Q_{\text{расх}}.$$

Сравнение полученных дифференциальных уравнений для различных регулируемых объектов показывает их идентичность.

Их дальнейшее преобразование возможно на основе знания физики общих свойств газа, теплопередачи и т.п. Например, известно, что состояние идеальных газов описывается уравнением Клапейрона

$$pV = GTR,$$

где p – давление газа в ресивере; V – объем ресивера; G – количество газа, находящегося в ресивере; T – абсолютная температура газа; R – газовая постоянная.

Из уравнения состояния следует

$$G = \frac{pV}{RT}.$$

Тогда уравнение динамического равновесия ресивера преобразуется к виду

$$\frac{V}{RT} \frac{dp}{dt} = \Delta G_{np} - \Delta G_{pacx}.$$

Описание приращений в правой части рассматриваемых уравнений также требует предварительного знания их связи со свойствами объекта.

Введение в уравнения динамического равновесия зависимостей, описывающих приращения, часто приводит к повышению порядка дифференциальных уравнений. Однако при некотором упрощении порядок дифференциальных уравнений удается снизить. Таким упрощением в большинстве случаев является пренебрежение инерционностью объекта или линеаризация приращений. Для линеаризации последних часто используют разложение функции в ряд Маклорена.

Пусть, например, крутящий момент объекта (вала двигателя), имеющего вращательное движение, зависит от угловой скорости вращения ω и h органа управления топливного насоса двигателя, т.е.

$$M_D = f(\omega, h).$$

Для определения приращения крутящего момента двигателя в зависимости от приращения ω и h полученную функцию следует разложить в ряд Маклорена:

$$M_D + \Delta M_D = M_D + \frac{\partial M_D}{\partial \omega} d\omega + \frac{\partial^2 M_D}{\partial \omega^2} \frac{d\omega^2}{2!} + \dots + \frac{\partial M_D}{\partial h} dh + \frac{\partial^2 M_D}{\partial h^2} \frac{dh^2}{2!} + \dots.$$

Если в разложении заменить бесконечно малые величины $d\omega$ и dh величинами $\Delta\omega$ и Δh конечными, но достаточно малыми, то ряд будет иметь вид

$$M_D + \Delta M_D = M_D + \frac{\partial M_D}{\partial \omega} \Delta\omega + \frac{\partial^2 M_D}{\partial \omega^2} \frac{\Delta\omega^2}{2!} + \dots + \frac{\partial M_D}{\partial h} \Delta h + \frac{\partial^2 M_D}{\partial h^2} \frac{\Delta h^2}{2!} + \dots.$$

Производные в этом разложении должны подсчитываться в положении равновесия режима работы двигателя, при котором выполняется условие статического равновесия.

При малых конечных приращениях $\Delta\omega$ и Δh и неразрывности функции $M_D = f(\omega, h)$ можно отбросить без внесения существенной ошибки все члены ряда с $\Delta\omega$ и Δh в степенях выше первой, т.е. практически произвести замену действительной функции $M_D = f(\omega, h)$ ее касательной в точке равновесного режима (такая замена обычно называется линеаризацией). После линеаризации приращение крутящего момента принимает вид

$$\Delta M_D = \frac{\partial M_D}{\partial \omega} \Delta\omega + \frac{\partial M_D}{\partial h} \Delta h.$$

Пусть момент сопротивления является лишь функцией угловой скорости $M_C = f(\omega)$.

Тогда после разложения этой функции в ряд Маклорена и линеаризации получим

$$\Delta M_C = \frac{dM_C}{d\omega} \Delta\omega.$$

После подстановки преобразованных приращений в уравнение динамического равновесия объекта получаем

$$J \frac{\partial \omega}{\partial t} + \left(\frac{\partial M_C}{\partial \omega} - \frac{\partial M_D}{\partial \omega} \right) \Delta\omega = \frac{\partial M_D}{\partial h} \Delta h.$$

Любые дифференциальные уравнения – это модель целого класса явлений, т.е. совокупность явлений, характеризуемых одинаковыми процессами. При интегрировании уравнений получают большое количество решений, удовлетворяющих исходному дифференциальному уравнению. Чтобы получить из множества возможных решений одно, удовлетворяющее только рассматриваемому процессу, необходимо задать дополнительные условия дифференциальному уравнению. Они должны четко выделить изучаемое явление из всего класса явлений.

Условия, которые раскрывают все особенности данного уравнения, называются условиями однозначности и характеризуются следующими признаками: геометрией системы (форма и размер тела); физическими свойствами тела (теплопроводность, влагопроводность, упругость и т.д.); начальными условиями, т.е. состоянием системы в начальный момент; граничными условиями, т.е. условиями взаимодействия системы на границах с окружающей средой. Начальные и граничные условия называют краевыми.

Рассмотренные примеры выбора вида модели объекта имеют отношение лишь к таким объектам, которые могут рассматриваться как детерминированные.

Рассмотрим способы выбора вида математических моделей для вероятностных объектов. Как и ранее, для статических объектов под стационарностью входа будем понимать постоянное его значение. Если входной сигнал принимает несколько значений, то его будем считать нестационарным.

Пусть имеется одномерно-одномерная схема взаимодействия объекта с внешней средой. Если воздействие на входе объекта постоянно во времени, то в качестве математической модели статического вероятностного объекта может выступать некоторый закон распределения выходной величины. Если входное воздействие может принимать различные значения и каждому значению соответствует ряд значений выходной величины объекта, то в качестве модели вероятностного объекта принимается набор законов распределения выходной величины для всех значений входного воздействия.

При моделировании вероятностных объектов помимо законов распределения входных и выходных величин существенна связь между ними. Поэтому в состав модели включают коэффициенты взаимной корреляции и функции

$$H_m = f(x); \quad R = f(x); \quad y_{\text{cp}} = f(x); \quad \sigma = f(x),$$

где x – входное воздействие; H_m – максимальная энтропия выходных характеристик; R – относительная организация выходных характеристик; y_{cp} – среднее значение выходной величины; σ – среднеквадратическое отклонение выходных величин.

Максимальная энтропия выходных характеристик оценивается по формуле

$$H_m = \log_2 n,$$

где n – число состояний объекта.

Для оценки числа состояний объекта используется формула

$$n = \frac{y_{\text{max}} - y_{\text{min}}}{\Delta y},$$

где y_{max} , y_{min} – максимальное и минимальное значения выходной величины; Δy – точность измерения выходных величин.

Относительная организация выходных характеристик оценивается по формуле Ферстера:

$$R = 1 - \frac{H}{H_m},$$

где

$$H = - \sum_{i=1}^N \frac{m_i}{N} \log_2 \frac{m_i}{N};$$

здесь m_i – число появлений y_i значения выходной характеристики; N – полное число наблюдений выходных характеристик.

При многомерно-одномерной схеме взаимодействия статического вероятностного объекта с внешней средой задача математического моделирования сводится к одномерно-одномерной схеме для каждого сочетания постоянных входных воздействий.

Нестационарный случай отличается тем, что каждое входное воздействие может принимать несколько значений. При этом для каждого конкретного сочетания задача анализа

связи между входами и выходом может решаться аналогично задаче для многомерно-одномерной стационарной схемы.

Оценка степени связи выхода с входами проводится путем сопоставления статических параметров и вычисления коэффициентов взаимной корреляции.

Моделирование объекта при одномерно-многомерной и многомерно-многомерной схемах взаимодействия производится аналогично вышеизложенному.

Рассмотрим далее моделирование динамических режимов вероятностных объектов.

При одномерно-одномерной схеме взаимодействия объекта с внешней средой и многократном поступлении на его вход одной и той же функции времени $x(t)$ возможны два случая:

на выходе объекта наблюдается стационарный случайный процесс;

на выходе объекта наблюдается нестационарный случайный процесс.

В первом случае в качестве математической модели выходной величины объекта принимается закон распределения значений выходной величины, имеющий одни и те же параметры для всех срезов по времени. Срезы по времени принимаются с интервалом Δt . Модель дополняется зависимостями

$$H_m(i, \Delta t) = f(x); \quad R(i, \Delta t) = f(x); \quad i = 1, 2, 3, \dots$$

Эти зависимости могут быть представлены алгебраической функцией или дифференциальными уравнениями.

Во втором случае (нестационарный выход объекта) в качестве математической модели объекта принимаются функциональные связи

$$m_y(i, \Delta t) = f(x); \quad \sigma_y(i, \Delta t) = f(x); \quad i = 1, 2, 3, \dots$$

где $m_y(i, \Delta t)$ – математическое ожидание распределения выходных величин y по дискретным отрезкам времени с шагом Δt ; $\sigma_y(i, \Delta t)$ – среднеквадратическое отклонение.

Эти связи также описываются с помощью алгебраических или дифференциальных уравнений.

Если на вход объекта многократно поступают различные функции времени $x_i(t)$, то их рассматривают как случайные реализации случайного процесса. Для каждого сечения по времени $(i, \Delta t)$ определяются информационные и вероятностные характеристики $H_m^{y_i}$, R_{y_i} , m_{y_i} , σ_{y_i} .

В качестве математической модели объекта принимаются функциональные связи

$$\begin{aligned} H_m^{y_i}(i, \Delta t) &= f_{1i}[H_m^x(i, \Delta t)]; \\ R_{y_i}(i, \Delta t) &= f_{2i}[R_x(i, \Delta t)]; \\ m_{y_i}(i, \Delta t) &= f_{3i}[m_x(i, \Delta t)]; \\ \sigma_{y_i}(i, \Delta t) &= f_{4i}[\sigma_x(i, \Delta t)] \quad (i = 1, 2, 3, \dots, k). \end{aligned}$$

Иногда все массивы реализаций выходной величины рассматриваются как единый массив. В этом случае в качестве математической модели объекта принимаются функциональные связи между обобщенными параметрами выхода и параметрами входа.

При многомерно-одномерной схеме взаимодействия вероятностного объекта со средой в нестационарном режиме в качестве его математической модели принимаются функциональные зависимости:

$$\begin{aligned} H_m^y &= f_{11}(H_m^{x_1}); & m_y &= f_{31}(m_{x_1}); \\ & \dots & & \\ H_m^y &= f_{1m}(H_m^{x_n}); & m_y &= f_{3m}(m_{x_n}); \\ R_y &= f_{21}(R_{x_1}); & \sigma_y &= f_{41}(\sigma_{x_1}); \\ & \dots & & \end{aligned}$$

$$R_y = f_{2m}(R_{x_n}); \quad \sigma_y = f_{4m}(\sigma_{x_n}).$$

При одномерно-многомерной схеме взаимодействия объекта в нестационарном режиме в качестве математической модели объекта принимаются функциональные зависимости

$$\begin{aligned} H_m^{y_i} &= f_{1i}(H_m^x); \\ R_y^i &= f_{2i}(R_x); \\ M_{y_i} &= f_{3i}(m_x); \\ \sigma_{y_i} &= f_{4i}(\sigma_x) \quad (i = 1, 2, 3, \dots, n). \end{aligned}$$

При многомерно-многомерной схеме взаимодействия в нестационарном режиме устанавливаются функциональные зависимости информационных и вероятностных характеристик для каждого входа и выхода. При этом можно считать, что любой выход зависит от всех входов, и соответственно выбрать вид математических моделей.

Процесс выбора математической модели объекта заканчивается ее предварительным контролем. При этом осуществляются следующие виды контроля: размерностей; порядков; характера зависимостей; экстремальных ситуаций; граничных условий; математической замкнутости; физического смысла; устойчивости модели.

Контроль размерностей сводится к проверке выполнения правила, согласно которому приравняться и складываться могут только величины одинаковой размерности.

Контроль порядков направлен на упрощение модели. При этом определяются порядки складываемых величин и явно малозначительные слагаемые отбрасываются.

Контроль характера зависимостей сводится к проверке направления сходимости и скорости изменения одних величин при изменении других. Направления и скорость, вытекающие из математической модели, должны соответствовать физическому смыслу задачи.

Контроль экстремальных ситуаций сводится к проверке наглядного смысла решения при приближении параметров модели к нулю или бесконечности.

Контроль граничных условий состоит в том, что проверяется соответствие математической модели граничным условиям, вытекающим из смысла задачи. При этом проверяется, действительно ли граничные условия поставлены и учтены при построении искомой функции, и что эта функция на самом деле удовлетворяет таким условиям.

Контроль математической замкнутости сводится к проверке того, что математическая модель дает однозначное решение.

Контроль физического смысла сводится к проверке физического содержания промежуточных соотношений, используемых при построении математической модели.

Контроль устойчивости модели состоит в проверке того, что варьирование исходных данных в рамках имеющихся данных о реальном объекте не приведет к существенному изменению решения.

5.3 Аналитические методы

Следующим этапом решения практических задач математическими методами является выбор метода исследования модели. Выбор метода исследования математической модели непосредственно связан с такими понятиями, как внешнее и внутреннее правдоподобие исследования.

Под внешним правдоподобием исследования понимается ожидаемая степень адекватности математической модели реальному объекту по интересующим исследователя свойствам.

Под внутренним правдоподобием исследования понимается ожидаемая степень точности решения полученных уравнений, которые приняты за математическую модель объекта.

Если вид модели уже выбран, то внешнее правдоподобие модели считается фиксированным, и выбор метода исследования будет целиком определяться необходимой степенью внутреннего правдоподобия.

В подавляющем большинстве случаев при выборе метода исследования руководствуются принципом соответствия внешнего и внутреннего правдоподобия, аналогичным известному правилу приближенных вычислений: степень точности вычислений должна соответствовать степени точности исходных данных. Однако в зависимости от условий и задач исследования возможны отклонения от принципа. Перечислим некоторые из них:

1) если речь идет о разработке нового единого метода исследований, который предполагается применять к широкому, заранее не фиксированному, классу моделей, то нужно стремиться к максимальному внутреннему правдоподобию исследования независимо от уровня внешнего правдоподобия;

2) если осуществляется проверка внешнего правдоподобия модели, то внутреннее правдоподобие избранного метода проверки должно быть максимальным;

3) если модель настолько проста, что для нее легко получить точное решение, то искусственно понижать строгость решения бессмысленно.

В других случаях предпочтение отдается «принципу равного правдоподобия».

Выбор метода исследования тем эффективнее, чем больше имеется сведений о конечном решении задачи. Такие сведения могут быть получены путем прикидочных исследований модели и ее элементов.

В процессе прикидочных исследований осуществляется сравнение величин отдельных членов уравнений в изучаемом диапазоне изменения переменных и параметров задачи. Относительно малые слагаемые отбрасываются, нелинейные зависимости заменяются на линейные. Некоторые из компонентов модели аппроксимируются грубыми уравнениями. Все это позволяет быстро получить грубое решение задачи.

Знание, хотя бы самое грубое, качественных и количественных характеристик искомого решения помогает при выборе точности метода исследования. Иногда даже грубое решение оказывается достаточным. В качестве примера можно привести задачу о поиске экспериментального значения функции. Если точка экстремума является стационарной, то даже грубая ошибка в ее отыскании мало скажется на подсчете этого значения. Поэтому применение высокоточных методов поиска такого экстремума неэкономично. Громоздкие точные вычисления в этом случае создают лишь иллюзию точности. В случае применения грубой математической модели не следует применять громоздкие вычислительные методы.

Выбор метода исследования математической модели во многом предопределен ее видом.

Статические системы, представленные при помощи алгебраических уравнений, исследуются с помощью определителей, метода итераций, методов Крамера и Гаусса. В случае затруднений с аналитическими решениями используются приближенные методы: графический метод; метод хорд; метод касательных; метод итераций. В последнем случае, который требует контроля точности (числа значащих цифр) в зависимости от грубости вычислительного метода, целесообразно применение ЭВМ.

Исследование динамических режимов функционирования объекта, представленных в классе дифференциальных уравнений, также предопределяется классом, к которому относится решаемое уравнение.

Если в результате решения алгебраического уравнения получаются числа, то при решении дифференциальных уравнений получаются функции.

Для решения дифференциальных уравнений широко используются метод разделения переменных, метод подстановки, метод интегрирующего множителя, метод качественного

анализа и т.п. Для получения приближенных решений используют метод последовательных приближений, метод функциональных рядов, метод Рунге-Кутты, численные методы интегрирования и т.п.

Для подробного изучения моделей динамических систем, построенных в классе дифференциальных уравнений, используется качественная теория дифференциальных уравнений.

Качественная теория дифференциальных уравнений позволяет изучить все возможные решения – регулярные и особые.

В основе качественной теории лежит понятие *фазового портрета системы*. Построение фазового портрета иллюстрируется следующим примером. Пусть рассматривается система, описываемая следующим дифференциальным уравнением:

$$\frac{d^2 y}{dt^2} + \omega^2 y = 0;$$

при начальных условиях

$$\begin{aligned} y(0) &= y_0, \\ \frac{dy(0)}{dt} &= \dot{y}_0. \end{aligned}$$

Частное решение этого уравнения представляется в виде

$$y = y_0 \cos \omega t + \frac{\dot{y}_0}{\omega} \sin \omega t.$$

Принимая $\frac{dy}{dt}$ за новую искомую функцию и вводя обозначения

$$\begin{aligned} y &= z_1, \\ \frac{dy}{dt} &= z_2, \end{aligned}$$

преобразуем исходное дифференциальное уравнение в систему уравнений первого порядка:

$$\begin{aligned} \dot{z}_1 &= z_2; \\ \dot{z}_2 &= -\omega^2 z_1; \end{aligned}$$

при начальных условиях

$$z_1(0) = y_0, \quad z_2(0) = \dot{y}_0.$$

Частное решение имеет вид

$$\begin{aligned} z_1 &= y_0 \cos \omega t + \frac{\dot{y}_0}{\omega} \sin \omega t; \\ z_2 &= -y_0 \omega \sin \omega t + \dot{y}_0 \cos \omega t. \end{aligned}$$

Из полученной системы уравнений для z_1 и z_2 , исключая t , имеем

$$\frac{z_1^2}{\rho_0^2} + \frac{z_2^2}{(\omega \rho_0)^2} = 1;$$

где

$$\rho_0 = \sqrt{y_0^2 + \frac{\dot{y}_0^2}{\omega^2}} > 0.$$

Последнее уравнение описывает эллипс на плоскости $z_1 z_2$. Следовательно, частное решение для z_1 и z_2 выражается зависимостью от времени текущих координат точки $M(t)$, которая начинает свое движение в момент $t = 0$ от точки $M_0(y_0, \dot{y}_0)$ и движется по эллипсу (рисунок 5.6).

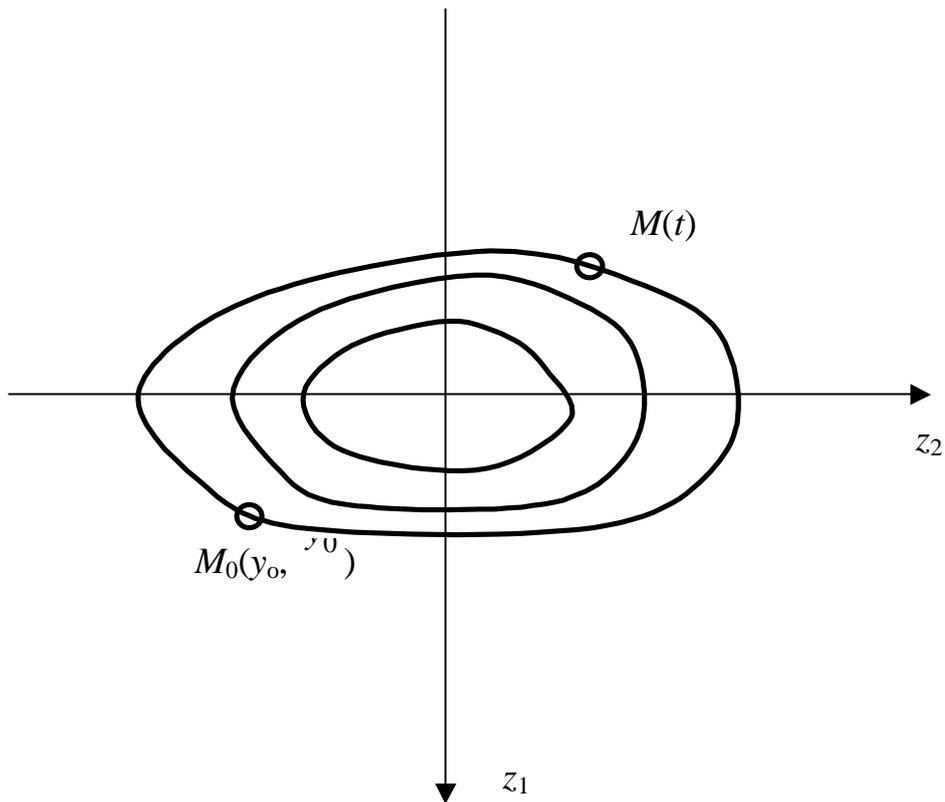


Рисунок 5.6 – Фазовый портрет системы

Точка $M(t)$ называется изображающей точкой. Траектория такой точки называется фазовой траекторией.

Изменяя начальные условия, можно получить семейство фазовых траекторий, которое называется фазовым портретом, а плоскость z_1z_2 , на которой расположено это семейство, – фазовой плоскостью.

Иногда ряд задач исследуется с помощью интегральных уравнений, содержащих искомую функцию $\varphi(s)$ под знаком интеграла:

$$h(x)\varphi(x) - \lambda \int_a^s k(x, s)\varphi(s)ds = f(x),$$

где $h(x)$, $\varphi(x)$ – неизвестные функции x ; λ – постоянный параметр, который называют собственным числом; $k(x, s)$ – заданная функция, которую называют ядром интегрального уравнения.

Общего метода решения интегральных уравнений даже линейного типа $h(x) = 0$, $\varphi(x) = 0$ не существует. Интегральное уравнение является решением дифференциального. Например, решением дифференциального уравнения первой степени

$$p = E_y s_y + \eta \frac{ds}{dt}$$

является интегральное уравнение

$$s = e^{-E_y t / \eta} \left(s_y + \frac{1}{\eta} \int_0^t p e^{-E_y t / \eta} dt \right).$$

Если $p = p_0 = \text{const}$, то имеем

$$s = \frac{p_0}{E_y} \left(1 - e^{-E_y t / \eta} \right).$$

Это позволяет сводить решение дифференциальных уравнений к решению интегральных и наоборот.

Многие задачи исследуются с помощью вариационного исчисления. Чтобы сформулировать задачу вариационного исчисления, вводят понятия **функционала**. Имеем плоскую кривую $y = f(x)$ с областью определения $x_0 \leq x \leq x_1$ (рисунок 5.7). Нетрудно видеть, что длина кривой s_1 , площадь p криволинейной трапеции, объем тела вращения v зависят от вида заданной кривой $y = f(x)$

$$s_1 = \int_{x_0}^{x_1} \sqrt{1 + [y'(x)]^2} dx; \quad p = \int_{x_0}^{x_1} y(x) dx; \quad v = \pi \int_{x_0}^{x_1} [y(x)]^2 dx.$$

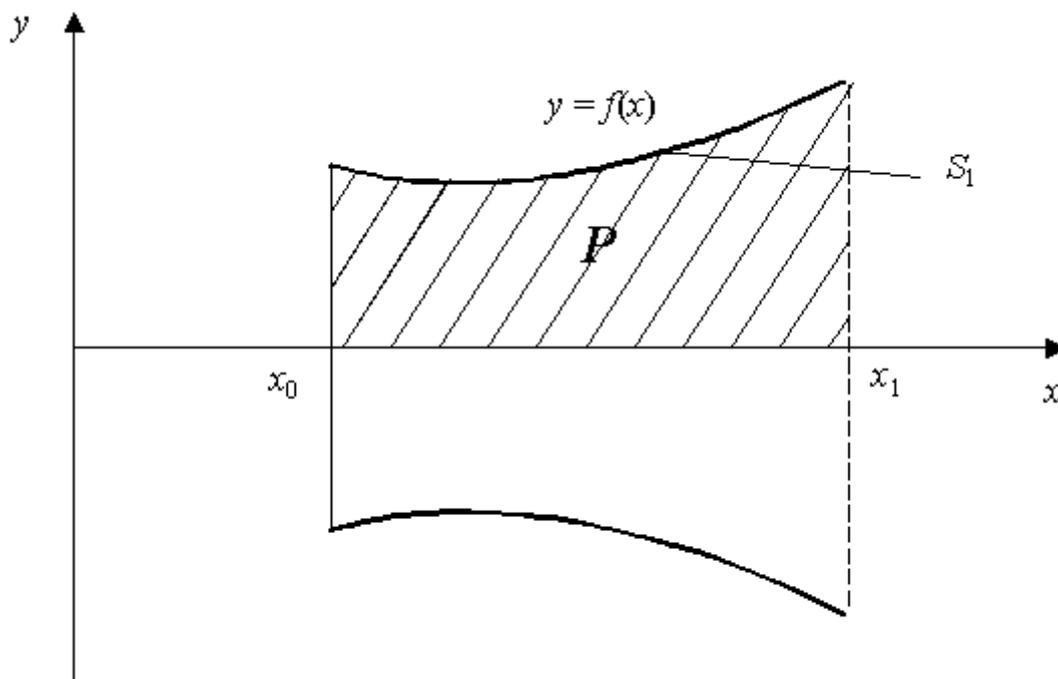


Рисунок 5.7 – Схема к понятию функционала

Таким образом, функция $y = f(x)$ однозначно определяет величины s_1 , p , v , т.е. она играет роль своеобразного «аргумента».

В этом случае величины s_1 , p , v называют функционалами относительно функции $y = f(x)$.

Суть задачи вариационного исчисления состоит в том, что если задан функционал $F(y')$ в области $x_0 \leq x \leq x_1$, то требуется найти такую функцию $y = f(x)$ в заданной области определения функционала $F(y')$, при которой этот функционал принимает минимальное или максимальное значение.

При исследовании процессов методами вариационного исчисления находят такие закономерности, при которых их развитие энергетически наиболее экономно. Очень часто они описываются экспоненциальными функциями, удовлетворяющими принципам вариационного исчисления.

При теоретических исследованиях широко используется теория функций комплексной переменной. В основе этой теории лежит положение о конформном преобразовании, согласно которому две пересекающиеся кривые z_1z_2 и z_1z_3 из области z всегда можно перенести в область ω соответственно кривым $\omega_1\omega_2$ и $\omega_2\omega_3$, сохраняя

равенство углов между кривыми и в каждой паре. Это позволяет изменить координаты таким образом, чтобы упростить громоздкие математические преобразования.

Рассмотренные аналитические методы, как правило, позволяют успешно решать лишь относительно простые задачи. В то же время все чаще возникает необходимость использования сложных дифференциальных уравнений или их систем со сложными начальными и граничными условиями (часто нелинейными). Их решение весьма сложно, в этих случаях прибегают к тем или иным приближенным вычислениям с помощью численных методов.

Идея численных методов (методы конечных разностей или сеток) заключается в следующем:

1. В плоской области G , в которой разыскивается решение, строится сеточная область G_h , состоящая из одинаковых ячеек (рисунок 5.8) и приближающаяся к области G .
2. Заданное дифференциальное уравнение заменяется в узлах построенной сетки соответствующим конечно-разностным уравнением.
3. На основании граничных условий устанавливаются значения искомого решения в граничных узлах области G_h .

Решив полученную систему конечно-разностных уравнений (для чего необходимо решить алгебраическую систему с большим числом неизвестных), найдем значения искомой функции в узлах сетки, т.е. будем иметь численное решение поставленной задачи. Выбор сеточной области производится в зависимости от конкретной задачи, но во всех случаях контур сеточной области G_h следует выбирать так, чтобы он возможно лучше аппроксимировал контур заданной области G . Сеточная область может состоять из квадратных, прямоугольных, треугольных и других клеток.

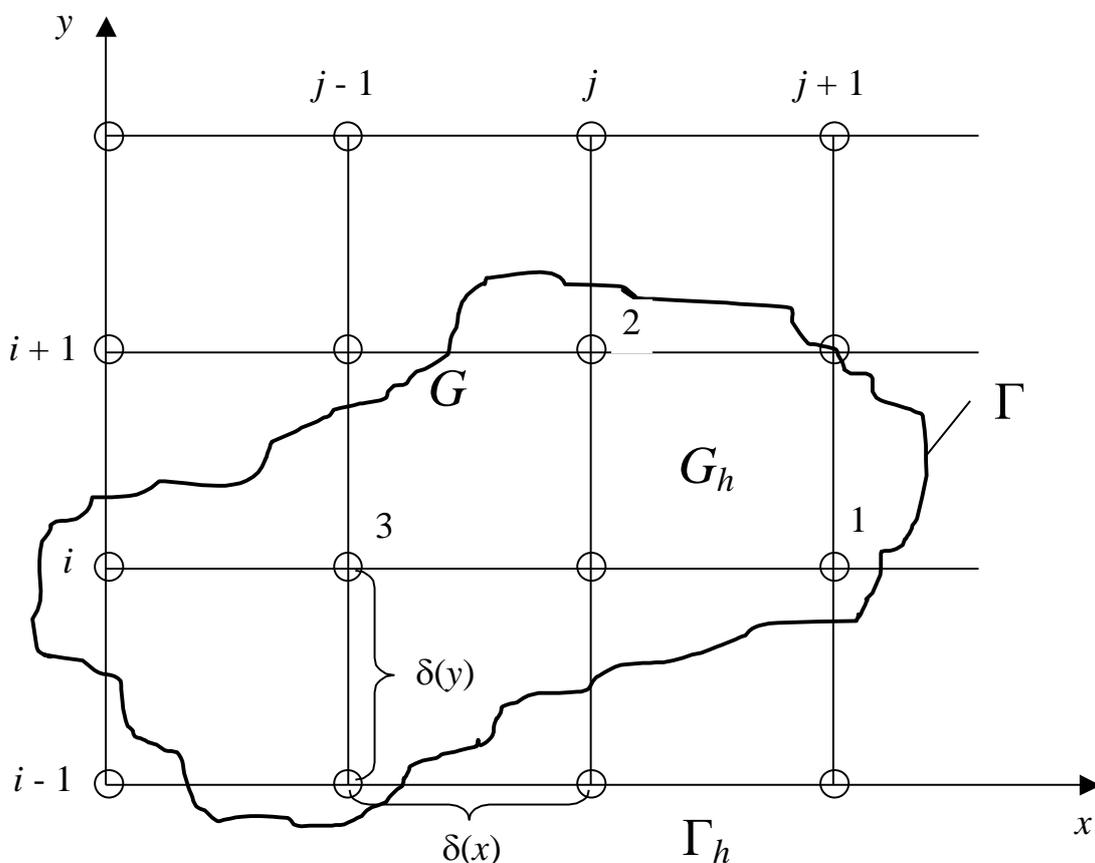


Рисунок 5.8 – Сеточная область G_h с контуром Γ для плоской области G , в которой производится решение двумерного дифференциального уравнения

В качестве примера использования численных методов может служить решение задачи виброформования бетонной смеси. Бетонная смесь при вибрации моделируется в виде сплошной упруговязкой среды, а математическая модель процесса виброформования представлена в виде

$$\rho \frac{dv_x}{dt} = E \frac{\partial^2 l}{\partial x^2} + \eta \left(\frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_y}{\partial y^2} \right),$$

где ρ – плотность бетонной смеси; t – время; x, y – декартовы координаты; v_x – составляющая скорости бетонной смеси вдоль направления распространения колебаний; E – модуль упругости; l – смещение элемента бетонной смеси; η – коэффициент динамической вязкости.

Первый член в правой части уравнения, описывающего процесс виброформования, представляет собой упругую составляющую среды, а второй – вязкую.

Решение полученного уравнения аналитическими методами является чрезвычайно сложным и обычно в литературе не приводится. Для практического использования оно может быть решено с использованием ЭВМ, для чего его необходимо представить в конечно-разностном (безразмерном) виде. Для этого введем следующие обозначения:

$$u = l/A\omega, \quad W = v_x/A\omega, \quad dx = dy = cdt,$$

где c – скорость распространения колебаний в бетонной смеси; $A\omega$ – амплитуда колебаний.

Тогда математическая модель процесса виброформования бетонной массы выразится уравнением

$$\frac{\partial W}{\partial t} = \frac{c^2}{\omega} \frac{\partial^2 u}{\partial x^2} + \nu \left(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} \right).$$

При численном интегрировании на ЭВМ это уравнение представляют в разностном виде, для чего производные заменяются конечно-разностными отношениями. Выбрав шаг δx по оси x и δy по оси y , построим сетку $x_i = x_0 + i\delta x$; $y_j = y_0 + j\delta y$ ($i, j = 0, \pm 1, \pm 2, \dots$) (рисунок 4.3). Тогда для первой производной функции $f(x)$ получим следующие варианты записи:

$$\frac{\partial f}{\partial x} \approx \frac{f_{i+1} - f_i}{\delta x}; \quad \frac{\partial f}{\partial x} \approx \frac{f_i - f_{i-1}}{\delta x}; \quad \frac{\partial f}{\partial x} \approx \frac{f_{i+1} - f_{i-1}}{2\delta x};$$

для второй производной

$$\frac{\partial^2 f}{\partial x^2} \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{\delta x^2},$$

где f_{i+1}, f_i, f_{i-1} – значения функции f в узлах интегрирования соответственно.

Значение функции в данном нулевом узле через шаг интегрирования по времени δt обозначается через $W_0^{\delta t}$, где W_0 – значение функции в нулевом узле в данный момент времени. С учетом принятых обозначений модель виброформования бетонной массы примет вид

$$W_0^{+\delta t} = W_0 + \frac{1}{\delta t} \left[\frac{\nu\omega}{c^2} (\Delta_x^2 W + \Delta_y^2 W) + \Delta_x^2 u \right],$$

где

$$\begin{aligned} \frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} &\approx \frac{\Delta_x^2 W}{\delta x^2} + \frac{\Delta_y^2 W}{\delta y^2}; & \Delta_x^2 u &\approx \partial^2 u; \\ \frac{\Delta_x^2 W}{\delta x^2} &\approx \frac{W_{i,j+1} - 2W_{i,j} + W_{i,j-1}}{\delta x^2}; \\ \frac{\Delta_y^2 W}{\delta y^2} &\approx \frac{W_{i+1,j} - 2W_{i,j} + W_{i-1,j}}{\delta y^2}; \end{aligned}$$

$$\frac{\Delta_x^2 u}{\delta x^2} \approx \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\delta x^2},$$

$\delta t + \delta t\omega$ – безразмерное время;

$$\frac{\partial W}{\partial t} = \frac{W_0^{+\delta t} - W_0}{\delta t}.$$

Обозначим через $\delta W = W_0^{+\delta t} - W_0$ приращение безразмерной скорости за время δt и, разделив левую и правую части уравнения на δt , получим

$$a = \frac{\delta W}{\delta t} = \frac{1}{\delta t^2} \left[\bar{k} (\Delta_x^2 W + \Delta_y^2 W) + \Delta_x^2 u \right],$$

где a – безразмерное ускорение элементарного объема среды; $\bar{k} = \frac{v\omega}{c^2} = \frac{\eta\omega}{E}$ – безразмерный критерий подобия, представляющий отношение силы вязкостного трения к динамическому давлению.

Реализация предложенного алгоритма заключается в следующем. На основании последнего уравнения определяется безразмерное ускорение в заданном узле сетки интегрирования с использованием известных значений безразмерной скорости и перемещения в данный момент времени. Значения безразмерных перемещений и скорости в заданном узле сетки интегрирования в последующий момент времени определяется по формулам

$$u^{+\delta u} = u + W\delta t + \frac{1}{2}a\delta t^2; \quad W^{+\delta t} = W + a\delta t.$$

Описанный процесс повторяется по всем узлам сетки интегрирования для данного момента времени. Это позволяет определять значения безразмерного ускорения для последующего момента времени по всем узлам сетки интегрирования.

С учетом начальных условий $u = -1$, $W = 0$ и граничных, выраженных в виде $u_r = f_1(t)$ и $W_r = f_2(t)$, можно последовательно получить численные значения скоростей и перемещений по всем узлам сетки интегрирования в функции координат и времени. Использование полученных значений u и W позволяет сравнительно просто определить напряженное состояние вибрируемой бетонной смеси (рисунок 5.9). Естественно, что вычисления в этом случае громоздки и необходимо использование специальных программ и ЭВМ.

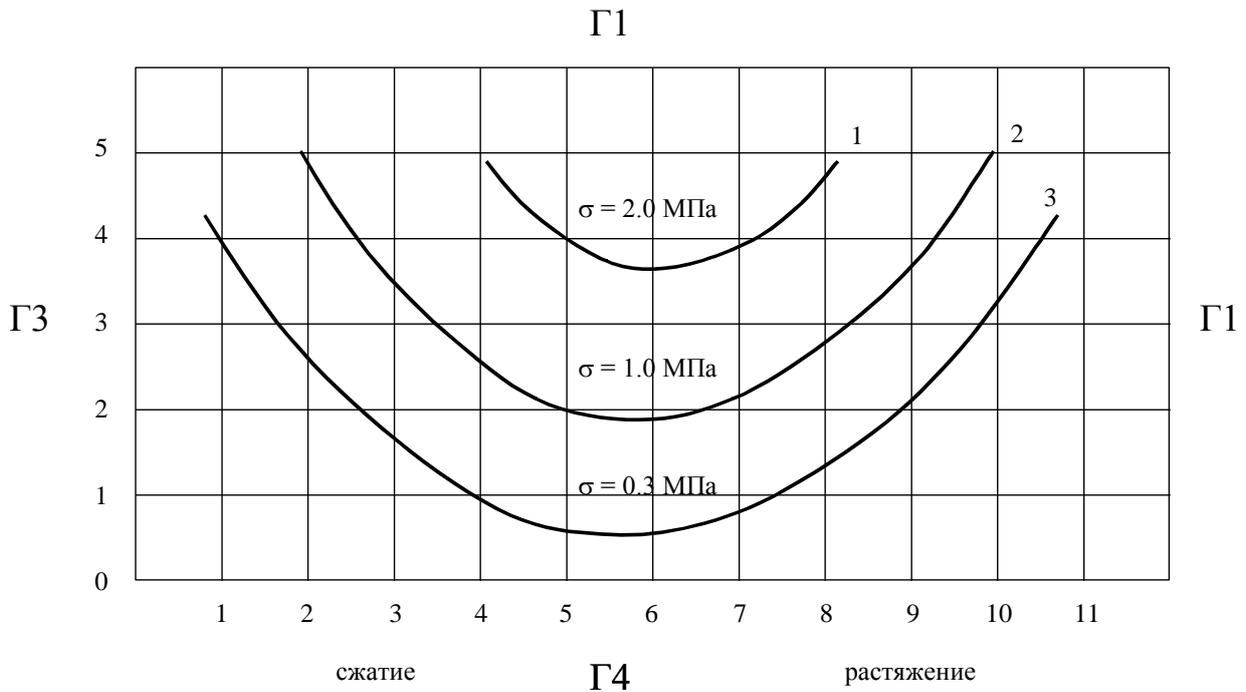
Использование аналитических методов решения математических задач является основным методом современного научного исследования. Однако громоздкость моделей и прямых методов решения уравнений затрудняет получение конечных решений. Поэтому в решении практических задач нашли широкое применение **методы преобразования исходных уравнений (логарифмирование, преобразований Лапласа, Фурье и т.д.)**.

Логарифмирование уравнений является простейшим способом преобразований.

Пусть нам необходимо получить решение простейшего уравнения

$$y = a^{0.2},$$

которое называется оригиналом функции.



Фазовый угол $\alpha = 0^\circ$, частота колебаний 20 Гц, амплитуда – 3.5 мм;
 1 – $\sigma = 2.0$ МПа; 2 – $\sigma = 1.0$ МПа; 3 – $\sigma = 0.3$ МПа

Рисунок 5.9 – Распределение кривых равного значения динамического давления в модели формы с бетонной смесью при верхней свободной поверхности

Возведение числа a в степень 0.2 прямыми методами затруднительно. Поэтому осуществляется преобразование данного уравнения при помощи логарифмирования

$$\log y = 0.2 \log a$$

– это уравнение называется изображением функций.

При логарифмировании функция переводится из пространства оригиналов в пространство изображений и операция возведения в степень сводится к умножению чисел 0.2 и $\log a$, что не встречает никаких затруднений.

При помощи антилогарифмирования полученный результат переводится из пространства изображений в пространство оригиналов.

Приведенный пример хорошо иллюстрирует и является аналогом преобразований Лапласа и преобразований Фурье.

Смысл указанных преобразований аналогичен логарифмированию. Например, в преобразованиях Лапласа исходная функция переводится из пространства оригиналов в пространство изображений при помощи интеграла

$$F(p) = \int_0^{\infty} f(t)e^{-pt} dt,$$

где p – оператор Лапласа $\left(p = \frac{d}{dt}\right)$.

Перевод функции из пространства изображений в пространство оригиналов осуществляется при помощи интеграла

$$f(t) = \frac{1}{2\pi} \int_{c-i_{\infty}}^{c+i_{\infty}} F_p e^{pt} dp.$$

Величина c выбирается так, чтобы обеспечить сходимость интеграла.

Преобразование Лапласа широко используется при решении дифференциальных и интегральных уравнений. В процессе решения этих уравнений широко используются таблицы преобразований функции, так же как это делается в случае логарифмирования.

Основываясь на методе преобразования функций, решаются задачи анализа переходных процессов в системах управления. В процессе анализа оперируют передаточными функциями.

Под передаточной функцией понимается отношение преобразования Лапласа выходной координаты линейной системы к преобразованию входной координаты при нулевых начальных условиях:

$$W(p) = \frac{y(p)}{x(p)},$$

где $W(p)$ – передаточная функция; $x(p)$ – преобразование Лапласа входного сигнала; $y(p)$ – преобразование Лапласа выходного сигнала.

Вид передаточной функции может быть получен из дифференциального уравнения системы управления путем замены операции дифференцирования по времени оператором Лапласа p , а операции интегрирования по времени – заменой $1/p$.

Например, если система управления описывается дифференциальным уравнением

$$\frac{dy(t)}{dt} + ay(t) = kx(t),$$

то, произведя замену $\frac{d}{dt}$ на p и переходя к изображениям, получаем

$$py(p) + ay(p) = kx(p).$$

Такое представление дифференциального уравнения называется операционным.

Для нахождения передаточной функции этой системы достаточно произвести несложные алгебраические преобразования. В результате получим

$$W(p) = \frac{y(p)}{x(p)} = \frac{k}{p+a}.$$

Кроме метода передаточных функций для анализа систем управления широко используется метод частотных характеристик, который составляет теоретическую базу обобщенного гармонического анализа.

Под частотной характеристикой системы понимают отношение комплексных изображений выходной и входной амплитуд в установившемся режиме гармонических колебаний

$$W(j\omega) = \frac{A_2(\omega)}{A_1} e^{j\varphi(\omega)},$$

где A_1 – амплитуда синусоидального входного сигнала; ω – круговая частота; $A_2(\omega)$ – амплитуда колебаний выходной координаты; $\varphi(\omega)$ – сдвиг по фазе синусоидальных колебаний выходной координаты.

Аналитически $W(j\omega)$ можно получить из передаточной функции заменой параметра преобразования Лапласа p на $j\omega$ с последующим выделением модуля комплексного числа и фазового сдвига.

Частотные характеристики систем управления используют при анализе устойчивости, качества переходных процессов и динамической точности, синтеза корректирующих устройств.

Кроме перечисленных при решении задач широко используются: метод пространства состояний; метод компараментального анализа; информационные методы.

Рассмотрим **примеры** по составлению моделей.

На рисунке 5.10 приведена модель деформирования. Модель 1, представленная пружиной, характеризует упругие свойства и подчиняется закону Гука – величина

деформации прямо пропорциональна прилагаемой нагрузке P . Такой закон деформирования характерен для твердых упругих тел.

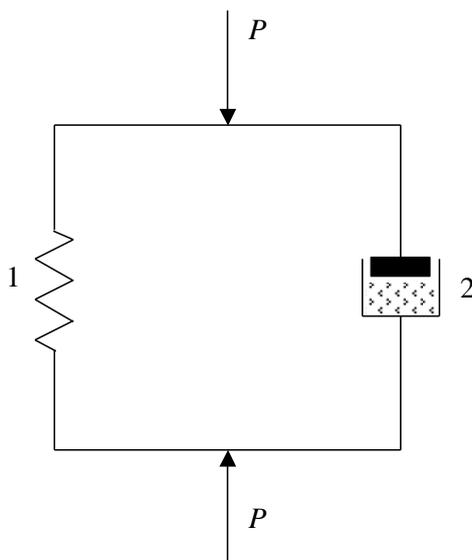


Рисунок 5.10 – Модель деформирования упруговязких материалов

Модель 2, представленная движением поршня в заполненном вязкой жидкостью цилиндре, характеризует вязкие свойства тел. Деформация тел в данном случае происходит медленно, развиваясь во времени, и подчиняется закону Ньютона – сопротивление пропорционально скорости демпфирования.

При параллельном соединении двух моделей в единое целое имеем модель деформирования упруговязкого тела. Такое деформирование подчиняется закону Кельвина.

Математическая модель, соответствующей физической модели (рисунок 5.10), может быть представлена в виде

$$P = P_y + P_B = E_y S_y + \eta \frac{dS}{dt}, \quad (5.1)$$

где P_y , P_B – упругое сжатие пружины и вязкое сопротивление жидкости; E_y , S_y – модуль упругости и относительная деформация пружины; η – коэффициент вязкости; $\frac{dS}{dt}$ – скорость деформирования.

Решая (5.1) при $t = 0$, $S = 0$, имеем

$$S = \frac{P}{E_y} \left[1 - \exp\left(-\frac{E_y t}{\eta}\right) \right]. \quad (5.2)$$

Зависимость (5.2) в ряде случаев хорошо согласуется с экспериментом и позволяет изучить законы деформирования упруговязких материалов.

Приведенный пример иллюстрирует процесс познания в соответствии с формулой – от живого созерцания (наблюдение за поведением материала) к абстрактному мышлению (физическая модель – рисунок 5.10 и уравнение (5.1)) и от него к практике – уравнение (5.2).

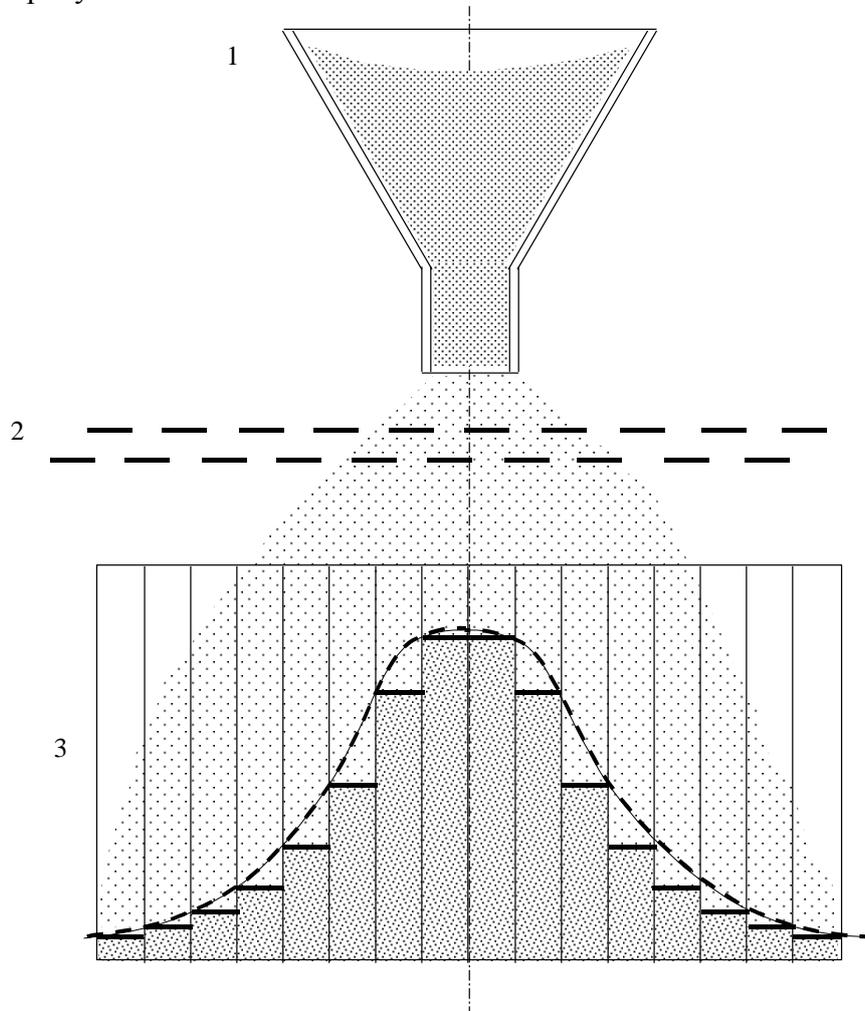
Рассмотренная модель соответствует функциональной зависимости, когда одному значению аргумента соответствует только одно значение функции. Однако в природе встречаются процессы, когда одному значению аргумента соответствует несколько значений функции вследствие действия на явление различных случайных факторов.

На рисунке 5.11 приведена физическая модель, характеризующая закон вероятностного распределения песка, который вытекает непрерывной струей из лейки через решето в ящик с вертикальными секциями. Наблюдения показывают, что распределение песка в ящике подчиняется закону нормального распределения:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad (5.3)$$

где y – ордината, частота распределения песка или количество песка в секции; x – абсцисса, номер секции в ящике, отсчитываемой от середины; σ – среднеквадратическое отклонение.

Выражение (5.3) является математической моделью вероятностного процесса, приведенного на рисунке 5.11.



1 – воронка; 2 – решето; 3 – ящик с секциями
Рисунок 5.11 – Закономерность распределения песка

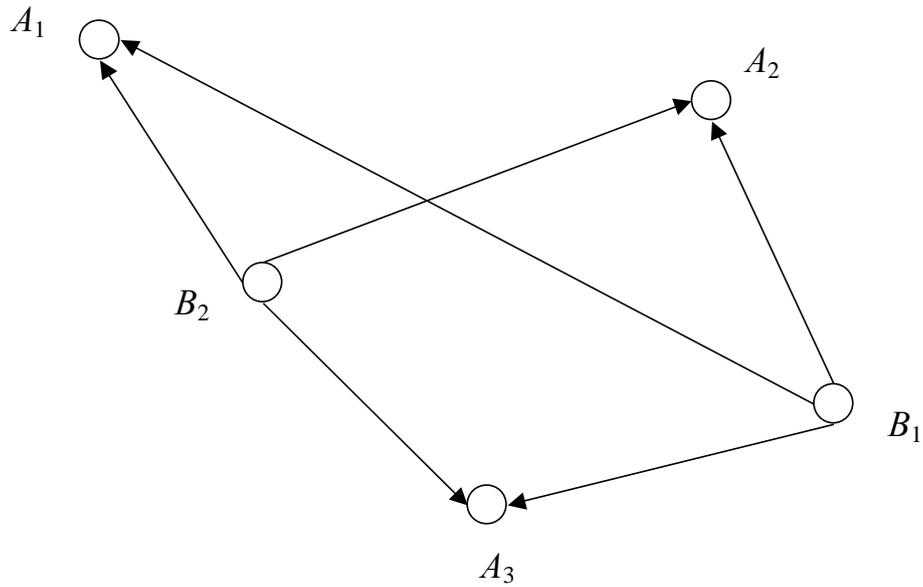
Широкое распространение получили модели, обеспечивающие оптимизацию технологических процессов и их управления. В связи с этим рассмотрим, так называемую, транспортную задачу. Пусть имеется A_1, A_2, A_3 объектов строительства, потребляющих соответственно a_1, a_2, a_3 щебня (рисунок 5.12). В местах B_1 и B_2 есть притрассовые карьеры с запасами щебня b_1 и b_2 . При этом $a_1 + a_2 + a_3 = b_1 + b_2$.

Стоимость единицы продукции из карьера B_1 на объект A_1 равна C_{11} , A_2 – C_{12} на объект A_3 – C_{13} . Количество щебня x_{ij} , транспортируемое на объект A_i из карьера B_j , взаимосвязано с другими величинами системой уравнений:

$$\begin{aligned} x_{11} + x_{21} &= a_1; \\ x_{12} + x_{22} &= a_2; \\ x_{13} + x_{23} &= a_3; \end{aligned} \quad (5.4)$$

$$x_{11} + x_{12} + x_{13} = b_1;$$

$$x_{21} + x_{22} + x_{23} = b_2.$$



A – объекты строительства; *B* – карьеры
Рисунок 5.12 – Схема транспортных связей

В системе (5.4) первое уравнение означает количество щебня, транспортируемое из карьеров B_1 и B_2 на объект A_1 ; второе – на объект A_2 . Последнее уравнение – количество щебня, доставляемое на объекты A_1, A_2, A_3 из карьера B_2 , и т.д.

Все исходные данные сведены в матрицу условия задачи (таблица 5.1). Задача имеет много решений, так как есть шесть неизвестных в системе (5.4), состоящих из пяти уравнений. Требуется определить наиболее выгодный (экономичный) вариант перевозки щебня. В этом случае численными методами с помощью линейного программирования и ЭВМ находят функцию, которая удовлетворяет условию

$$C = \sum_{i=1}^n \sum_{j=1}^m C_{ij} x_{ij} = \min. \quad (5.5)$$

Уравнения (5.4), (5.5) – математическая модель, позволяющая оптимизировать транспортный процесс. Физическая модель изображена на рисунке 5.12.

Таблица 5.1 – Исходные данные

Карьеры	Объекты			Запасы
	A_1	A_2	A_3	
B_1	C_{11} x_{11}	C_{12} x_{12}	C_{13} x_{13}	b_1
B_2	C_{21} x_{21}	C_{22} x_{22}	C_{23} x_{23}	b_2
Общая потребность	a_1	a_2	a_3	---

5.4 Метод аналогии. Модели подобия

Явления, процессы изучаются не изолированно друг от друга, а комплексно. Различные объекты с их специфическими переменными величинами объединяются в комплексы, характеризуемые едиными законами. Это позволяет распространить анализ одного явления на другие или целый класс аналогичных явлений. При таком принципе исследований уменьшается число переменных величин, они заменяются обобщенными критериями. В результате упрощается искомое математическое выражение. На этом принципе основаны методы сочетания аналитических и экспериментальных исследований с методами аналогии, подобия, размерностей, являющихся разновидностью методов моделирования.

Суть *метода аналогии* рассмотрим на примере. Тепловой поток зависит от температурного перепада (закон Фурье):

$$q_T = -\lambda \frac{dt}{dx}. \quad (5.2)$$

Здесь λ – коэффициент теплопроводности. Массоперенос или перенос вещества (газа, пара, влаги) определяется перепадом концентрации вещества c (закон Фика):

$$q_B = -\mu \frac{dc}{dx}, \quad (5.3)$$

μ – коэффициент массопереноса. Перенос электричества по проводнику с погонным сопротивлением обуславливается перепадом напряжения (закон Ома):

$$q_{\mathcal{E}} = -\frac{1}{\rho} \frac{du}{dx}, \quad (5.4)$$

где ρ – коэффициент электропроводности. Все три рассматриваемые явления характеризуются различными физическими процессами, но имеют идентичные математические выражения, т.е. их можно исследовать методом аналогий.

В зависимости от того, что принимается за оригинал и модель, могут быть различные виды моделирования методом аналогии. Так, если тепловой поток q_T изучают на модели с движением жидкости, то моделирование называют гидравлическим; если тепловой поток исследуют на электрической модели, моделирование называют электрическим. Моделирование может быть механическим, акустическим и др. Целесообразность выбора вида моделирования зависит от сложности изучаемого процесса и модели, ее стоимости и эксплуатации, возможности постановки различных экспериментов, точности результатов.

Идентичность математических выражений процессов оригинала и модели не означает, что эти процессы абсолютно аналогичны. Для того чтобы на модели максимально моделировать изучаемый процесс оригинала, необходимо соблюдать *критерий аналогий*. Так, сравнивать q_T и $q_{\mathcal{E}}$, коэффициенты теплопроводности λ и электропроводности ρ , температуру и напряжение u нет смысла. Для устранения этой несопоставимости оба уравнения необходимо представить в безразмерных величинах: каждую переменную величину Π_{Π} представить в виде произведения постоянной размерности на переменную безразмерную $\Pi_{\mathcal{B}}$:

$$\Pi = \Pi_{\Pi} \Pi_{\mathcal{B}}. \quad (5.5)$$

Имея в виду (5.5), выражения для q_T и $q_{\mathcal{E}}$ запишем так:

$$q_T = q_{T\Pi} q_{T\mathcal{B}}; \quad \lambda = \lambda_{\Pi} \lambda_{\mathcal{B}}; \quad t = t_{\Pi} t_{\mathcal{B}}; \quad x = x_{\Pi} x_{\mathcal{B}}; \quad q_{\mathcal{E}} = q_{\mathcal{E}\Pi} q_{\mathcal{E}\mathcal{B}};$$

$$\gamma = \gamma_{\Pi} \gamma_{\mathcal{B}}; \quad u = u_{\Pi} u_{\mathcal{B}}.$$

После простых преобразований с учетом (5.2) – (5.4) имеем:

$$\left[\frac{q_{ТПХП}}{\lambda_{ПТ}} \right] q_{ТБ} = \lambda_{Б} \frac{dt_{Б}}{dx_{Б}}; \quad \left[\frac{q_{ЭПХП}}{\gamma_{П^u}} \right] q_{ЭБ} = \gamma_{Б} \frac{du_{Б}}{dx_{Б}}.$$

Оба выражения записаны в безразмерном виде и их можно сравнивать. Уравнения будут идентичны, если

$$\left[\frac{q_{ТПХП}}{\lambda_{ПТ}} \right] = \left[\frac{q_{ЭПХП}}{\gamma_{П^u}} \right]. \quad (5.6)$$

Это равенство называют критерием аналогий. С его помощью устанавливают параметры модели по исходному уравнению объекта. Количество критериев аналогии на единицу меньше числа членов изучаемого исходного выражения. Поскольку число неизвестных больше числа уравнений, то некоторыми параметрами модели задаются. Обычно это время наблюдения или протекания процесса на модели. Оно должно быть удобным для наблюдения исследователю.

В настоящее время широко распространено *электрическое моделирование*. Рассмотрим пример его. Необходимо изучить закономерности колебания массы m , подвешенной параллельно упругой пружине и демпфером к плоскости. Для этой системы дифференциальное уравнение имеет вид:

$$m \frac{d^2 s}{dT^2} + \alpha \frac{ds}{dT} + \frac{1}{\beta} s = F(T), \quad (5.7)$$

где α – коэффициент демпфирования; s – механическое перемещение; β – коэффициент, характеризующий упругость пружины (деформация пружины при действии единицы силы); $F(T)$ – сила, прилагаемая к системе. Чтобы определить параметры m , α , β уравнение (5.7) можно исследовать методом электрических аналогий. Для электрической модели цепи уравнение имеет вид

$$c_1 \frac{d^2 f}{dt_1^2} + \frac{1}{R} \frac{df}{dt_1} + \frac{1}{l} f = i(t_1), \quad (5.8)$$

где c_1 – емкость конденсатора; f – магнитный поток; t_1 – время процесса в электросети; R , l – сопротивление, индуктивность; i – ток электросети. После соответствующих преобразований безразмерные уравнения запишем так:

$$\left[\frac{m_p}{\alpha_p t_p} \right] = \left[\frac{c_1 R_p}{t_{1p}} \right]; \quad \left[\frac{t_p}{\alpha_p \beta_p} \right] = \left[\frac{R_p t_{1p}}{l_p} \right]; \quad \left[\frac{f_p t_p}{s_p \alpha_p} \right] = \left[\frac{R_p i_p t_{1p}}{f_p} \right]. \quad (5.9)$$

Выбор критериев (5.9) представляет определенные трудности. Чтобы упростить построение модели, пользуются системой масштабных уравнений.

Поскольку механический (оригинал) и электрический (модель) процессы аналогичны, то переменные величины этих систем изменяются во времени закономерно в определенном соотношении – масштабе. Масштабный коэффициент той или иной переменной величины представляет собой отношение переменных величин модели и оригинала:

$$M_f = \frac{f}{s}; \quad M_t = \frac{t_1}{t}; \quad M_i = \frac{i}{F}$$

где M_f , M_t , M_i – масштабы переменных величин.

С учетом масштабных переменных уравнения для модели и оригинала получим следующие:

$$\frac{c_1 R}{M_t} \frac{d^2 s}{dt^2} + \frac{ds}{dt} + \frac{R M_t}{l} S = \frac{R M_t M_i}{M_f} F; \quad \frac{m}{\alpha} \frac{d^2 s}{dt^2} + \frac{ds}{dt} + \frac{1}{\alpha \beta} S = \frac{1}{\alpha} F.$$

Эти уравнения тождественны, если

$$\left[\frac{c_1 R}{M_t} \right] = \left[\frac{m}{\alpha} \right]; \quad \left[\frac{R M_t}{l} \right] = \left[\frac{1}{\alpha \beta} \right]; \quad \left[\frac{R M_t M_i}{M_f} \right] = \left[\frac{1}{\alpha} \right]. \quad (5.10)$$

Масштабные системы (5.10) идентичны критериям аналогии (5.9), но в более простой форме.

С помощью системы масштабных уравнений (5.10) вычисляют параметры модели, а на основе предельных отклонений переменных величин оригинала и модели – масштабные коэффициенты. Задаваясь средними значениями параметров оригинала, по (5.10) вычисляют средние значения параметров модели и проектируют электрическую цепь. Далее оригинал исследуют на модели. Варьируя l, R, c на модели изучают параметры m, α, β оригинала.

С помощью электрического моделирования можно изучать и анализировать различные физические процессы, которые описываются математическими зависимостями. Это моделирование универсально, простое в эксплуатации, не требует громоздкого оборудования.

Среди методов электрического моделирования имеются и другие: метод сплошных сред, электрических сеток, электрогидродинамическая аналогия и др. Для простых задач применяют обычно метод сплошных сред с использованием электропроводящей бумаги (плоская задача) или электролитических ванн (объемная задача). Модель изготавливают из токопроводящей бумаги одинаковой электропроводимости. Геометрию объекта моделируют в определенном масштабе. К концам фигуры присоединяют электроды, моделирующие краевые условия. При моделировании процессов с токопроводными жидкостями (электролитами) ванны заполняют слабыми растворами солей, кислот, щелочей и др. Неоднородное поле моделируют с применением электролита разной концентрации. Метод сплошных сред предназначен для решения задач теплопроводности, распределения напряжений и др. Он прост, но ограничен решением краевых задач Лапласа.

В методе электрических сеток дифференциальные уравнения преобразуют в систему линейных, решаемых способом конечных разностей. С помощью сеточных моделей на электроинтеграторах можно исследовать стационарные и нестационарные задачи.

Широко распространенным методом моделирования является электрогидродинамическая аналогия. Она основана на электрическом моделировании движения жидкости, пара или газа.

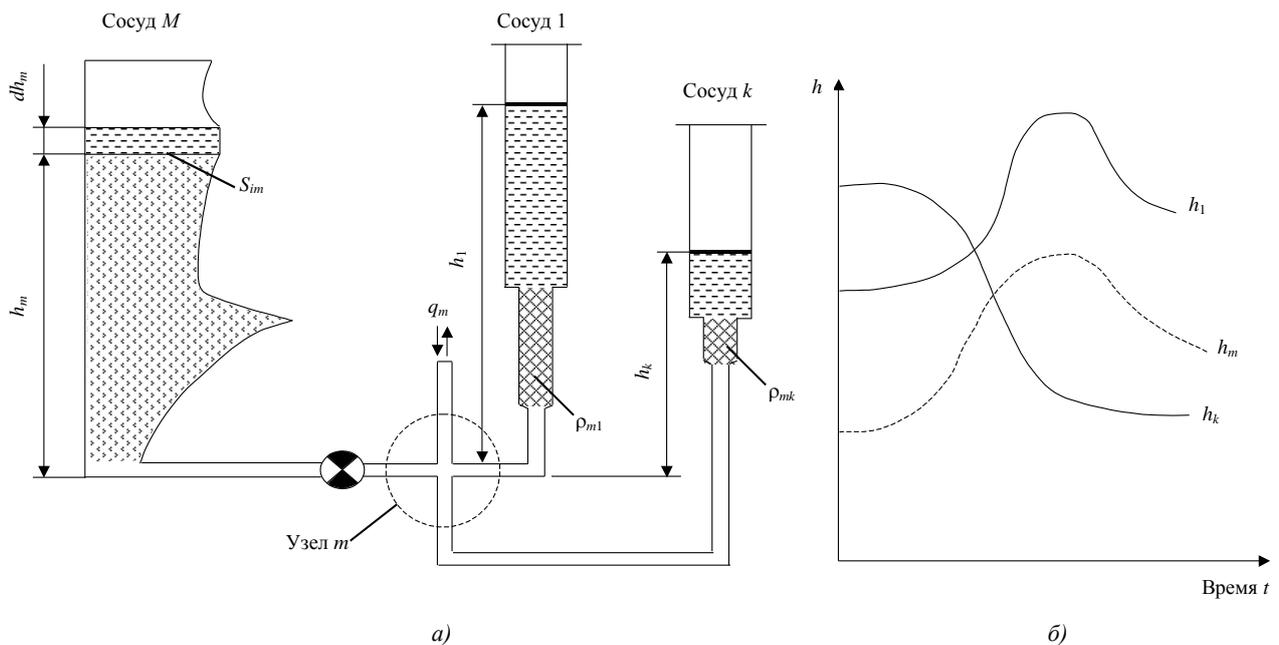
Иногда пользуются методом гидравлического моделирования на гидроинтеграторах. Гидроинтеграторы – это приборы, в которых вода передвигается по системе соединенных между собой трубок и узлов. Изучаемые постоянные и переменные величины моделируются напорами, уровнями и расходами воды в сосудах (рисунок 5.13). Метод гидравлического моделирования позволяет решать различные задачи: стационарные и нестационарные; одно-, двух- и трехмерные; с постоянными и переменными коэффициентами; для однородного и неоднородного поля, т.е. является универсальным.

Данный метод характеризуется доступностью программирования, простотой решения сложных задач, хорошей наглядностью протекаемых процессов, достаточно высокой точностью расчетов, возможностью остановить и повторить процесс на модели. Однако оборудование для этого метода громоздко, выпускается в ограниченном количестве.

В последние годы широкое распространение получили программные комплексы для анализа схмотехнических процессов в электрических схемах, например, Electronics Workbench, MicroCAP, PSpice, DesignLab, OrCAD.

Используя метод аналогии можно сводить задачу моделирования различных физических процессов к составлению эквивалентной электрической схемы рассматриваемого процесса и с помощью программ схмотехнического моделирования проводить анализ этого процесса, используя масштабные коэффициенты.

На рисунках 5.14–5.15 приведены примеры составленных эквивалентных электрических схем для различных случаев моделирования физических процессов.



M – сосуд с переменным сечением; 1, 2, ..., k – сосуды, соединенные с сосудом M через узел m ; S_{im} – сечение сосуда M на высоте h_1 ; ρ_{m1} – сопротивление течения в жидкости в сосуде 1; q_m – расход жидкости в сосуде M

Рисунок 5.13 – Схема гидроинтегратора (а) и изменения уровня жидкости в сосудах во времени (б)

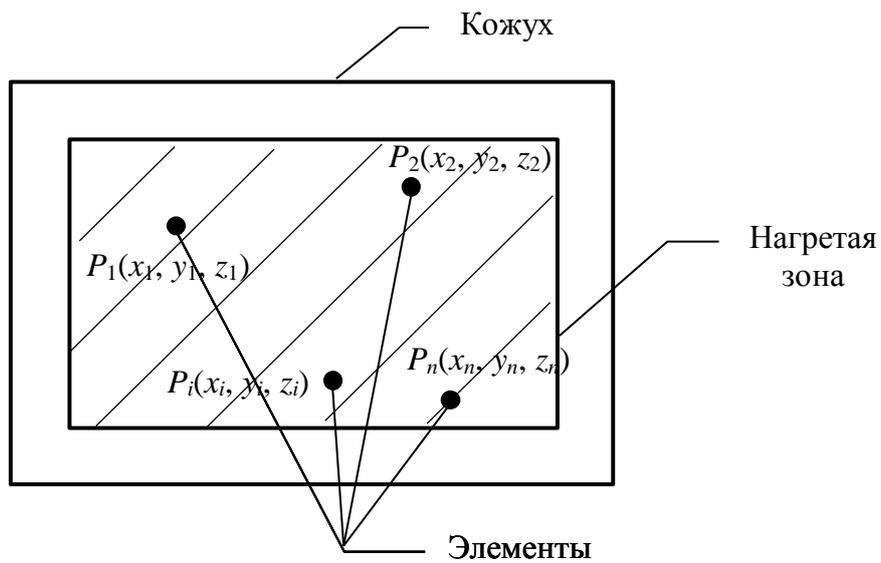
В процессе электрического моделирования указанных процессов необходимо сформировать массив исходных данных для каждой схемы в виде численных значений электрических параметров с учетом выбранных масштабных коэффициентов. Для формирования указанного массива необходимо детально рассмотреть каждый процесс с физических позиций и задать (выбрать) исходный вариант анализируемой конструкции. Причем, для каждого вида процесса составление массива исходных данных требует особых подходов. Так, основным теплофизическим параметром, определяющим процесс теплообмена в блоке радиоаппарата (рисунок 5.14), является коэффициент теплоотдачи:

$$\alpha_{\Sigma} = \alpha_K + \alpha_L,$$

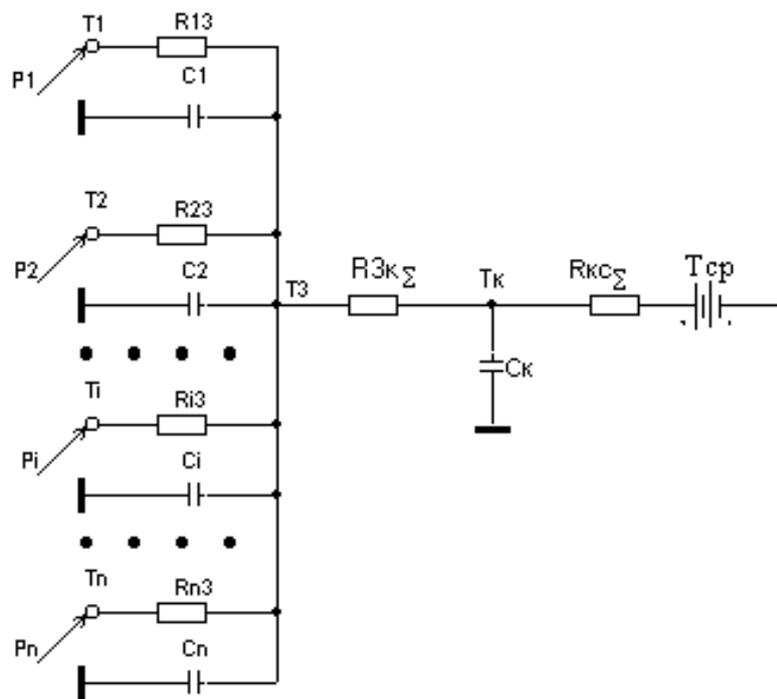
где α_K – конвективный коэффициент теплоотдачи; α_L – коэффициент теплоотдачи излучением.

Для ориентировочного определения величины α_{Σ} можно использовать хорошо зарекомендовавшие себя эмпирические методы расчета тепловых режимов радиоаппаратуры, например, коэффициентный метод. Если использовать результаты расчета теплового режима анализируемого блока как исходную итерацию, можно с достаточной для практики инженерного проектирования точностью определить искомые $\alpha_{\Sigma i}$ для всех элементов блока, участвующих в теплообмене.

При электромеханическом моделировании системы амортизации (рисунок 5.15) вибрацию можно моделировать синусоидальным гармоническим колебанием, а удар – импульсным источником различной формы.

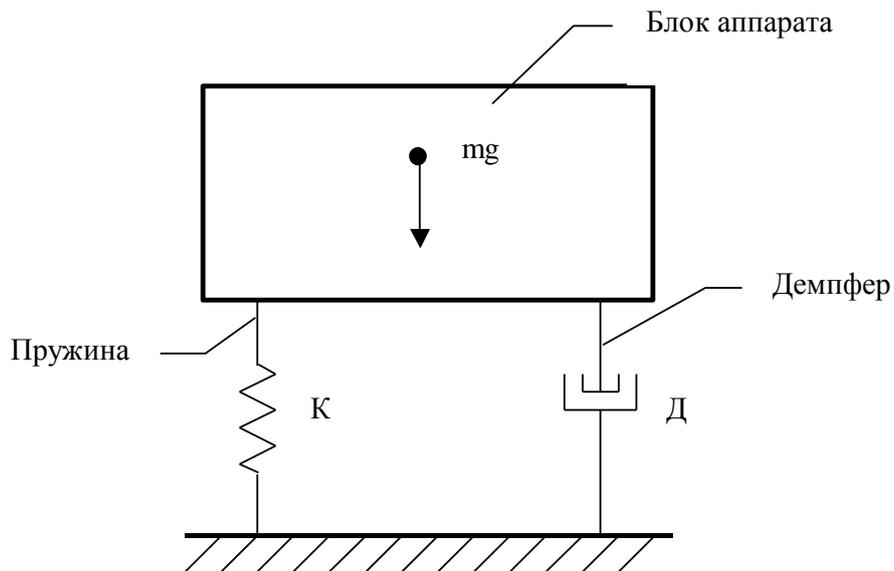


а) физическая модель

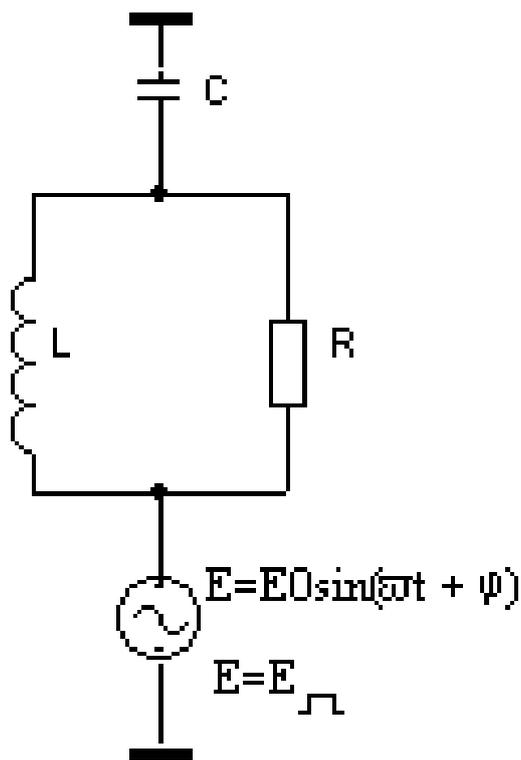


б) электротепловая модель

Рисунок 5.14 – Эквивалентная схема процесса теплообмена в блоке радиоаппарата



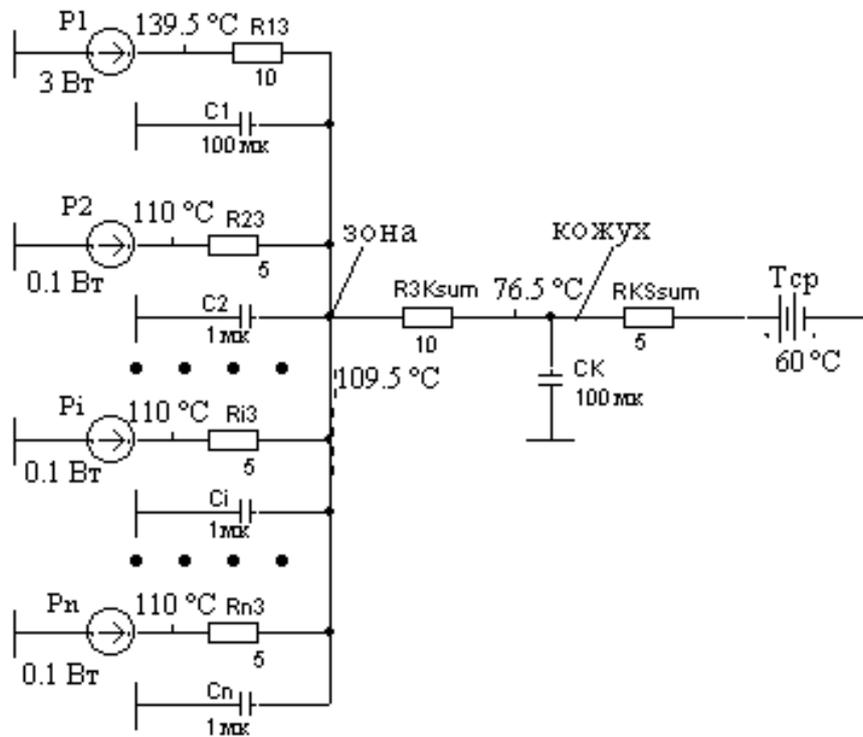
а) физическая модель



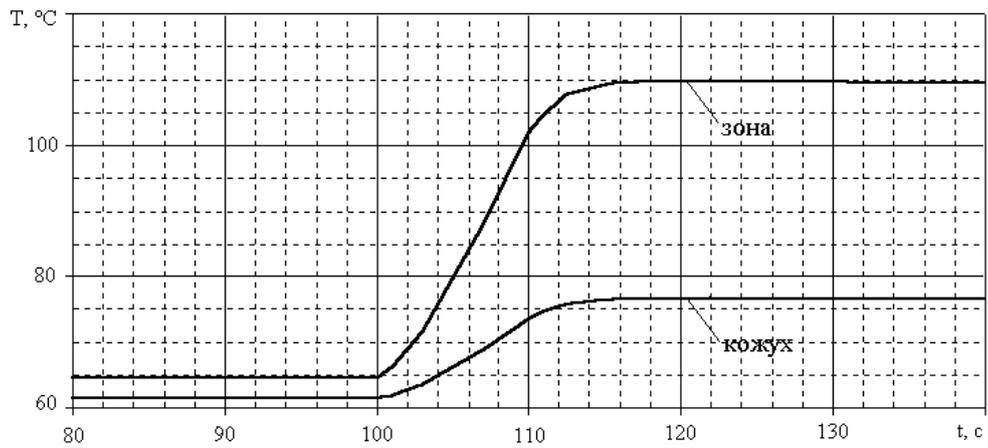
б) электромеханическая модель

Рисунок 5.15 – Электромеханическое моделирование системы амортизации

Моделирование с использованием программных комплексов схемотехнического моделирования показывает, что результаты анализа целесообразно представить в статическом и динамическом видах. Примеры результатов моделирования по схемам, приведенным на рисунках 5.14–5.15, представлены на рисунках 5.16–5.17, соответственно.

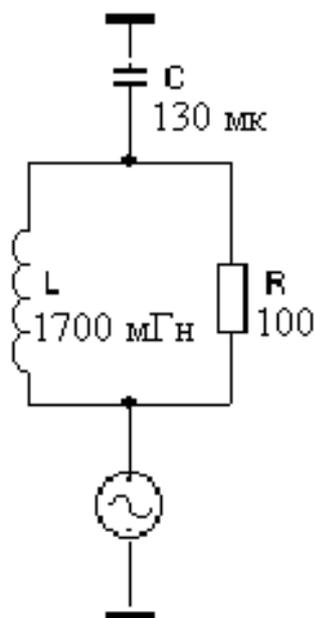


а) распределение температур

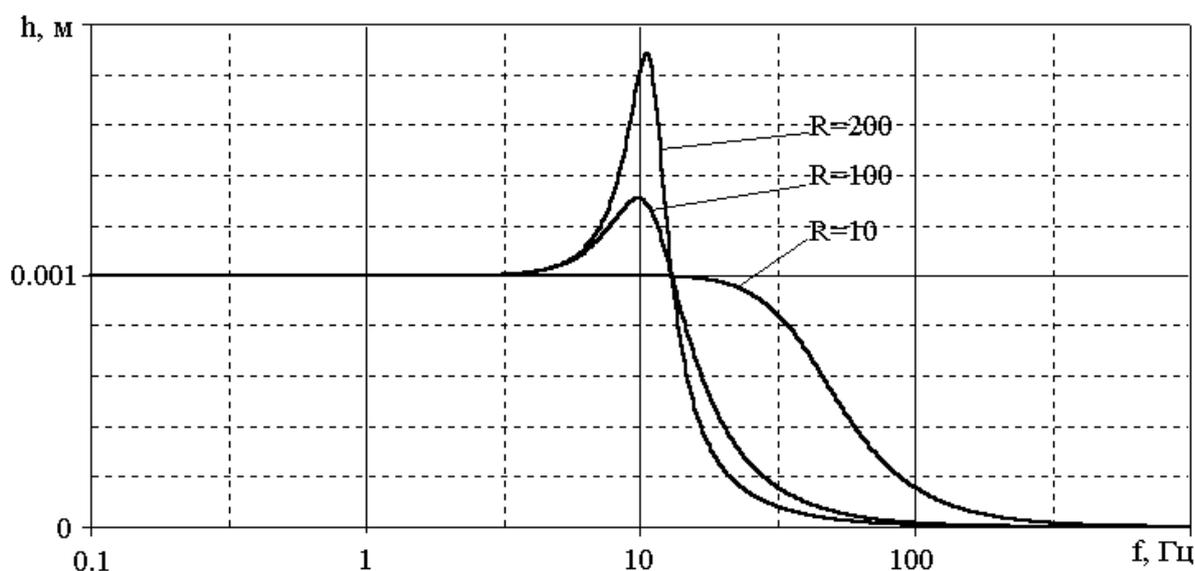


б) переходные характеристики

Рисунок 5.16 – Анализ переходных процессов теплообмена в блоке радиоаппарата



а) система амортизации



б) амплитудно-частотная характеристика системы амортизации

Рисунок 5.17 – Анализ системы амортизации

Анализ различных физических процессов с использованием программных комплексов схемотехнического моделирования является весьма удобным с точки зрения представления результатов. Результаты моделирования могут быть использованы в последующем процессе оптимального синтеза конструктивно-технологических решений, а также в прикладных исследованиях, направленных на синтез автоматизированной методологии проектирования различных устройств.

Теория подобия – это учение о подобии явлений. Она наиболее эффективна в том случае, когда на основе решения дифференциальных уравнений зависимости между переменными отыскать невозможно. Тогда необходимо произвести предварительный эксперимент и, воспользовавшись его данными, составить с применением метода подобия

уравнение (или систему уравнений), решение которого можно распространить за пределы границ эксперимента. Этот метод теоретического исследования явлений и процессов возможен лишь на основе комбинирования с экспериментальными данными.

Суть теории подобия рассмотрим на простом примере. Пусть имеется ряд прямоугольников. Это класс плоских фигур, поскольку они объединены общими свойствами – имеют по четыре стороны и четыре прямых угла. Из этого класса можно выделить единичную фигуру, которая имеет конкретное значение сторон l_1 и l_2 . Численные значения l_1 и l_2 определяют из условия однозначности. Если стороны l_1 и l_2 умножить на K_l , которому можно придать любое значение, то получим серию подобных плоских фигур, объединяемых в определенную группу:

$$\frac{l_1'}{l_2'} = \frac{l_1''}{l_2''} = \frac{l_1'''}{l_2'''} \dots = K_l.$$

K_l называют критерием подобия.

Такой способ приведения подобия применим не только для плоских объединенных фигур, но и для различных физических величин: времени $K_m = \frac{T''}{T'}$, давлений $K_p = \frac{p''}{p'}$,

вязкостей $K_\mu = \frac{\mu''}{\mu'}$, температуропроводности $K_a = \frac{a''}{a'}$ и т.д.

Критерии подобия создают внутри данного класса явлений группы путем преобразования условий однозначности в подобные системы. Все явления, входящие в одну группу, подобны и отличаются только масштабами. Таким образом, любое дифференциальное уравнение с граничными условиями и критериями подобия характерно только для группы подобных явлений. Если граничные условия представлены без критерия подобия, то дифференциальное уравнение можно применить для анализа только частного случая.

Теория подобия базируется на трех теоремах.

Теорема 1 (М.В.Кирпичева и А.А.Гухмана). Два физические явления подобны, если они описываются одной и той же системой дифференциальных уравнений и имеют подобные (граничные) условия однозначности, и их определяющие критерии подобия – численно равны.

Теорема 2. Если физические процессы подобны, то критерии подобия этих процессов равны между собой.

Теорема 3. Уравнения, описывающие физические процессы, могут быть выражены дифференциальной связью между критериями подобия.

В группе подобных между собой явлений, отличающихся только масштабом, можно распространять результаты единичного эксперимента. При использовании теории подобия удобно оперировать критериями подобия, которые обозначаются двумя латинскими буквами фамилий ученых.

Рассмотрим некоторые подобия.

Изучая потоки жидкостей, применяют критерий Рейнольдса

$$Re = \frac{\omega l}{\nu}, \quad (5.11)$$

где ν – динамическая вязкость; ω - скорость движения; l - характерный размер (расстояние, толщина, диаметр). Критерий Re является показателем отношения сил инерции к силам трения. Критерий Эйлера

$$Eu = \frac{\Delta p}{\rho \omega^2}. \quad (5.12)$$

Здесь Δp – падение давления при движении жидкости в трубопроводе вследствие трения; ρ - плотность. В тепломассопереносе применяют различные критерии. Критерий Фурье

$$Fo = \frac{aT}{l^2}, \quad (5.13)$$

где a – коэффициент температуро- или влагопроводности; T - время; l - характерный размер тела (длина, радиус).

Этот критерий характеризует скорость выравнивания тепла в данном теле. Критерий Лыкова

$$Lu = \frac{a_1}{a}. \quad (5.14)$$

Здесь a , a_1 – коэффициенты тепло- и массопереноса. Данный критерий характеризует интенсивность изменения массопереноса (влаги, пара) относительно теплопереноса. Он изменяется в широких пределах (от 0 до 1000). Критерий Кирпичева

$$Ki = \frac{q(T)l}{\lambda\Delta T}, \quad (5.15)$$

где $q(T)$ – поток тепла.

Этот критерий характеризует отношение потока тепла, подводимого к поверхности тела, к потоку тепла, отводимого внутрь тела. Все приведенные, а также другие критерии имеют безразмерный вид. Они независимы друг от друга, поэтому их сочетание дает новые критерии.

При исследовании явлений и процессов удобно использовать критерии подобия. Экспериментальные данные обрабатывают в виде обобщенных безразмерных переменных и составляют уравнения в критериальной форме, т.е. в дифференциальные уравнения вместо переменных l , a , t , q и т.д. ставят критерии подобия. Далее приступают к решению теоретического уравнения в критериальном виде. Полученное аналитическое решение позволяет распространить результаты единичного опыта на группу подобных явлений и анализировать переменные величины за пределами эксперимента.

Критерии подобия применяются и для решения дифференциальных уравнений со многими переменными. В этом случае уравнения и граничные условия целесообразно представлять в критериальном безразмерном виде, хотя это иногда и нелегко. Решение уравнений в безразмерном виде менее трудоемко, поскольку число переменных уменьшается, анализ аналитических выражений упрощается, а объем расчетов существенно снижается. Все это упрощает составление графиков и номограмм. Поэтому умение составлять дифференциальные уравнения в критериальном виде, решать их и анализировать представляет большой интерес для научного работника.

В ряде случаев встречаются процессы, которые не могут быть непосредственно описаны дифференциальными уравнениями. Зависимость между переменными величинами в таких процессах, в конечном счете, можно установить лишь экспериментально. Чтобы ограничить эксперимент и отыскать связь между основными характеристиками процесса, эффективно применять метод анализа размерностей.

Суть метода размерностей поясним на простом примере. Допустим, что в результате исследований получена функция

$$Fo = f(V^x, S^y, \rho^z),$$

где V – скорость, м/с; S – площадь, м²; ρ - плотность, кгс·с²/м⁴; Fo – сила, кгс.

Необходимо определить x , y , z . Если известно, что все коэффициенты безразмерны, то искомые величины определяются из системы уравнений, получаемых из условия равенства размерности правой и левой частей уравнения: кгс: $1 = z$; м: $0 = x + 2y - 4z$; с: $0 = -x + 2z$. Отсюда $z = 1$; $x = 2$; $y = 1$.

В более сложных случаях метод анализа размерностей позволяет составить функциональные зависимости в критериальном виде. Пусть известна в общем виде функция F для какого-либо сложного процесса:

$$F = f(n_1, n_2, \dots, n_k), \quad (5.16)$$

содержащая K неизвестных постоянных или переменных размерных величин. Необходимо отыскать F и найти ее зависимость от основных величин.

Значения n_1, n_2, \dots, n_k имеют определенную размерность единиц измерения. Метод размерностей предусматривает выбор из числа k трех основных независимых друг от друга единиц измерения. Остальные $k - 3$, входящие в функциональную зависимость (5.16), выбирают так, чтобы они были представлены в функции F как безразмерные, в критериях подобия. В этом случае преобразования производят с помощью основных, выбранных единиц измерения.

При этом функция (5.16) принимает вид

$$\frac{F}{a^x b^y c^z} = f\left(1, 1, 1, \frac{A}{a^{x_1} b^{y_1} c^{z_1}}, \frac{B}{a^{x_2} b^{y_2} c^{z_2}}, \frac{C}{a^{x_3} b^{y_3} c^{z_3}}\right).$$

Три единицы означают, что первые три числа являются отношением n_1, n_2, n_3 к соответственно равным значениям a, b, c . Выражение (5.16) анализируют по размерностям величин. В результате устанавливают численные значения показателей степени $x \dots x_3, y \dots y_3, z \dots z_3$ и определяют критерии подобия. Например, при обтекании опоры моста водой со скоростью V (м/с) на поверхность площадью S (м²) действует сила F_0 (кГс). Плотность воды ρ (кГс·с²/м⁴).

Функциональную зависимость можно записать так:

$$F_0 = f\left(V, S, \rho, \mu, g, p, \frac{l}{B}\right),$$

где $\frac{l}{B}$ – отношение высоты к ширине опоры моста (величина безразмерная); μ – вязкость воды $\left(\frac{\text{кГс} \cdot \text{с}^2}{\text{м}^4}\right)$; g – ускорение (м/с²); p – давление (кГс/см²).

Переменные размерные значения μ, g, p как и V, S, ρ , подлежат изучению при условии, что функция F_0 будет представлена в критериальном виде. При использовании метода анализа размерностей возможны только три безразмерные величины. Применительно к (5.16) запишем:

$$\frac{F_0}{V^x S^y \rho^z} = f\left(1, 1, 1, \frac{\mu}{V^{x_1} S^{y_1} \rho^{z_1}}, \frac{g}{V^{x_2} S^{y_2} \rho^{z_2}}, \frac{p}{V^{x_3} S^{y_3} \rho^{z_3}}, \frac{l}{B}\right).$$

В качестве основных размерностей принимаем м/с, кГс, кГс·с²/м⁴, т.е. для V, S, ρ . При этом $[F] = [V^x, S^y, \rho^z]$ или кГс = (м/с)^x(м²)^y(кГс·с²/м⁴)^z. Из этого выражения находим показатели степеней, принимая числитель со знаком +, знаменатель – со знаком –, показатель кГс – 1 = z; показатель м – 0 = x + 2y – 4z; показатель с – 0 = –x + 2z.

Решая эти уравнения, имеем $z = 1, x = 2, y = 1$. Таким же образом $x_1 = 1, y_1 = 0.5, z_1 = 1$; $x_2 = 2, y_2 = -0.5, z_2 = 0$; $x_3 = 2, y_3 = 0, z_3 = 1$. Отсюда запишем $n = \frac{F}{V^2 S \rho}, n_4 = \frac{\mu}{V \rho \sqrt{S}}$,

$n_5 = \frac{g \sqrt{S}}{V^2}, n_6 = \frac{p}{\rho V^2}$, это определяется условием равенства размерностей числителя и знаменателя.

Выражения n_4 и n_6 представляют собой критерии подобия Рейнольдса и Эйлера, а n_5 – критерий Фруда Fr.

В результате исследуемая функция принимает вид:

$$\frac{F}{V^2 S_\rho} = f\left(\text{Re}, \text{Fr}, \text{Eu}, \frac{l}{B}\right).$$

Эта формула позволяет исследовать процесс обтекания опоры моста в различных вариантах размеров l , B , скоростей V при условии равенства критериев подобия. Ее можно также использовать для анализа процесса методом теории подобия на моделях.

5.5 Вероятностно-статистические методы

Во многих случаях необходимо исследовать не только детерминированные, но и случайные, вероятностные (стохастические) процессы. Те или иные события могут произойти или не произойти. В связи с этим приходится анализировать случайные, вероятностные или стохастические связи, в которых каждому аргументу соответствует множество значений функции. Наблюдения показали, что, несмотря на случайный характер связи, рассеивание имеет вполне определенные закономерности. Для таких статистических законов теория вероятностей позволяет представить исход не одного какого-либо события, а средний результат случайных событий и тем точнее, чем больше число анализируемых явлений. Это связано с тем, что, несмотря на случайный характер событий, они подчиняются определенным закономерностям, рассматриваемым в теории вероятностей.

Теория вероятностей изучает случайные события и базируется на следующих основных показателях. Совокупность множества однородных событий случайной величины x составляет первичный статистический материал. Совокупность, содержащая самые различные варианты массового явления, называют генеральной совокупностью или большой выборкой N . Обычно изучают лишь часть генеральной совокупности, называемой выборочной совокупностью или малой выборкой N_1 . Вероятностью $p(x)$ события x называют отношение числа случаев $N(x)$, которые приводят к наступлению события x к общему числу возможных случаев N :

$$p(x) = N(x) / N.$$

Теория вероятностей рассматривает теоретические распределения случайных величин и их характеристики. Математическая статистика занимается способами обработки и анализа эмпирических событий. Эти две родственные науки составляют единую математическую теорию массовых случайных процессов, широко применяемую в научных исследованиях.

В математической статистике важное значение имеет понятие о частоте события $\bar{y}(x)$, представляющего собой отношение числа случаев $n(x)$, при которых имело место событие к общему числу событий n :

$$\bar{y}(x) = n(x) / n.$$

При неограниченном возрастании числа событий частота $y(x)$ стремится к вероятности $p(x)$. Частота $y_{i0} = n(x) / \sum n(x)$ характеризует вероятность появления случайной величины и представляет собой ряд распределения (рисунок 5.18), а плавная кривая – функцию плотности распределения $f(x)$.

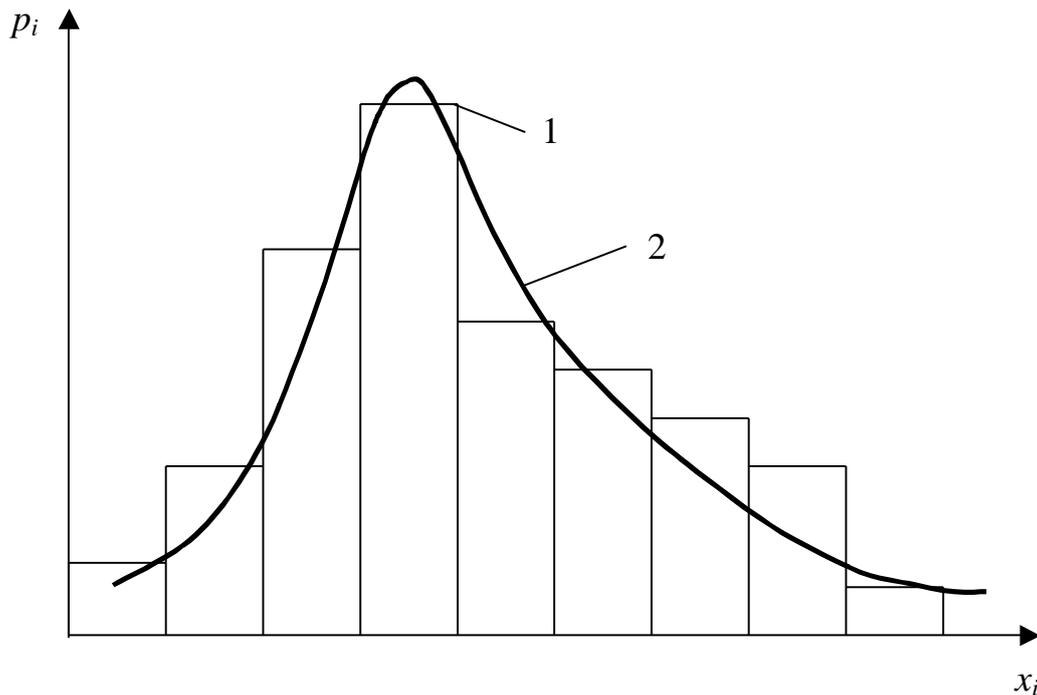
Вероятность случайной величины (события) – это количественная оценка возможности ее появления. Достоверное событие имеет вероятность $p = 1$, невозможное событие $p = 0$. Следовательно, для случайного события $0 \leq p(x) \leq 1$, а сумма вероятностей всех возможных значений

$$\sum_0^n p_i = 1.$$

В исследованиях иногда недостаточно знать функцию распределения. Необходимо еще иметь ее характеристики: среднеарифметическое и математическое ожидания, дисперсию, размах ряда варьирования.

Пусть среди n событий случайная величина x_1 повторяется n_1 раз, величина $x_2 - n_2$ раза и т.д. Тогда среднеарифметическое значение x имеет вид

$$\bar{x} = \sum_1^n (x_i n_i) / n.$$



1- гистограмма; 2 – плотность распределения
Рисунок 5.18 – Общий вид распределения случайных величин

Размах можно использовать для ориентировочной оценки вариации ряда событий:

$$R = x_{\max} - x_{\min},$$

где x_{\max} , x_{\min} – максимальное и минимальное значения измеренной величины или погрешности.

Если вместо эмпирических частот y_1, \dots, y_n принять их вероятности p_1, \dots, p_n , то это даст важную характеристику распределения – математическое ожидание:

$$m(x) = \sum_1^n x_i p_i. \quad (5.17)$$

Пусть, например, имеется пять измерений одной выборки: $x_1 = 1$; $x_2 = 2$; $x_3 = 3$; $x_4 = 4$; $x_5 = 5$ с вероятностями $p_1 = 0.10$; $p_2 = 0.15$; $p_3 = 0.45$; $p_4 = 0.30$; $p_5 = 0$. В этом случае среднее значение $x = 15/5 = 3.0$, а математическое ожидание составит в соответствии с формулой (5.17) $m(x) = 1 \times 0.10 + 2 \times 0.15 + 3 \times 0.45 + 4 \times 0.30 + 5 \times 0 = 2.95$.

Для непрерывных случайных величин математическое ожидание определяется интегралом

$$m(x) = \int_{-\infty}^{+\infty} xp(x)dx,$$

т.е. оно равно действительному значению x_d наблюдаемых событий. Таким образом, если систематические погрешности измерений полностью исключены, то истинное значение измеряемой величины равно математическому ожиданию, а соответствующая ему абсцисса называется центром распределения. Площадь, расположенная под кривой распределения (рисунок 5.18), соответствует единице вследствие того, что кривая охватывает все результаты измерений. Для одной и той же площади можно построить большое количество кривых распределения, т.е. они могут иметь различное рассеяние. Мерой рассеяния (точности измерений) является дисперсия или среднеквадратическое отклонение. Таким образом, дисперсия характеризует рассеивание случайной величины по отношению к математическому ожиданию и вычисляется с помощью формулы

$$D(x) = \sum_1^n (x_i - m(x))^2 p_i.$$

Для рассмотренного выше примера $D(x) = (1 - 2.95)^2 \cdot 0.10 + (2 - 2.95)^2 \cdot 0.15 + (3 - 2.95)^2 \cdot 0.45 + (4 - 2.95)^2 \cdot 0.30 + (5 - 2.95)^2 \cdot 0 = 0.85$.

Важной характеристикой теоретической кривой распределения является среднеквадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)}.$$

Коэффициент вариации

$$k_B = \sigma/m(x)$$

применяется для сравнения интенсивности рассеяния в различных совокупностях, определяется в относительных единицах ($k_B < 1$).

Выше были рассмотрены основные характеристики теоретической кривой распределения, которые анализирует теория вероятностей. В статистике оперируют с эмпирическими распределениями. Основной задачей статистики является подбор теоретических кривых по имеющемуся эмпирическому закону распределения. Пусть в результате n измерений случайной величины получен ряд ее значений $x_1, x_2, x_3, \dots, x_n$. При первичной обработке таких рядов их вначале группируют в интервалы и устанавливают для каждого из них частоты x_i и \bar{y}_{0i} . По значениям x_i и \bar{y}_{0i} строят ступенчатую гистограмму частот и вычисляют характеристики эмпирической кривой распределения. Основными характеристиками эмпирического распределения являются среднеарифметическое значение

$$\bar{x} = \sum_1^n x_i / n, \text{ дисперсия } D = \sum_1^n (x_i - \bar{x})^2 / n \text{ и среднеквадратическое отклонение } \sigma(x) = \sqrt{D}.$$

Значения этих величин соответствуют величинам \bar{x} , $D(x)$ и $\sigma(x)$ теоретического распределения.

В исследованиях наиболее часто применяется закон нормального распределения (рисунок 5.19)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{[x - m(x)]^2}{2\sigma^2}\right].$$

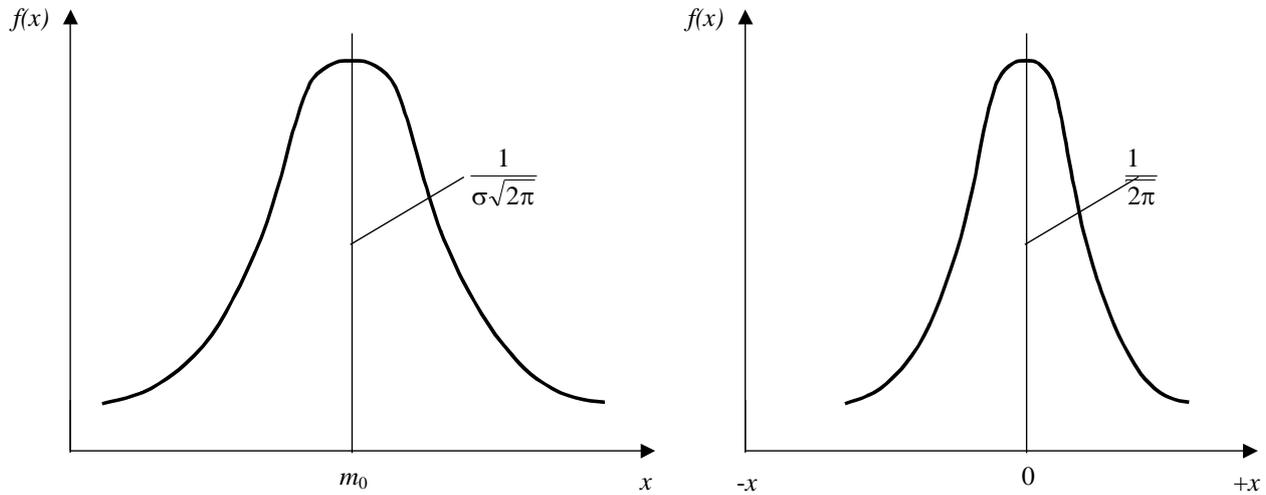
Это уравнение соответствует функции нормального распределения при $m(x) \neq 0$ (рисунок 5.19, а). Если совместить ось ординат с точкой m , т.е. $m(x) = 0$ (рисунок 5.19, б), и принять $\sigma^2 = 1$, то закон нормального распределения описывается зависимостью

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right],$$

если за единицу масштаба принять дисперсию σ^2 .

a)

б)



$a - m(x) \neq 0$; $b - m(x) = 0$

Рисунок 5.19 – Общий вид кривой нормального распределения

Для оценки рассеяния обычно пользуются величиной σ . Чем меньше σ , тем меньше рассеяние, т.е. большинство наблюдений мало отличается друг от друга (рисунок 5.20). С увеличением σ рассеяние возрастает, вероятность появления больших погрешностей увеличивается, а максимум кривой распределения (ордината, равная $\frac{1}{\sigma\sqrt{2\pi}}$) уменьшается.

Поэтому величину $y = \frac{1}{\sigma\sqrt{2\pi}}$ при $\sigma = 1$ или $y = \frac{1}{\sqrt{2\pi}}$ называют мерой точности.

Таким образом, чем меньше σ , тем больше сходимость результатов измерений, а ряд измерений более точен, среднеквадратическое отклонение определяет закон распределения. Отклонения $+\sigma$ и $-\sigma$ соответствует точкам перегиба кривой (заштрихованная площадь на рисунке 5.20). Вероятность того, что случайные события не выйдут за эти пределы, составляет 0.683. В общем случае для предела $\pm t\sigma$ вероятность того, что событие x_i попадает в данный предел, вычисляется по распределению Лапласа

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-x_i^2} dx.$$

При анализе многих случайных дискретных процессов пользуются распределением Пуассона.

Так, вероятность появления числа событий $x = 1, 2, 3, \dots$ в единицу времени определяется законом Пуассона (рисунок 5.21) и подсчитывается по формуле

$$p(x) = \frac{m^x}{x!} e^{-m} = \frac{(\lambda t)^x}{x!} e^{-\lambda t},$$

где x – число событий за данный отрезок времени t ; λ – плотность, т.е. среднее число событий за единицу времени; λt – число событий за время t , $\lambda t = m$.

Распределение Пуассона относят к редким событиям, т.е. $p(x)$ – вероятность того, что событие в период какого-то испытания произойдет x раз при очень большом числе

измерений m . Для закона Пуассона дисперсия равна математическому ожиданию числа наступления события за время t , т.е. $\sigma^2 = m$. Пуассоновский процесс можно задать параметрами x и m .

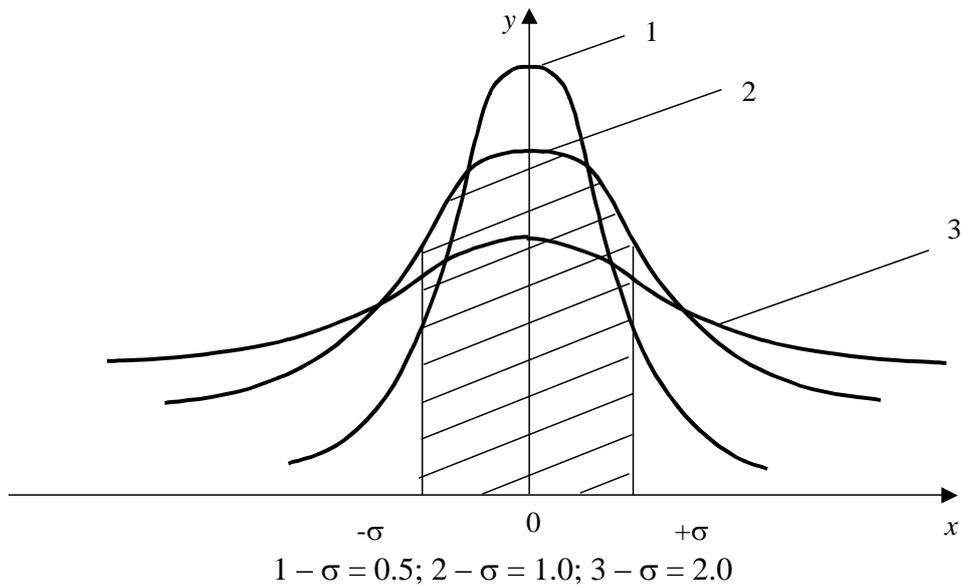


Рисунок 5.20 – Характер рассеяния кривой нормального распределения

Например, в процессе наблюдений установлено, что за 5 мин на конвейер в среднем поступает шесть комплектующих узлов автомобиля. Какова вероятность поступления 10 узлов автомобиля за 5 мин?

В этом случае $x = 10$, $\lambda t = 6$, $p(x) = \frac{6^{10} e^{-6}}{10!} = 0.041$. Как видно, эта вероятность очень мала.

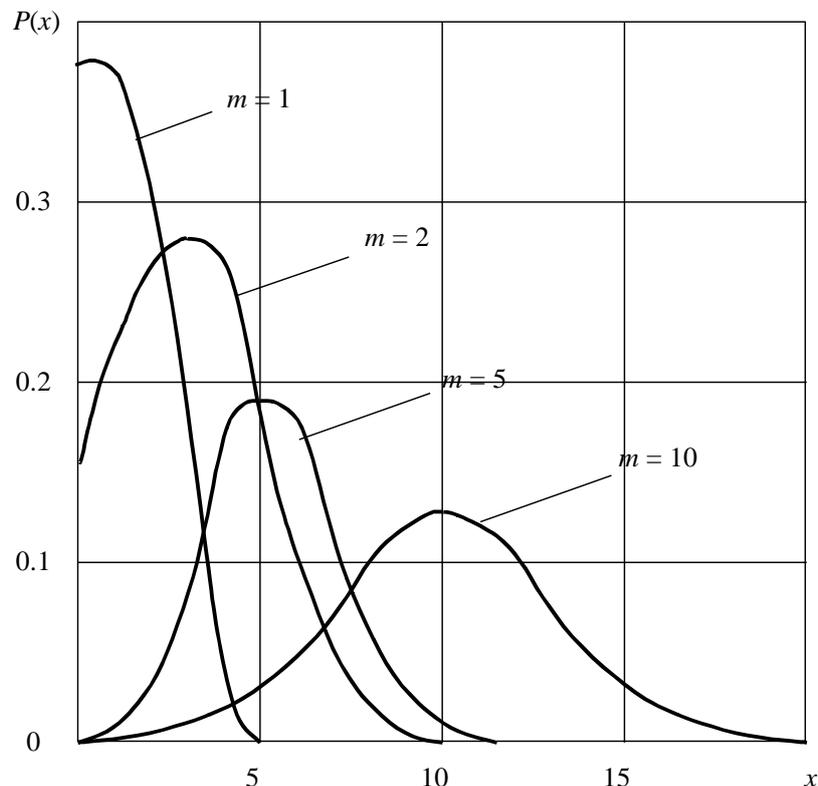


Рисунок 5.21 – Общий вид кривой распределения Пуассона

Для исследования количественных характеристик некоторых процессов (время обслуживания автомобилей на станции технического обслуживания, время отказов машин и изделий, длительность телефонных разговоров и т.д.) можно применять показательный закон распределения (рисунок 5.22, а). Плотность вероятности показательного закона выражается зависимостью $f(x) = \lambda e^{-\lambda x}$. Здесь плотность является величиной, обратной математическому ожиданию $\lambda = 1/m(x)$, кроме того, $\sigma^2 = [m(x)]^2$.

В различных областях исследований широко применяется закон распределения Вейбулла (рисунок 5.22, б) $f(x) = n\mu^n x^{n-1} e^{-\mu^n x^n}$, где n, μ – параметры закона; x – аргумент (чаще принимаемый как время).

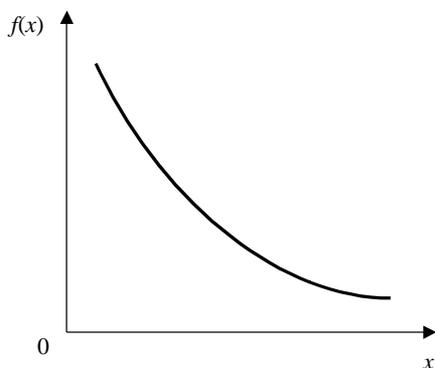
Исследуя процессы, связанные с постепенным снижением параметров (ухудшением свойств материалов по времени, деградация конструкций, процессы старения, износные отказы в машинах и др.), применяют закон γ -распределения (рисунок 5.22, в)

$$f(x) = (\lambda^\alpha / \alpha!) x^{\alpha-1} e^{-\lambda x};$$

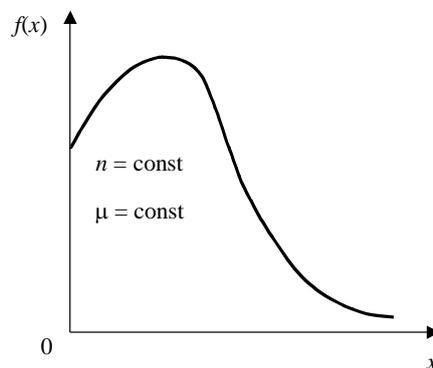
где λ, α – параметры. Если $\alpha = 1$, γ -функция превращается в показательный закон (рисунок 5.22, а).

При исследовании многих процессов, связанных с установлением расчетных характеристик, материалов, элементов и т.п., используют закон распределения Пирсона (рисунок 5.22, г), чаще всего представляемый в виде $f(x) = a e^{dx} \left(1 + \frac{x}{b}\right)^{-db}$, где a – максимальная ордината; d, b – соответственно расстояния от максимальной ординаты до центра распределения C и начала координат 0 .

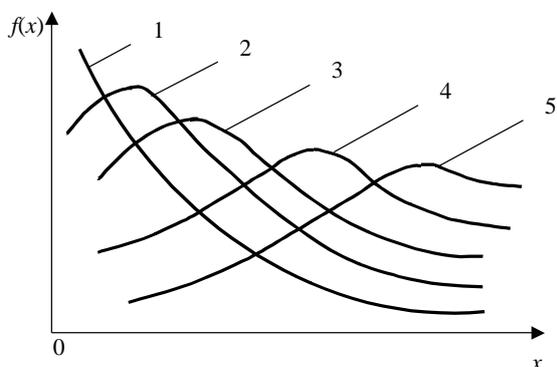
а)



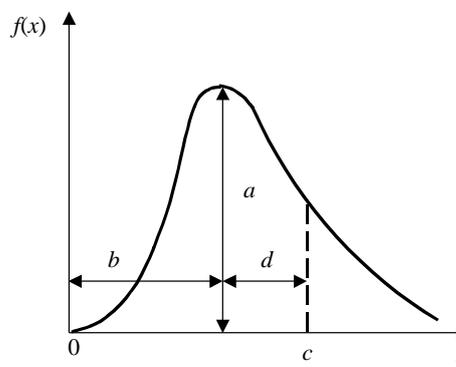
б)



в)



г)



a – показательное; $б$ – Вейбулла; $в$ – γ -распределения ($1 - \alpha = 1; \lambda = 1;$
 $2 - \alpha = 3; \lambda = 1; 3 - \alpha = 4; \lambda = 1.5; 4 - \alpha = 5; \lambda = 2; 5 - \alpha = 6$); $г$ – Пирсона

Рисунок 5.22 – Кривые распределения

Кроме приведенных выше, применяют и другие виды распределений, например, Рэлея, β -распределение, Шарлье, Гудрича.

При исследовании вероятностных систем широкое распространение получили дисперсионный, регрессионный, корреляционный и спектральный анализы, а также их различные комбинации (например, корреляционно-спектральный анализ).

В исследованиях часто возникает необходимость выявления факторов или их комбинаций, существенно влияющих на исследуемый процесс, так как при измерении какой-либо величины результаты обычно зависят от многих факторов. Практика показывает, что основными факторами, как правило, являются техническое состояние прибора и внимание оператора. Для установления основных факторов и их влияния на исследуемый процесс используется дисперсионный одно- и многофакторный анализ. Суть однофакторного дисперсионного анализа рассмотрим на примере. Пусть необходимо проверить степень точности группы m приборов и установить, являются ли их систематические ошибки одинаковыми, т.е. изучить влияние одного фактора-прибора на погрешность измерения. Каждым прибором выполнено n измерений одного и того же объекта, а всего nm измерений. Отдельное значение x_{ij} , где i – номер прибора, имеющий значения от 1 до m ; j – номер выполненного на этом приборе измерения, измеряющийся от 1 до n . Дисперсионный анализ допускает, что отклонения подчиняются нормальному закону распределения, в соответствии с которым вычисляют для каждой серии измерений среднеарифметическое значение и среднюю из показаний первого прибора и т.д. для каждого из n_i измерений и m_i приборов. В результате расчетов устанавливают величину Q_1

$$Q_1 = n \sum_{i=1}^m (\bar{x}_i - \bar{x})^2,$$

называемую суммой квадратов отклонений между измерениями серий. Она показывает степень расхождения в систематических погрешностях всех m приборов, т.е. характеризует рассеивание исследуемого фактора между приборами.

Здесь \bar{x}_i – среднеарифметическое для n измерений; \bar{x} – среднеарифметическое для всех серий измерений, т.е. общее среднее значение.

Определяется также величина Q_2 по формуле

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2,$$

где x_{ij} – отдельное i -е измерение на j -м приборе.

Величину Q_2 называют суммой квадратов отклонений внутри серии. Она характеризует остаточное рассеивание случайных погрешностей одного прибора.

При таком анализе допускается, что центры нормальных распределений случайных величин равны (или равны с определенной степенью точностью), в связи с чем все nm измерения можно рассматривать как выборку из одной и той же нормальной совокупности. Чтобы убедиться в возможности такого допущения, вычисляют критерий

$$J = \frac{Q_1 / (m - 1)}{Q_2 / m(n - 1)},$$

числитель и знаменатель которого представляют собой дисперсии σ^2 для m и mn наблюдений. В зависимости от значений $k_1 = m - 1$ и $k_2 = m(n - 1)$ числа степеней свободы и вероятности p (например, 0.95; 0.99 и др.) составлены табличные значения J_T . Если $J \leq J_T$, то считается, что в данном примере все приборы имеют одинаковые (допустимые) систематические ошибки.

Дисперсионный анализ является многофакторным, если он имеет два фактора и более. Суть его принципиально не отличается от однофакторного, но существенно увеличивается количество расчетов.

Методы теории вероятностей и математической статистики часто применяются в теории надежности, широко используемой в различных отраслях науки и техники. Под надежностью понимают свойство изделия (объекта) выполнять заданные функции (сохранять установленные эксплуатационные показатели) в течение требуемого периода времени. Обеспечение надежности (исключение отказов, нарушений работоспособности) продукции стало одной из основных народнохозяйственных задач. В теории надежности отказы рассматривают как случайные события. Для количественного описания отказов применяются математические модели – функции распределения вероятностей интервалов времени. Наиболее часто применяются законы нормального и экспоненциального распределения, закон Вейбулла и некоторые другие.

Основной задачей теории надежности является прогнозирование (предсказание с той или иной вероятностью) различных показателей безотказной работы (долговечности, срока службы и т.д.), что связано с нахождением вероятностей.

Для исследования сложных процессов вероятностного характера применяют метод Монте-Карло, с помощью которого отыскивается наилучшее решение из множества рассматриваемых вариантов. Этот метод статистического моделирования или статистических испытаний основан на использовании случайных чисел, моделирующих вероятностные процессы. Результаты решения метода позволяют установить эмпирические зависимости исследуемых процессов. Математической основой метода является закон больших чисел, разработанный П.Л.Чебышевым, который формулируется так: при большом числе статистических испытаний вероятность того, что среднеарифметическое значение случайной величины стремится к ее математическому ожиданию, равна 1, т.е.

$$\lim_{n \rightarrow \infty} p \left\{ \left| \frac{\sum x_i}{n} - m(x) \right| < \varepsilon \right\} \rightarrow 1,$$

где ε – любое малое положительное число.

Из этой формулы видно, что по мере увеличения числа испытаний n среднеарифметическое неограниченно (асимптотически) приближается к математическому ожиданию.

Последовательность решения задач методом Монте-Карло сводится к сбору, обработке и анализу статистических наблюдений исследуемого процесса: отбору главных, отбрасыванию второстепенных факторов и составлению адекватной математической модели (уравнений, графиков, циклограмм и т.д.); составлению алгоритмов и решению задачи на ЭВМ.

Для решения задач методом Монте-Карло необходимо иметь статистический ряд, знать закон его распределения, среднее значение \bar{x} и математическое ожидание $m(\bar{x})$, среднеквадратическое отклонение. С помощью метода можно получить сколько угодно заданную точность решения, т.е. $\bar{x} \rightarrow m(\bar{x})$. При нормальном законе распределения точность результатов, полученных методом Монте-Карло, оценивается по формуле $p|\bar{x} - m(x)| < \frac{3\sigma}{\sqrt{n}}$.

Пусть, например, по условию задачи задана допустимая ошибка ε_D . Если при имеющемся числе ряда n_1 и σ_1 ошибка ε_{D_1} окажется больше, чем ε_D , то необходимо увеличить число испытаний до n_2 и вычислить новое значение ошибки ε_{D_2} и т.д., пока не будет соблюдаться условие $\varepsilon_{D_i} \leq \varepsilon_D$ (i – число испытаний). Следует при этом подчеркнуть, что решение задач методом Монте-Карло эффективно лишь с использованием быстродействующих ЭВМ.

При исследованиях процессов и объектов в последнее время стали применять методы, основанные на теории массового обслуживания (ТМО), в целях отыскания условия наибольшей эффективности работы системы «требование – обслуживание». Под обслуживанием понимают удовлетворение какой-либо заявки. Таким образом, в ТМО система состоит из числа (потока) требований, обслуживающего прибора (аппарата) и выходящего потока. В зависимости от условий функционирования системы число требований создает очередь на обслуживание.

Основными характеристиками ТМО являются: интенсивность поступления требований или заявок на обслуживание λ ; интенсивность обслуживания (пропускная способность прибора обслуживания) μ ; коэффициент использования системы $\varphi = \lambda/\mu$; время ожидания в очереди до обслуживания в системе t_0 ; длительность обслуживания t_1 ; время обслуживания в системе $t_{об}$; число требований в очереди n ; математическое ожидание числа требований в системе n_c . Эти характеристики имеют следующие соотношения: $t_1 = 1/\mu$; $t_{об} = \bar{t}_0 + \bar{t}_1$; $n = \bar{n}_c \varphi$; $\bar{t}_0 = n/\lambda$.

Знак «—» означает, что при расчете принимаются средние случайные значения λ , t_0 , t_1 , $t_{об}$, n . Распределение времени обслуживания по длительности чаще всего выражается показательным законом.

В ТМО интенсивность обслуживания всегда выше интенсивности требования, т.е. $\varphi < 1$. Тем не менее, несмотря на то, что $\mu > \lambda$, возникают очереди на обслуживание, поскольку $t_{об}$ по ряду причин величина переменная, а интервал между обслуживанием неритмичен. Задачей ТМО, в конечном счете, является установление наиболее достоверных зависимостей между интенсивностью потока требований и производительностью (пропускной способностью, количеством и эффективностью обслуживания) системы. Показателями эффективности функционирования могут быть t_0 , t_1 , $t_{об}$, приведенная стоимость и др.

Теория массового обслуживания базируется на анализе случайных процессов. При решении тех или иных практических задач в каждом случае должны приниматься индивидуальные решения.

Для оптимизации различных процессов используются методы теории игр, которая рассматривает развитие процессов в зависимости от случайных ситуаций. Теорию игр можно назвать математической теорией конфликтов, связанных с тем, что интересы двух сторон не совпадают. Примером конфликтной ситуации являются, например, спортивные игры. Как правило, теория игр рассматривает конфликтные ситуации при частичном или полном отсутствии данных об обстановке. Поэтому могут быть и случайные ходы, эффект от которых можно оценить в среднем математическим ожиданием. Методы теории игр применяются также не только при исследовании действительно конфликтных ситуаций, но и при решении таких задач, в которых в качестве «противника» выступает, например, природа. Такие задачи обычно возникают при строительстве различных сооружений, в сельском хозяйстве, метеорологии и др. С помощью теории игр можно оценивать наиболее благоприятные и неблагоприятные ситуации и на их основе принимать оптимальное для данных условий решение.

При анализе математического результата, полученного в результате теоретического исследования, часто ставятся задачи оптимизации исследуемых процессов, для чего используются методы оптимизации с математическим программированием: аналитические, градиентные, автоматические с самонастраивающимися моделями.

Оптимизация аналитическими методами состоит в определении экстремального значения некоторой функции $\varphi(x_1, x_2, \dots, x_n)$ в области значений параметров x_1, x_2, \dots, x_n . Однако при оптимизации сложных реальных процессов классические аналитические методы используются редко, и вместо них применяется метод наискорейшего (градиентного) спуска и подъема. Суть метода можно понять из следующего примера. Допустим, что необходимо найти экстремум целевой функции $f(x_1, x_2)$, описывающей некоторую поверхность (рисунок

5.23). Для нахождения экстремума выбирается любая точка поверхности $A_0(x_{01}, x_{02})$, затем определяется наиболее крутое направления подъема или спуска, которое называют градиентом и обозначают \bar{g} . По направлению градиента начинается движение с шагом $c\bar{g}$ к оптимуму (c – постоянная величина, зависящая от точности измерения). В результате достигается новая точка $A_1(x_{11}, x_{22})$, в которой повторяют описанную процедуру до тех пор, пока не определится точка с действительным экстремумом.

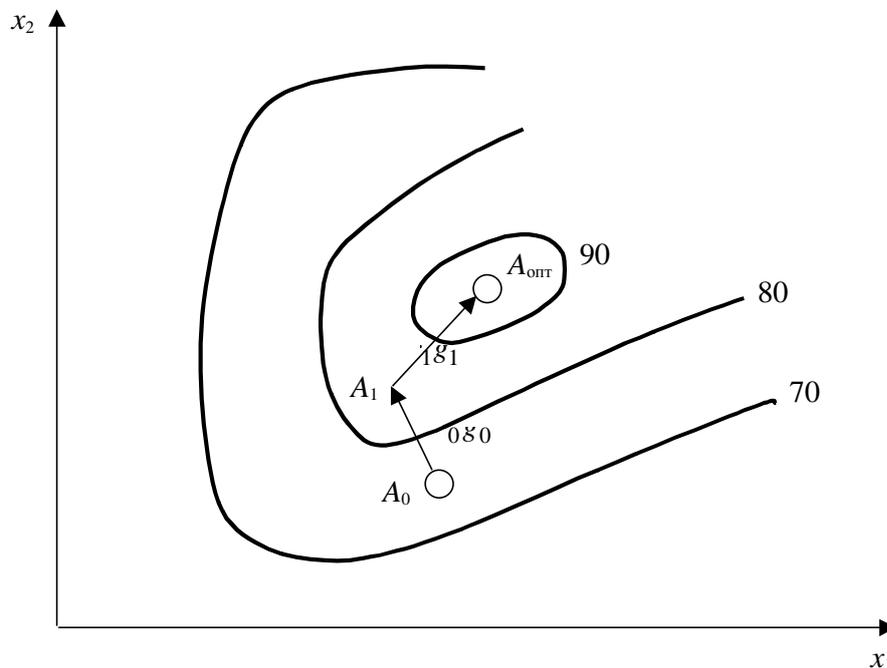


Рисунок 5.23 – Схема движения к оптимуму по градиенту (крутое восхождение)

На практике встречаются задачи оптимизации, когда при нахождении экстремума целевая функция f и граничные уравнения ее области s оказываются линейными. При решении задачи такого класса чаще всего применяются методы линейного программирования, заключающиеся в нахождении экстремума критерия оптимальности в задачах с линейными уравнениями. Целевая функция в таких случаях выражается в виде

суммы $f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n c_i x_i \rightarrow \min(\max)$. Ограничения задаются в виде линейных

неравенств

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{im}x_m \geq b_i;$$

$$x_1, x_2, \dots, x_n \geq 0; i = 1, 2, \dots, m,$$

где a_{ij}, b_i, c_i – константы; $x_1, x_2, x_3, \dots, x_m$ – независимые переменные.

В настоящее время задачи линейного программирования изучены достаточно полно и для многих из них имеются стандартные программы для ЭВМ.

В некоторых случаях приходится использовать нелинейное программирование. Целевая функция в таких случаях записывается в виде суммы линейных слагаемых

$$f(x_1, x_2, \dots, x_n) = \sum_{j=1}^n c_j x_j + \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_i x_j.$$

Среди задач нелинейного программирования встречаются и такие, в которых ограничения не имеют дискретных переменных. В них функции $f(x_1, x_2, \dots, x_n)$ непрерывные и выражаются частными производными. Эти задачи иногда называют классическими задачами оптимизации, поскольку решаются классическими методами на основе дифференциального исчисления.

Среди задач нелинейного программирования встречаются задачи целочисленного линейного программирования. В этом случае в качестве ограничений выставляется особое требование о целостности переменных значений. Задача представляется в виде суммы

$$f(x_1, x_2, \dots, x_n) = \sum_{j=1}^n a_{ij} x_j, \quad i = 1, 2, \dots, m,$$

где $x_j \geq 0, j = 1, 2, \dots, n$ – целые числа.

Решение большого количества технических задач методами линейного и нелинейного программирования обеспечивает определенный экономический эффект.

Некоторые производственные процессы непрерывно изменяются. К числу таких могут быть отнесены процессы управления производством. В связи с изменением условий производства приходится рассматривать все новые ситуации. Решение таких практических задач с учетом различных ситуационных изменений можно осуществить с помощью метода динамического программирования.

Динамическое программирование («динамическое планирование») представляет собой математический метод оптимизации решений, специально приспособленный к многошаговым (или многоэтапным) операциям. Пусть, например, какая-то исследуемая операция представляет собой процесс, развивающийся во времени и распадающийся на ряд «шагов» или «этапов». Некоторые операции расчлняются на шаги естественно (например, при планировании хозяйственной деятельности – хозяйственный год), в других операциях разделение на шаги приходится вводить искусственно (например, подъем температуры, изотермическая выдержка, остывание).

В основу задач динамического программирования положены принципы оптимального управления процессом в соответствии с заданной целью и состоянием системы в рассматриваемый период времени независимо от изменившихся условий, которые привели систему в данное состояние.

Целевая функция выражается в таких случаях суммой

$$\omega = \sum_{k=0}^{N-1} f_0[x(k), u(k)] = \min(\max),$$

где N – общее число интервалов (шагов); $u(k)$ – управляющее воздействие; $x(k)$ – значения координаты в дискретные моменты времени t .

При оптимальном управлении данный функционал должен быть минимизирован (или максимизирован). Оптимальный процесс станет известен, если будут найдены соответствующие значения управляющего воздействия u_0, u_1, \dots, u_{N-1} во все дискретные моменты времени $k = 0, 1, \dots, N - 1$, имеющие определенные ограничения. Чтобы решить задачу динамического программирования, необходимо отыскать минимум (максимум) сложной дискретной функции большого количества переменных. Метод динамического программирования сводит эту задачу к минимизации простых функций в обратном порядке – от конца к началу процесса.

Для оптимизации процесса методами линейного или динамического программирования нет стандартных решений. В каждом конкретном случае применяют свой метод.

Следует иметь в виду, что при решении задач оптимизации могут возникнуть случаи, когда вследствие оптимизации какого-либо одного процесса ухудшится другой. Поэтому при оптимизации необходимо соблюдать так называемую комплексность решения, при которой испытываются все особенности процесса.

Рассматривая задачу по этапам, необходимо оценивать обстановку в целом, которая может меняться в результате оптимизации исследуемого процесса.

Выше были отмечены особенности лишь некоторых математических методов теоретических исследований. Детальное их изучение и получение практического опыта

применения возможно путем ознакомления со специальной литературой в зависимости от профиля исследования.

* * *

Стремление к математической формализации в особенности проявляется в тех областях знания, где прямой эксперимент, позволяющий собрать достаточно полную и объективную информацию об исследуемой реальности, практически невозможен.

Побуждающим стимулом к созданию новой теории является обычно небольшое число фундаментальных фактов; увеличение числа экспериментальных данных, как правило, ничего принципиального не добавляет к нашим представлениям и не облегчает формулировку новой теоретической концепции; после того как концепция сформулирована и модель явления построена, этот «дополнительный» экспериментальный материал в лучшем случае может быть использован для проверки возникшей теории, в худшем случае – оказаться бесполезным.

Теории, не позволяющие осуществить их экспериментальную проверку, должны рассматриваться как находящиеся за пределами собственно науки. Чтобы теория была составной частью науки, необходимо сознавать, что можно в принципе получить факты, побуждающие сомневаться в этой теории. Чтобы теория была научной, она должна быть опровержимой.

Причины выдвижения отрицательного свойства опровержимости в качестве критерия научности теории, возможно, носят психологический характер. Люди, в том числе и ученые, консервативны. Какие бы теории они не имели, они стремятся сохранить их, признавать лишь факты, подтверждающие, а не отрицающие их теории.

Подтверждаемость и опровергаемость – два свойства научной теории. Теории могут пройти множество экспериментальных проверок и все-таки оказаться неверными. Сколько бы проверок ни прошла теория, всегда сохраняется возможность того, что новый эксперимент обнаружит ее недостатки. Поэтому, если говорить о проверяемости теорий, можно утверждать лишь, что успешная проверка увеличивает уверенность. «Доказательство» правильности невозможно.

Отсутствует однозначность и в обратном направлении: теория, не удовлетворившая экспериментальной проверке, отвергается часто, но не всегда. Ни одна из существующих сегодня теорий не удовлетворяет полностью всем экспериментам, проведенным для ее проверки. В каждый момент времени исследователь удовлетворяется той теорией, которая представляется лучшей, и это связано со значительной долей субъективной оценки. Нередко предлагается вполне разумное объяснение того, почему хорошая теория не срабатывает в каких-то конкретных условиях.

Большинство научных гипотез формулируются таким образом, что нет возможности просто и непосредственно проверить их. Вместо этого приходится полагаться на то, что можно назвать косвенным свидетельством. Выводятся определенные следствия из гипотезы, которые могут наблюдаться непосредственно, и затем анализируют, наблюдаются ли они.

Степень уверенности в теории нельзя измерить количественно: она субъективна, и каждый оценивает ее сам. Но легко можно заметить, когда наблюдение оказывает воздействие на силу уверенности исследователя. Иногда это изменение невелико: делается вывод, что теория менее правдоподобна, чем полагали, когда еще не знали о новом экспериментальном факте. Иногда это изменение оказывается резким: от уверенности переходят к полному отрицанию.

6. ИНФОРМАТИКА В НАУЧНЫХ ИССЛЕДОВАНИЯХ

6.1 Основные направления информатики

Термин «*Информатика*» обозначает совокупность дисциплин, изучающих свойства информации, а также способы представления, накопления, обработки и передачи информации с помощью компьютера. Ядро информатики – информационные технологии как совокупность конкретных технических и программных средств, с помощью которых выполняются разнообразные операции по обработке информации.

Использование возможностей научно-технического прогресса во многом зависит от своевременного обеспечения предприятий, учреждений и организаций страны оперативной и полной информацией о достижениях науки и техники; эффективного ее использования в научно-исследовательском, проектно-конструкторском, производственном процессах и при принятии решения на всех уровнях управления.

При создании новой техники в случае неполноты, недостаточной достоверности или неоперативности информации практически невозможно составить представление о лучших мировых и отечественных образцах, в связи с чем уже на стадии проектирования может быть заложена техническая отсталость.

Не менее важное значение имеет задача обеспечения научных исследований удобной для восприятия информацией о важнейших научных достижениях, полученных в прошлом. Методы информатики успешно применяются для создания эффективных информационных систем и составляют основу для автоматизации научных исследований, проектирования, различных производственных процессов.

В настоящее время сформировалось понятие информатики как важной отрасли научного знания, включающей в себе несколько научных дисциплин, связанных с проблемой общения человека с ЭВМ, с созданием компьютерных систем.

Современную информатику рассматривают:

во-первых, как фундаментальную естественную междисциплинарную науку об информации, информационных процессах, системах и технологиях; в этом смысле она расширилась до границ создания общей научной методологии разработки информационного обеспечения процессов управления материальными объектами и интеллектуальными процессами любой природы;

во-вторых, как прикладную дисциплину, так как исторически она возникла и развивалась как инженерная дисциплина, в состав которой входили:

- разработка методов и правил рационального проектирования устройств и систем обработки информации;

- разработка технологии использования этих устройств и систем для решения научных и практических задач;

- разработка методов взаимодействия человека с этими устройствами и системами;

в-третьих, как отрасль народного хозяйства, так как в ней, как и в любой отрасли, можно выделить три относительно автономные составные части («создать», «преобразовать», «потребить»):

- производство, обработка и преобразование информации;

- производство технических средств обработки информации (hardware) и программных средств и систем (software);

- маркетинг, продажа и потребление информационных продуктов и услуг.

Рассматривая информатику как теоретическую и прикладную междисциплинарную науку, можно выделить следующие восемь основных направлений.

Теоретическая информатика. Это – математическая дисциплина. Она использует методы математики для построения и изучения моделей обработки, передачи и использования информации, создает тот теоретический фундамент, на котором строится все здание информатики.

С достаточной степенью условности выделяют пять классов дисциплин, изучаемых теоретической информатикой.

а) Дисциплины, опирающиеся на математическую логику. В них разрабатываются методы, позволяющие использовать достижения логики для анализа процессов переработки информации с помощью компьютеров (теория алгоритмов, теория параллельных вычислений), а также методы, с помощью которых можно на основе моделей логического типа изучать процессы, протекающие в самом компьютере во время вычислений (теория автоматов, теория сетей Петри).

б) Компьютеры оперируют с числами, т.е. с информацией, представленной в дискретной форме. А сами процедуры, реализуемые компьютером, – это алгоритмы, описанные в виде программ. Поэтому необходимы специальные приемы решения задач, допускающие программирование, т.е. перевод в компьютерные программы. Так появились вычислительная математика и вычислительная геометрия.

в) Изучением информации как таковой (т.е. в виде абстрактного объекта, лишенного конкретного содержания), выявлением общих свойств информации, законов, управляющих ее рождением, развитием и уничтожением, занимается теория информации. К этой науке близко примыкает теория кодирования, т.е. изучение тех форм, в которые может быть «отлито» содержание любой конкретной информационной единицы. В теории информации также имеется раздел, специально занимающийся теоретическими вопросами передачи информации по различным каналам связи.

г) Вместо реальных объектов в компьютерах используются их модели – абстрактные, формализованные описания той физической среды, в которой «живет» информация в реальном мире. Переход от реальных объектов к моделям требует знания особых приемов. Их изучает системный анализ. Эта наука занимает пограничное положение между теоретической информатикой и кибернетикой. Такое же положение занимает имитационное моделирование (вычислительный эксперимент), позволяющее воспроизводить реальные процессы на компьютерных моделях, и теория массового обслуживания, которая изучает широкий класс моделей передачи и переработки информации – системы массового обслуживания.

д) Дисциплины, ориентированные на использование информации для принятия решений в самых различных ситуациях. Это – теория принятия решений, изучающая общие схемы, используемые людьми при выборе нужного им решения из множества альтернативных возможностей. Часто такой выбор происходит в условиях конфликта или противоборства. Модели такого типа изучаются в теории игр. Всегда хочется среди всех возможных решений выбрать наилучшее или близкое к нему. Проблемы, возникающие при этом, изучает математическое программирование. При организации поведения, ведущего к нужной цели, решения приходится принимать многократно. Поэтому выбор отдельных решений должен подчиняться единому плану. Способами построения таких планов и их использования занимается научная дисциплина исследование операций. Если же решения принимаются не единолично, а в коллективе, то возникает немало специфических ситуаций: образование партий, коалиций, появление соглашений и компромиссов. Эти проблемы частично изучаются в теории игр, но в последнее время активно развивается теория коллективного поведения, изучающая проблемы коллективного принятия решений.

Кибернетика. Возникла в 40-х годах, когда Норберт Винер выдвинул идею о том, что системы управления в живых, неживых и искусственных системах обладают общими чертами. Установление аналогий обещало создание «общей теории управления». Эта гипотеза не выдержала проверку временем, но накопленная в кибернетике информация принесла большую пользу, а именно: перенос идей и моделей из одних областей в другую

привел к взаимообогащению различных научных направлений, появилась возможность «говорить» на одном, кибернетическом языке различным специалистам. На стыке различных наук возникли: структурная лингвистика, бионика, нейрокибернетика, теория автоматического управления, гомеостатика, распознавание образов и т.д. Кибернетику можно рассматривать как прикладную информатику в области создания и использования АСУ различной степени сложности – от управления станком до управления отраслью промышленности, банковскими системами, системами связи и даже сообществами людей.

Программирование. Это научное направление своим появлением полностью обязано вычислительным машинам. С накоплением опыта программирования появилась основа для создания теоретического программирования: теория языков программирования, системное программирование (создание трансляторов, операционных систем и т.п.), проблемно-ориентированное программирование (создание пакетов прикладных программ, банков и баз данных).

Искусственный интеллект. Пожалуй, именно искусственный интеллект определяет стратегические направления развития информатики. Основная цель работ в области искусственного интеллекта – проникнуть в тайны творческой деятельности людей, в их способности к овладению навыками, знаниями и умениями. Такая цель тесно связывает этот раздел информатики с психологией, лингвистикой, семиотикой. (Семиотика изучает общие свойства различных знаковых систем, способных описывать явления окружающего мира).

Информационные системы. Начало этому направлению положили исследования в области анализа научно-технической документации еще до появления компьютеров. Сейчас в рамках этого направления решаются несколько основных задач:

а) изучение потоков документов с целью их минимизации, стандартизации и эффективной обработки на ЭВМ; изучение потоков информации через средства масс-медиа и ее влияние на общество;

б) исследование способов представления и хранения информации; создание банков данных;

в) создание информационно-поисковых систем;

г) создание сетей хранения, обработки и передачи информации;

д) создание корпоративных информационных систем.

Это направление тесно связано с прикладной лингвистикой и теорией информации.

Вычислительная техника. Развитие современной информатики немыслимо без компьютеров – эффективного инструмента для работы с разнообразной информацией. С другой стороны, развитие и эффективное использование компьютеров невозможно без знания их архитектуры и принципов функционирования. Создание операционных систем, вообще всего программного обеспечения (software), разработка современной элементной базы вычислительных машин (hardware) требует знания теоретической информатики, программирования, искусственного интеллекта и других разделов информатики.

Информатика в обществе. Развитие информационной техники и технологий стало генеральным фактором, определяющим развитие общества. Современное общество можно назвать информационным. Влияние процессов информатизации на человека, на его взаимоотношения изучается в этом разделе информатики.

Информация в природе. Это направление изучает информационные процессы, протекающие в биологических системах. Можно выделить три составляющих этого направления:

- биокибернетика (исследование информационно-управляющих процессов в живых организмах, диагностика заболеваний, оценка биологической активности химических соединений);

- бионика (поиск аналогий между живыми и неживыми системами);

- биогеоэкология (разработка системно-информационных моделей поддержания и сохранения равновесия природных систем и поиск таких воздействий на них, которые стабилизируют разрушающее воздействие человеческой цивилизации на биомассу Земли).

6.2 Понятие информации

Информация – одно из самых фундаментальных понятий в современной науке, наряду с веществом, энергией, пространством, временем. А фундаментальное, т.е. первичное, понятие невозможно строго определить через вторичные, или произвольные понятия.

Например, под информацией в быту понимают любые сведения об окружающем мире и протекающих в нем процессах, воспринимаемые человеком (с помощью органов слуха, зрения, осязания, обоняния, вкуса) или специальными устройствами.

Под информацией в технике понимают сообщения, передаваемые в форме знаков или сигналов.

Под информацией в теории информации понимают нелюбое сообщение или сведение, а только такое сообщение, которое уменьшает существующую до этого неопределенность в той предметной области, к которой оно относится. Позже мы покажем связь информации с мерой неопределенности – энтропией.

Под информацией в теории управления понимают те сообщения (сигналы), которые используются для совершения активного действия, например, управленческого решения.

Под информацией в семантике понимают сведения, обладающие новизной. (Семантика исследует содержание, смысловые аспекты информации).

Под информацией в документоведении понимают все то, что зафиксировано в знаковой форме в виде документа.

Таким образом, в узком смысле информацией можно назвать сведения о предметах, фактах, понятиях некоторой предметной области.

С середины XX века информация рассматривается в широком смысле как общенаучное понятие, включающее в себя как совокупность сведений об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, так и обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом, обмен сигналами между живой и неживой природой, в животном и растительном мире, а также генетическую информацию.

Имея в виду это определение, информацию можно подразделить на:

1) структурную (или связанную), присущую объектам неживой и живой природы естественного или искусственного происхождения. Это объекты (орудия труда, предметы быта, произведения искусства, научные теории и т.п.) возникают путем опредмечивания циркулирующей информации, то есть благодаря и в результате целенаправленных управленческих процессов;

2) оперативную (или рабочую), циркулирующую между объектами материального мира и используемую в процессах управления в живой природе, в человеческом обществе. Поэтому говорят, что человека кроме гравитационного и электромагнитного полей окружает и информационное поле, со стороны которого на него воздействуют различные потоки информации: сенсорный поток, который воспринимается органами чувств через первую сигнальную систему; вербальный (от лат. *verbalis* – словесный) поток устных и письменных слов и структурный поток, компонентами которого являются вода и пища, вдыхаемый воздух.

Преимущество этих двух форм информации очевидна. Генетически структурная информация неживой природы явилась необходимой предпосылкой возникновения оперативной информации и функциональных систем живой природы. Субъект извлекает из объектов неживой природы информацию, включает ее в контур познания и управления. При этом выявляется содержание информации, она приобретает ценность, то есть раскрывается семантический и прагматический аспекты информации.

Оперативная информация всегда предусматривает существование ее источника и ее потребителя (получателей).

На основе представления информации как отражения существующего мира существует и другое – хронологическое – деление информации на виды:

- физическая информация – присуща процессам отражения в неорганической природе (информация механических, физических, химических, ядерных движений материи);

- социальная информация – передается в человеческом обществе в процессе коммуникаций между людьми (массовая информация: мировоззренческая, публицистическая, бытовая, эстетическая, религиозная и т.п.; специальная информация: научная, техническая, экономическая (статистическая), технологическая и т.п.).

Информация, представленная в формализованном виде, удобном для ее сбора, пересылки, хранения, обработки получила название «данные».

Взаимосвязь этих понятий «сообщение», «информация», «данные» можно проиллюстрировать на кругах Эйлера (рисунок 6.1).

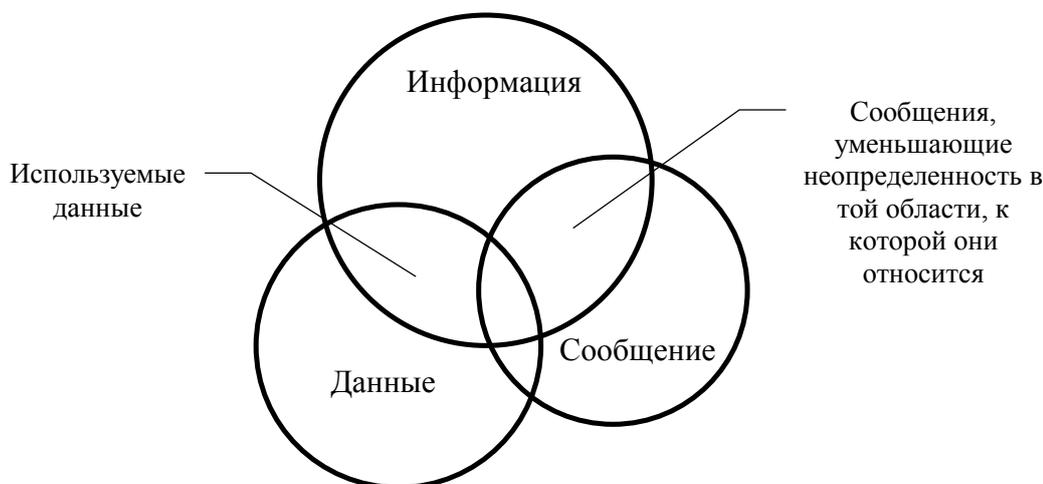


Рисунок 6.1 – Круги Эйлера

В некоторых случаях данные могут оказаться полезными и в их первоначальном, необработанном виде и, таким образом, служить информацией для принятия решения. Но обычно данные предварительно обрабатываются, систематизируются, реорганизуются (часто с проведением необходимых вычислений) и представляются пользователю в удобном виде.

Более того, если раньше побеждал тот, кто умел получать недоступную другим информацию, то теперь ситуация изменилась. Информация наступает на нас со всех сторон, и «иметь информацию» еще не означает «владеть информацией». Сейчас впереди оказывается тот, кто умеет обработать ее быстрее и лучше других.

Совокупность полезной информации, правил и процедур ее обработки, необходимая для получения новой информации о какой-либо предметной области называют знанием.

Знания – результат подготовительной деятельности человеческого интеллекта; их характерной чертой является динамичность и творческий характер. Если же знания отчуждаются от человека, то приобретают автономный закрытый характер, лишаются творческого начала, обретают статичность, тогда они превращаются в информацию (прибегая к образному изложению, информация – это знания минус человек).

Схематично понятие информации и связанных с ней информационных процессов показано на рисунках 6.2 и 6.3.

Знаниями, преобразованными в информацию, оперируют системы искусственного интеллекта, выдавая новые решения. Возвращаясь к человеку и поступая в его распоряжение, достигнутые с помощью компьютера решения, обретают все свойственные знанию качества. Таким образом, между знаниями и информацией происходит неизменный процесс взаимодействия и преобразования.

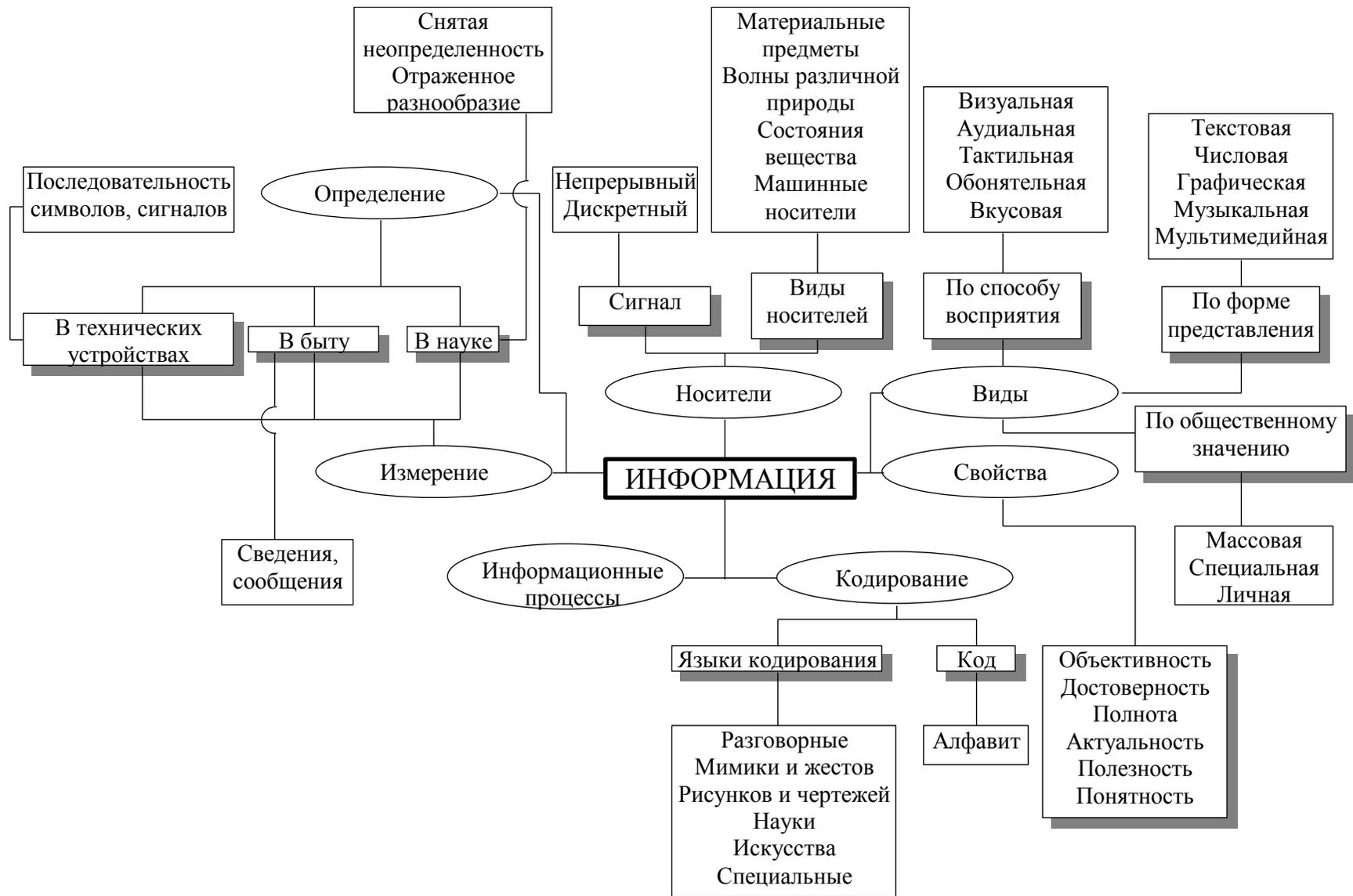


Рисунок 6.2 – Понятие информации

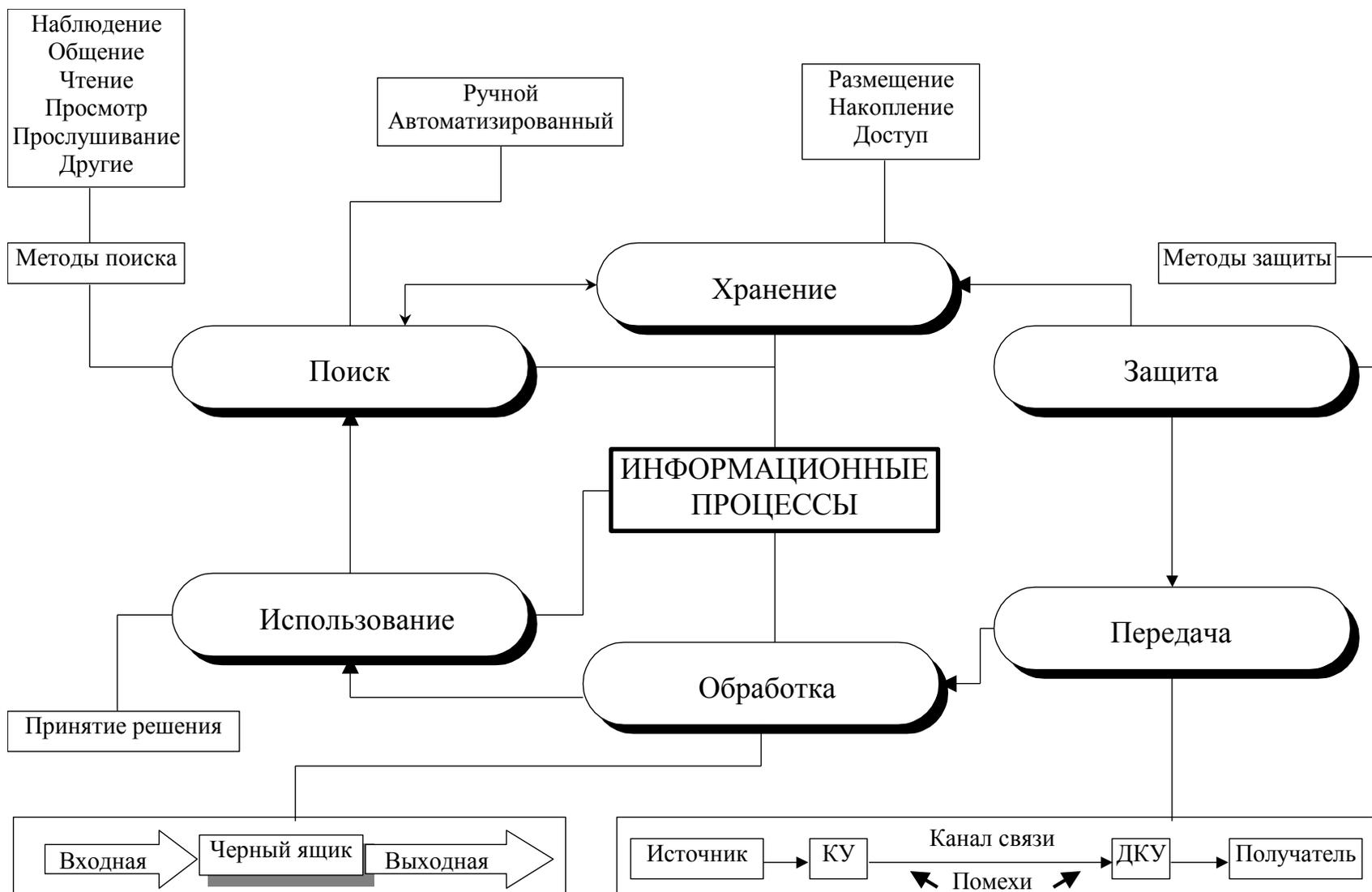


Рисунок 6.3 – Информационные процессы

6.3 Цели информатизации

Анализируя жизнь общества на разных ступенях его развития с точки зрения того, что определяло в тот или иной период его выживание и прогрессивное движение вперед, можно заметить, что вплоть до XVI века деятельность общества в целом и каждого человека в отдельности была направлена на овладение веществом, т.е. познание свойств вещества и изготовление сначала примитивных, а потом все более сложных орудий труда, вплоть до механизмов и машин, позволяющих изготавливать потребительские ценности.

Затем, в процессе становления индустриального общества на первый план вышла проблема овладения энергией – сначала тепловой, затем электрической и, наконец, в XX веке – атомной. Овладение энергией позволило перейти к массовому производству потребительских ценностей и, как следствие, повышению уровня жизни людей и изменению характера их труда.

При этом на протяжении тысячелетий человечество стремилось постичь тайны мироздания, составляя его модели, выделяя общие закономерности, пытаясь увидеть некоторое единство в разнообразии материальных объектов. Затем физические, химические, биологические процессы стали рассматриваться с позиций передачи и преобразования энергии. Но постепенно наука стала перед фактом невозможности детально описать сложные системы в технике, биологии, обществе на языке материально-энергетических моделей.

Фундаментальной чертой современной цивилизации явился лавинообразный рост производства, потребления и накопления информации (так называемый информационный взрыв) и осознание того, что информация наряду с веществом и энергией является основой мироздания. Вещество отражает постоянство материи, энергия отражает движение, изменение материи, информация отражает структуру, организацию материи.

В настоящее время наиболее развитые страны мира находятся на завершающей стадии индустриального этапа развития общества. В них осуществляется переход к следующему этапу, который, как и соответствующее общество, назван информационным. В России этот переход от индустриального общества к информационному получил название информатизации.

Информатизация – это реализация комплекса мер, направленных на обеспечение полного и своевременного использования достоверной информации и знаний во всех общественно значимых видах человеческой деятельности на основе современных информационных технологий.

К этим мерам относятся:

- информационно-вычислительное обеспечение экономических и социальных процессов;
- информационная поддержка процессов принятия решений;
- формирование и развитие информационных потребностей населения;
- формирование и поддержка условий, обеспечивающих развитие процесса информатизации общества.

Научные основы информатизации развиваются в двух направлениях.

Первое направление – фундаментальные и прикладные исследования, связанные с созданием новых информационных средств и технологий. Теоретическую основу этого направления формирует современная информатика как фундаментальная междисциплинарная наука об информации, информационных процессах, системах и технологиях.

Второе направление формируется в настоящее время. Оно объединяет исследования, позволяющие анализировать и прогнозировать пути и последствия информатизации общества.

Отметим, что следует различать информатизацию и компьютеризацию. Второе более технично, хотя современное содержание информатизации немислимо без компьютеризации.

Компьютеризация – это процесс развития индустрии компьютерных продуктов и услуг и их широкого использования в обществе; оснащение предприятий, учреждений и учебных заведений страны вычислительной техникой и повышение общеобразовательного уровня населения в области ее применения.

К основным сферам информатизации общества относятся: социальная сфера, материальное производство и управление.

Социальная сфера – это главный объект информатизации. Информатизация этой сферы непосредственно направлена на формирование и удовлетворение потребностей населения, информационное обеспечение социальных процессов, улучшение быта всех членов общества и повышение качества предоставляемых услуг, а в конечном результате – на коренное улучшение общественной деятельности и жизни человека.

Основной целью информатизации сферы материального производства является информационное обеспечение технического перевооружения отраслей общественного производства путем внедрения высоконадежных, эффективных автоматизированных средств труда, комплексной автоматизации на их базе технологических и производственных процессов, создания гибких перестраиваемых модулей, участков и производств. Информатизация должна охватывать все стадии жизненного цикла создаваемой продукции: исследование, проектирование, производство, сбыт и эксплуатация. Современный этап развития научно-технической революции характеризуется тем, что производство и продукция становятся не только наукоемкими, но и информационно емкими.

Информатизация сферы управления играет особую роль, так как она не только повышает эффективность управления на всех его уровнях, но и позволяет увеличить эффективность целенаправленной деятельности человека в других сферах, в том числе такой, как информатизация общества. Информатизация позволяет систематизировать и учесть многие факторы, которые прежде игнорировались: их нельзя было предоставить в обозримой форме, так что смысл информатизации заключается не в высвобождении работников, а в оптимизации принимаемых на основе информации управленческих решений.

Информатизация процессов управления на любом уровне территориальных, отраслевых и межотраслевых структур позволяет более полно учитывать как интересы региона, области, городов, территорий, отдельных предприятий и отраслей, так и интересы страны в целом, как единого народнохозяйственного и социально-политического образования, что особенно актуально в наше непростое время.

В сфере управления необходимо выделить управление действиями в чрезвычайных ситуациях, для которого информатизация является необходимым условием его эффективности, так как позволит:

- осуществлять непрерывное слежение за экологической обстановкой, состоянием потенциально опасных зон и объектов;
- заблаговременно прогнозировать моменты возникновения чрезвычайных ситуаций, ход их развития и последствия;
- вырабатывать рекомендации по предотвращению, локализации чрезвычайных ситуаций и ликвидации их последствий;
- оперативно решать задачи оптимального перераспределения ресурсов для предупреждения и локализации чрезвычайных ситуаций и ликвидации их последствий с учетом изменяющейся обстановки;
- использовать автоматические роботы и телеоператоры для выполнения операций в экстремальных условиях при ликвидации чрезвычайных ситуаций и их последствий.

Таким образом, развитие потребности в информатизации общества и ее необходимость формируется объективным развитием самого производства и в особенности тех сфер, где развиваются новые технологии, где высока потребность в применении научных знаний, где по этой причине информационный процесс является необходимым условием нормального функционирования производства.

6.4 Основные проблемы информатизации

По своей глубокой сути информатизация представляет собой процесс преобразования человеком среды своего существования, биосферы в ноосферу, результатом которого будет создание высокоразвитой инфосферы как области производства и потребления информации. Этот процесс затрагивает как среду обитания, так и общество, самого человека. Глубина совершаемых преобразований порождает проблемы, от своевременного решения которых зависит не только ход информатизации, но и при неблагоприятном исходе – существование общества и человека как биологического вида.

Рассмотрим основные проблемы информатизации.

Проблемы индустриализации производства и обработки информации. Это проблемы создания и развития крупного машинного производства в информационной среде, которые порождены противоречием между необходимостью своевременного использования во всех сферах человеческой деятельности больших объемов высококачественной информации и невозможностью оперативно формировать такие объемы с помощью традиционных информационных средств, технологий и средств связи. В эту группу входят материально-техническая, технологическая проблемы и проблема связи.

Материально-техническая проблема заключается в преодолении разрыва между существующим состоянием материально-технического обеспечения информационной сферы и необходимым уровнем этого обеспечения.

Технологическая проблема обусловлена отсталостью не только информационных технологий, но и технологий в тех областях экономики, которые должны обеспечивать процесс развития инфосферы.

Проблема связи порождается противоречием между необходимостью в информационном обществе связывать каждого с каждым, обеспечивая тем самым своевременную высококачественную передачу каждому члену общества всей необходимой информации, и невозможностью выполнить это при современном состоянии сетей связи в регионах.

Психологические проблемы. Готовность населения к информатизации, к использованию получаемых в ходе информатизации результатов. Существование этой проблемы обусловлено следующими факторами:

- практическим отсутствием готовности населения к информатизации ввиду низкого уровня информационной культуры, и, прежде всего, компьютерной грамотности населения;
- недостаточным уровнем общей культуры населения и, как следствие, низкими информационными потребностями и отсутствием желания их развивать;
- консерватизмом значительной части населения, нежеланием затрачивать усилия на восприятие нового.

Психофизиологическая проблема, т.е. психическая и физиологическая совместимость человека и новой информационной техники, проблема воздействия на человека информационных технологий.

Правовые проблемы возникают в связи с превращением информации в основной ресурс развития общества, необходимостью правовой регламентации производства, обработки и использования этого ресурса и отсутствием таковой в настоящее время.

В эту группу входят:

- общие проблемы правового регулирования в информационной сфере: проблемы правовой охраны программных средств, правового регулирования договорных отношений в информационной сфере, экономического и правового режима информационных ресурсов;

- проблемы правонарушений в информационной сфере: проблемы имущественной ответственности в информационной сфере, применения новых информационных средств и технологий в доказательном праве «компьютерных преступлений»;

- проблемы развития инфосферы в области права: проблемы применения экспертных систем в этой области, криминалистического моделирования преступной деятельности, компьютерной подготовки юристов и повышения их информационной культуры.

Экономические проблемы возникают в связи с переходом к экономике иного вида, экономике «информационного» общества. Информация стала товаром, причем товаром особым, который не исчезает у продавца при продаже и которого не становится меньше при его потреблении. Это приводит к особым отношениям между субъектами рынка информации.

Социальные проблемы обусловлены коренным изменением образа жизни членов общества под воздействием информатизации.

Рассмотрим основные из них.

Коммуникационные. Высокоразвитая инфосфера, несомненно, активизирует коммуникационный процесс в обществе, основные функции которого состоят в достижении социальной общности при сохранении индивидуальности каждого ее индивида. Человек как личность формирует себя только в процессе совместной деятельности и общения с другими людьми.

Поэтому усиление коммуникативного аспекта, предоставляемого новыми информационными технологиями, создает возможности для дальнейшего развития человека как личности. С другой стороны, тревогу социологов и психологов вызывает обратный процесс, постепенно набирающий силу в связи развитием средств мультимедиа и сетевых технологий – десоциализация личности, уход в виртуальный мир.

Проблема гуманизации возникает в связи с тем, что развитие инфосферы имеет не только позитивные, но и негативные последствия. Информационный экстремизм и информационное вторжение чужих идеалов и ценностей приобретает особый размах, особенно в связи с взрывообразным расширением Интернета и увеличением числа домашних компьютеров. Кстати, проблема «культурного империализма» широко обсуждается в Европе, где принят ряд законодательств, запрещающих некоторые виды кино-, видео-, теле-, книгопродукции (пропаганда фашизма, расового превосходства, насилия и жестокости).

Так как наша страна – страна крайностей (очень часто работает «принцип маятника» – оптимальное положение проходит с максимальной скоростью), то в России после снятия «железного занавеса» воцарился информационный беспредел. Правда, в последнее время, хотя и с трудом, информационный процесс законодательными мерами стараются направить в цивилизованное русло.

По мнению исследователей в России ведущая роль в развитии информатизации принадлежит социокультурным факторам. Перечислим некоторые чисто российские феномены, тормозящие развитие информатики в стране.

1) Информационное безразличие и нетребовательность, отсутствие привычки к информационной обеспеченности. Информатизация – это осознанная организация

информационных процессов, то есть надо хотеть, иметь интерес их организовывать. Информационные потребности удовлетворяются на основе свободного выбора, к которому мы не были приучены.

2) Слабое осознание населением социальной значимости информации. В командно-административной системе знания и профессионализм обществом не поощрялись, воспринимались как нечто несущественное с точки зрения карьеры и оплаты труда, что привело к девальвации знаний как таковых.

3) Существование (особенно в доперестроечной России) эффекта тезаврирования информации, превращение ее в «сокровища»: много информации хранится, но мало используется. Это относится, в первую очередь, к работникам сферы управления, которым было невыгодно делиться информацией.

4) Другой, связанный с этим феномен: в нашей культуре мало развита прагматическая функция знаний. Французский философ Пьер Тейяр де Шарден в книге «Феномен Человека» (1951 г.) выделил две функции знаний: «знать, чтобы знать» и «знать, чтобы мочь». Эта вторая функция знания сознательно подавлялась в ориентированной на закрытость командно-административной системе, она плохо развивалась и в традиционной православной культуре. В результате функция использования знаний слабо развита в обществе, усложняется ситуация и тем, что структуры потребления знаний устойчивы и консервативны.

Перечисленные проблемы порождают, в свою очередь, кадровые, финансовые, организационные и научные проблемы.

Кадровые проблемы связаны с необходимостью не только готовить кадры для развития инфосферы и эффективного использования получаемых результатов, но и проводить профессиональную переориентацию работников тех профессий, которые окажутся излишними в процессе создания высокоразвитой инфосферы.

Финансовые проблемы возникают в связи с высокой стоимостью информатизации, отсутствием централизованных средств на ее проведение и необходимостью искать и создавать источники средств, способные поддерживать желаемые темпы создания высокоразвитой инфосферы.

Организационные проблемы связаны с необходимостью создания таких структур и механизмов, которые на практике обеспечивали бы организацию и проведение процесса развития инфосферы.

Научные проблемы обусловлены неразработанностью научного фундамента информатизации и, в первую очередь, концептуальных основ методов научного обоснования и экспертиз программ и проектов развития инфосферы, научного сопровождения этого процесса в регионах.

Информатизация общества и решение порождаемых этим процессом проблем могут осуществляться различными путями.

Стихийная самоорганизация процесса информатизации. Этот путь характерен для общественных процессов, связанных с изменением условий жизни и адаптацией общества к новым условиям. Такая адаптация требует организационных перестроек в обществе, затрагивающих его моральные и нравственные основы. Эти основы относятся к наиболее консервативным элементам общества, и их изменение воспринимается членами общества достаточно болезненно. Включение стихийных механизмов регуляции позволяет несколько сгладить остроту восприятия изменений, но делает сам процесс длительным и, как правило, приводит к значительному перерасходу ресурсов.

Централизованное управление процессом информатизации. Этот путь не может быть реально осуществлен, так как рассматриваемый процесс является несколько сложным, что практически относится к неуправляемым объектам.

Направляемая информатизация. В этом случае саморазвитие процесса протекает в условиях действия системы ограничений и стимулов, определяющих границы

существования процесса и желательные направления его развития. Это позволяет, сохраняя все преимущества самоорганизации и саморазвития процесса, сократить время его протекания и избежать при этом излишних затрат.

Развитые страны и регионы, первыми начавшие переход к информационному обществу, могут позволить себе первый путь решения возникающих при этом проблем, так как для них фактор времени не столь существенен. Они и так являются первыми в этой сфере.

Регионы, отставшие в развитии информационной сферы, должны выбирать третий путь, так как значительное отставание в создании высокоразвитой инфосферы может привести к безнадежному отставанию «навсегда» в социально-экономическом развитии региона и исключения его из общего процесса цивилизации, по крайней мере, в обозримом будущем.

Выбор третьего пути информатизации общества предопределяет необходимость правильного понимания и осознания сущности, целей и принципов; выделения основных направлений, этапов и сфер информатизации; формирования ее организационно-экономических механизмов; прогноза результатов информатизации и предотвращения или нейтрализации ее негативных последствий.

Таким образом, глобальной целью информатизации является обеспечение уровня информированности населения, обусловленного целями социально-экономического развития страны. Уровень информированности населения определяется полнотой, точностью, достоверностью и своевременностью предоставления информации, необходимой каждому члену общества в процессе выполнения им всех общественно значимых видов человеческой деятельности.

Исходя из целей и проблем, основными направлениями информатизации общества следует считать:

- проведение фундаментальных и прикладных исследований по информатике и информатизации;
- создание и развитие материально-технической базы информатизации;
- совершенствование существующих, разработка, развитие и применение новых информационных технологий, создание технологической базы информатизации;
- перевод производства программных средств на промышленную основу, создание индустрии программных средств;
- создание и развитие информационной инфраструктуры;
- создание и развитие индустрии переработки информации;
- информатизация всей системы общего и специального образования;
- подготовку, введение и корректировку правовых и хозяйственных норм, обеспечивающих требуемые темпы и направления информатизации;
- участие в международном сотрудничестве и разделении труда в сфере информатизации.

6.5 Новые информационные технологии

Для того чтобы из материальных ресурсов получить конечный продукт, необходимо подвергнуть их переработке по соответствующей технологии.

Технология – это совокупность методов обработки, изготовления, изменения состояния, свойств, формы сырья, материалов в процессе производства. Аналогичное определение можно дать и применительно к процессу обработки данных.

Информационная технология – совокупность средств и методов обработки данных для получения информации о состоянии объекта, процесса, явления. На всех этапах развития общества (таблица 6.1) информационные технологии обеспечивали

информационный обмен между людьми, коллективами, институтами общества, отражали соответствующий уровень и возможности систем регулирования, хранения, обработки и передачи информации и, по существу, являлись синтезом методов оперирования человека с информацией в интересах той или иной сферы его деятельности.

Таблица 6.1 – Основные этапы развития информационных технологий

Основные вехи ИТ	Время от наших дней (годы)	Общая характеристика технологий	Уровень технологических знаний	Типовой производственный процесс
Наскальные изображения	20-30 тыс.	Домашний этап: первые инструменты счета из кости и камня, календарь, домеханические часы, компас, бумага, книги	Ручная технология: регистрация знаний	Ремесленное производство уникальных и малосерийных изделий
Первая информационная технология				
Книгопечатание	500	Машинный этап: печатный станок, механические часы, телеграф, фото, телефон, фонограф, радио, кино, магнитофон, телевидение, ЭВМ	Машинная технология: тиражирование и распространение	Механизирование массового производства стандартизируемых изделий
Телеграф	150			
Телефон	---			
Фото	---			
Радио	---			
Телевидение	90			
ЭВМ	50			
Вторая информационная революция				
Персональные ЭВМ	20	Симбиотический этап: «игровая компонента» ПЭВМ; индивидуальные рабочие станции, локальные сети, системы индивидуального доступа к отраслевым региональным и мировым информационным ресурсам	Человеко-машинная технология: автоматизация знаний	Гибкое автоматизирование производства уникальных и малосерийных изделий
Третья информационная революция				
Новые информационные и телекомму-	10	Интеграционный этап: мировая информационная сеть Интернет,	Технология «клиент – сеть»	Распределенная обработка информации

никационные технологии		корпоративные информационные системы, интегрированные пакеты прикладных программ		
------------------------	--	--	--	--

В настоящее время основным техническим средством переработки информации служит персональный компьютер. Его внедрение в информационную среду ознаменовало новый этап развития информационных технологий и, как следствие, изменение ее названия на «новые информационные технологии», причем «новое» подчеркивает не эволюционный характер этой технологии, а новаторский, революционный.

Новая информационная технология (НИТ) – это информационная технология, использующая персональный компьютер, компьютерные сети и средства связи, для которой характерно наличие «дружественной» среды работы пользователя.

С появлением компьютера появилась возможность автоматизации элементов умственного труда в результате освобождения человека от рутинных операций и замыкания части информационных потоков на компьютер. Современный компьютер дает возможность накапливать необходимую информацию: записывать в памяти рефераты книг, статей, доклады, отчеты, результаты исследований и извлекать эту информацию по мере необходимости.

На базе методов информатики – как теоретической, так и прикладной, – создаются компьютерные сети с базами данных, автоматизированные информационные системы, системы электронной почты и телеконференций, автоматизированные системы управления, системы комплексной автоматизации административно-управленческой деятельности, составляющие существо новых информационных технологий. Современные компьютерные сети и информационные системы – как локальные, так и мировые – позволяют осуществлять глобальный поиск информации и распределенную ее обработку.

Основу новых информационных технологий составляют три технических достижения:

- 1) появление новой среды накопления информации на машиночитаемых носителях;
- 2) развитие средств связи, обеспечивающих доставку информации практически в любую точку земного шара без существенных ограничений во времени и расстоянии, широкий охват населения средствами связи;
- 3) возможности автоматизированной обработки информации с помощью компьютера по заданным алгоритмам.

Естественно, новые информационные технологии возникли не на пустом месте. Научно-технический прогресс XX века, подстегнутый информационным взрывом, инициировал самые разнообразные исследования, приведшие, в конечном счете, к созданию наиболее сложной из современных технологий – информационной.

Эти исследования можно объединить в несколько групп.

К первой из них относятся работы по математической логике, начатые в XIX веке и приведшие к созданию в настоящее время обширного пакета компьютерных логик, с помощью которых разрабатываются языки программирования и интегральные схемы.

Ко второй группе относятся исследования по структурной грамматике, фонологии и лингвистической семантике. Они, в конечном счете, предоставили нам возможность общения с компьютером на естественном языке, создание систем автоматического перевода и разработки языков программирования высшего уровня.

К третьей группе относятся исследования по радио- и электротехнике. К ним, прежде всего, принадлежат создание радио, телефонной связи, телевидения и оптоволоконной технологии.

К четвертой группе относятся исследования в области физики твердого тела, физики полупроводников, кристаллографии, нелинейной оптики, сверхпроводимости.

К пятой относятся теоретические исследования и практическое создание сверхчистых материалов.

К шестой группе – исследования по психологии мышления, чувственного восприятия, синтезу зрительных и слуховых образов.

К седьмой – исследования по теории управления, теории информации и теории связи, проектирования больших систем.

И, наконец, к восьмой – теоретическая и практическая разработка ряда вспомогательных технологий, обеспечивающих создание беспредельных по сложности изготовления и вместе с тем дешевизне элементов аппаратных систем.

Приведем примеры формирования НИТ.

На основе сочетания компьютера, устройств долговременной памяти, электронной коммутируемой телефонной сети и модифицированных телевизионных приемников организована система видеографического информационного обслуживания, известная под названием видеотекст.

Другим примером информационной технологии может служить система электронной почты, для формирования которой необходимы компьютер, устройства долговременной памяти, система факсимильной передачи изображений и сеть передачи данных. Электронная почта является новым видом информационного обслуживания, способным заменить в будущем традиционную почту.

В полной мере вобрала в себя все черты современной информатики техника телеконференций. В отличие от обычного видеотелефона участники телеконференции могут пользоваться необходимыми базами данных, передавать друг другу любой вид информации и выполнять в случае необходимости вычислительные работы для обоснования принимаемых решений. Организация телеконференций невозможна без широкополосных линий связи, телевидения, компьютерных информационных сетей. Все эти достижения науки и техники позволили организовать новую информационную технологию, раздвинувшую стены конференц-зала и значительно повысившую эффективность конференций, совещаний и т.п.

Еще один пример эффективных НИТ – ситуационные центры в сфере высшего руководства. Это комплексы, позволяющие визуализировать, представив в графическом виде, огромные объемы информации, которые человек должен в течение короткого времени осмыслить, проанализировать и принять решение.

Сегодня в мире существует несколько сотен таких центров. В России они созданы в резиденции Президента, в Совете безопасности при Президенте, в МЧС, при Правительстве РФ, создается около 20 региональных центров.

Что входит в их состав? Например, в ситуационном центре в резиденции Президента (введен в феврале 1996 г.) три экрана 1.5×2 м, более десятка рабочих станций (студии нелинейного монтажа, графические станции, компьютеры для подготовки презентаций), мощный сервер данных, хранящий огромные объемы информации, а также набор различных инструментальных средств, позволяющих обрабатывать информацию и предоставлять ее Президенту.

Информационные технологии, используемые в таких центрах, можно разделить на три группы. Первая – системы обработки символьной информации, в частности системы сбора данных и оперативной аналитической обработки. Вторая – технологии, в которых применяются модели и методы, присущие конкретным предметным областям. Третья – когнитивно – графические технологии, основанные на визуализации образов.

СПИСОК ЛИТЕРАТУРЫ

1. Грушко И.М., Сиденко В.М. Основы научных исследований. Учебное пособие для вузов. Харьков: Вища школа. Изд-во при Харьк. ун-те, 1977. – 199 с.
2. Основы научных исследований. Учебное пособие. Под ред. В.И.Крутова. М.: Высшая школа, 1989. – 400 с.
3. Чкалова О.Н. Основы научных исследований. Киев: Вища школа, 1976. – 195 с.
4. Голдстейн М., Голдстейн И. Как мы познаем. М.: Знание, 1984. – 256 с.
5. Коваленко Е.С., Киселев О.Н., Шарыгин Г.С. Основы научных исследований: Учебное пособие для вузов. Томск: Изд-во Томского ун-та, 1988. – 192 с.
6. Зельдович Я.Б., Хлопов М.Ю. Драма идей в познании природы (частицы, поля, заряды). – М.: Наука. Гл. ред. физ.-мат. лит., 1988. – 240 с. – (Б-чка «Квант»; Вып. 67).
7. Мышкис А.Д. Элементы теории математических моделей. – М.: Физматлит, 1994. – 192 с.
8. Горстко А.Б. Познакомьтесь с математическим моделированием. – М.: Знание, 1991. – 160 с.