КИПР
ТУСУР

Кафедра конструирования
и производства радиоаппаратуры

В.П. Алексеев, Д.В. Озёркин

ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ И ПАТЕНТОВЕДЕНИЕ

Учебное пособие

ТОМСК 2012

Министерство образования и науки Российской Федерации

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

УТВЕРЖДАЮ
Заведующий кафедрой КИПР
_____________В.Н. ТАТАРИНОВ
“___” __________20___ г.

В.П. Алексеев, Д.В. Озёркин

ОСНОВЫ НАУЧНЫХ ИССЛЕДОВАНИЙ И ПАТЕНТОВЕДЕНИЕ

Учебное пособие

2012

Рецензент: профессор, д.ф.-м.н. Кузнецов Г.В.

Эксперт СибРОУМО: зав. кафедрой приборостроения Красноярского государственного технического университета, д.т.н., профессор Громыко А.И.

Алексеев В.П., Озёркин Д.В.
Основы научных исследований и патентоведение. Учебное пособие для студентов, обучающихся по специальностям 210201 «Проектирование и технология РЭС», 160905 «Техническая эксплуатация транспортного радиооборудования».
Томск: Томский государственный университет систем управления и радиоэлектроники, 2012. – 171 с.

В пособии изложен лекционный материал по методологии и организации научных исследований, а также по вопросам оформления заявок на патенты и по защите авторских прав изобретателей. В процессе подготовки пособия использованы монографии, учебники, учебные пособия и дидактические материалы, опубликованные ранее в российских изданиях различных ВУЗов. Основой для пособия послужил опыт преподавания указанного курса на радиоконструкторском факультете ТУСУР в период с 1977 г. по настоящее время в рамках программы целевой интенсивной подготовки (ЦИПС) для студентов-радиоконструкторов-технологов РЭС. При подготовке пособия использовались литературные источники, приведенные в разделе «Рекомендуемая литература», а также результаты научных исследований, полученные авторами.

СОДЕРЖАНИЕ

1 МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ПОЗНАНИЯ И ТВОРЧЕСТВА ПРИ ПРОЕКТИРОВАНИИ РЭС

1.1 Основные понятия и определения науки

1.1.1 Понимание

Данное учебное пособие может оказаться полезным как студентам, так и молодым ученым, давая возможность взглянуть на научный метод познания под более широким углом зрения, чем тот, который они имеют в своей практической работе в конкретной научной области. Каждому живущему в XXI веке, когда наука играет такую важную роль, необходимо иметь представление о научном методе: как делаются научные открытия, как возникают теории, как осуществляется их проверка, и почему они принимаются или отвергаются.

Наука является сферой деятельности, характеризуемой тремя свойствами:

1. Она представляет собой поиск понимания, т.е. чувства, что найдено удовлетворительное объяснение какого-то аспекта реальности.

2. Понимание достигается посредством формулировки общих законов или принципов – законов, приложимых к возможно более широкому классу явлений.

3. Законы или принципы могут быть проверены и подтверждены экспериментально.

Поиск законов и закономерностей развития природы, техники и общества – основная цель науки. Большое значение при поиске закономерностей природы имеет понимание. Оно, очевидно, субъективно: что удовлетворит одного, не удовлетворит другого; разные культуры обладают разными стандартами того, что можно считать удовлетворительным объяснением; то, что удовлетворяло людей 100 лет назад, может теперь оказаться непригодным. Несмотря на всю расплывчатость и неопределенность этого понятия, субъективное чувство удовлетворения от понимания аспекта действительности оказывается очень сильным и является одной из серьезных побудительных причин для занятий наукой. Понимание – субъективно, тогда как открытые законы природы – объективны и приносят пользу всему человечеству, независимо от понимания их отдельным индивидуумом.

Понимание, которого мы ожидаем от науки, выражается в виде законов или принципов, позволяющих предсказывать, что произойдет, и выяснять причину произошедшего. Под общностью подразумевается свойство принадлежности к возможно более широкому классу явлений. Как можно меньше законов должны покрывать как можно больше явлений.

Наука – это поиск единства в разнородном, сходства в том, что кажется совершенно несхожим. Чем более общими будут наши законы, тем большее единство мы вскрыли.

Отличительная черта науки заключается в необходимости подвергать наши объяснения экспериментальной проверке.

Именно экспериментальная проверка, признание того, что мы можем изменить свое мнение, если нас к этому принудят факты, является особенностью науки.

Чтобы факты могли нас заставить изменить мнение, прежде всего, должны существовать сами факты: заинтересованные наблюдатели должны иметь возможность прийти к согласию в том, что является фактом, а что нет (эта проблема далеко не так проста, как кажется). Далее, факты должны влиять на степень нашей уверенности в правильности теории. 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.

1.1.2 Наука и гуманитарное знание

Не преуменьшая различия между использованием критериев оценки теории и критериев точной экспериментальной проверки, отметим три фактора, делающие эти различия менее резкими, чем может показаться.

Во-первых, исследования в гуманитарных областях так же неукоснительно основаны на фактах, как и исследования в «точных» науках.

Во-вторых, принятие той или иной из конкурирующих теорий даже в естественных и технических науках не всегда происходило на основе одних лишь экспериментов, по крайней мере, не в том идеализированном смысле, в котором понимается экспериментальная проверка. Несомненно, если две теории согласуются во многих областях, но в некоторых не согласуются и эксперименты показывают, что там, где они не согласуются, одна из них всегда дает удовлетворительный ответ, а другая – всегда неверный, тогда легко принять решение в пользу первой из теорий. На практике, однако, ни одна теория не объясняет все возможные экспериментальные факты, и всегда имеется достаточно большая свобода в определении того, какие эксперименты важны для проверки теорий. Аналогичным образом в гуманитарном знании разрешение спорного вопроса ищется в консенсусе (договоренности) специалистов, использующих критерии объяснительной силы, непротиворечивости и эффективности.

В-третьих, как наука, так и гуманитарное знание требуют постоянной критической оценки результатов. Критерии экспериментальной проверки в науке представляют собой осознание постоянного требования подходить критически к своим воззрениям, всегда задаваться вопросами: как мы познаем? Почему в этом уверены? Не можем ли мы ошибаться? Если мы ошибаемся, то как можно это установить? И в тех дисциплинах, которые не относятся к научному знанию, и критерием их истинности является не экспериментальная проверка, а более неясные и трудно формулируемые требования, необходимо ставить эти же вопросы. Как в науке, так и в гуманитарном знании мы должны постоянно критически оценивать результаты и искать более верное и глубокое объяснение. Результат исследований часто проявляется в виде полученных фактов.

Факты окружают нас – это вещи вокруг нас, которые мы можем 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.

Существует совокупность понятий, применяемых в науке, которая используется при анализе отдельных научных открытий. В дополнение к не осознаваемой культурной компоненте в восприятии исследователя большинство научных фактов содержит компоненту, сознательно выбранную и доступную для анализа, а именно – существующие знания и теории.

Таким образом, имеется существенное различие между общепринятым взглядом на «факты» как на жесткие, неизбежные, неизменные вещи и реальным положением в науке, где вещи, называемые фактами, менее определены. Факты имеют обусловленную культурой компоненту и до некоторой степени создаются имеющимися теориями и поэтому подвержены изменениям, если меняются сами теории.

1.1.3 Использование фактов в научных исследованиях

Результатом обыденного представления о фактах как о фиксированных и неизменных вещах, является неверный взгляд на то, как факты используются в науке. Часто говорят о том, что, например, теории должны согласовываться со всеми фактами; если этого нет, то теория неверна. Это мнение не имеет ничего общего с тем, как на самом деле развивается наука. Это утверждение не принимает во внимание, что реальный мир состоит из бесчисленного количества фактов: каждая песчинка на берегу отличается от другой; ни одна из них не обладает одинаковыми массой или формой. Тем не менее, масса и форма каждой песчинки – это отдельные факты. Ни от одной теории нельзя ожидать, что она объяснит такое огромное количество вещей; приходится удовлетворяться гораздо меньшим.

Если предположить, что выбрана группа фактов из их бесконечной совокупности, существующей в мире, – должна ли теория объяснить, по крайней мере, все эти факты? Ответ будет также отрицательным. Те факты, которые выбраны, в большей степени определяются теорией или предубеждением относительно того, какие факты важны, а какие – нет. Если бы выбор теории состоял в простом выяснении, какая из них согласуется с ограниченным количеством фактов, то тогда конкурирование теорий закончилось бы очень быстро. Причина того, что оно продолжается так долго, в том, что разные ученые имеют разные мнения о том, какие факты важно объяснить.

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. Споры, что важнее – теория или эксперимент – давно уже прекращены, так как развитие науки возможно только при гармоничном сочетании теории с экспериментом.

1.1.4 С чего начинается наука

Фундаментальная наука начинается с желания раскрыть тайны Природы; прикладная наука начинается с выявления проблемы и веры в возможность ее решения.

Начинается ли после выявления и формулировки проблемы сбор фактов? Когда-то полагали, что науку можно свести к точной методологии, приложимой к фактам чисто механически каждым, кто попытается выявить причины явления. Те обстоятельства, которые отсутствуют в некоторых из случаев, не могут быть причиной этого явления. И наоборот, у явления имеется единственная причина, только то обстоятельство, которое присутствует всякий раз, когда явление происходит, и отсутствует всякий раз, когда оно не происходит.

Приведенное выше утверждение вполне правдоподобно, но бесполезно. Оно бесполезно, поскольку количество обстоятельств, сопутствующих любому явлению, бесконечно, и нет никаких критериев отбора нужных. Такой отбор требует предварительной гипотезы, а предлагаемый подход не дает никаких способов выдвижения гипотез. Лишь с помощью выдвижения гипотез и последующего их отбора возможна наука.

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 них 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.

При этом возможно, что все подготовленные и заинтересованные наблюдатели могут ошибаться. Так было неоднократно и так будет. Именно это делает науку возможной.

1.2 Классификация - начало науки

Из различий исследуемых объектов вытекают важные практические следствия: наши наблюдения и выводы в отношении этих объектов очень субъективны и неточны, поэтому требуется их классификация. Классификация является одним из видов научной деятельности, так же как наблюдения, интуиция, логика, синтез, анализ.

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 же 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.

Таким образом, перед началом научных исследований необходимо:

- сформулировать проблему;
- собрать и классифицировать имеющиеся факты в отношении данной проблемы;
- выбрать метод исследования;
- провести исследование;
- обобщить результаты;
- довести их до научной общественности (опубликовать).

Далее рассмотрим известные методы научных исследований.

2 ЭКСПЕРИМЕНТАЛЬНЫЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ

2.1 Понятие научного знания

Известно, что «практика – критерий истины». Поэтому метод наблюдения за изучаемым объектом является самым простым и эффективным методом исследования. Задача экспериментального исследования состоит в том, чтобы на основании наблюдений за изучаемым объектом получить новое научное знание.

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

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. Следует отметить, что 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.

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. Действительно, 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) обусловливает другое (следствие).

Умозаключения делятся на две категории: дедуктивные и индуктивные. Дедуктивные умозаключения представляют собой выведение частного случая из какого-нибудь общего положения. В индуктивных умозаключениях на основании частных случаев приходят к общему положению.

Умозаключения подразделяются также 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.

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.

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, обусловливающая 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.

Как уже отмечалось, 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.

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 слагается 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 заключает в себе не только знания основных законов, но и объяснение фактов на их основе. Теория позволяет открывать новые законы и предсказывать будущее.

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 естественнонаучную 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?

2.2 Эксперимент как основа метода

Эксперимент – одна из сфер человеческой практики, в которой подвергается проверке истинность выдвигаемых гипотез или выявляются закономерности объективного мира. В процессе эксперимента исследователь вмешивается в изучаемый процесс с целью познания, при этом одни условия опыта изолируются, другие исключаются, третьи усиливаются или ослабляются. Экспериментальное изучение объекта или явления имеет определенные преимущества по сравнению с наблюдением, так как позволяет изучать явления в «чистом виде» при помощи устранения побочных факторов; при необходимости испытания могут повторяться и организовываться так, чтобы исследовать отдельные свойства объекта, а не их совокупность.

Экспериментальное исследование – один из основных способов получить новые научные знания. 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.

Эксперимент состоит из наблюдения, сравнения, счета, измерения, обобщения, абстрагирования, формализации, анализа и синтеза.

Наблюдение – это способ познания объективного мира, основанный на непосредственном восприятии предметов и явлений при помощи органов чувств без вмешательства в процесс со стороны исследователя.

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.

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

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

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.

Абстрагирование – это мысленное отвлечение от несущественных свойств, связей, отношений предметов и выделение нескольких сторон, интересующих исследователя. Оно, как правило, осуществляется в два этапа. На первом этапе определяются несущественные свойства, связи и т.д. На втором – исследуемый объект заменяют другим, более простым, представляющим собой упрощенную модель, сохраняющую главное в сложном.

Различают следующие виды абстрагирования: отождествление (образование понятий путем объединения предметов, связанных по своим свойствам в особый класс); изолирование (выделение свойств, неразрывно связанных с предметами); конструктивизация (отвлечение от неопределенности границ реальных объектов) и, наконец, допущение потенциальной осуществимости.

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.

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.

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. Различают следующие виды анализа и синтеза: прямой или эмпирический метод (используют для выделения отдельных частей объекта, обнаружения его свойств, простейших измерений и т.п.); возвратный или элементарно-теоретический метод (базирующийся на представлениях о причинно-следственных связях различных явлений); структурно-генетический метод (включающий вычленение в сложном явлении таких элементов, которые оказывают решающее влияние на все остальные стороны объекта).

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 выводного 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.

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.

При исследованиях сложных систем с многообразными связями, характеризуемыми как непрерывностью и детерминированностью, так и дискретностью и случайностью, используются системные методы (исследование операций, теория массового обслуживания, теория управления, теория множеств и др.). В настоящее время такие методы получили широкое распространение в значительной степени в связи с развитием ЭВМ.

При анализе явлений и процессов в сложных системах возникает потребность рассматривать большое количество факторов (признаков), среди которых важно уметь выделять главное при помощи метода ранжирования и исключения второстепенных факторов, не влияющих существенно на исследуемое явление. Следовательно, этот метод допускает усиление основных и ослабление второстепенных факторов, т.е. размещение факторов по определенным правилам в ряд убывающей или возрастающей последовательности по силе фактора.

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, определимости 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.

2.3 Основы теории случайных ошибок и методов оценки случайных погрешностей в эксперименте

2.3.1 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.

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 $k_в$:

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

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

Доверительным называется интервал значений x_i, в который попадает истинное значение $x_д$ 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 $a \le x_д \le b$. Эта величина определяется в долях единицы или в процентах. Доверительная вероятность $p_д$ описывается выражением

$$p_Д = p(a \le x_Д \le b) = (1/2)[\varphi(b - \bar{x})/\sigma - \varphi(a - \bar{x})/\sigma],$$

где φ(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.

$$t = \mu/\sigma, \tag{2.2}$$

где t – гарантированный коэффициент;

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

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

$$\mu = \sigma \arg \varphi(p_д) = \sigma t, \tag{2.3}$$

где $\arg \varphi(p_д)$ – 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 корпуса блока РЭА при среднем модуле упругости корпуса E = 170 МПа и вычисленном значении среднеквадратического отклонения σ = 3.1 МПа.

Таблица 2.1 - Интегральная функция Лапласа

t	$p_д$	t	$p_д$	T	$p_д$
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 ($p_д$ = 0.9; 0.95; 0.9973), приняв значения t по таблице 2.1. В этом случае соответственно μ = ±3.1·1.65 = 5.1; ±3.1·2.0 = 6.2; ±3.1·3.0 = 9.3 МПа. Следовательно, для данного средства и метода доверительный интервал возрастает примерно в два раза, если увеличить $p_д$ только на 10%.

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

Значение (1 – $p_д$) называют уровнем значимости. Из него следует, что при нормальном законе распределения погрешность, превышающая доверительный интервал, будет встречаться один раз из $n_и$ измерений, где

$$n_и = p_д/(1 - p_д), \qquad (2.4)$$

или иначе приходится браковать одно из $n_и$ измерений.

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

N	$p_Д$					
	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 $p_Д$ = 0.9 определяется $n_И$ = 0.9(1 – 0.9) = 9 измерений. При $p_Д$ равной 0.95 и 0.9973, соответственно 19 и 369 измерений.

2.3.2 Определение минимального количества измерений

Для проведения опытов с заданной точностью и достоверностью необходимо знать то количество измерений, при котором экспериментатор уверен в положительном исходе. 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:

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

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 $p_д$.

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

$$\mu = \sigma \arg \varphi(p_Д) = \sigma_0 / \sqrt{n}t.$$

При $N_{\min} = n$ получаем

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

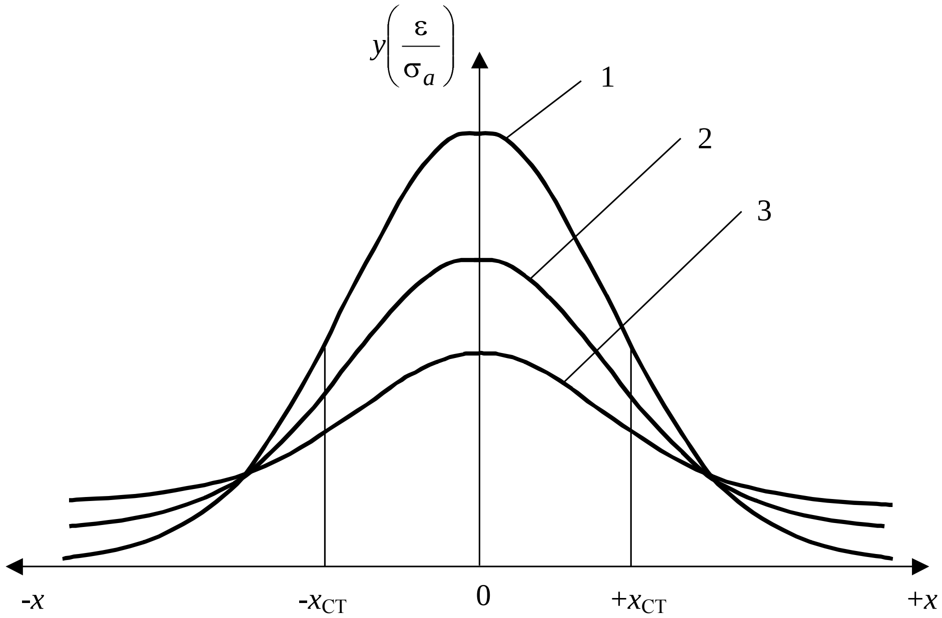
здесь $k_в$ – коэффициент вариации (изменчивости), %; Δ – точность измерений, %.

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; допускаемое отклонение параметра ±0.1 Ом. Если предварительно вычисленное значение σ = 0.4 Ом, то можно определить, с какой достоверностью оценивается данный параметр.

Согласно инструкции Δ = 0.1 Ом. Из формулы (2.6) можно записать $t = \sqrt{n} \cdot \dfrac{\Delta}{\sigma} = \sqrt{25} \cdot \dfrac{0.1}{0.4} = 1.25$. В соответствии с таблицей 2.1 доверительная вероятность для t = 1.25 $p_д$ = 0.79. Это низкая вероятность. Погрешность, превышающая доверительный интервал 2μ = 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 $p_д$, 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 $p_д = 0.90$ и 64 измерения при $p_д = 0.95$, 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 $n\to\infty$ (практически при $n > 20$) переходят в кривые нормального распределения (рисунок 2.1).



1 – n → ∞; *2 – n* = 10; *3 – n* = 2

Рисунок 2.1 – Кривые распределения Стьюдента для различных значений объема выборки

Для малой выборки доверительный интервал

$$\mu_{ст} = \sigma_0\, \alpha_{ст}, \tag{2.7}$$

где $\alpha_{ст}$ – коэффициент Стьюдента, принимаемый по таблице 2.2 в зависимости от значения доверительной вероятности $p_д$.

Зная $\mu_{ст}$, можно вычислить действительное значение изучаемой величины для малой выборки

$$x_Д = \bar{x} \pm \mu_{CT}\,. \tag{2.8}$$

Возможна и иная постановка задачи. По n известным измерениям малой выборки необходимо определить доверительную вероятность $p_д$ при условии, что погрешность среднего значения не выйдет за пределы $\pm\mu_{ст}$. Задачу решают в такой последовательности: вначале вычисляется среднее значение $\bar{x}$, σ_0

и $\alpha_{ст} = \mu_{ст}/\sigma_0$. С помощью величины $\alpha_{ст}$, 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

$$\beta_1 = (x_{\max} - \bar{x}) / \sigma\sqrt{(n-1)/n}\,;$$
$$\beta_2 = (\bar{x} - x_{\min}) / \sigma\sqrt{(n-1)/n}\,, \qquad (2.9)$$

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

В таблице 2.3 приведены в зависимости от доверительной вероятности максимальные значения $\beta_{\max}$, 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 $p_д$ 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 $p_д$ и найти доверительные интервалы $\mu_{ст}$ из (2.7); окончательно установить действительное значение измеряемой величины $x_д$ по формуле (2.8).

Таблица 2.3 - Критерий появления грубых ошибок

N	β_{max} при $p_д$			n	β_{max} при $p_д$		
	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 $k_в$; 5) при большой выборке задаются доверительной вероятностью $p_д = \varphi(t)$ или уравнением значимости $(1 - p_д)$ и по таблице 2.1 определяют t; 6) при малой выборке ($n \le 30$) в зависимости от принятой доверительной вероятности $p_д$ и числа членов ряда n принимают коэффициент Стьюдента $\alpha_{ст}$; 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 $p_д$:

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

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

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

Формулой (2.10) следует пользоваться при $\alpha_{ст}\sigma_0 \le 3B_{пр}$. Если же $\alpha_{ст}\sigma_0 \ge 3B_{пр}$, 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	Σ = -3	Проверка Σ = 0	Σ = 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 $k_B = \frac{6.58}{74.83}100 = 8.8\%$. 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 $p_д$ = 0.99 и n = 18 β_{max} = 2.90. Поскольку 2.68 < β_{max}, измерение 92 не является грубым промахом. Если $p_д$ = 0.95, β_{max} = 2.58, то значение 92 следует исключить.

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

В случае, когда измерение 92 исключается, $\bar{x} = 73.82$, σ = 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 $n = 18$ $\mu_{ст} = \pm1.55\cdot2.11 = 3.3$; при $n = 17$ $\mu_{ст} = \pm1.25\cdot2.12 = 2.7$. Действительное значение изучаемой величины: при $n = 18$ $x_д = 74.8\pm3.3$; при $n = 17$ $x_д = 73.8\pm2.7$. Относительная погрешность результатов серии измерений: при $n = 18$ $\delta = (3.3\cdot100)/74.8 = 4.4\%$; $n = 17$ $\delta = (2.7\cdot100)/73.8 = 3.7\%$. 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 $\sigma = 6.58$; $k_в = 8.8\%$. Если задана точность $\Delta = 5$ и 3% при доверительной вероятности $p_д = 0.95$, $\alpha_{ст} = 2.11$. Следовательно, при $\Delta = 5\%$ $N_{\min} = (8.8^2\cdot2.11^2)/5^2 = 14$, а при $\Delta = 3\%$ $N_{\min} = (8.8^2\cdot2.11^2)/3^2 = 40$.

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

$$y = f(x_1, x_2, \ldots, x_n). \tag{2.11}$$

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:

$$\varepsilon_{пр} = \pm\varepsilon_x f'(x),$$

$$\delta_{пр} = \pm \mathrm{d}\ln(x),$$

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

Если исследуется функция многих переменных, то

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

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

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 $x_д \pm \varepsilon$ каждого переменного измерена, следовательно, абсолютные ошибки для аргументов известны, т.е. ε_{x1}, ε_{x2}, …, ε_{xn}. Затем вычисляют относительные ошибки независимых переменных:

$$\delta_{x1} = \varepsilon_{x1}/ x_д;\ \delta_{x2} = \varepsilon_{x2}/ x_д \quad ;\ \ldots;\ \delta_{xn} = \varepsilon_{xn}/ x_д.$$

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 $\delta_{пр} = \delta_{пр\ min}$. 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, …, 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 $f(x_1, x_2, \ldots, x_n)$, 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 РЭС до закалки $R_1 = \overline{R}_1 \pm \sigma_1 = 20 \pm 0.5$ МПа и прочность этих образцов после закалки $R_2 = \overline{R}_2 \pm \sigma_2 = 23 \pm 0.6$ 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

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

В данном случае проверяется разница $\overline{x} = R_1 - R_2 = 3$ 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):

$$k_{KP} = \max D_i / \sum_{1}^{m} D_i \,, \tag{2.14}$$

где max D_i – наибольшее значение дисперсий из числа рассматриваемых параллельных серий m; $\sum_{1}^{m} D_i$ – сумма дисперсий m серий. Рекомендуется принимать $2 \le m \le 4$. Опыты считают воспроизводимыми при

$$k_{\text{кр}} \le k_{\text{кт}},$$

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

Таблица 2.6 - Критерий Кохрена $k_{\text{кт}}$ при $p_{\text{д}}$ = 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 блока РЭА (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 блока РЭА 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 $k_{\text{кт}} = 0.74$. Так как $0.55 < 0.74$, то измерения в эксперименте следует считать воспроизводимыми. Если бы оказалось наоборот, т.е. $k_{\text{кр}} > k_{\text{кт}}$, то необходимо было бы увеличить число серий m или число измерений n.

2.3.3 Регрессионный анализ в экспериментальном методе исследований

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.

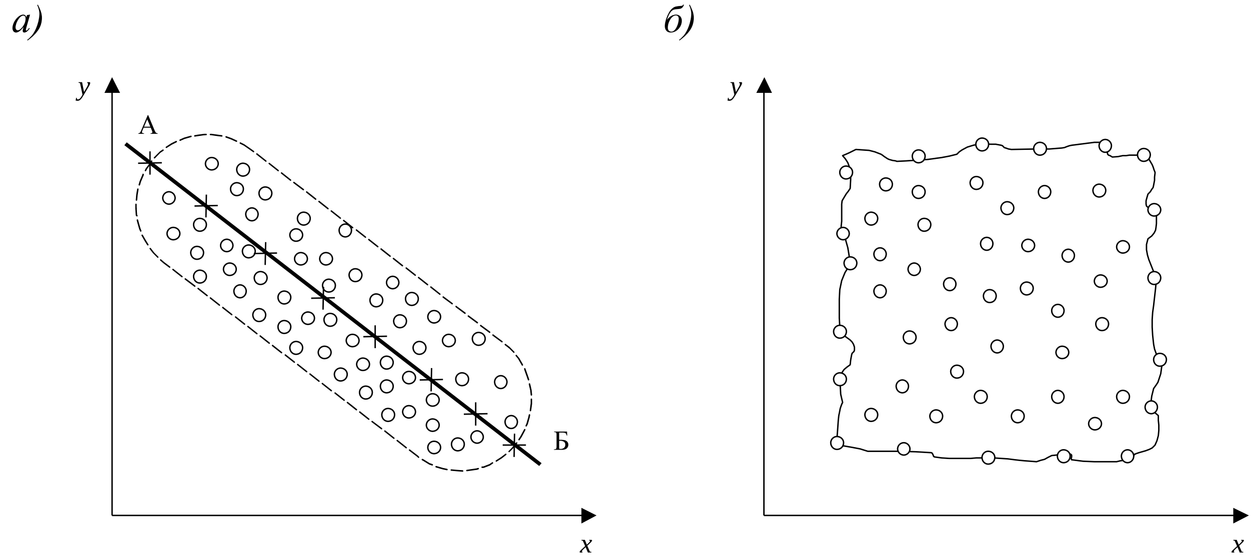


Рисунок 2.2 – Корреляционное поле

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 *АБ* (см. рисунок 2.2, *а*).

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:

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

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 *АБ* (см. рисунок 2.2, *а*) наилучшим образом выравнивает значения постоянных коэффициентов a и b, т.е. коэффициентов уравнения регрессии. Их вычисляют по выражениям

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

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

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

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

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 $r = 0.8 \ldots 0.85$. Для определения процента разброса (изменчивости) искомой функции y относительно ее среднего значения, определяемого изменчивостью фактора x, вычисляют коэффициент детерминации

$$k_д = r^2. \qquad (2.19)$$

Уравнение регрессии прямой можно представить выражением

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

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
$y_э$	7,1	10,6	14,2	17,7	21,8	24,8	28,3	31,9	35,4	39,0

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

Коэффициент корреляции согласно (2.18)

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

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 $k_д = 0.99^2 = 0.98$, 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 $x_1, x_2, \ldots, x_n$ 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.

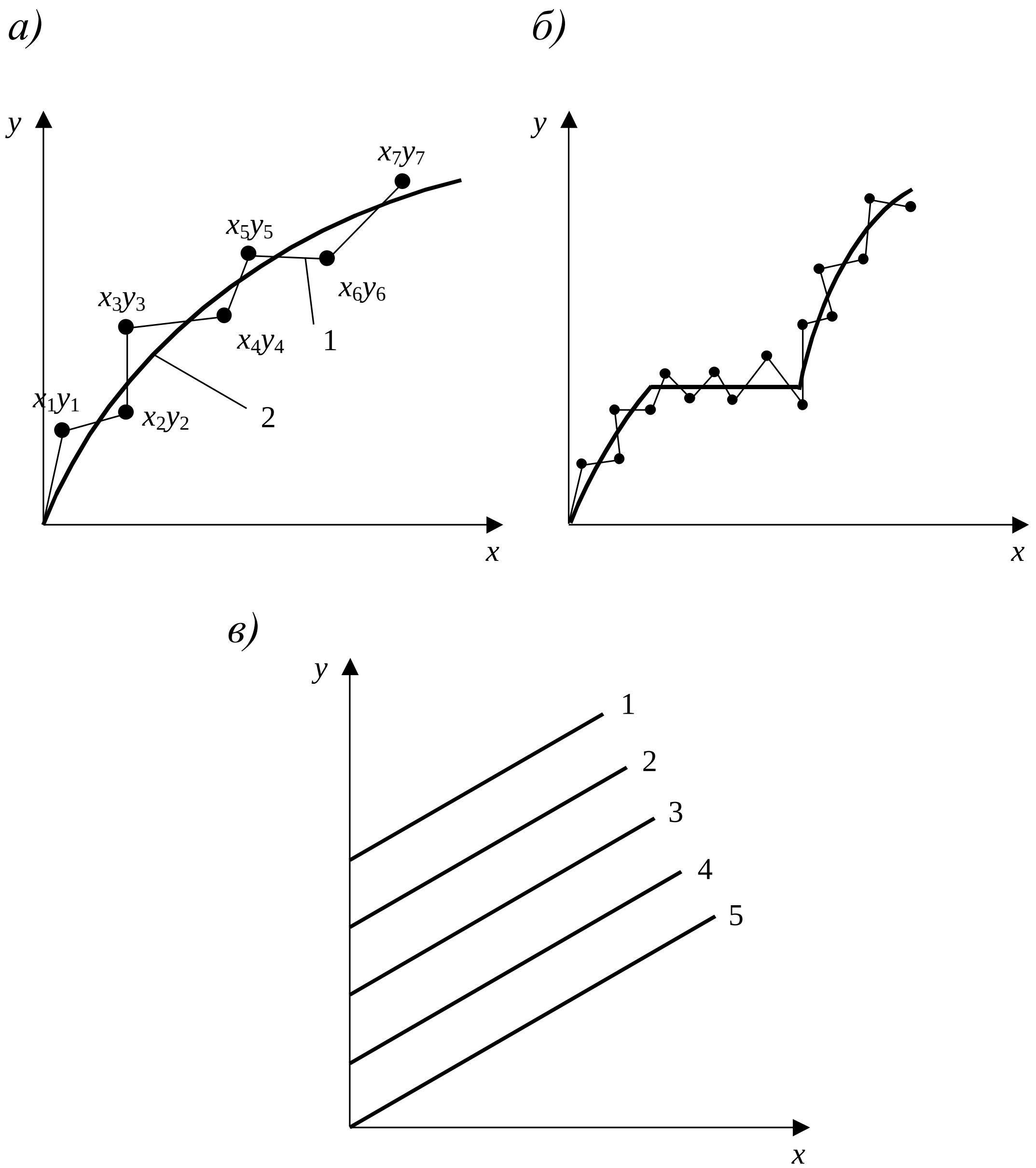
Теоретическую модель множественной регрессии можно получить методами математического планирования, т.е. активным экспериментом, а также пассивным, когда точки факторного пространства выбираются в процессе эксперимента произвольно.

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 $x_1y_1, x_2y_2, \ldots, x_ny_n$ (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.



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

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

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.

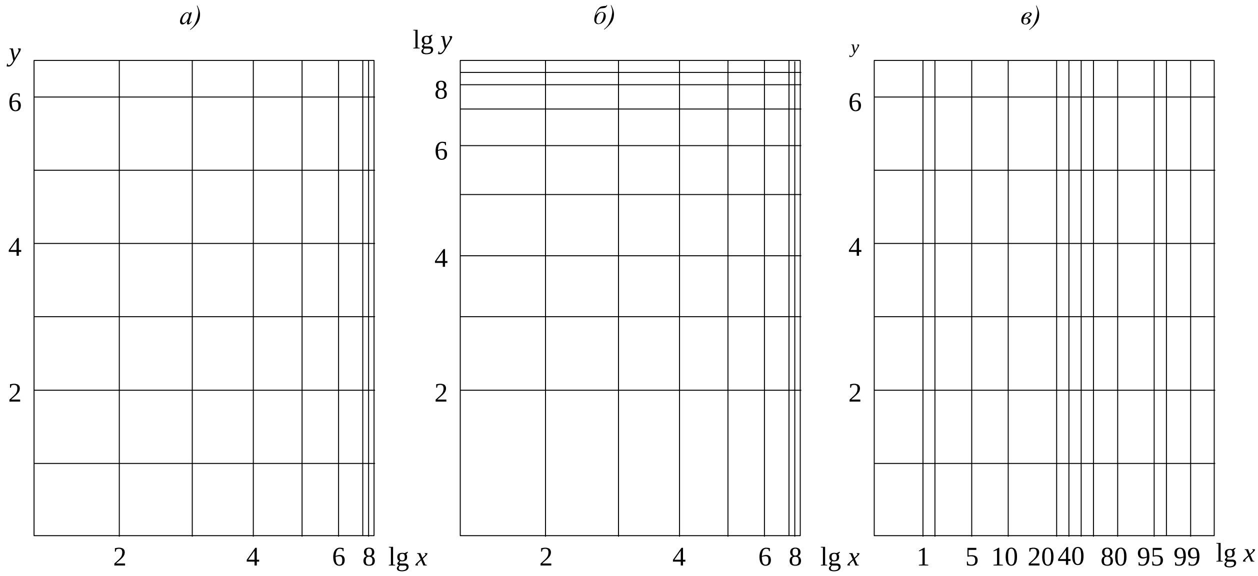
Иногда при построении графика одна-две точки резко удаляются от кривой. В таких случаях вначале следует проанализировать физическую сущность явления, и если нет основания полагать наличие скачка функции, то такое резкое отклонение можно объяснить грубой ошибкой или промахом. Это может возникнуть тогда, 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.

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 $b_1, b_2, \ldots, b_n = \text{const}$ 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

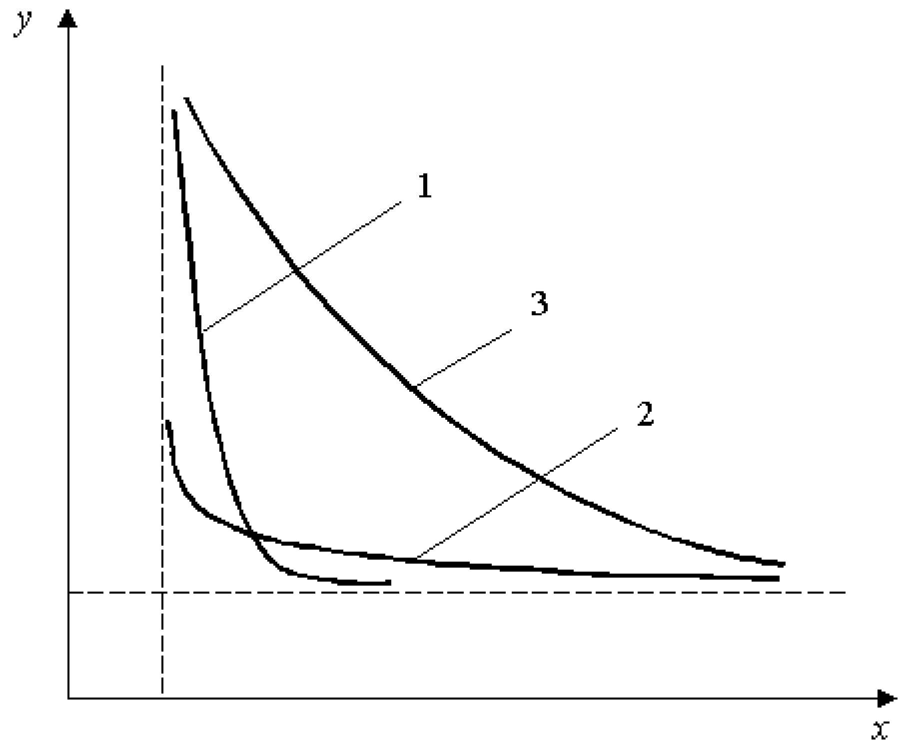
Рисунок 2.4 – Координатная сетка

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 ±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 $f(x_1, x_2, \ldots, x_n)$ 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 $y_1, y_2, \ldots, y_n$ соответствует определенное значение аргумента $x_1, x_2, \ldots, x_n$.

На основе экспериментальных данных можно подобрать алгебраические выражения функции

$$y = f(x), \tag{2.20}$$

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

$$y = a + bx, \tag{2.21}$$

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:

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

В искомом уравнении они должны быть связаны линейной зависимостью

$$Y = a + bX. \tag{2.23}$$

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): $b = \mathrm{tg}\,\alpha = (Y_i - a)/X_i$.

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.

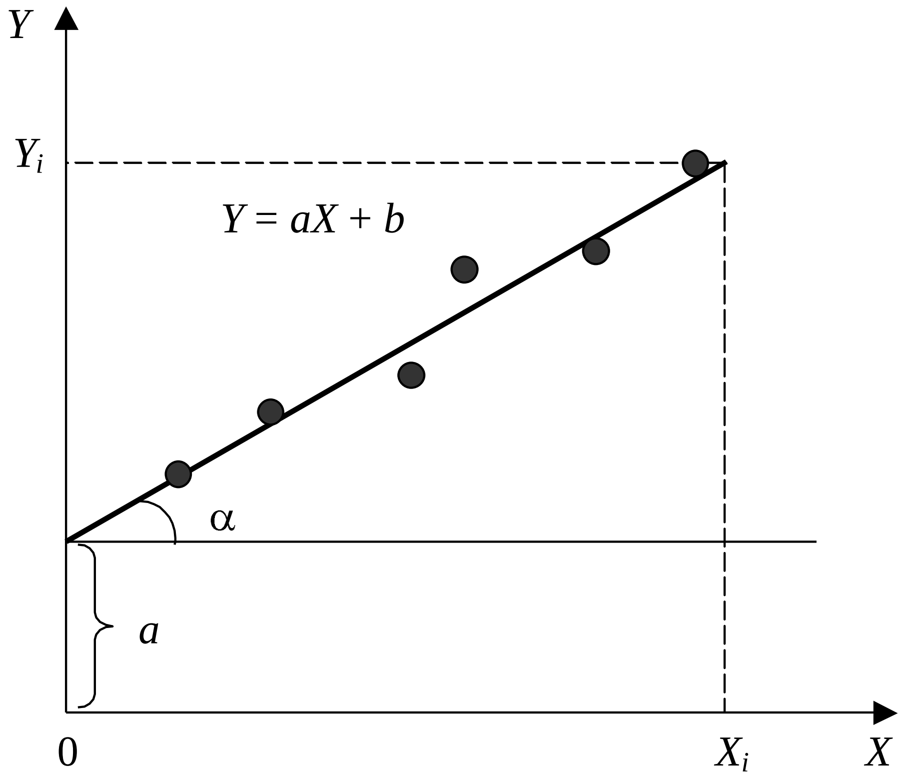


Рисунок 2.6 – Графическое определение параметров *x* и *y*

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

$$y = ax^b. \tag{2.24}$$

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

$$y = ae^{bx}. \tag{2.25}$$

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

$$y = c + ax^b. \tag{2.26}$$

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

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

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

$$y = c + ae^{bx}. \tag{2.28}$$

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

$$y = a + b/x. \tag{2.29}$$

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

$$y = 1/(a + bx). \tag{2.30}$$

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

Аналогично, уравнению

$$y = \frac{1}{a + bx + cx^2} \tag{2.31}$$

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

Сложную степенную функцию

$$y = ae^{nx + mx^2} \tag{2.32}$$

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*).

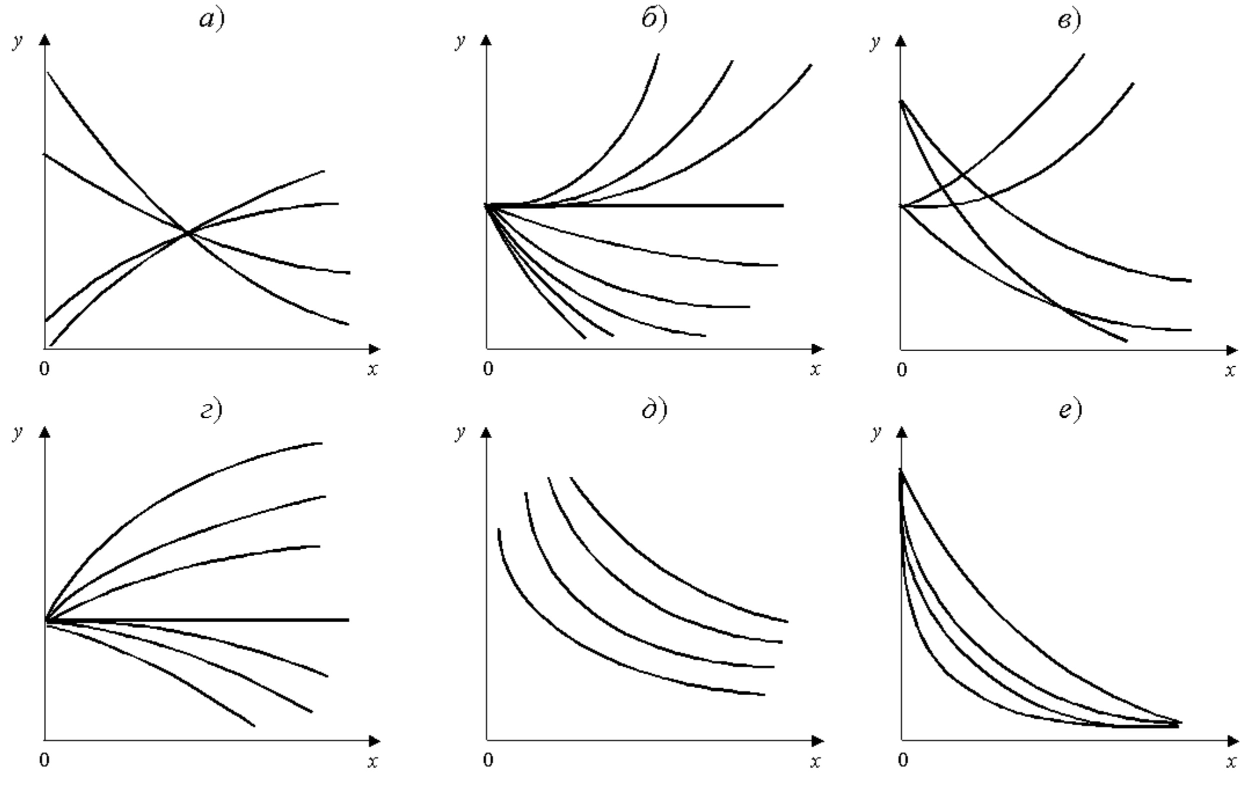


Рисунок 2.7 – Основные виды графиков эмпирических формул

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,

$$\lg a + b \lg e = 1.182;$$
$$\lg a + 4.5b \lg e = 2.070,$$

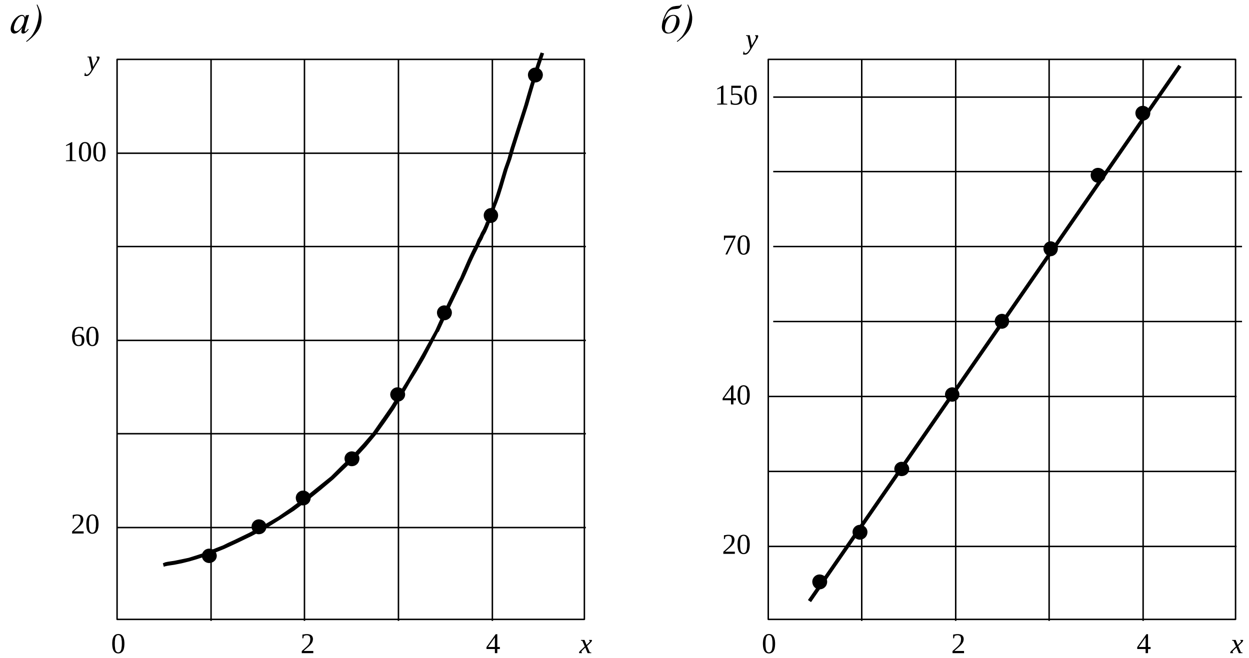
откуда $b = 0.888/(3.5 \lg e) = 0.584$; $\lg a = 1.182 - 0.254 = 0.928$; $a = 8.47$. Окончательно эмпирическая формула получит вид

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

При подборе эмпирических формул широко используются полиномы

$$y = A_0 + A_1 x + A_2 x^2 + A_3 x^3 + \ldots + A_n x^n, \tag{2.33}$$

где A_0, A_1, A_2, …, 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.



а - 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:

$$A_0 + A_1 x_1 + A_2 x_1^2 + \ldots + A_n x_1^n - y_1 = \varepsilon_1;$$

$$A_0 + A_1 x_2 + A_2 x_2^2 + \ldots + A_n x_2^n - y_2 = \varepsilon_2; \tag{2.34}$$

………………………………………

$$A_0 + A_1 x_m + A_2 x_m^2 + ... + A_n x_m^n - y_m = \varepsilon_m .$$

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 $\Sigma\varepsilon^2$ = 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

$$y = A_0 + A_1 x + A_2 x^2.$$

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

1. $A_0 + 4A_1 + 16A_2 - 10.2 = \varepsilon_1$;
2. $A_0 + 5A_1 + 25A_2 - 6.7 = \varepsilon_2$;
3. $A_0 + 6A_1 + 36A_2 - 4.8 = \varepsilon_3$;
4. $A_0 + 7A_1 + 49A_2 - 3.6 = \varepsilon_4$;
5. $A_0 + 8A_1 + 64A_2 - 2.7 = \varepsilon_5$;
6. $A_0 + 9A_1 + 81A_2 - 2.1 = \varepsilon_6$;
7. $A_0 + 10A_1 + 100A_2 - 1.7 = \varepsilon_7$.

Сложение уравнений в каждой подгруппе дает

1-я группа $2A_0 + 9A_1 + 41A_2 = 16.9$;

2-я группа $2A_0 + 13A_1 + 85A_2 = 8.4$;

3-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:

$$1.7604 = A + \frac{3}{2.303}B; \qquad 1.2201 = A + \frac{15}{2.303}B;$$
$$1.6222 = A + \frac{6}{2.303}B; \qquad 1.0864 = A + \frac{18}{2.303}B;$$
$$1.4914 = A + \frac{9}{2.303}B; \qquad 0.9494 = A + \frac{21}{2.303}B;$$
$$1.3560 = A + \frac{12}{2.303}B; \qquad 0.8129 = A + \frac{24}{2.303}B;$$

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

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, …, 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, …, a_n) необходимо решить систему линейных уравнений:

$$y_1 = a_1x_1 + a_2u_1 + \ldots + a_nz_1;$$
$$y_2 = a_1x_2 + a_2u_2 + \ldots + a_nz_2;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$y_n = a_1x_m + a_2u_m + \ldots + a_nz_m,$$

где y_1, y_2, …, 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, …, x_m и последующего их сложения, затем умножения соответственно на u_1, u_2, …, u_m. Это позволяет получить так называемую систему нормальных уравнений

$$\sum_1^m yx = a_1\sum_1^m xx + a_2\sum_1^m xu + ... + a_n\sum_1^m xz;$$

$$\sum_1^m yu = a_1\sum_1^m ux + a_2\sum_1^m uu + ... + a_n\sum_1^m uz;$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\sum_1^m yz = a_1\sum_1^m zx + a_2\sum_1^m zu + ... + a_n\sum_1^m zz,$$

решение которой и дает искомые коэффициенты.

Пусть, например, необходимо определить коэффициенты a_1 и a_2 в уравнении $k_р = a_1 + a_2u$. 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 $y = k_р$; $x_1 = 1$; $x_2 = u$.

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 $k_р = 0.78 + 0.0025u$.

Таблица 2.10 - Результаты опытов

$x_2 = u$	$y = k_р$	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 $k_р = a_1 + a_2u$ 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, …) 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 $p_д$ (обычно $p_д$ = 0.95). Если $k_{фэ} < k_{фт}$ – модель адекватна; если $k_{фэ} \geq k_{фт}$ – модель неадекватна. Опытный критерий Фишера вычисляют по формуле

$$k_{фэ} = D_а/D_{ср}, \tag{2.35}$$

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

$$D_a = \frac{\sum_1^n \left(y_{iT} - \bar{y}_{iЭ}\right)^2}{n-d}; \tag{2.36}$$

$$D_{cp} = \frac{\sum_1^m \sum_1^n \left(y_{iT} - y_{iЭ}\right)^2}{mn}.$$

Здесь $y_{iт}$ – 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) $y_{i\text{Т}}$ вычисляют по теоретической регрессии для фактора x_i; $\bar{y}_{i\text{Э}}$ – как среднее из m серий измерений, т.е.

$$\bar{y}_{i\text{Э}} = \frac{1}{m}(y_{1\text{Э}} + y_{2\text{Э}} + ... + y_{m\text{Э}}).$$

Таблица 2.11 - Критерий Фишера

q_1	Значения $k_{\text{фт}}$ при $p_{\text{д}} = 0.95$ 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 $D_а = 5.36/(7 - 1) = 0.89$. Здесь значение $d = 1$, поскольку в теоретическом выражении один значащий член x. Дисперсия $D_{ср}$ вначале вычисляется построчно для m строк. Для первой строки $D_1 = \Sigma(y_{iТ} - y_{iЭ})^2/m = 1/5[(12 - 16)^2 + (17 - 16)^2 + (15 - 16)^2 + (14 - 16)^2 + (16 - 16)^2] = 4.4$; 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 $\bar{y}_{iэ} = \frac{\sum_1^m y_{iэ}}{m}$	$y_{iт}$	$y_{iT} - \bar{y}_{iэ}$	$(y_{iT} - \bar{y}_{iэ})^2$	$\frac{\sum_1^m (y_{iT} - \bar{y}_{iэ})^2}{m}$
		$y_{1э}$	$y_{2э}$	$y_{3э}$	$Y_{4э}$	$y_{5э}$					
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

$$q = m - s, \tag{2.37}$$

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

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

где $y_{Эi}$, $y_{Тi}$ – 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: $x_1…x_2$; $x_3…x_4$; $x_5…x_6$ и т.д. По данным измерений в каждом разряде может оказаться $y_Э$ измерений. Так, в диапазоне $x_1…x_2$ имеется $y_{Э1}$ измерений (частота); в $x_3…x_4$ 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

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

$$+\frac{(\,-1)^2}{1}=2.56.$$

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.

Таблица 2.13 - Критерий Пирсона

χ^2	Значения критерия Пирсона $p(x^2, q)$ 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

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.

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:

$$D_0 = \max (\Sigma y_{\text{Э}i} - \Sigma y_{\text{Т}i}).$$

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

$$b = \pm(x_{\text{н}} - x_{\text{д}}),$$

где $x_{\text{н}}$ – показание прибора (номинальное значение измеряемой величины); $x_{\text{д}}$ – 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 %):

$$b_{OT} = \pm \frac{x_И - x_Д}{x_Д} 100.$$

Иногда применяют понятие приведенной погрешности

$$b_{np} = \pm \frac{x_И - x_Д}{x_{np}} 100,$$

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

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

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.

Поверка 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 применим для данных измерений.

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

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.

Важным моментом в организации эксперимента является выбор средств измерений. Желательно также максимально использовать средства измерений с автоматической записью.

3 ТЕОРЕТИЧЕСКИЙ МЕТОД НАУЧНЫХ ИССЛЕДОВАНИЙ

3.1 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).

Метод декомпозиции предложен французским философом и естествоиспытателем Р.Декартом. В своей работе «Правила для руководства ума» он пишет: «Освободите вопрос от всех излишних представлений и сведите его к простейшим элементам». В процессе декомпозиции выделяются существенные и несущественные параметры, основные элементы и связи между ними. Следует, однако, отметить, что каждый объект можно разделить разными способами и это существенно влияет на проведение теоретических исследований, так как в зависимости от способа разделения процесс изучения объекта может упроститься или наоборот усложниться. После разделения объекта изучается вид взаимосвязи элементов и осуществляется моделирование этих элементов. Наконец, элементы объединяются в сложную модель объекта. Декомпозиция – это, по сути, анализ объекта, а композиция – синтез.

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.

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 удается 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 3.1). 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.

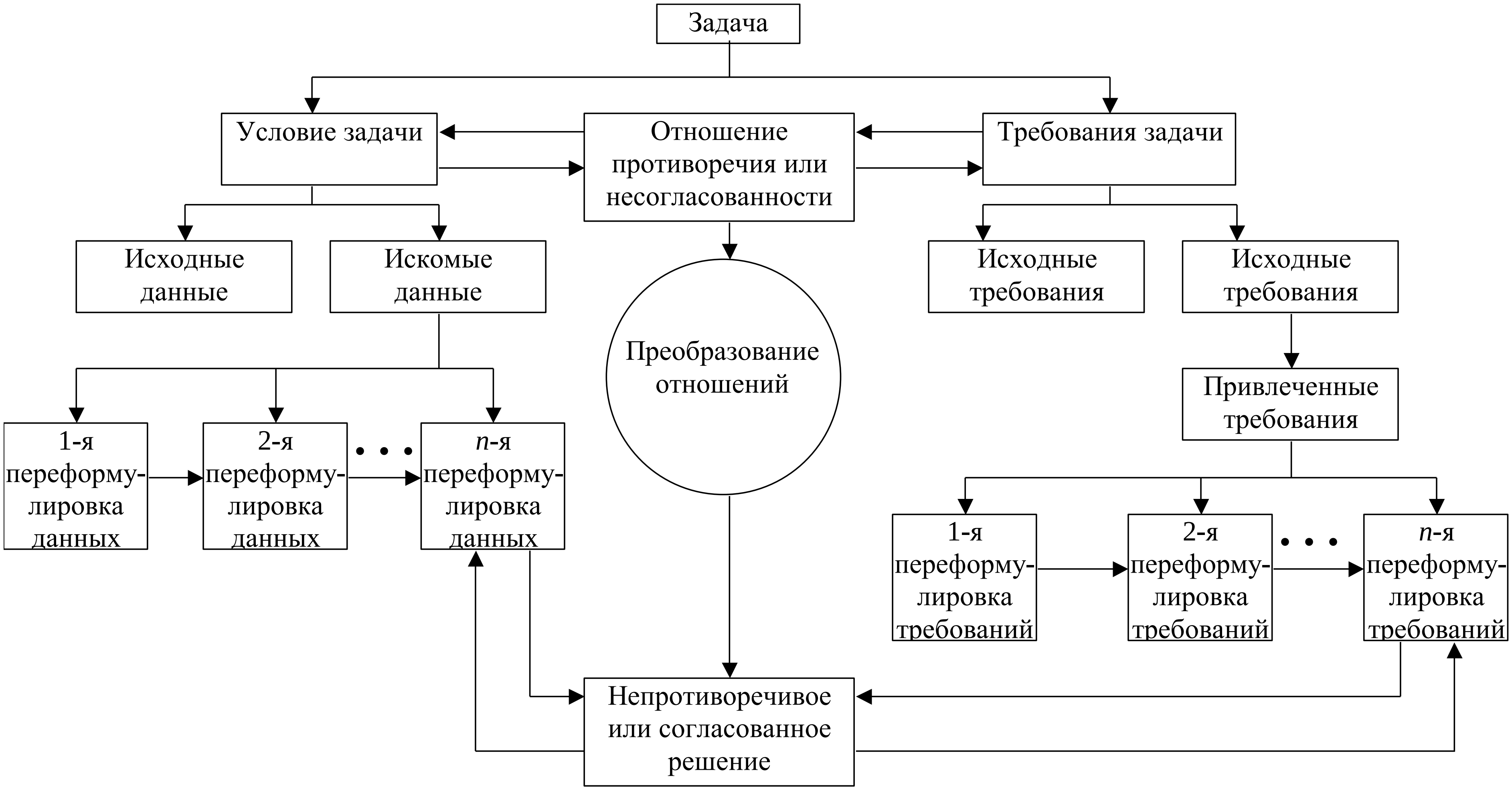


Рисунок 3.1 – Структурные компоненты решения задачи

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, привходящих 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.

3.2 Использование математических методов в теоретических исследованиях

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.

Математическая модель представляет собой систему математических соотношений – формул, функций, уравнений, систем уравнений, описывающих те или иные стороны изучаемого объекта, явления, процесса.

Первым этапом математического моделирования 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 моделирования 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.

Линейность устанавливается по характеру статической характеристики исследуемого объекта. Под статической характеристикой объекта понимается связь между величиной внешнего воздействия на объект (величиной входного сигнала) и максимальной величиной его реакции на внешнее воздействие (максимальной амплитудой выходной характеристики системы). Под выходной характеристикой системы понимается изменение выходного сигнала системы во времени. Если статическая характеристика исследуемого объекта оказывается линейной, то моделирование этого объекта осуществляется с использованием линейных функций. Нелинейность статической характеристики и наличие запаздывания в реагировании объекта на внешнее воздействие являются яркими признаками нелинейности объекта. В этом случае для его моделирования должна быть принята нелинейная математическая модель

Применение линейной математической модели значительно упрощает ее дальнейший анализ, поскольку такая модель позволяет пользоваться принципом суперпозиции. Принцип суперпозиции утверждает, что когда на линейную систему воздействуют несколько входных сигналов, то каждый из них фильтруется системой так, как будто никакие другие сигналы на нее не действуют. Общий выходной сигнал линейной системы по принципу суперпозиции образуется в результате суммирования ее реакции на каждый входной сигнал.

Установление динамичности или статичности осуществляется по поведению исследуемых показателей объекта во времени. Применительно к детерминированной системе можно говорить о стохастичности или динамичности по характеру ее выходной характеристики. Если среднее арифметическое значение выходного сигнала по разным отрезкам времени не выходит за допустимые пределы, определяемые точностью методики измерения исследуемого показателя, то это свидетельствует о статичности объекта. Применительно к вероятностным системам их статичность устанавливается по изменчивости уровня ее относительной организации. Если изменчивость этого уровня не превышает допустимые пределы, то система определяется как статичная.

Весьма важным является выбор отрезков времени, на которых устанавливается статичность или динамичность объекта. Если объект на малых отрезках времени оказался статичным, то при увеличении этих отрезков результат не изменится. Если же статичность установлена для крупных отрезков времени, то при их уменьшении результат может измениться и статичность объекта может перейти в динамичность. При выборе типа (класса) модели вероятностного объекта важно установление его стационарности. Обычно о стационарности или нестационарности вероятностных объектов судят по изменению во времени параметров законов распределения случайных величин. Чаще всего для этого используют среднее арифметическое случайной величины $M(\tau_i)$ и среднее квадратическое отклонение случайных величин σ_i ($i = 1, 2, \ldots, n$) от среднего арифметического и среднего квадратического отклонения во времени.

Из ряда средних арифметических $M(\tau_1)$, $M(\tau_2)$, …, $M(\tau_i)$ выбирается минимальное значение $M(\tau_{\min})$ и строятся интервалы с границами

$$M(\tau_{\min}) + \Delta x, \qquad M(\tau_{\min}) - \Delta x,$$

где Δx – точность методики измерения исследуемого показателя.

Если значение $M(\tau_i)$ укладывается в этот интервал, то объект определяется как стационарный по среднему арифметическому $M(\tau)$.

Аналогично определяется стационарность по среднему квадратическому отклонению.

Граничные значения σ при установлении стационарности определяются по формулам

$$\sigma_1 = \sqrt{\frac{D_1}{n}}, \quad \sigma_2 = \sqrt{\frac{D_2}{n}},$$

где

$$D_1 = \left\{ \sum_{i=1}^{n} \left[x_i - \left(M(\tau_{\min}) + \Delta x \right) \right]^2 \right\} \frac{1}{n-1};$$

$$D_2 = \left\{ \sum_{i=1}^{n} \left[x_i - \left(M(\tau_{\min}) - \Delta x \right) \right]^2 \right\} \frac{1}{n-1};$$

здесь n – число наблюдений.

Если все значения σ укладываются в интервал $\sigma_1 \ldots \sigma_2$, то объект считается стационарным. В противном случае объект определяется как вероятностный нестационарный, даже если величина среднего арифметического M не изменяется во времени.

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 3.2. 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.

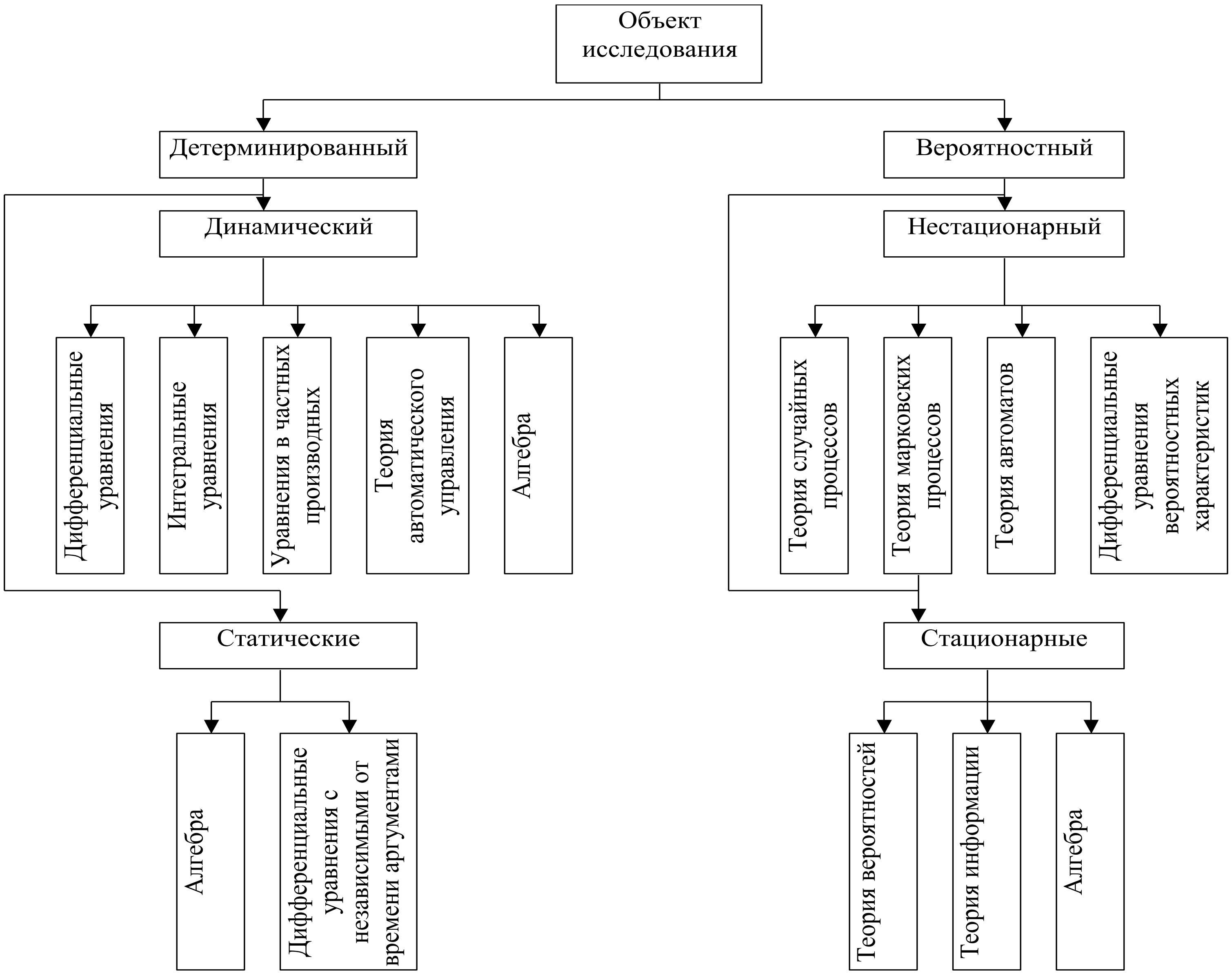


Рисунок 3.2 – Математический аппарат для построения математической модели

В непрерывных объектах все сигналы представляют собой непрерывные функции времени. В дискретных объектах все сигналы квантуются по времени и амплитуде. Если сигналы квантуются только по времени, т.е. представляются в виде импульсов с равной амплитудой, то такие объекты называют дискретно-непрерывными.

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. Выбор вида математической модели в данном их классе является третьим этапом математического моделирования. Данный этап 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 иди дифференциальных), для качественных – используются табличные способы заданий 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:

одномерно-одномерная схема (рисунок 3.3, *а*) – на объект воздействует только один фактор, а его поведение рассматривается по одному показателю (один выходной сигнал);

одномерно-многомерная схема (рисунок 3.3, *б*) – на объект воздействует один фактор, а его поведение оценивается по нескольким показателям;

многомерно-одномерная схема (рисунок 3.3, *в*) – на объект воздействует несколько факторов, а его поведение оценивается по одному показателю;

многомерно-многомерная схема (рисунок 3.3, *г*) – на объект воздействует множество факторов и его поведение оценивается по множеству показателей.

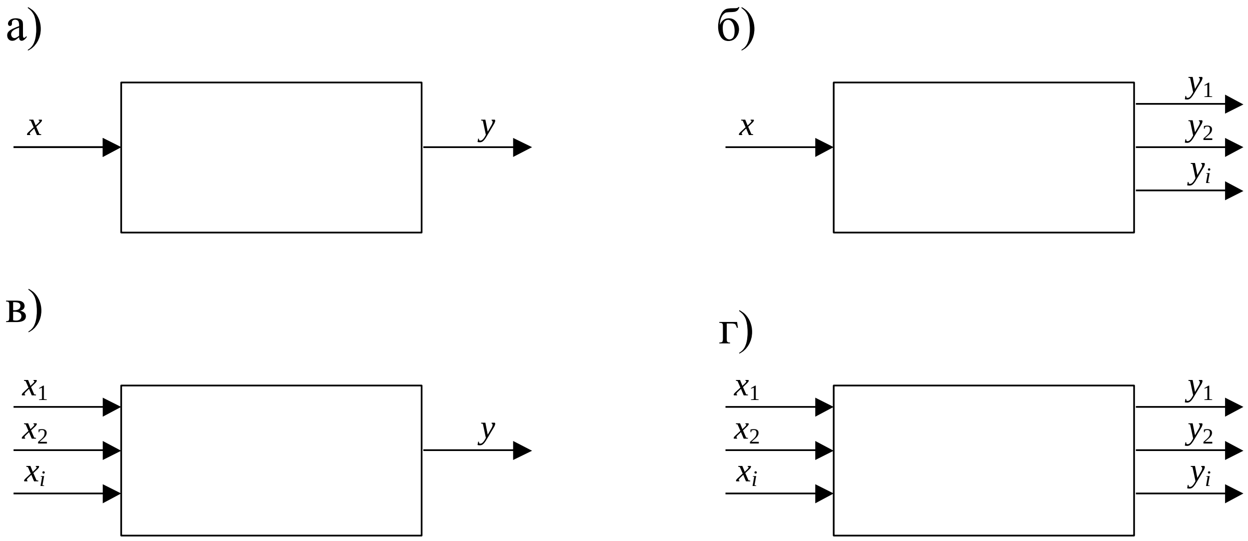


Рисунок 3.3 – Схемы воздействия объекта с внешней средой

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:

$$y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m_1} \sum_{j=1}^{m_1} a_{ij} x_i x_j + \sum_{i=1}^{m_2} \sum_{j=1}^{m_2} \sum_{\nu=1}^{m_2} a_{ij\nu} x_i x_j x_\nu + \ldots,$$

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:

$$\frac{d\bar{y}}{dx} + a\bar{y} = a\bar{y}^m \qquad (\bar{y} = 0 \text{ при } x = 0),$$

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 3.4, *а*). Такое изменение выхода определяется решением дифференциального уравнения

$$\frac{dy}{dt} = kx, \qquad y_0 = 0,$$

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 3.4, *б*. Однако дифференциальное уравнение остается неизменным.

Более сложный вид реакции объекта на ступенчатое входное воздействие (рисунок 3.4, *в*) может быть описан полным неоднородным дифференциальным уравнением первого порядка:

$$\frac{dy}{dt} + a_0 y = kx, \qquad y(0) = y_0,$$

где a_0 – коэффициент дифференциального уравнения.

Реакция объекта, соответствующая рисунку 3.4, *г*, позволяет использовать в качестве математической модели объекта дифференциальное уравнение второго порядка:

$$\frac{d^2y}{dt^2} + a_1\frac{dy}{dt} + a_0 y = kx, \qquad y(0) = y_0.$$

Рассмотренные виды математических моделей соответствуют постоянному входному воздействию (x = 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.

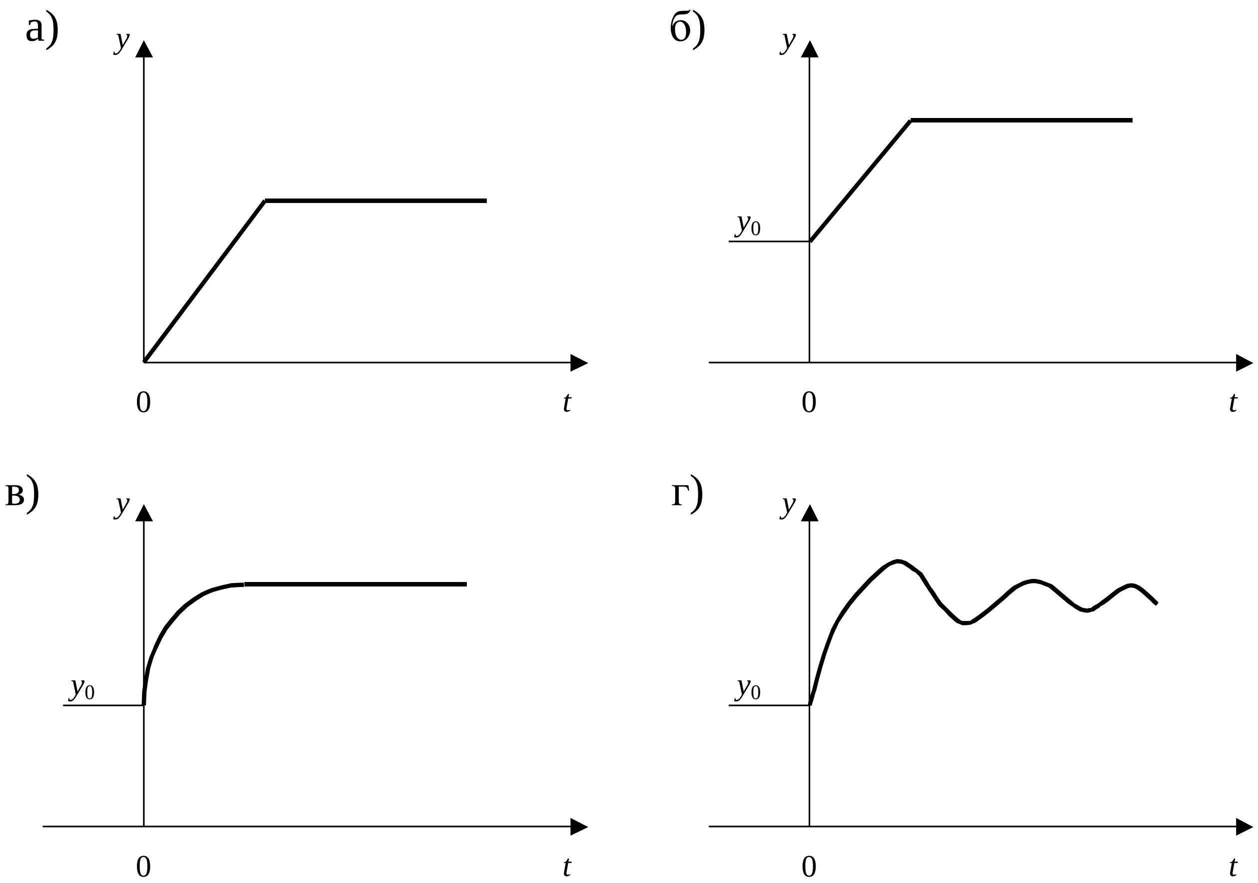


Рисунок 3.4 – Выходные характеристики детерминированного объекта при ступенчатом внешнем воздействии

Если допущение о независимости действия факторов неправомочно, 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:

1) дифференциальные уравнения в дифференциалах;
2) дифференциальные уравнения в производных;

3) 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 3.5).

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

$$dr_1 + dr_2 = 0.$$

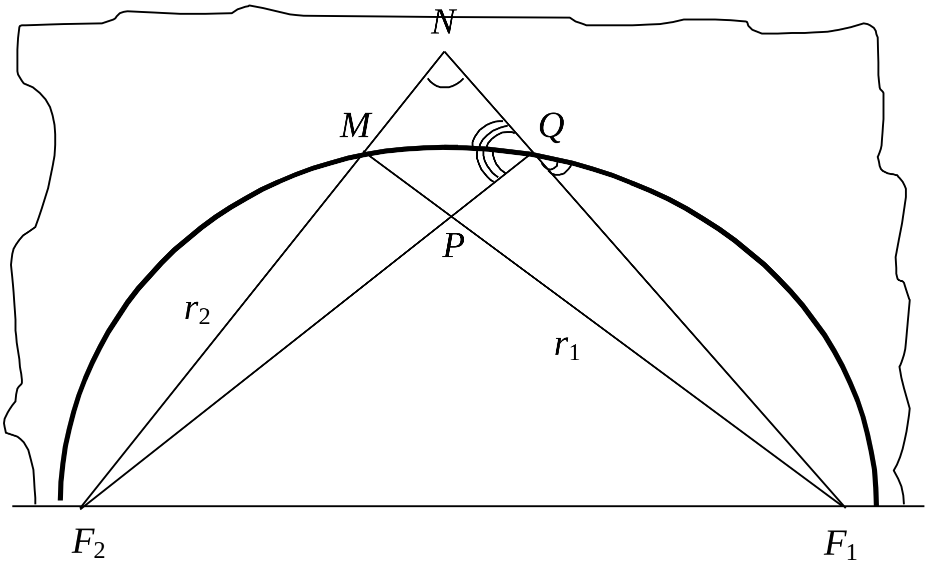


Рисунок 3.5 – Расчетная схема

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

$$\frac{1}{y}\frac{dy}{dt} = k \qquad (k > 0),$$

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 точки блока РЭА массы m, если известно, что работа действующей на точку 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 $A = \int\limits_{s_0}^{s} F(u)du$, где $F(u)$ – действующая на точку сила.

По условию задачи

$$A = kt.$$

Сравнивая оба выражения для A, находим

$$\int\limits_{s_0}^{s} F(u)du = kt.$$

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 $\dfrac{ds}{dt}$ 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 $F_Д$ равняются силам сопротивления $F_С$. Уравнение статического равновесия принимает вид

$$F_Д - F_С = 0.$$

В том случае, когда объект исследования (ротор электродвигателя электропривода управления движением летательного объекта) совершает вращательное движение, то условием статического равновесия является равенство крутящего момента $M_Д$ моменту сопротивления $M_С$, т.е. $M_Д - M_С = 0$.

При выполнении этого условия вращение вала будет равномерным.

Если регулируемым объектом является объем блока РЭА, в котором должна поддерживаться постоянная температура, то условие статического равновесия получает вид

$$Q_{пр} - Q_{расх} = 0,$$

где $Q_{пр}$ – количество теплоты, поступающей в объем за единицу времени; $Q_{расх}$ – 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

$$F_Д + \Delta F_Д \neq F_С + \Delta F_С;$$

при вращательном движении

$$M_Д + \Delta M_Д \neq M_С + \Delta M_С;$$

при нарушении теплового режима

$$Q_{пр} + \Delta Q_{пр} \neq Q_{расх} + \Delta Q_{расх}.$$

Полученные неравенства могут быть упрощены, если учесть в них условие статического равновесия:

$$\Delta F_Д \neq \Delta F_С; \ \Delta M_Д \neq \Delta M_С; \ \Delta Q_{пр} \neq \Delta Q_{расх}.$$

Таким образом, при нарушении условий статического равновесия в исследуемом объекте возникает избыток или недостаток движущих сил (или моментов), 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

$$ma = \Delta F_{Д} - \Delta F_{С},$$

или, так как

$$a = \frac{dv}{dt},$$

где v – скорость движения; t – время, то

$$m\frac{dv}{dt} = \Delta F_{Д} - \Delta F_{C}.$$

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

$$J\frac{d\omega}{dt} = \Delta M_{Д} - \Delta M_{C}.$$

Нарушение теплового режима некоторого объема приводит к изменению количества теплоты, сосредоточенной в выбранном объеме, на dQ за интервал времени dt, поэтому

$$dQ = (\Delta Q_{пр} - \Delta Q_{расх})dt,$$

или

$$\frac{dQ}{dt} = \Delta Q_{пр} - \Delta Q_{расх}.$$

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.

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.

$$M_{Д} = f(\omega, h).$$

Для определения приращения крутящего момента двигателя в зависимости от приращения ω и h полученную функцию следует разложить в ряд Маклорена:

$$M_Д + \Delta M_Д = M_Д + \frac{\partial M_Д}{\partial\omega}d\omega + \frac{\partial^2 M_Д}{\partial\omega^2}\frac{d\omega^2}{2!} + \ldots + \frac{\partial M_Д}{\partial h}dh + \frac{\partial^2 M_Д}{\partial h^2}\frac{dh^2}{2!} + \ldots .$$

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

$$M_Д + \Delta M_Д = M_Д + \frac{\partial M_Д}{\partial\omega}\Delta\omega + \frac{\partial^2 M_Д}{\partial\omega^2}\frac{\Delta\omega^2}{2!} + \ldots + \frac{\partial M_Д}{\partial h}\Delta h + \frac{\partial^2 M_Д}{\partial h^2}\frac{\Delta h^2}{2!} + \ldots .$$

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 $M_Д = f(\omega, h)$ 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 $M_Д = f(\omega, h)$ ее касательной в точке равновесного режима (такая замена обычно называется линеаризацией). После линеаризации приращение крутящего момента принимает вид

$$\Delta M_Д = \frac{\partial M_Д}{\partial\omega}\Delta\omega + \frac{\partial M_Д}{\partial h}\Delta h .$$

Пусть момент сопротивления является лишь функцией угловой скорости

$$M_С = f(\omega).$$

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

$$J\frac{\partial\omega}{\partial t} + \left(\frac{\partial M_C}{\partial\omega} - \frac{\partial M_Д}{\partial\omega}\right)\Delta\omega = \frac{\partial M_Д}{\partial h}\Delta h .$$

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

$$H_m = f(x); \quad R = f(x); \quad y_{\text{ср}} = f(x); \quad \sigma = f(x),$$

где x – входное воздействие; H_m – максимальная энтропия выходных характеристик; R – относительная организация выходных характеристик; $y_{\text{ср}}$ – 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

$$n = \frac{y_{\max} - y_{\min}}{\Delta y},$$

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

$$H_m^{y_i}(i, \Delta t) = f_{1i}\left[H_m^x(i, \Delta t)\right];$$
$$R_{yi}(i, \Delta t) = f_{2i}\left[R_x(i, \Delta t)\right];$$
$$m_{yi}(i, \Delta t) = f_{3i}\left[m_x(i, \Delta t)\right];$$
$$\sigma_{yi}(i, \Delta t) = f_{4i}\left[\sigma_x(i, \Delta t)\right] \qquad (i = 1, 2, 3, \ldots, k).$$

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:

$$H_m^y = f_{11}(H_m^{x_1}); \quad m_y = f_{31}(m_{x_1});$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$H_m^y = f_{1m}(H_m^{x_n}); \; m_y = f_{3m}(m_{x_n});$$
$$R_y = f_{21}(R_{x_1}); \quad \sigma_y = f_{41}(\sigma_{x_1});$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$R_y = f_{2m}(R_{x_n}); \quad \sigma_y = f_{4m}(\sigma_{x_n}).$$

При одномерно-многомерной схеме взаимодействия объекта в нестационарном режиме в качестве математической модели объекта принимаются функциональные зависимости

$$H_m^{y_i} = f_{1i}(H_m^x);$$
$$R_y^i = f_{2i}(R_x);$$
$$M_{y_i} = f_{3i}(m_x);$$
$$\sigma_{y_i} = f_{4i}(\sigma_x) \qquad (i = 1, 2, 3, \ldots, n).$$

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 приравниваться 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.

3.3 Вероятностно-статистические методы в теоретических исследованиях

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 появлений случайной величины и представляет собой ряд распределения (рисунок 3.6), а плавная кривая – функцию плотности распределения $f(x)$.

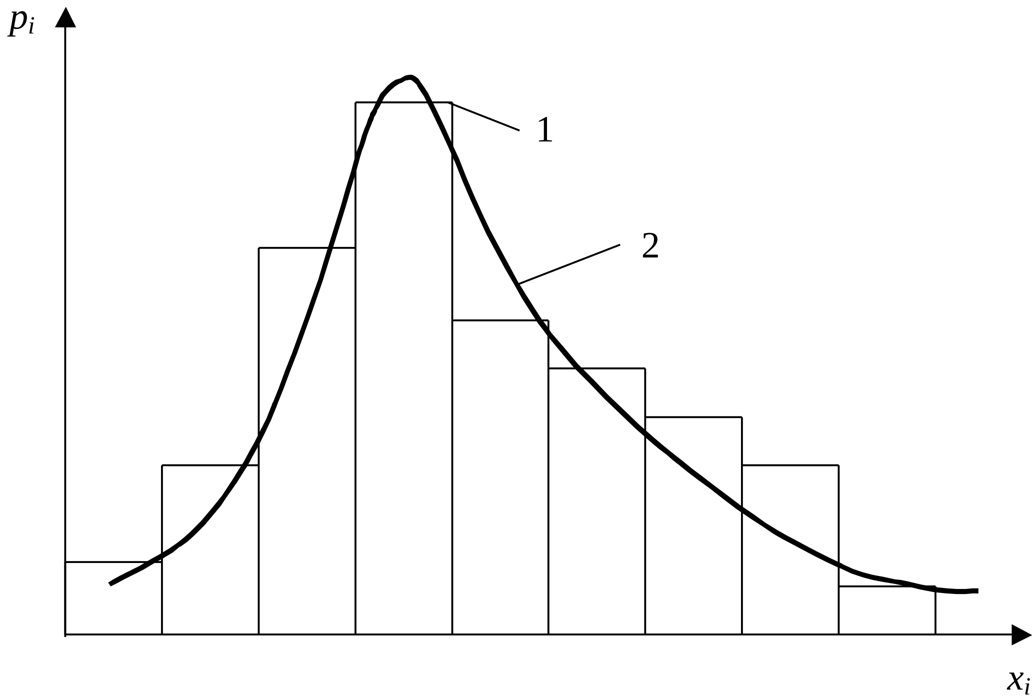
Вероятность случайной величины (события) – это количественная оценка возможности ее появления. Достоверное событие имеет вероятность $p = 1$, невозможное событие $p = 0$. Следовательно, для случайного события $0 \le p(x) \le 1$, а сумма вероятностей всех возможных значений

$$\sum_{0}^{n} p_i = 1.$$

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 3.6 – 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 $y_1, \ldots, y_n$ принять их вероятности $p_1, \ldots, p_n$, то это даст важную характеристику распределения – математическое ожидание:

$$m(x) = \sum_{1}^{n} x_i p_i \,. \qquad (3.1)$$

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 (3.1) $m(x) = 1\times0.10 + 2\times0.15 + 3\times0.45 + 4\times0.30 + 5\times0 = 2.95$.

Для непрерывных случайных величин математическое ожидание определяется интегралом

$$m(x) = \int_{-\infty}^{+\infty} x p(x) dx \,,$$

т.е. оно равно действительному значению $x_д$ 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 3.6), 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

$$D(x) = \sum_{1}^{n} \left(x_i - m(x)\right)^2 p_i \, .$$

Для рассмотренного выше примера $D(x)$ = (1 – 2.95)^2·0.10 + (2 – 2.95)^2·0.15 + (3 – 2.95)^2·0.45 + (4 – 2.95)^2·0.30 + (5 – 2.95)^2·0 = 0.85.

Важной характеристикой теоретической кривой распределения является среднеквадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)} \, .$$

Коэффициент вариации

$$k_в = \sigma/m(x)$$

применяется для сравнения интенсивности рассеяния в различных совокупностях, определяется в относительных единицах ($k_в < 1$).

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, …, 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 $\bar{x} = \sum_{1}^{n} x_i / n$, дисперсия $D = \sum_{1}^{n} (x_i - \bar{x})^2 / n$ 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 3.7)

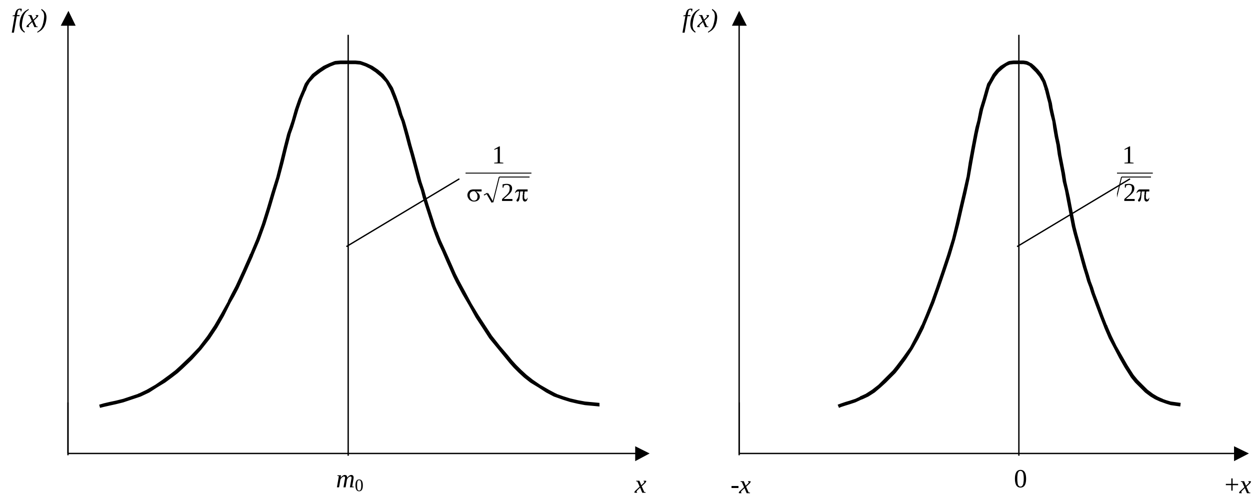
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - m(x))^2}{2\sigma^2} \right].$$

Это уравнение соответствует функции нормального распределения при $m(x) \neq 0$ (рисунок 3.7, *а*). Если совместить ось ординат с точкой m, т.е. $m(x) = 0$ (рисунок 3.7, *б*), и принять $\sigma^2 = 1$, то закон нормального распределения описываются зависимостью

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2} \right],$$

если за единицу масштаба принять дисперсию σ^2.

а) *б)*



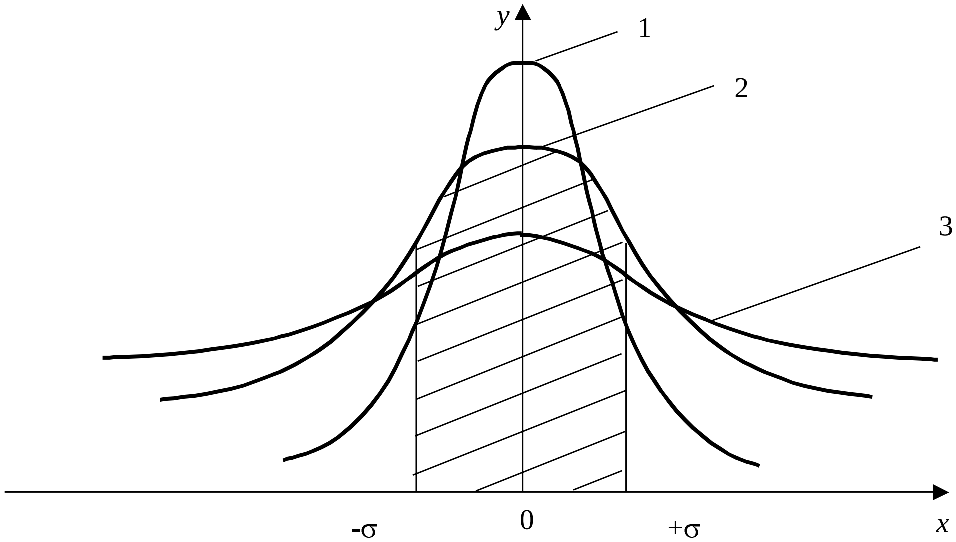
$а – m(x) \neq 0$; $б – m(x) = 0$

Рисунок 3.7 – 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 3.8). 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 σ = 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 +σ и –σ соответствует точкам перегиба кривой (заштрихованная площадь на рисунке 3.8). 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

$$\Phi(x) = \frac{2}{2\pi}\int_0^t e^{-x_i^2}\,dx.$$



$1 - \sigma = 0.5$; $2 - \sigma = 1.0$; $3 - \sigma = 2.0$

Рисунок 3.8 – Характер рассеяния кривой нормального распределения

При анализе многих случайных дискретных процессов пользуются распределением Пуассона.

Так, вероятность появления числа событий $x = 1, 2, 3, \ldots$ в единицу времени определяется законом Пуассона (рисунок 3.9) 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.

Например, в процессе наблюдений установлено, что за 5 мин на конвейер в среднем поступает шесть комплектующих узлов РЭА. Какова вероятность поступления 10 узлов РЭА за 5 мин?

В этом случае $x = 10$, $\lambda t = 6$, $p(x) = \dfrac{6^{10}e^{-6}}{10!} = 0.041$. Как видно, эта вероятность очень мала.

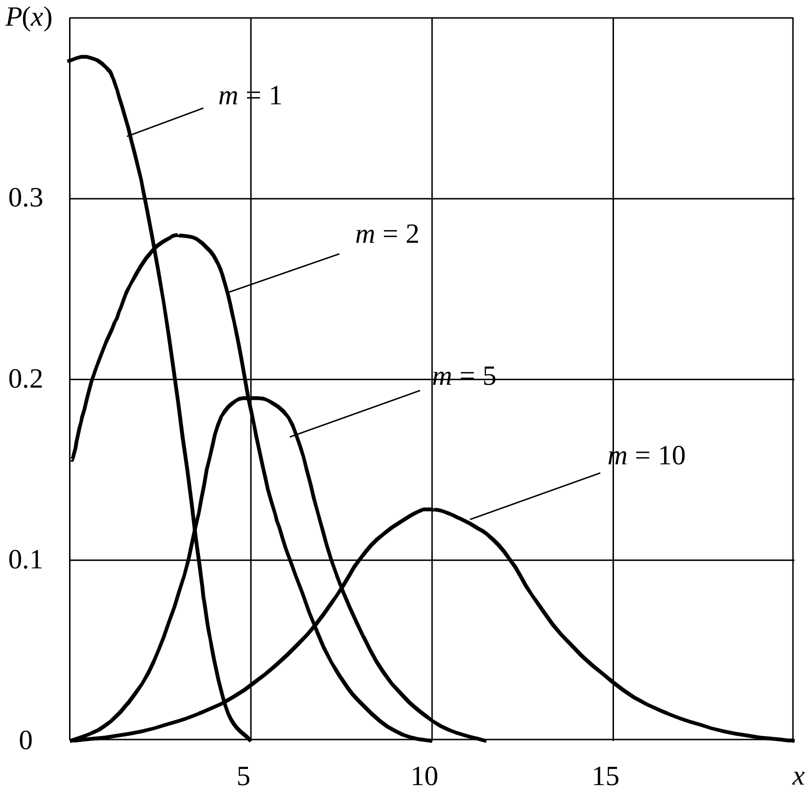


Рисунок 3.9 – Общий вид кривой распределения Пуассона

Для исследования количественных характеристик некоторых процессов (время сервисного обслуживания компьютеров на стадии предпродажной подготовки, время отказов узлов и блоков РЭА, длительность телефонных разговоров и т.д.) можно применять показательный закон распределения (рисунок 3.10, *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 3.10, *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).

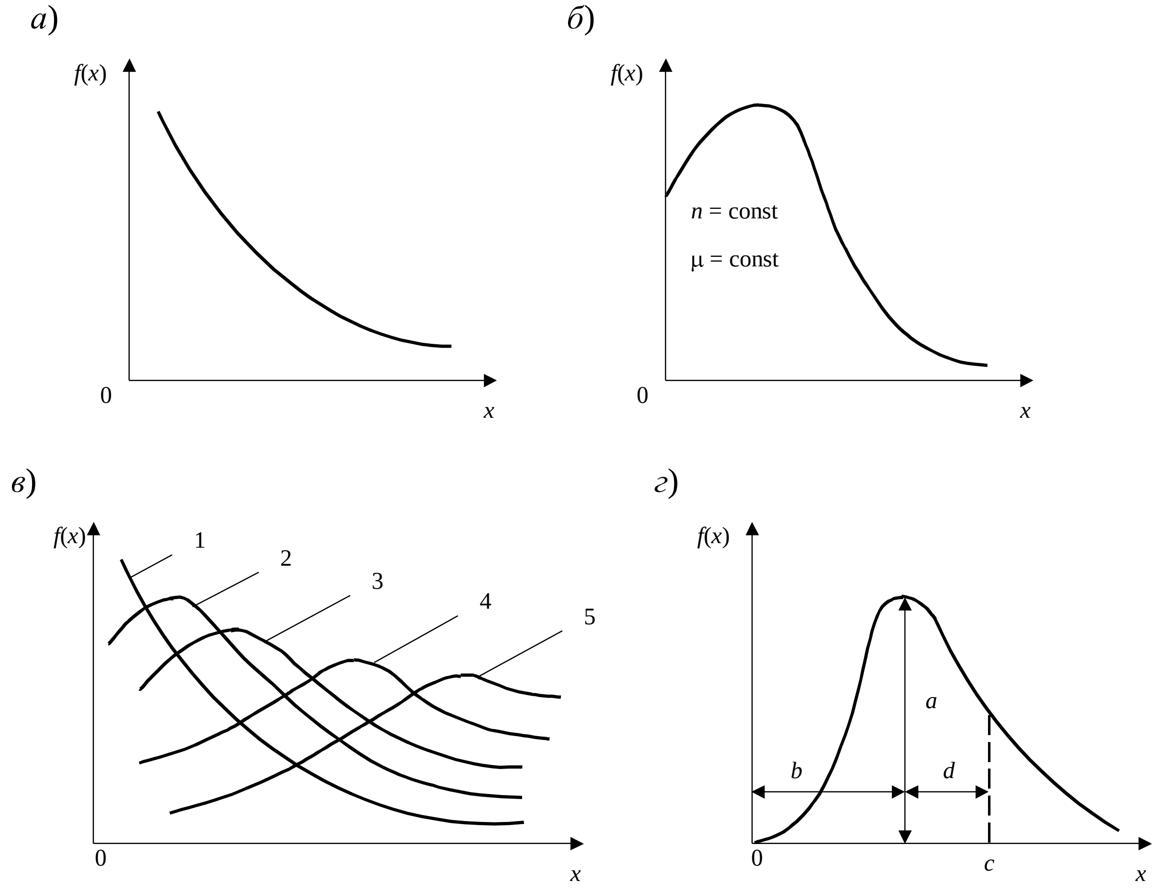
Исследуя процессы, связанные с постепенным снижением параметров (ухудшением свойств материалов по времени, деградация конструкций, процессы старения, износовые отказы в блоках РЭА и др.), применяют закон γ-распределения (рисунок 3.10, *в*)

$$f(x) = (\lambda^{\alpha}/\alpha!)x^{\alpha-1}e^{-\lambda x};$$

где λ, α – параметры. Если $\alpha = 1$, γ-функция превращается в показательный закон (рисунок 3.10, *а*).

При исследовании многих процессов, связанных с установлением расчетных характеристик, материалов, элементов и т.п., используют закон распределения Пирсона (рисунок 3.10, *г*), чаще всего представляемый в виде

$f(x)=ae^{dx}\left(1+\dfrac{x}{b}\right)^{db}$, где a – максимальная ордината; d, b – соответственно расстояния от максимальной ординаты до центра распределения C и начала координат 0.



а – показательное; *б* – Вейбулла; *в* – γ-распределения (*1* – α = 1; λ = 1; *2* – α = 3; λ = 1; *3* – α = 4; λ = 1.5; *4* – α = 5; λ = 2; *5* – α = 6); *г* – Пирсона

Рисунок 3.10 – 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

$$Q_1 = n\sum_{i=1}^{m}\left(\bar{x}_i - \bar{x}\right)^2,$$

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

$$Q_2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2,$$

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 $J \le J_T$, 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.

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.

$$\lim_{n\to\infty} p\left\{\left|\frac{\sum x_i}{n} - m(x)\right| < \varepsilon\right\} \to 1,$$

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. $\bar{x} \to m(\bar{x})$. При нормальном законе распределения точность результатов, полученных методом Монте-Карло, оценивается по формуле $p\left|\bar{x} - m(x)\right| < \frac{3\sigma}{\sqrt{n}}$. 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 $\varepsilon_{D_i} \le \varepsilon_D$ (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 $n_с$. 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 $\varphi(x_1, x_2, \ldots, x_n)$ в области значений параметров $x_1, x_2, \ldots, x_n$. 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 3.11). 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 $\vec{g}$. По направлению градиента начинается движение с шагом $c\vec{g}$ 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.

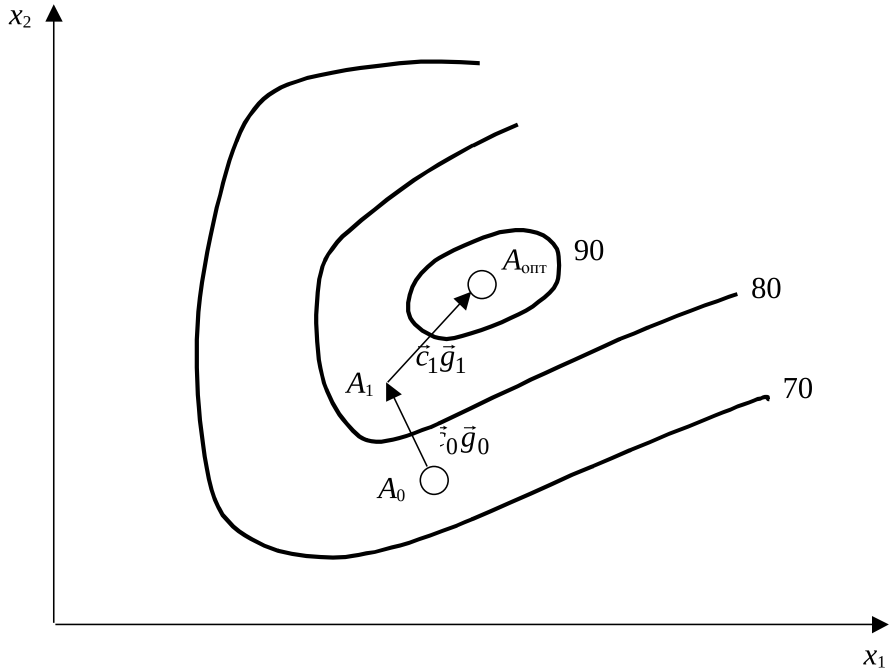


Рисунок 3.11 – Схема движения к оптимуму по градиенту (крутое восхождение)

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 $f(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} c_i x_i \to \min(\max)$. Ограничения задаются в виде линейных неравенств

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{im}x_m \ge b_i;$$
$$x_1, x_2, \ldots, x_n \ge 0;\ i = 1, 2, \ldots, m,$$

где a_{ij}, b_i, c_i – константы; x_1, x_2, x_3, …, 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

$$f(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_i x_j .$$

Среди задач нелинейного программирования встречаются и такие, в которых ограничения не имеют дискретных переменных. В них функции $f(x_1, x_2, \ldots, x_n)$ 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

$$f(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} a_{ij} x_j ,\ i = 1, 2, \ldots, m,$$

где $x_i \ge 0$, $j = 1, 2, \ldots, n$ – 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 расчленяются 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

$$\omega = \sum_{k=0}^{N-1} f_0 \left[x(k), u(k)\right] = \min(\max) \ ,$$

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, …, u_{N-1} во все дискретные моменты времени $k = 0, 1, \ldots, N-1$, 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.

4 ОСОБЕННОСТИ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ПРОЕКТИРОВАНИЯ И ПРОИЗВОДСТВА РЭС

4.1 Роль математического моделирования в проектировании и технологии РЭС

Современная технологическая система – это совокупность взаимосвязанных потоков энергии, материалов и информации, действующая как единое целое, в котором осуществляется определенная последовательность технологических процессов.

Технологическим системам, которым соответствуют отдельные пределы, производства и технологические цеха современных предприятий, свойственны все характерные признаки больших систем: определенная целенаправленность или наличие общей цели функционирования всей системы (все технологические аппараты и потоки объединены для выпуска продукции); большие размеры как по числу элементов, составляющих систему, так и по числу параметров, характеризующих процесс ее функционирования (большое число аппаратов, связанных технологическими потоками); сложность поведения системы, проявляющаяся в большом числе переплетающихся взаимосвязей между ее переменными (изменение режима работы одного аппарата может оказывать влияние на работу производства в целом); выполнение системой в процессе ее функционирования некоторой сложной и многофакторной целевой функции.

В современной практике проектирования больших промышленных систем часто используется эмпирический подход. Это объясняется тем, что большую систему принципиально невозможно точно описать и точно предсказать ее поведение. Единственный метод, позволяющий облегчить проектирование (а часто и эксплуатацию) такой системы, – это моделирование и в первую очередь – математическое.

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 системы. 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.

Математические модели используют при прогнозировании поведения моделируемых объектов. Например, строить современный реактивный самолет лишь для определения его летных характеристик экономически нецелесообразно, если они могут быть предсказаны средствами моделирования. На математических моделях выполняют контролируемые эксперименты в тех случаях, когда экспериментирование на реальных объектах практически невозможно из-за отсутствия последних или возникающей во время экспериментов опасности (сети энергоснабжения, химические производства).

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.

4.2 Аналитические методы в моделировании

Результатом применения экспериментального или теоретического методов исследования является синтез математической модели. 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 нерационально. 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-Кутта, 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:

$$\frac{d^2y}{dt^2}+\omega^2 y=0;$$

при начальных условиях

$$y(0)=y_0,$$
$$\frac{dy(0)}{dt}=y_0.$$

Частное решение этого уравнения представляется в виде

$$y=y_0\cos\omega t+\frac{y_0}{\omega}\sin\omega t.$$

Принимая $\frac{dy}{dt}$ за новую искомую функцию и вводя обозначения

$$y=z_1,$$
$$\frac{dy}{dt}=z_2,$$

преобразуем исходное дифференциальное уравнение в систему уравнений первого порядка:

$$\dot z_1=z_2;$$
$$\dot z_2=-\omega^2 z_1;$$

при начальных условиях

$$z_1(0)=y_0,\quad z_2(0)=\dot y_0.$$

Частное решение имеет вид

$$z_1=y_0\cos\omega t+\frac{\dot y_0}{\omega}\sin\omega t;$$
$$z_2=-y_0\omega\sin\omega t+\dot y_0\cos\omega t.$$

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 4.1).

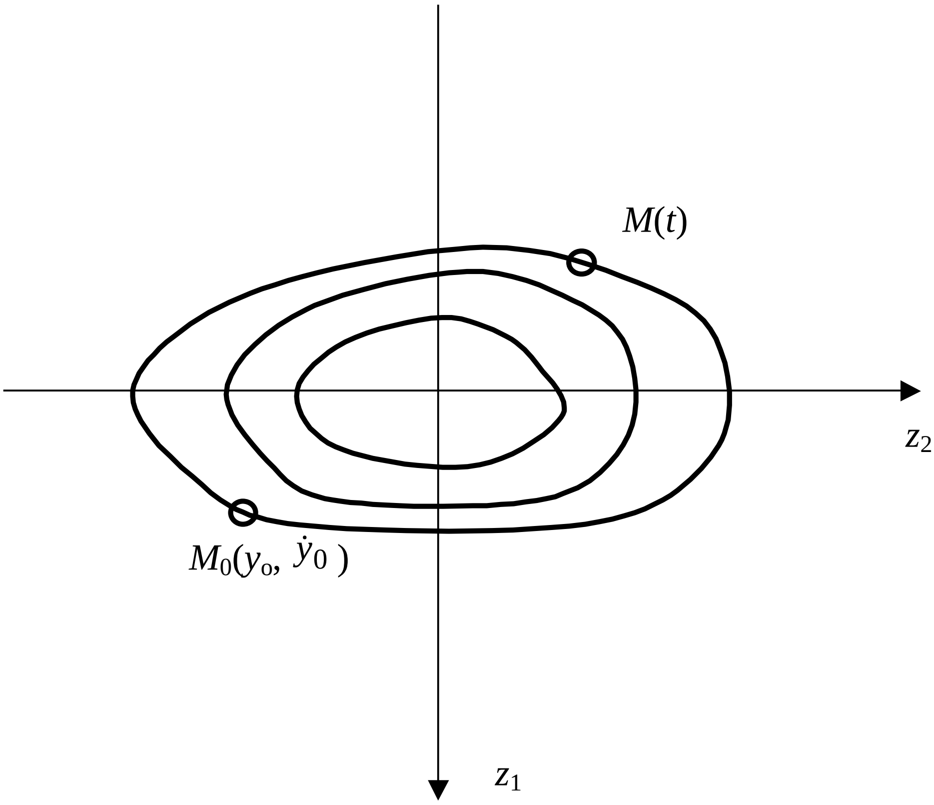


Рисунок 4.1 – Фазовый портрет системы

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 $z_1 z_2$, 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 $x_0 \le x \le x_1$ (рисунок 4.2). Нетрудно видеть, что длина кривой s_1, площадь p криволинейной трапеции, объем тела вращения v зависят от вида заданной кривой $y = f(x)$

$$s_1 = \int_{x_0}^{x_1} \sqrt{1 + [y'(x)]^2} dx; \quad p = \int_{x_0}^{x_1} y(x) dx; \quad v = \pi \int_{x_0}^{x_1} [y(x)]^2 dx.$$

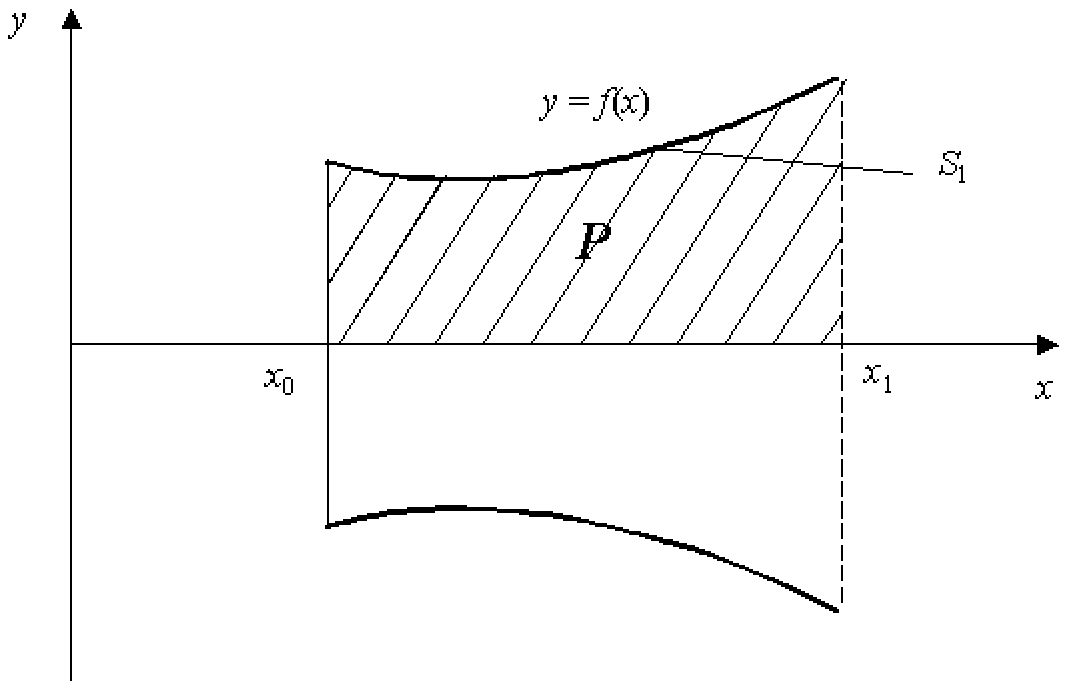


Рисунок 4.2 – Схема к понятию функционала

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 $x_0 \le x \le x_1$, 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 4.3) 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.

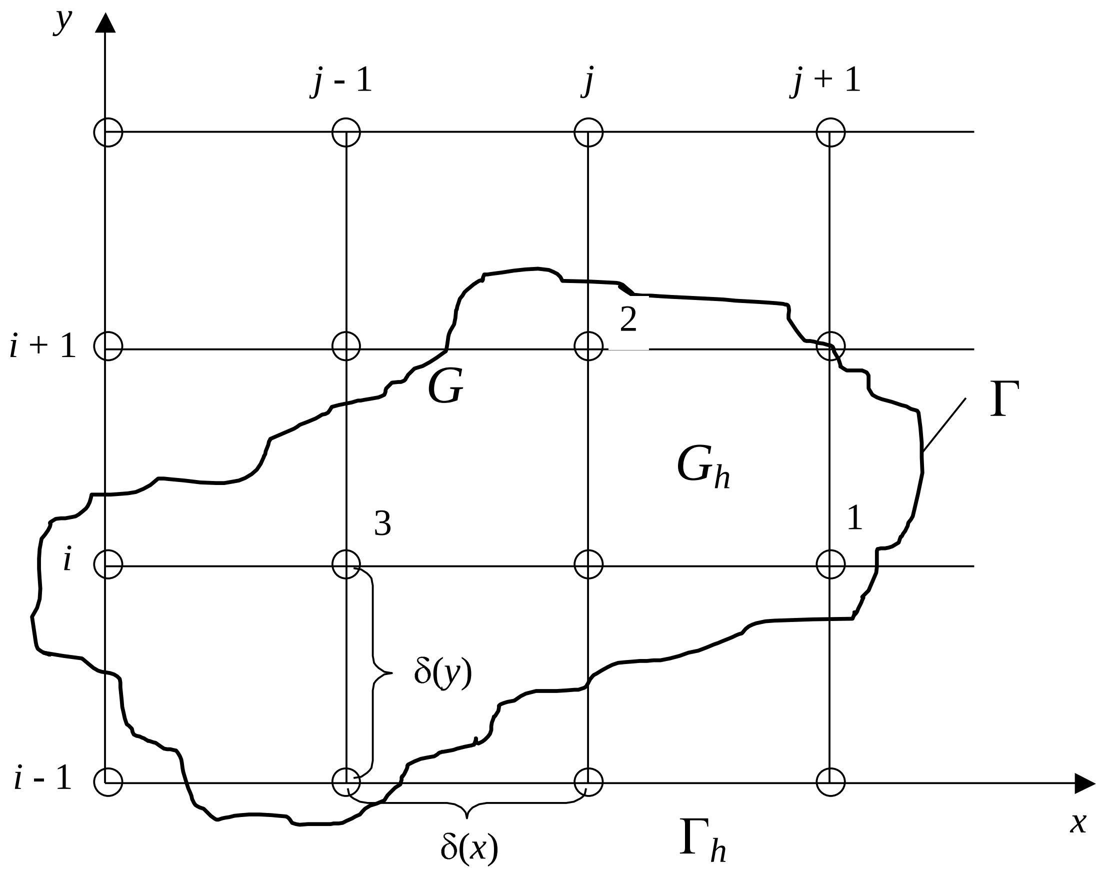


Рисунок 4.3 – Сеточная область G_h с контуром Г для плоской области *G*, в которой производится решение двумерного дифференциального уравнения

В качестве примера использования численных методов может служить решение задачи виброформования компаунда для заливки блока РЭА. Компаундная 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 компаундной смеси; t – время; x, y – декартовы координаты; v_x – составляющая скорости компаундной смеси вдоль направления распространения колебаний; E – модуль упругости; l – смещение элемента компаундной 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:

$$u = l/A\omega, \qquad W = v_x/A\omega, \qquad dx = dy = cdt,$$

где c – скорость распространения колебаний в компаундной смеси; $A\omega$ – амплитуда колебаний.

Тогда математическая модель процесса виброформования компаундной массы выразится уравнением

$$\frac{\partial W}{\partial t} = \frac{c^2}{\omega}\frac{\partial^2 u}{\partial x^2} + v\left(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2}\right).$$

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 $x_i = x_0 + i\delta x$; $y_i = y_0 + j\delta y$ ($i, j = 0, \pm 1, \pm 2, \ldots$) (рисунок 4.3). Тогда для первой производной функции $f(x)$ получим следующие варианты записи:

$$\frac{\partial f}{\partial x} \approx \frac{f_{i+1} - f_i}{\delta x}; \qquad \frac{\partial f}{\partial x} \approx \frac{f_i - f_{i-1}}{\delta x}; \qquad \frac{\partial f}{\partial x} \approx \frac{f_{i+1} - f_{i-1}}{2\delta x};$$

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 компаундной массы примет вид

$$W_0^{+\delta t} = W_0 + \frac{1}{\bar{\delta} t}\left[\frac{v\omega}{c^2}\left(\Delta_x^2 W + \Delta_y^2 W\right) + \Delta_x^2 u\right],$$

где

$$\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} \approx \frac{\Delta_x^2 W}{\delta x^2} + \frac{\Delta_y^2 W}{\delta y^2}; \qquad \Delta_x^2 u \approx \partial^2 u;$$

$$\frac{\Delta_x^2 W}{\delta x^2} \approx \frac{W_{i,j+1} - 2W_{i,j} + W_{i,j-1}}{\delta x^2};$$

$$\frac{\Delta_y^2 W}{\delta y^2} \approx \frac{W_{i+1,j} - 2W_{i,j} + W_{i-1,j}}{\delta y^2};$$

$$\frac{\Delta_x^2 u}{\delta x^2} \approx \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\delta x^2};$$

$\delta t + \delta t\omega$ – 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

$$a = \frac{\delta W}{\bar{\delta} t} = \frac{1}{\bar{\delta} t^2}\left[\left(\Delta_x^2 W + \Delta_y^2 W\right) + \Delta_x^2 u\right],$$

где a – безразмерное ускорение элементарного объема среды; $\bar{k} = \frac{\nu\omega}{c^2} = \frac{\eta\omega}{E}$ – 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

$$u^{+\delta u} = u + W\delta t + \frac{1}{2}a\delta t^2; \qquad W^{+\delta t} = W + a\delta t.$$

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 $u_г = f_1(t)$ и $W_г = f_2(t)$, 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 компаундной смеси (рисунок 4.4). Естественно, что вычисления в этом случае громоздки и необходимо использование специальных программ и ЭВМ.

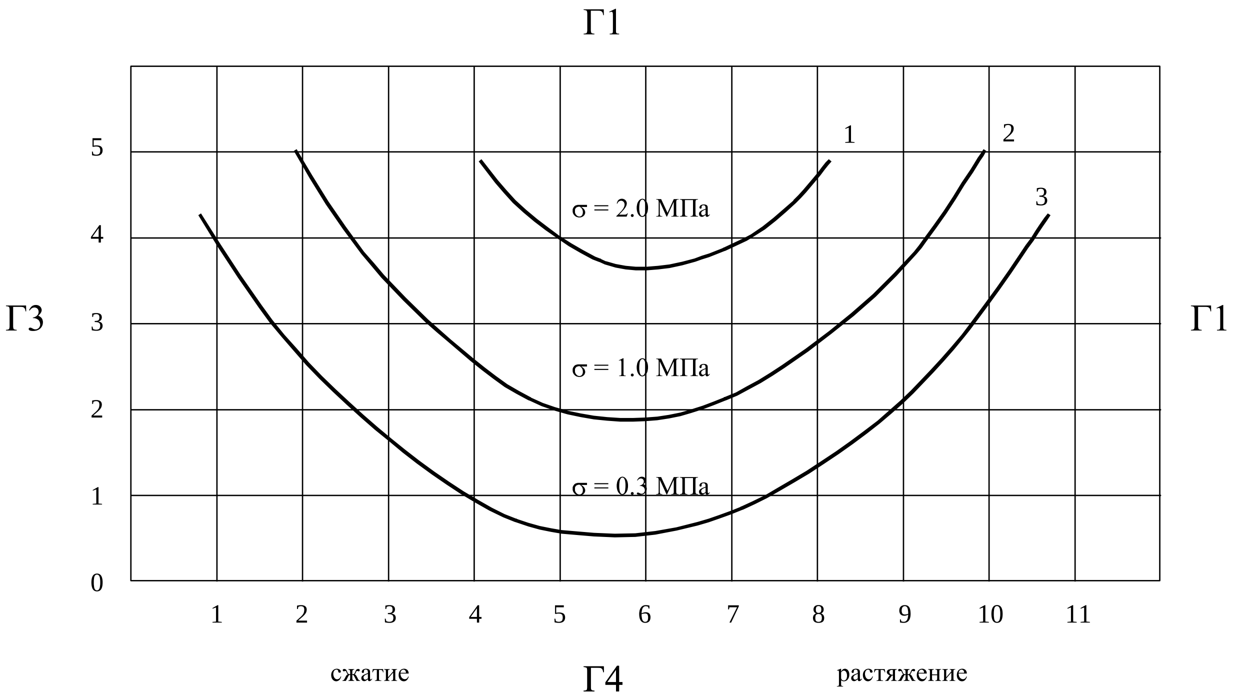
Использование аналитических методов решения математических задач является основным методом современного научного исследования. Однако громоздкость моделей и прямых методов решения уравнений затрудняет получение конечных решений. Поэтому в решении практических задач нашли широкое применение методы преобразования 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 $\alpha = 0°$, частота колебаний 20 Гц, амплитуда – 3.5 мм;
1 – σ = 2.0 МПа; 2 – σ = 1.0 МПа; 3 – σ = 0.3 МПа

Рисунок 4.4 – Распределение кривых равного значения динамического давления в модели формы с компаундной 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.

4.3 Физическое подобие и моделирование

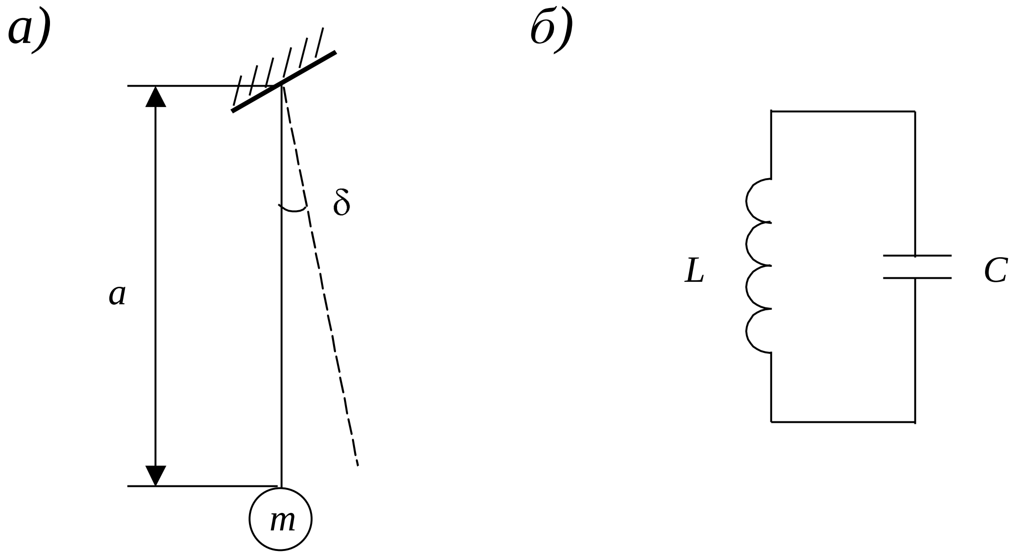
Физические модели, описывающие различные процессы в конструкциях РЭС, как правило, представляют собой уравнения или системы уравнений, сходные по виду. Для их решения, как это показано выше, могут быть использованы, соответственно, сходные методы и преобразования. При этом механические, тепловые, гидродинамические процессы могут быть на основании использования теории подобия представлены моделями, аналогичными моделям, описывающим электрические процессы. Соответственно, мы будем использовать электромеханические, электротепловые, электрогидродинамические аналогии в описании рассматриваемых процессов.

Например, 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 $\varepsilon \to 0$ 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 4.5, *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 4.5 – Простейшие динамические объекты

Аналогично анализ процессов в электрическом колебательном контуре (рисунок 4.5, *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

$$a_0\frac{d^2y}{dt^2}+a_1\frac{dy}{dt}+a_2y=0.$$

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.

В последние годы широкое распространение получили пакеты прикладных программ (ППП) для анализа схемотехнических процессов в электрических схемах, например, Electronics Workbench, MicroCAP, PSpice, DesignLab, OrCAD.

Используя теорию электрического подобия можно сводить задачу моделирования различных физических процессов к составлению эквивалентной электрической схемы рассматриваемого процесса и с помощью программ схемотехнического моделирования проводить анализ этого процесса, используя масштабные коэффициенты подобия.

В электротепловом подобии:

$$t_i = k_t \cdot U_i;\ \ P_i = k_P \cdot I_i;\ \ C_{Ti} = k_C \cdot C_{\text{ЭЛ}\ i};\ \ R_{Ti} = k_R \cdot R_{\text{ЭЛ}\ i},$$

где t_i, P_i, C_{Ti}, R_{Ti} – температура, мощность, теплоемкость, тепловое сопротивление i-ого элемента, соответственно; U_i, I_i, $C_{\text{ЭЛ}i}$, $R_{\text{ЭЛ}i}$ – напряжение, ток, электрическая емкость и электрическое сопротивление i-ого элемента замены, соответственно.

В электромеханическом подобии:

$$m_i = k_m \cdot C_{\text{ЭЛ}\ i};\ \ K_i = k_K \cdot L_i;\ \ D_i = k_D \cdot R_{\text{ЭЛ}\ i};\ \ A_i = k_A \cdot U_{\ i},$$

где m_i, K_i, D_i – масса, жесткость, коэффициент демпфирования i-ого элемента, L_i – индуктивность i-ого элемента замены.

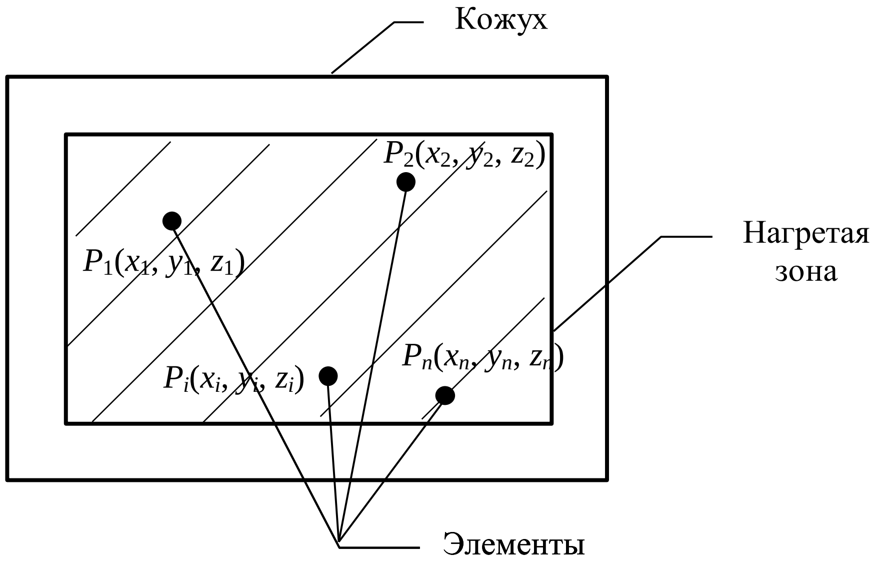
В обоих случаях k_t, k_P, k_C, k_R, k_m, k_K, k_D, k_A - соответствующие масштабные коэффициенты. При моделировании процессов внутренней электромагнитной совместимости для анализа составляется полная эквивалентная электроконструктивная схема радиоэлектронного устройства с учетом паразитных емкостей, индуктивностей, взаимоиндукций и магнитных связей, определяемых исходным вариантом конструкции.

На рисунках 4.6–4.9 приведены примеры составленных эквивалентных электрических схем для различных случаев моделирования физических процессов.

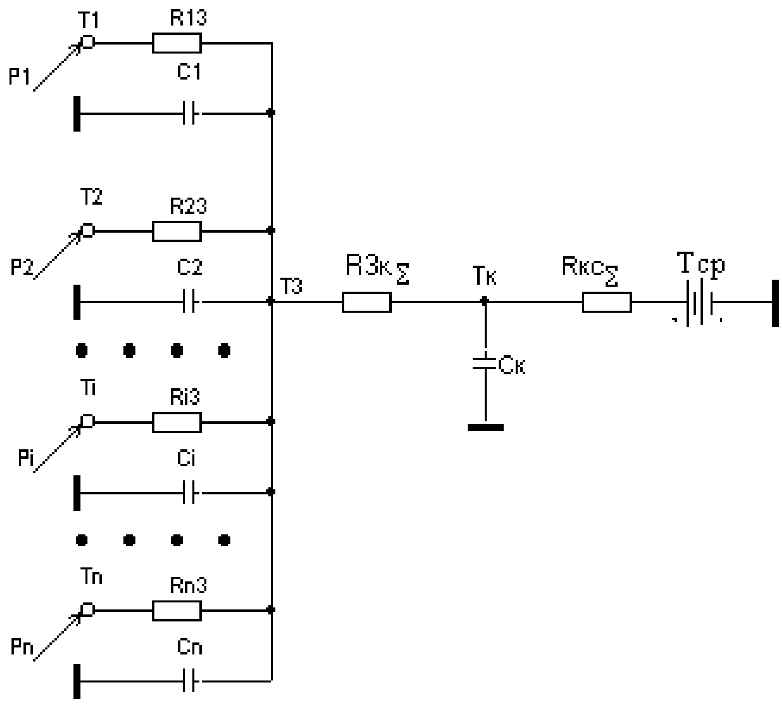
В процессе электрического моделирования указанных процессов необходимо сформировать массив исходных данных для каждой схемы в виде численных значений электрических параметров с учетом выбранных масштабных коэффициентов. Для формирования указанного массива необходимо детально рассмотреть каждый процесс 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 РЭС (рисунок 4.6), является коэффициент теплоотдачи:

$$\alpha_\Sigma = \alpha_\text{К} + \alpha_\text{Л},$$

где $\alpha_К$ – конвективный коэффициент теплоотдачи; $\alpha_Л$ – коэффициент теплоотдачи излучением.

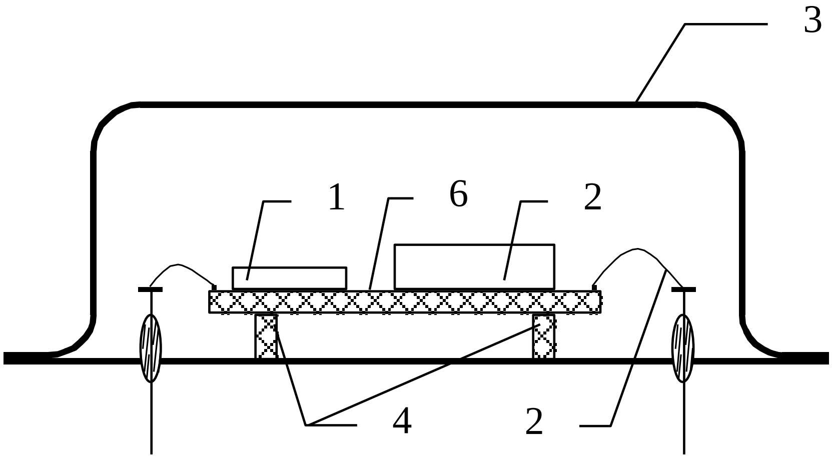


а) физическая модель

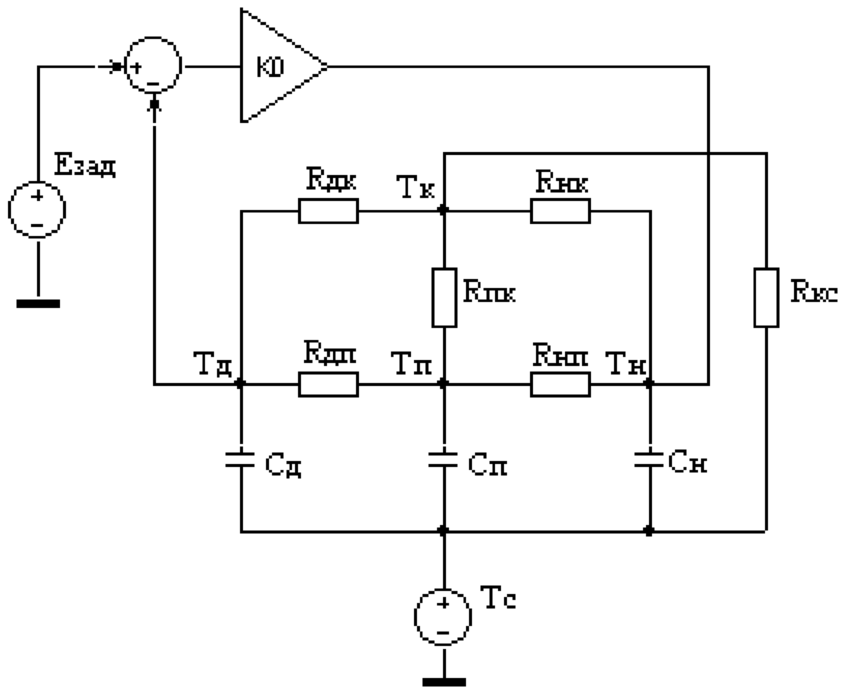


б) электротепловая модель

Рисунок 4.6 – Эквивалентная схема процесса теплообмена в блоке РЭС



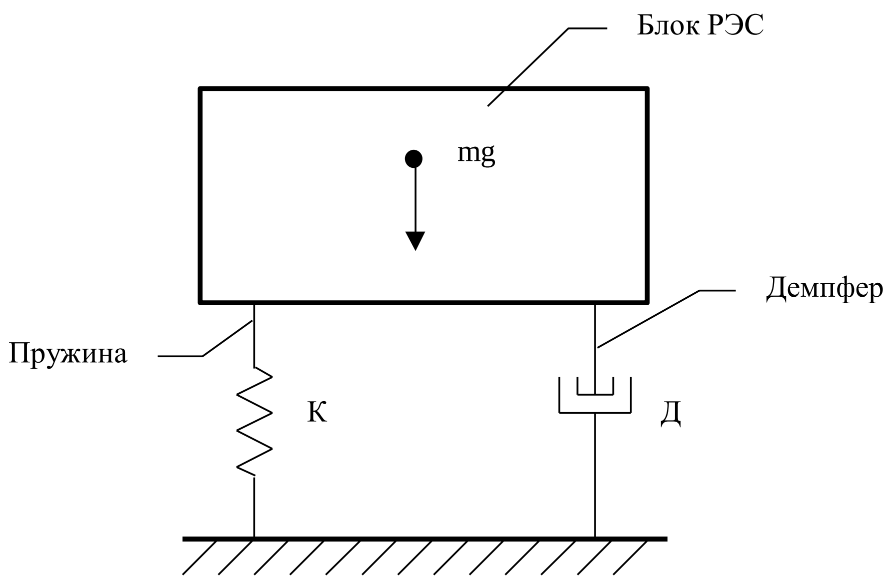
а) физическая модель



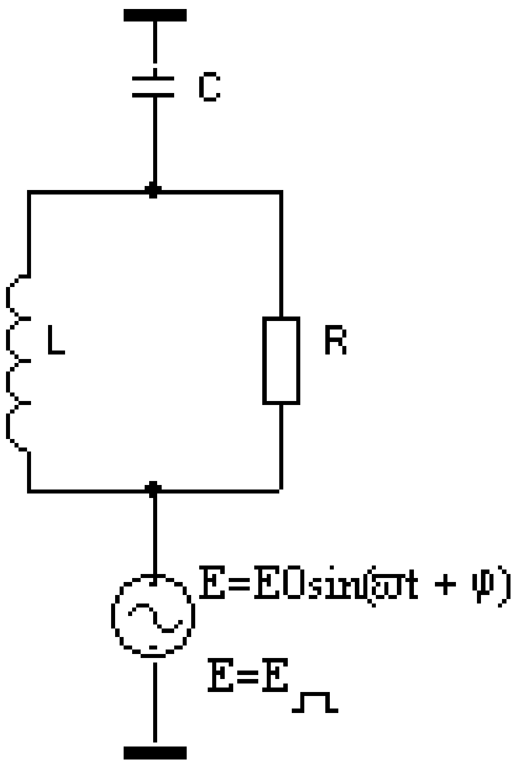
б) электротепловая модель

1 – термостатируемая схема; 2 – схема регулирования температуры; 3 – корпус; 4 – теплоизоляция; 5 – выводы; 6 – подложка

Рисунок 4.7 – Эквивалентная схема процесса регулирования температуры подложки в микротермостате

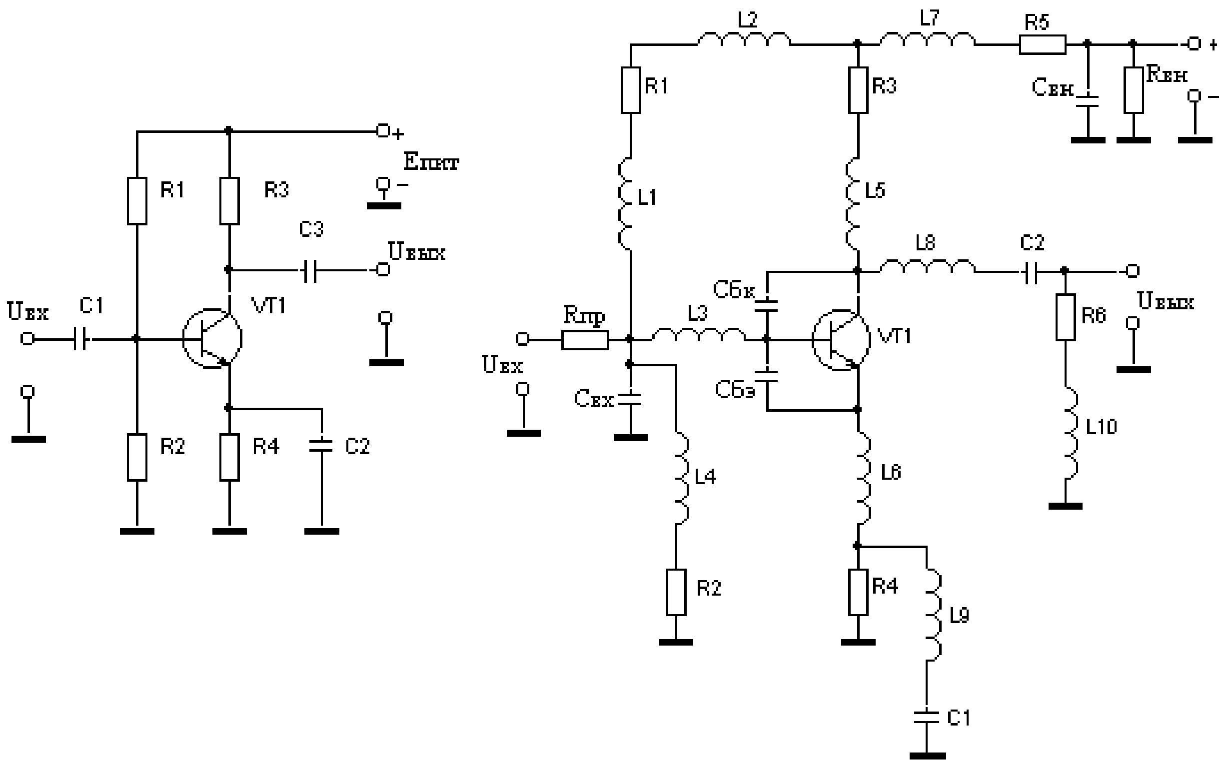


а) физическая модель



б) электромеханическая модель

Рисунок 4.8 – Электромеханическое моделирование системы амортизации



а) исходная схема б) эквивалентная схема

Рисунок 4.9 – Эквивалентная схема усилителя для определения паразитных параметров

Для ориентировочного определения величины α_Σ можно использовать хорошо зарекомендовавшие себя эмпирические методы расчета тепловых режимов РЭС, 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 моделировании процесса регулирования температуры подложки в микротермостате в виде гибридной интегральной схемы (ГИС) (рисунок 4.7) коэффициент передачи элементов схемы регулирования температуры и микротермостата (МТ) задается идеализированным операционным усилителем с коэффициентом усиления:

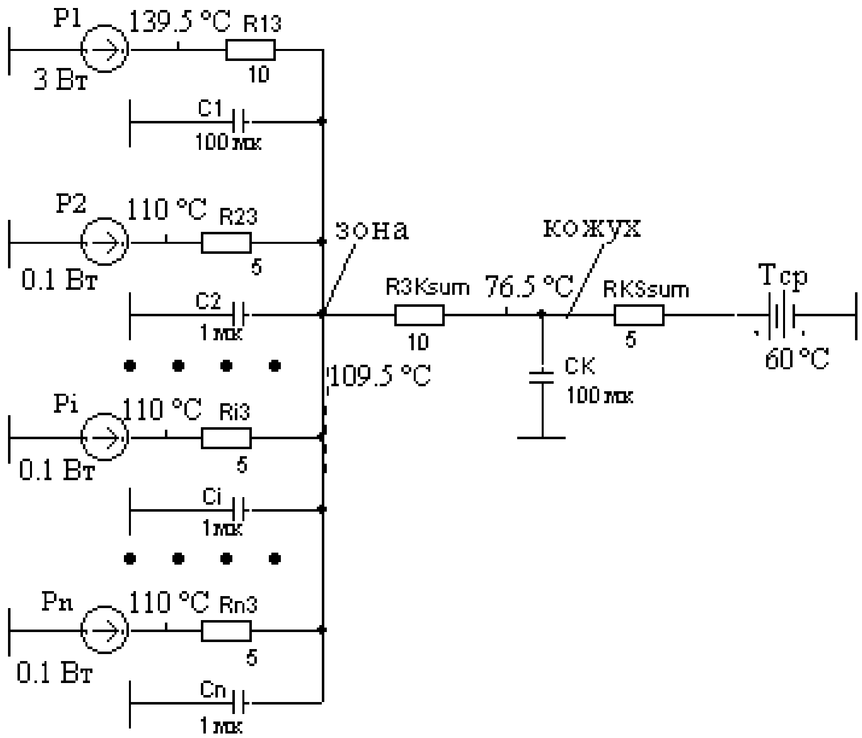
$$K_0 = \prod_{i=1}^{n} K_i ,$$

где n – число элементов схемы регулирования и МТ; K_i – коэффициент передачи i-ого элемента.

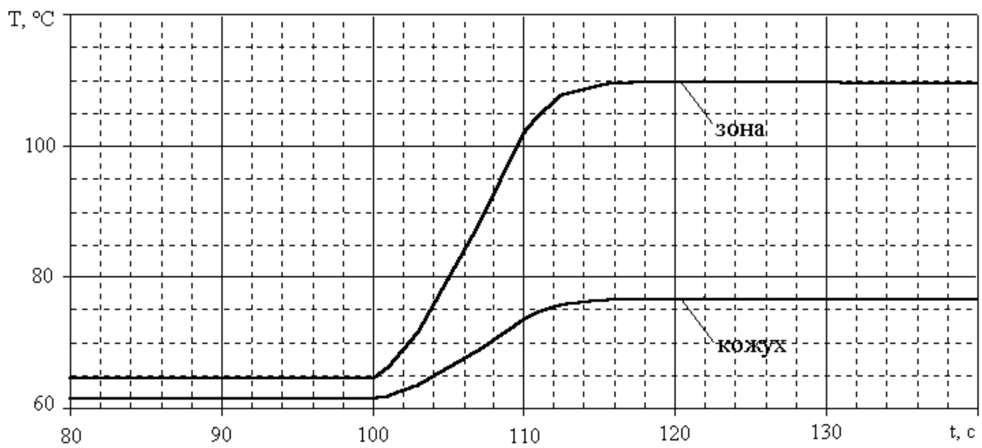
При электромеханическом моделировании системы амортизации (рисунок 4.8) вибрацию можно моделировать синусоидальным гармоническим колебанием, а удар – импульсным источником различной формы.

Для определения массива значений паразитных параметров L_i, C_j в схеме усилителя (рисунок 4.9) необходимо иметь вариант топологии рассматриваемого функционального узла.

Моделирование с использованием ППП схемотехнического моделирования показывает, что результаты анализа целесообразно представить в статическом и динамическом видах. Примеры результатов моделирования по схемам, приведенным на рис. 4.6–4.9, представлены на рисунках 4.10–4.13, соответственно.

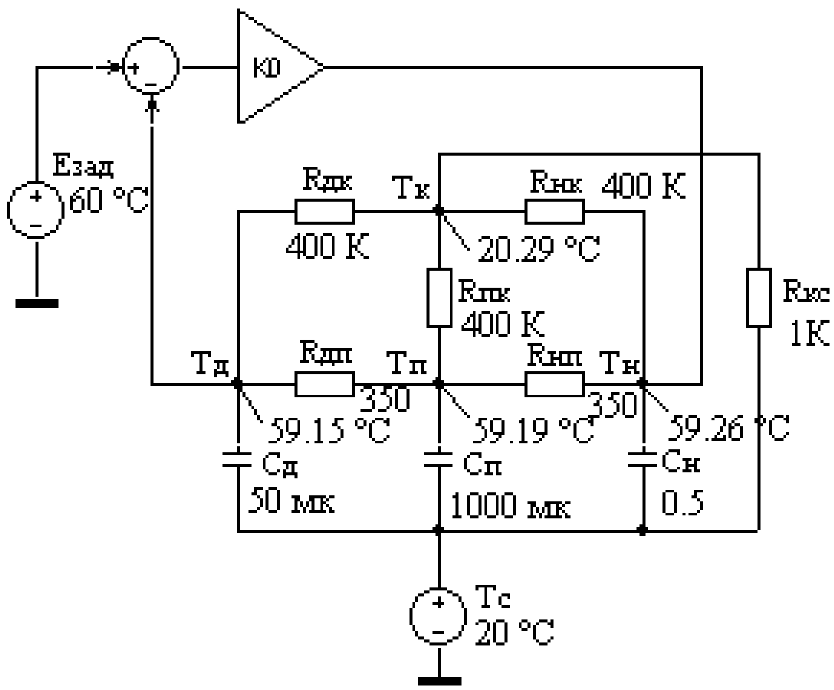


а) распределение температур

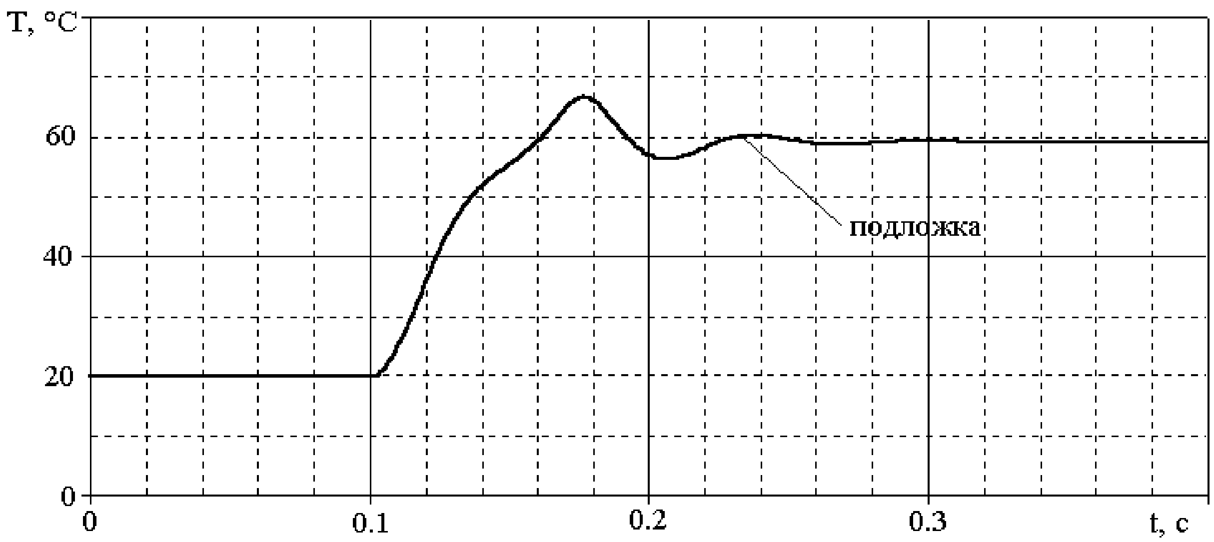


б) переходные характеристики

Рисунок 4.10 – Анализ переходных процессов теплообмена в блоке РЭС

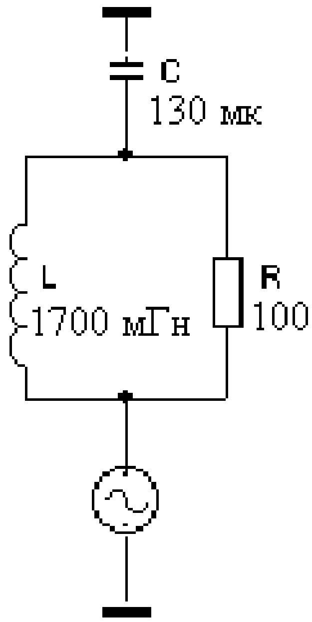


а) распределение температур

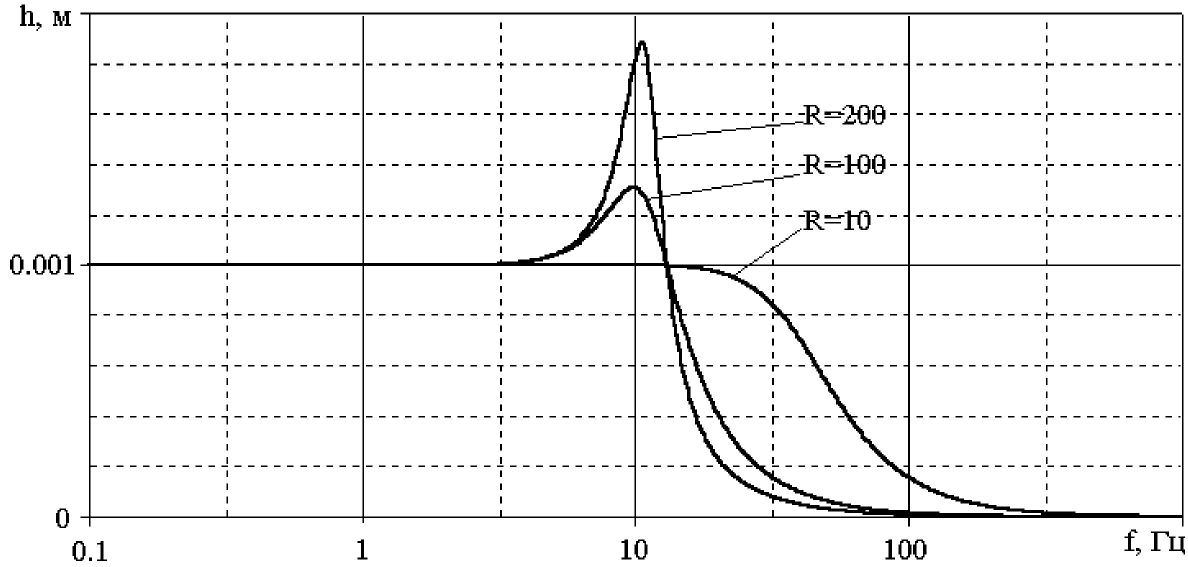


б) переходная характеристика

Рисунок 4.11 – Анализ пускового режима микротермостата

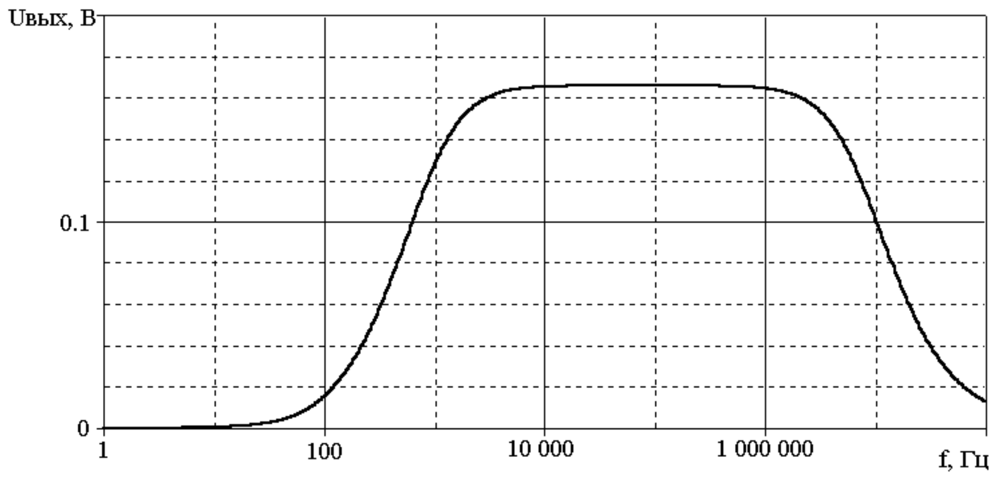


а) система амортизации



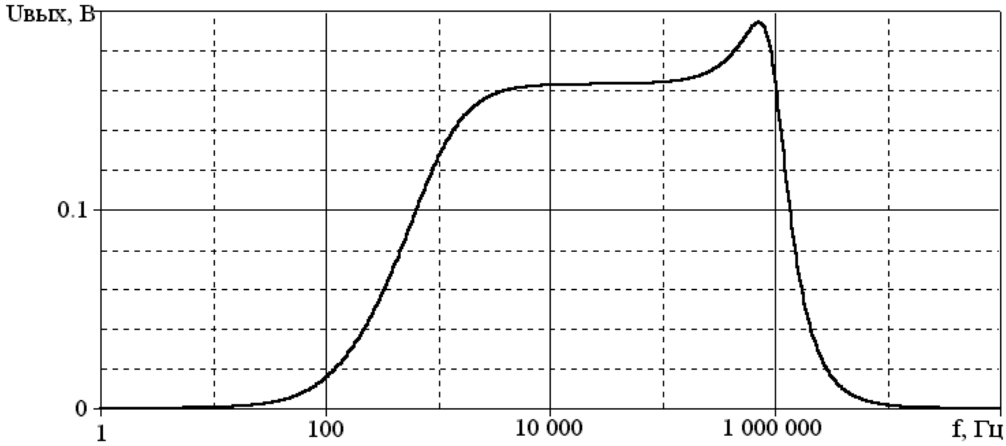
б) АЧХ системы амортизации

Рисунок 4.12 – Анализ системы амортизации блока РЭС



а) АЧХ исходной схемы

($L_Б$ = 0; $C_{БК}$ = 0; $C_{БЭ}$ = 0)



б) АЧХ эквивалентной схемы

($L_Б$ = 10 мкГн; $C_{БК}$ = 10 пФ; $C_{БЭ}$ = 10 пФ)

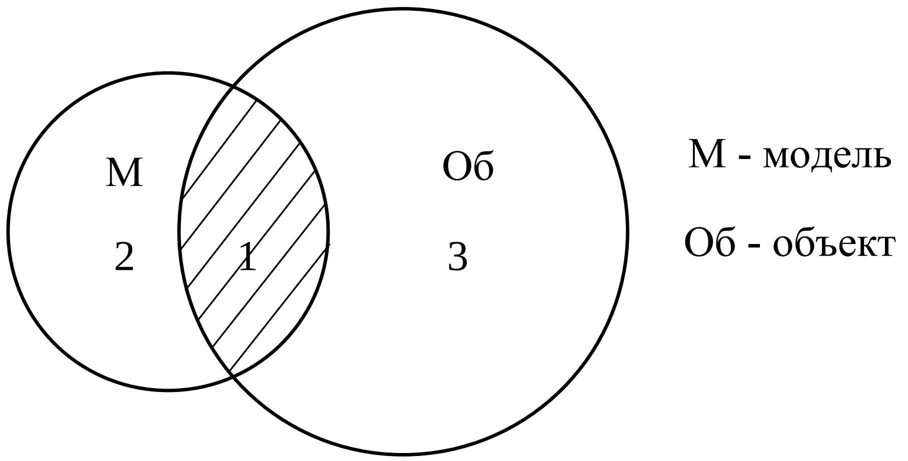
Рисунок 4.13 – Анализ паразитных эффектов эквивалентной схемы усилителя

Анализ различных физических процессов РЭС с использованием ППП 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 радиоэлектронных устройств.

В рассмотренных выше примерах применения программных продуктов для анализа электрических процессов в моделировании широко применяется вычислительная техника, а процесс анализа является, по сути, своеобразным вычислительным экспериментом. В таком эксперименте физическая модель исследуемого процесса заменяется алгоритмической, и исследование свойств рассматриваемого объекта проходит на этой алгоритмической модели. При этом одной из проблем, с которой сталкивается исследователь, является проблема виртуальности.

4.4 Проблема виртуальности в моделировании с использованием вычислительной техники

Виртуальность человеческого мышления существовала всегда. Она выражается, прежде всего, в условности моделей, которыми мы пользуемся, решая определенные задачи. Если рассматривать представления человечества о Космосе, то до сих пор можно считать, что они в большей степени виртуальны, чем реальны. Так же виртуальны и знания о микромире, поскольку «электрон также неисчерпаем, как и атом». В процессе научных исследований постоянно пользуются моделями различного уровня, которые, в общем случае, классифицируются как абстрактные и реальные, физические и математические, подлинники и оригиналы, строгие и приближенные, косвенного подобия и условного подобия и т.п. Степень приближения модели к объекту моделирования определяется ее адекватностью. На рисунке 4.14 представлена модель адекватности, определяющая наши представления об изучаемом или проектируемом объекте.



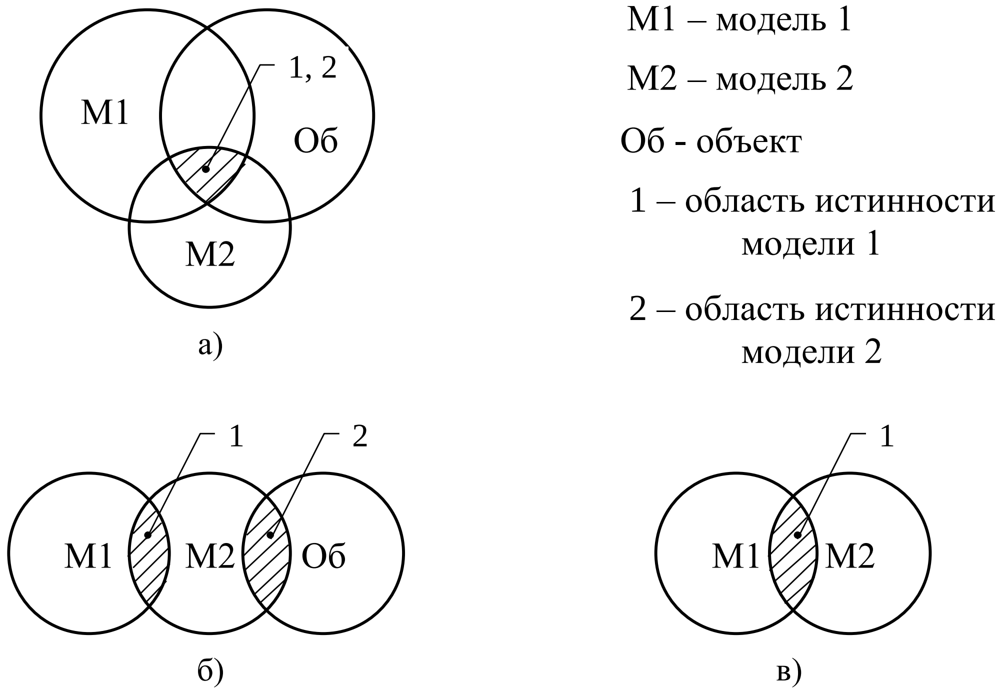
1- область истинности; 2 – область виртуальности; 3 – область недостаточности

Рисунок 4.14 – Адекватность модели и моделируемого объекта

Адекватность модели определяется тремя областями, а, точнее, их соотношением. Если *М = Об*, т.е. область истинности совпадает с моделируемым объектом, мы имеем дело с, так называемой, религиозной моделью, например, как абсолютное черное тело или вечный двигатель. Если *М – Об* = 0, то эта модель виртуальна и не адекватна. С помощью такой модели нельзя изучить объект, а, тем более, его правильно спроектировать. Также невозможна и эксплуатация технического объекта, если в распоряжении специалиста по эксплуатации нет модели с достаточной, на его уровне, областью истинности.

Проблема виртуальности инженерных знаний является в настоящее время весьма актуальной в связи с тем, что интенсивно развиваются компьютерная техника и компьютерное моделирование, информационные технологии, базы данных и вычислительные сети, а также телекоммуникационные системы и средства массовой информации (СМИ).

На рисунке 4.15 дана графическая интерпретация ситуации, связанной с увеличением числа источников информации в процессе получения знаний. Представим, что в процессе исследований формируются две модели *M1* и *M2*, получаемые из различных источников. В этом случае область истинности знаний уменьшается в варианте *а)* и теряется в варианте *б)*. Если считать моделью знаний *M1*, то на определенном этапе формирования мировоззрения, а, следовательно, и массива знаний у исследователя, происходит подмена изучаемого объекта его моделью, а, значит, переход в область виртуальных знаний. Реально это приводит к тому, что человек начинает мыслить условными категориями, жить в условном для него мире. Реальные цели также подменяются условными, достигая которых исследователь удовлетворяется, считая, что он выполнил свое предназначение. В результате синтезируются конструкции и технологические процессы, не обладающие свойством потребности, проектирование становится «бумажным», т.е. виртуальным. В технике это приводит к полному застою, прекращению выпуска конкурентоспособных изделий, увеличению сроков внедрения в производство.



а) с сохранением адекватности; *б)* с потерей адекватности;
в) с переходом в виртуальность

Рисунок 4.15 – Адекватность моделей знаний об изучаемом объекте

Говоря о преодолении виртуальности инженерного мышления, нельзя не остановиться на фундаментализации уровня исследований. Это один из основных и эффективных путей устранения условности творческой деятельности, направленный на синтез полезных технических решений. Только на основе фундаментальных и глубоких представлений об изучаемых объектах можно объективно оценить виртуальность полученных моделей. С помощью компьютерных технологий можно при рациональном подходе с фундаментальных позиций быстро получить нужный результат, но можно уйти в гипотетическую область виртуальных решений. Именно поэтому требуется не фетишизация компьютерных моделей, а их взвешенная оценка и грамотное применение в зависимости от уровня и степени сложности конкретной проблемной ситуации. Глубокая фундаментальная подготовка позволяет составить адекватную модель для изучения явлений, эффектов, закономерностей, а технологические знания дают возможность оценить степень виртуальности этой модели. Виртуальность как категория в области инженерного образования дает, при правильных методологических подходах, возможность наискорейшим путем приблизиться к решению поставленной задачи. Виртуальная модель дает множество решений, а тестовый эксперимент, прагматические знания в области технологии и такие субъективные категории, как опыт, интуиция и априорные сведения об исследуемом объекте дают возможность отсеять из множества возможных решений неприемлемые и выбрать оптимальные.

5 МЕТОД ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА В НАУЧНЫХ ИССЛЕДОВАНИЯХ

5.1 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.

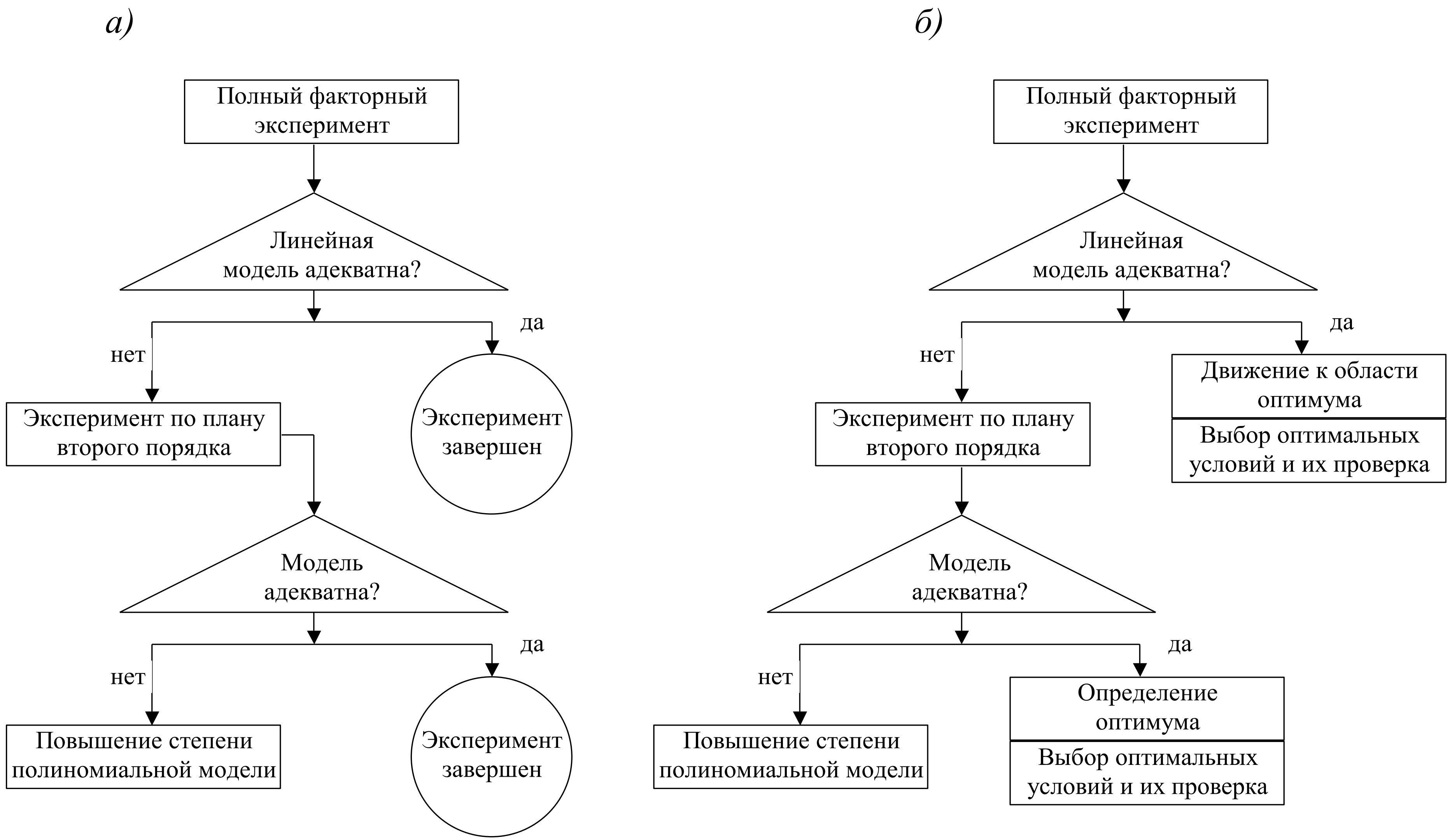
При последовательном проведении эксперимент выполняется не одновременно, а поэтапно, с тем, чтобы результаты каждого этапа анализировать и принимать решение о целесообразности проведения дальнейших исследований (рисунок 5.1).

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}, …, 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 5.1 – Структурная схема эксперимента

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

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \ldots + \beta_n z_n + \beta_{12} z_1 z_2 + \ldots +$$
$$+\beta_{(n-1)n} z_{n-1} z_n + \beta_{11} z_1^2 + \beta_{22} z_2^2 + \ldots + \beta_{nn} z_{nn}^2, \qquad (5.1)$$

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

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

Ряд Тейлора аналогичен уравнению регрессии:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + a_{12} x_1 x_2 + \ldots +$$
$$+a_{(n-1)n} x_{n-1} x_n + a_{11} x_1^2 + a_{22} x_2^2 + \ldots + a_{nn} x_{nn}^2. \qquad (5.2)$$

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

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

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 (5.2) вычисляются методами математической статистики и представляют собой приближенную оценку коэффициентов β_0, β_i, β_{ii} в уравнении (5.1). Следовательно, уравнение (5.2) описывает исследуемый объект (процесс) только с определенной степенью точности.

Уравнение (5.2) 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 (5.2), 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 (5.2) 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 (5.2) 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.

5.2 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:

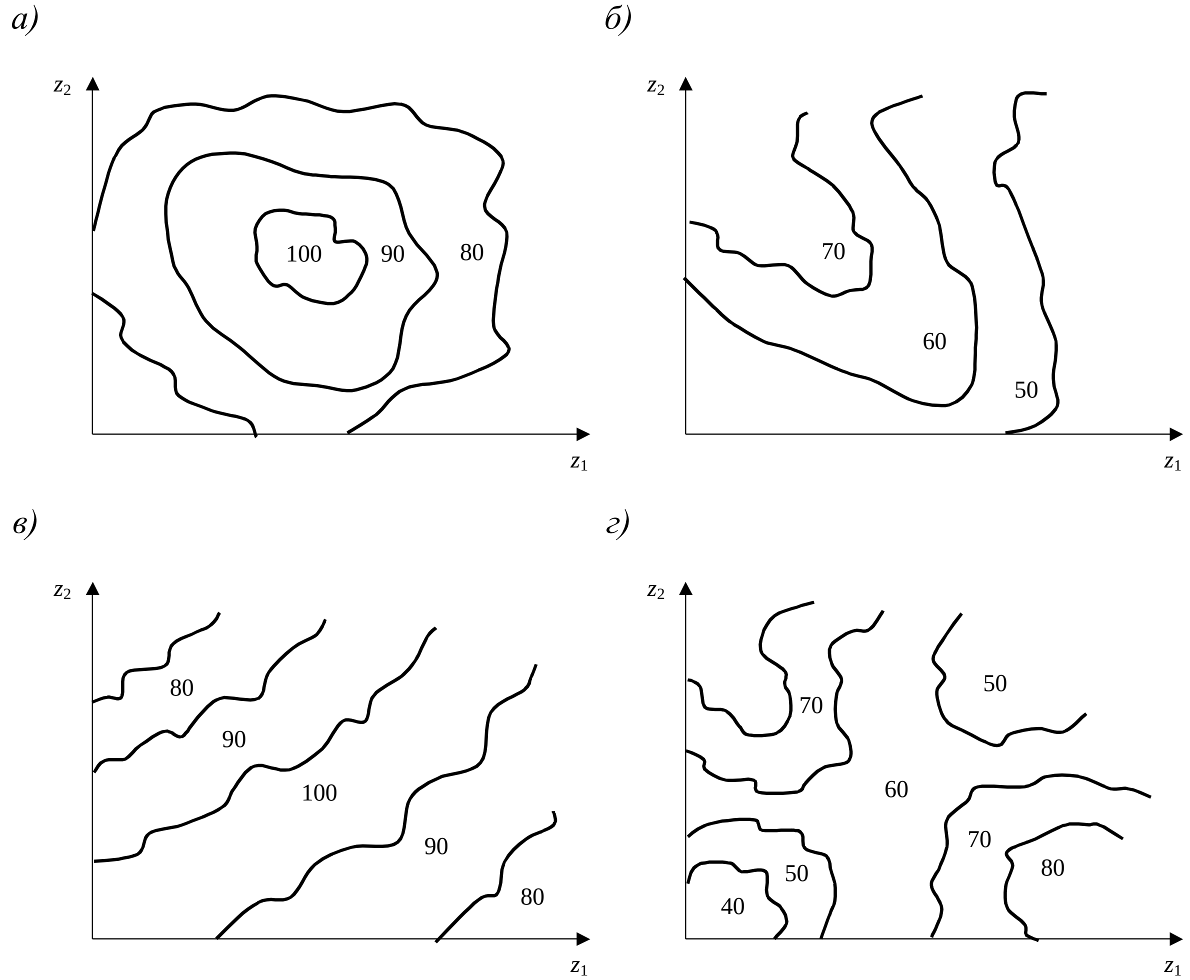
$$x_i = \frac{z_i - z_{0i}}{\Delta z_i}; \qquad z_{0i} = \frac{z_{i\max} + z_{i\min}}{2}; \qquad \Delta z_i = z_{i\max} - z_{0i} = z_{0i} - z_{i\min};$$

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

$z_{i\max}$, $z_{i\min}$ – 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 5.2).



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

Рисунок 5.2 – Виды поверхностей отклика

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 5.3) и матрицу (таблица 5.1) двухфакторного эксперимента, в соответствии с которыми и проводят исследования.

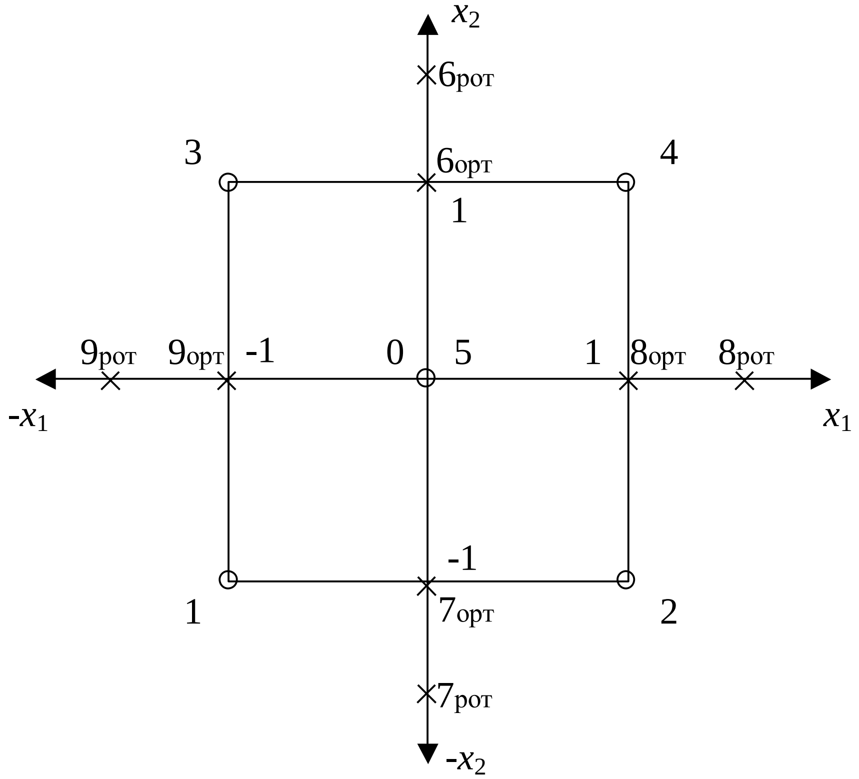


Рисунок 5.3 – Планы для функций $y = f(x_1, x_2)$

Таблица 5.1 - 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 5.4, матрица – в таблице 5.2.

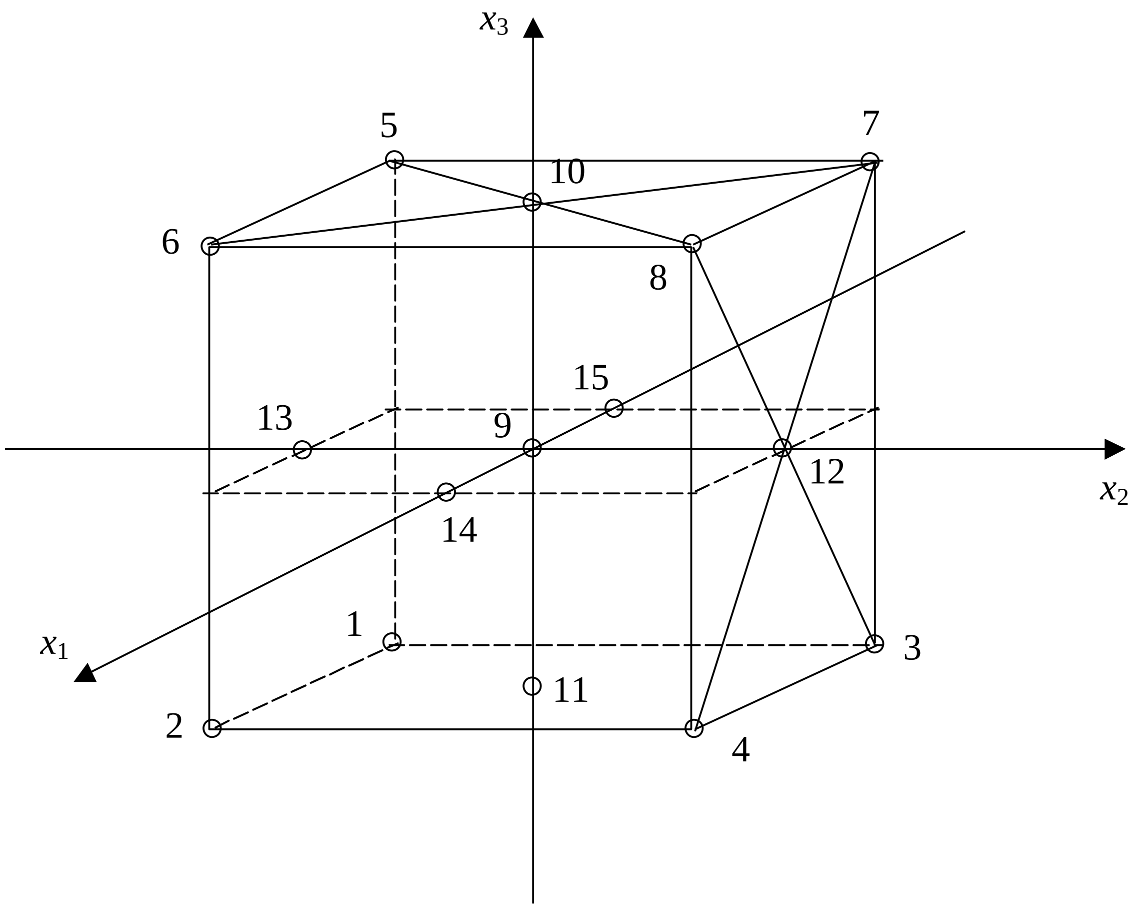


Рисунок 5.4 - Планы для функций $y = f(x_1, x_2, x_3)$

Таблица 5.2 - 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 5.1 и 5.2, 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:

$$\sum_{j=1}^{m} x_{ii} = 0; \quad \sum_{j=1}^{m} x_{ij}^2 = m; \quad \sum_{j=1}^{m} x_{iu} x_{ju} = 0. \tag{5.3}$$

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

Свойство, выраженное уравнениями (5.3), называется ортогональностью, а матрица – ортогональной (прямоугольной). Это свойство обеспечивает относительную простоту вычисления коэффициентов регрессии, которые определяют по формулам:

$$a_0 = \frac{1}{m} \sum_{j=1}^{m} y_i; \qquad a_i = \frac{1}{m} \sum_{j=1}^{m} x_{ji} y_j; \qquad a_{iu} = \frac{1}{m} \sum_{j=1}^{m} x_{ji} x_{ju} y_j. \tag{5.4}$$

Формулы (5.4) 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

$$|a| \geq D_{\text{a}} t, \tag{5.5}$$

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}, …, $a_{(n-1)n}$, 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

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3,$$

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 5.4). В этом случае основу матрицы планирования составляет матрица двухфакторного эксперимента, варьирование третьего фактора соответствует произведению x_1x_2 (таблица 5.3). Это преобразование допустимо, если коэффициент регрессии a_{12} незначим или равен нулю.

Таблица 5.3 - 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 5.3) обозначают 2^{3-1}.

Использование метода ДФЭ или дробных реплик дает возможность в несколько раз уменьшить число опытов при исследовании линейных моделей. В таблице 5.4 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 5.4 - 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 5.3). Тогда полное количество опытов можно выразить в виде

$$m = 2^n + 2n + N_0, \tag{5.6}$$

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:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 .$$

С учетом изложенного матрица ортогонального ЦКП для двух факторов представлена в виде таблицы 5.5.

В принятой матрице «0» показывает, что значения z_i приняты в начале координат (рисунок 5.3). Коэффициенты регрессии в этом случае вычисляются с помощью формул:

$$a_0 = \frac{1}{m}\sum_{j=1}^{m} y_j = \frac{a_{11}}{m}\sum_{j=1}^{m} x_{j1}^2 - \ldots - \frac{a_{uu}}{m}\sum_{j=1}^{m} x_{ju}^2 , \tag{5.7}$$

$$a_i = \frac{\sum_{j=1}^{m} x_{ij} y_j}{\sum_{j=1}^{m} (x_{ij})^2} \qquad (i \neq 0), \tag{5.8}$$

$$a_{iu} = \frac{\sum\limits_{j=1}^{m} x_{ij} x_{uj} y_j}{\sum\limits_{j=1}^{m} \left(x_{ij} x_{uj} \right)^2} \qquad (i \neq u), \tag{5.9}$$

$$a_{ii} = \frac{\sum\limits_{j=1}^{m} \left(x_{ij} - \frac{1}{m} \sum\limits_{j=1}^{m} x_{ij}^2 \right) y_j}{\sum\limits_{j=1}^{m} \left(x_{ij}^2 - \frac{1}{m} \sum\limits_{j=1}^{m} x_{ij}^2 \right)^2}, \tag{5.10}$$

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

Таблица 5.5 - Матрица планирования эксперимента

Система опытов	Номер опыта	x_1	x_2	$X_1 x_2$	y_j
Полный факторный эксперимент линейной модели	1 2 3 4	-1 +1 -1 +1	-1 -1 +1 +1	+1 -1 -1 +1	y_1 y_2 y_3 y_4
Опыт в центре плана	5	0	0	0	y_5
Опыты в звездных точках	6 7 8 9	0 0 +1 -1	+1 -1 0 0	0 0 0 0	y_6 y_7 y_8 y_9

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

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

$$D_{a_j} = \frac{D_y^2}{\sum\limits_{j=1}^{m} \left(x_{ij} \right)^2} \qquad (i \neq 0),$$

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

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

Коэффициенты a_0, a_i, a_{iu} значимы, если выполняется условие (5.5). Адекватность полученного уравнения регрессии проверяется с помощью критерия Фишера. Для двух- и трехфакторного эксперимента вычисление коэффициентов регрессии можно производить вручную:

$$a_{iu} = \frac{1}{c}\sum_{j=1}^{m}\gamma y_j\,, \tag{5.11}$$

где c и γ - коэффициенты, которые приведены в таблицах 5.6 и 5.7. При проведении опытов с четырьмя и более факторами для вычисления коэффициентов полинома по формулам (5.7-5.10) необходимо использовать ЭВМ.

Таблица 5.6 - 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

С учетом изложенного последовательность оптимального планирования эксперимента заключается в следующем.

1. Выбирают из n действующих в системе факторов наиболее значимые $z_1, z_2, \ldots, z_i, \ldots, z_n$.

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.

Таблица 5.7 - 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

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 5.8). Однако ротатабельное планирование не только более трудоемкое по сравнению с ортогональным, но и вычисление коэффициентов регрессии производится по относительно громоздким формулам.

Таблица 5.8 - 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 5.5, таблица 5.9). При таком расположении точек в факторном пространстве можно достаточно точно описать исследуемый объект. Количество опытов не превышает необходимого для ортогонального ЦКП (формула (5.6)).

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.

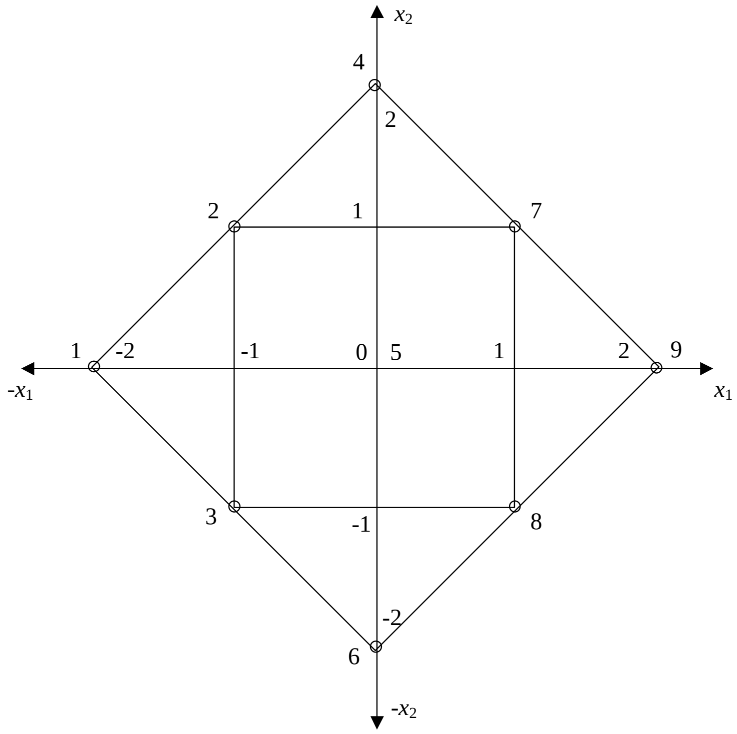


Рисунок 5.5 – *D*-оптимальный план для функций $f(x_1, x_2)$

Таблица 5.9 - 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

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.

5.3 Оптимизация технологических процессов с использованием планирования экспериментов

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 состояния 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

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

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

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.

6 АНАЛИЗ И ОФОРМЛЕНИЕ РЕЗУЛЬТАТОВ НАУЧНЫХ ИССЛЕДОВАНИЙ

6.1 Анализ теоретико-экспериментальных исследований и формулирование выводов и предложений

Основой совместного анализа теоретических и экспериментальных исследований является сопоставление выдвинутой рабочей гипотезы с опытными данными наблюдений.

Теоретические и экспериментальные данные сравнивают методом сопоставления соответствующих графиков. Критериями сопоставления могут быть минимальные, средние и максимальные отклонения экспериментальных результатов от данных, установленных расчетом на основе теоретических зависимостей. Возможно также вычисление среднеквадратического отклонения и дисперсии. Однако наиболее достоверными следует считать критерии адекватности соответствия теоретических зависимостей экспериментальным.

В результате теоретико-экспериментального анализа могут возникнуть три случая.

1. Установлено полное или достаточно хорошее совпадение рабочей гипотезы, теоретических предпосылок с результатами опыта. При этом дополнительно группируют полученный материал исследований таким образом, чтобы из него вытекали основные положения разработанной ранее рабочей гипотезы, в результате чего последняя превращается в доказанное теоретическое положение, в теорию.

2. Экспериментальные данные лишь частично подтверждают положение рабочей гипотезы и в той или иной ее части противоречат ей. В этом случае рабочую гипотезу изменяют и перерабатывают так, чтобы она наиболее

полно соответствовала результатам эксперимента. Чаще всего производят дополнительные корректировочные эксперименты с целью подтвердить изменения рабочей гипотезы, после чего она также превращается в теорию.

3. Рабочая гипотеза не подтверждается экспериментом. Тогда ее критически анализируют и полностью пересматривают. Затем проводят новые экспериментальные исследования с учетом новой рабочей гипотезы. Отрицательные результаты научной работы, как правило, не являются бросовыми, они во многих случаях помогают выработать правильные представления об объектах, явлениях и процессах.

После выполненного анализа принимают окончательное решение, которое формулируют как заключение, выводы или предложения. Эта часть работы требует высокой квалификации, поскольку необходимо кратко, четко, научно выделить то новое и существенное, что является результатом исследования, дать ему исчерпывающую оценку и определить пути дальнейших исследований. Обычно по одной теме не рекомендуется составлять много выводов (не более 5-10). Если же помимо основных выводов, отвечающих поставленной цели исследования, можно сделать еще и другие, то их формулируют отдельно, чтобы не затемнить конкретного ответа на основную задачу темы.

Все выводы целесообразно разделить на две группы: научные и производственные. В научных выводах необходимо показать, какой вклад внесен в науку в результате выполненных исследований (новые предложения, принципиальное различие существующих, опровержение некоторых известных положений и др.). В заключении нужно разработать план внедрения законченных НИР в производство и оценить прибыль. При выполнении научно-исследовательской работы заботятся о защите государственного приоритета (первенства в решении определенной научной или технической задачи) на изобретение или открытие.

6.2 Научные документы и их подготовка к опубликованию в печати

Структурной единицей, характеризующей информационные ресурсы и информационные продукты с количественной стороны, является научный документ, под которым понимается материальный объект, содержащий научно-техническую информацию и предназначенный для ее хранения и использования.

В зависимости от способа представления информации различают документы: текстовые (книги, журналы, отчеты и др.), графические (чертежи, схемы, диаграммы), аудиовизуальные (звукозаписи, кино- и видеофильмы), машиночитаемые (например, образующие базу данных, на микрофотоносителях) и др. Кроме того, документы подразделяются на первичные (содержащие непосредственные результаты научных исследований и разработок, новые научные сведения или новое осмысление известных идей и фактов) и вторичные (содержащие результаты аналитико-синтетической и логической переработки одного или нескольких первичных документов или сведения о них).

Первичные документы и издания. Как первичные, так и вторичные документы подразделяются на опубликованные (издания) и непубликуемые. С развитием информационной технологии это разграничение становится все менее существенным. В связи с наличием в непубликуемых документах ценной информации, опережающей сведения в опубликованных изданиях, органы научно-технической информации (НТИ) стремятся оперативно распространять эти документы с помощью новейших средств репродуцирования.

В числе первичных документов – книги (непериодические текстовые издания объемом свыше 48 страниц); брошюры (непериодические текстовые издания объемом свыше четырех, но не более 48 страниц). Книги и брошюры подразделяются на платные и бесплатные, а также на научные, учебные, официально-документальные, научно-популярные и, наконец, по отраслям науки и научным дисциплинам. Среди книг и брошюр важное научное значение имеют монографии, содержащие всестороннее исследование одной проблемы или темы и принадлежащие одному или нескольким авторам, и затем сборники научных трудов, содержащие ряд произведений одного или нескольких авторов, рефераты и различные официальные или научные материалы.

Для учебных целей издаются учебники и учебные пособия (учебные издания). Это непериодические издания, содержащие систематизированные сведения научного и прикладного характера, изложенные в форме, удобной для преподавания и изучения.

Некоторые издания, публикуемые от имени государственных или общественных организаций, учреждений и ведомств, называются официальными. Они содержат материалы законодательного, нормативного или директивного характера.

Наиболее оперативным источником НТИ являются периодические издания, выходящие через определенные промежутки времени, с постоянным для каждого года числом номеров. Традиционными видами периодических изданий являются газеты и журналы. К непериодическим относятся также продолжающиеся издания, выходящие через неопределенные промежутки времени, по мере накопления материала. Обычно это сборники научных трудов институтов, вузов, научных обществ, публикуемые без строгой периодичности под общим заглавием «Труды», «Ученые записки», «Известия» и др.

К специальным видам технических изданий принято относить нормативно-техническую документацию, регламентирующую научно-технический уровень и качество выпускаемой продукции (стандарты, инструкции, типовые положения, методические указания и др.). Стандарт – нормативно-технический документ, устанавливающий комплекс норм, правил, требования к объекту стандартизации и утвержденный компетентным органом. В зависимости от содержания стандарты включают: технические условия и требования; параметры и размеры; типы; конструкции; марки; сортаменты; правила приемки; методы контроля; правила эксплуатации и ремонта; типовые технологические процессы и т.п.

Важное значение для постановки научно-исследовательских работ имеет патентная документация, представляющая собой совокупность документов, содержащих сведения об открытиях, изобретениях и других видах промышленной собственности, а также сведения об охране прав потребителей. Патентная документация обладает высокой степенью достоверности, так как подвергается тщательной экспертизе на новизну и полезность.

Первичные непубликуемые документы могут быть размножены в необходимом количестве экземпляров и пользоваться правами изданий (рукописи и корректурные оттиски являются промежуточными этапами полиграфического процесса и не относятся к научным документам). К основным видам непубликуемых первичных документов относятся научно-технические отчеты, диссертации, депонированные рукописи, научные переводы, конструкторская документация, информационные сообщения о проведенных научно-технических конференциях, съездах, симпозиумах, семинарах.

Вторичные документы и издания подразделяют на справочные, обзорные, реферативные и библиографические.

В справочных изданиях (справочники, словари) содержатся результаты теоретических обобщений, различные величины и их значения, материалы производственного характера.

В обзорных изданиях содержится концентрированная информация, полученная в результате отбора, систематизации и логического обобщения сведений из большого количества первоисточников по определенной теме за определенный промежуток времени. Различают обзоры аналитические (содержащие аргументированную оценку информации, рекомендации по ее использованию) и реферативные (носящие более описательный характер). Кроме того, работники библиотек часто готовят библиографические обзоры, содержащие характеристики первичных документов как источников информации, появившихся за определенное время или объединенных каким-либо общим признаком.

Реферативные издания (реферативные журналы, реферативные сборники) содержат сокращенное изложение первичного документа или его части с основными фактическими сведениями и выводами. Реферативный журнал – это периодическое издание журнальной или карточной формы, содержащее рефераты опубликованных документов (или их частей). Реферативный сборник – это периодическое, продолжающееся или непериодическое издание, содержащее рефераты непубликуемых документов (в них допускается включать рефераты опубликованных зарубежных материалов).

Библиографические указатели являются изданиями книжного или журнального типа, содержащими библиографические описания вышедших изданий. В зависимости от принципа расположения библиографических описаний указатели подразделяются на систематические (описания располагаются по областям науки и техники в соответствии с той или иной системой классификации) и предметные (описания располагаются в порядке перечисления важнейших предметов в соответствии с предметными рубриками, расположенными в алфавитном порядке).

Вторичные непубликуемые документы включают регистрационные и информационные карты, учетные карточки диссертаций, указатели депонированных рукописей и переводов, картотеки «Конструкторская документация на нестандартное оборудование», информационные сообщения. К ним принято относить также вторичные документы, которые публикуются, но рассылаются по подписке (Бюллетени регистрации НИР и ОКР, сборники рефератов НИР и ОКР и др.).

Документные классификации. Традиционным средством упорядочения документальных фондов являются библиотечно-библиографические (документные) классификации. Наибольшее распространение получила Универсальная десятичная классификация (УДК), которая используется более чем в 50 странах мира и юридически является собственностью Международной федерации по документации (МФД), отвечающей за дальнейшую разработку таблиц УДК, их состояние и издание. В России УДК была введена с 1963 г. в качестве единой системы классификации всех публикаций по точным, естественным наукам и технике. УДК является международной универсальной системой, позволяющей детально представить содержание документальных фондов и обеспечить оперативный поиск информации, обладает возможностью дальнейшего развития и совершенствования. Отличительными чертами УДК являются охват всех отраслей знаний, возможность неограниченного деления на подклассы, индексация арабскими цифрами, наличие развитой системы определителей и индексов. Издаются полные, средние, отраслевые издания и рабочие схемы, а также методические пособия по классификации.

Закономерности производства НТИ. Анализ источников информации. Рассмотренные источники информации образуют систему научных документов и изданий, для которой характерны определенные закономерности, отражающие развитие науки. Установлен ряд общих закономерностей, характеризующих рост и старение документов.

Рост числа журналов и количества содержащихся в них статей в основном характеризуется экспоненциальной зависимостью с разными показателями для разных научных областей. Так, например, рост числа библиографических журналов за последние 200 лет характеризуется экспоненциальной зависимостью с удвоением за 18 лет, а журналов по математике – с удвоением за 28 лет.

Старение документов заключается в том, что с увеличением сроков со времени выпуска изданий они теряют ценность как источники информации и по этой причине все меньше используются учеными и специалистами.

Наряду с исследованием общих закономерностей роста и старения документов, анализ и статистическая обработка источников информации позволяют получить картину состояния и развития конкретных научных направлений (на основе анализа структуры документального потока) и выявить взаимосвязи между отдельными научными дисциплинами (направлениями), странами, школами, коллективами и учеными. Результаты анализа баз данных по запросам потребителей – специальный вид информационных продуктов, который по мере оснащения органов НТИ вычислительной техникой и соответствующим программным обеспечением получает все большее распространение.

Как правило, публикуют работы, содержащие новые научные результаты и конкретные предложения, имеющие важное теоретическое и практическое значение. Это издания, которые содержат теоретическую разработку проблем или научное исследование вопроса, или результаты исследований в области науки.

Подготовку материала исследования к печати необходимо проводить в такой последовательности. Составляют план-проспект и систематизируют материал исследования, при этом строго придерживаются положения о том, что второстепенные сведения или уже опубликованные ранее не следует помещать в подготавливаемое издание. Затем располагают подобранный материал по разделам и подразделам. Излагают материал в научном стиле, для которого характерны ясность изложения, точность словоупотребления, лаконизм, строгое соблюдение научной терминологии, позволяющей в возможно краткой и экономной форме давать четкие определения и характеристики научных фактов, понятий, процессов и явлений. Последовательное изложение принятой теоретической позиции, логичность, глубокая взаимосвязь теоретических положений, выразительность речи – характерные черты научного стиля.

Все цитаты приводят по первоисточникам с указанием подлинных авторов цитат и источников.

7 ОСНОВЫ ПАТЕНТОВЕДЕНИЯ В НАУЧНЫХ ИССЛЕДОВАНИЯХ

7.1 Интеллектуальная собственность и ее защита

Всякий интеллектуальный труд, связанный с научно-техническим творчеством, требует защиты авторских прав. Законодательство РФ о правах на результаты интеллектуальной деятельности внутренне дифференцировано. Оно распадается на несколько групп, самостоятельных, или, по крайней мере, автономных с точки зрения правовых средств.

Важной стороной вопроса об интеллектуальной собственности (ИС) является своевременное выявление в научной продукции, разработках и технологиях объектов ИС и обеспечение их охраны в стране и за рубежом, то есть, превращение в конкурентно способный на рынке товар, весьма дорогостоящий и приносящий при правильном его использовании высокую прибыль.

Этому в значительной степени способствуют принятые в нашей стране в конце 1992 года законы по охране интеллектуальной собственности:

- патентный закон РФ (вступил в силу с 14.10.1992 г.)
- закон обслуживания и наименований мест происхождения товаров (от 17.10.1992 г.)
- закон о правовой охране программ для ЭВМ и банков данных (с 20.10.1992 г.)
- закон о правовой охране топологий интегральных полупроводниковых микросхем (с 21.10.1992 г.)
- закон РФ «Об авторском праве и смежных правах» (с 3.08.1993 г.)

Одним из видов ИС являются объекты, которые представляют собой творения человеческого разума.

ИС признается (гражданский кодекс РФ, ч. 1, ст. 138) исключительное право гражданина или юридического лица на результаты интеллектуальной деятельности.

7.2 Объекты права интеллектуальной собственности

Использование результатов интеллектуальной деятельности может осуществляться третьими лицами только с согласия правообладателя. Конвенция об учреждении Всемирной организации интеллектуальной собственности (ВОИС), принятая в Стокгольме в 1967 г., предусматривает, что объектами права ИС являются (рисунок 7.1):

- литературные, художественные произведения и научные труды;
- исполнительная деятельность, фонограммы, радиопередачи;
- изобретения во всех областях человеческой деятельности;
- научные открытия;
- программное обеспечение;
- товарные знаки (ТЗ), знаки обслуживания, коммерческие наименования и обозначения;
- незаконная конкурентная деятельность.

Формы охраны ИС могут быть самыми разными:

1. Патенты исключительного права (на изобретения, полезную модель, промышленный образец).
2. Свидетельства (на ТЗ, знаки обслуживания).
3. Регистрация (программы для ЭВМ, наименования мест происхождения).
4. Сам факт выпуска в свет литературных произведений, картин и т.п.

Во всех случаях защита охраноспособного объекта интеллектуальной собственности должна быть оформлена надлежащим образом.

Рисунок 7.1 – Сферы прав интеллектуальной собственности

Авторское право распространяется на произведения науки, литературы и искусства независимо от формы, назначения и достоинства произведения, а также от способа его воспроизведения. Оно распространяется на произведения как выпущенные, так и не выпущенные в свет, но выраженные в какой-либо объективной форме, позволяющей воспроизводить результат творческой деятельности автора (рукопись, чертеж, изображение, публичное произнесение или исполнение, пленка, механическая или магнитная запись и т.д.).

Перечень предметов, которые охраняются авторским правом, не являются исчерпывающим – оно может охранять и другие произведения, если они отвечают общим требованиям, предъявляемым к объектам авторского права.

Произведение – это результат творческой деятельности автора, его творческого мышления, поэтому произведение охраняется авторским правом как нематериальный объект.

Результат творческой деятельности автора получает охрану по авторскому праву с того момента, когда он оказывается выражен в объективной форме.

В любом произведении следует различать содержание и форму. Для научной, технической, учебной литературы (кроме художественной) к элементам содержания относятся научные и технические факты, идеи, гипотезы, теории, систематизация элементов произведения. Форма такого произведения состоит в системе и последовательности изложения материала.

Авторское право охраняет только форму произведения, поэтому элементы содержания являются юридически несущественными. При этом для установления случаев нарушения авторских прав путем заимствования решающее значение имеют буквальные (например, для литературных произведений - словесные) совпадения; для этого составляются таблицы совпадений («текст на такой-то странице в одном произведении соответствует тексту на такой-то странице в другом произведении»).

В этой связи следует отметить, что даже если произведение отклонено, например, редакцией журнала, то и такое произведение охраняется авторским правом и никто, кроме автора, не может выдать это произведение за свое, использовать его отдельные части, переделать и т.п.

Для получения авторского права не нужно выполнять каких-либо формальностей. Это право возникает сразу после того, как произведение получит объективное выражение (например, как только автор запишет на бумаге сочиненное им). Никакой регистрации произведения не требуется, кроме подтверждения даты публикации (произнесения) текста.

После присоединения Российской Федерации к Всемирной конвенции об авторском праве на отечественных печатных изданиях стал проставляться знак охраны авторским правом (буква «C» в окружности ©, наименование владельца авторского права, год первого выпуска произведения в свет). Этот знак не имеет значения для признания (или не признания) произведения охраняемым в нашей стране. Он проставляется для информирования о том, что произведение охраняется в нашей стране и некоторых зарубежных странах (США, Мексика и др.).

Авторское право распространяет свое действие на произведение в целом и на любые его составные части (глава или абзац литературного, научного произведения, кадр кинофильма и др.).

После воплощения автором произведения в рукописи у него возникает и авторское право, и право собственности на рукопись. Следует различать действия по передаче (уступке) права собственности, например при дарении экземпляра произведения, и действия при передаче авторских прав. Сомнения могут возникать в отношении уникальных произведений, например произведений изобразительного искусства. Собственник произведения приобретает лишь право помещать его на публичных выставках, не выплачивая автору никакого дополнительного вознаграждения. С другой стороны, направление в издательство рукописи обычно свидетельствует о согласии автора передать свое авторское право на издание произведения.

Произведения, охраняемые авторским правом, делятся на две большие группы: самостоятельные произведения и производные. В самостоятельном произведении (именуемом также новым, творчески самостоятельным) все элементы формы создаются автором, без заимствования из других произведений. Авторское право на них принадлежит автору данного произведения. Производные – это те произведения, в которых использованы элементы формы других произведений. Типичными примерами производных произведений являются переработки, переводы, сборники. Главная особенность производных произведений состоит в том, что если включенные в них элементы форм других произведений охраняются авторским правом, то использование произведения (производного) может осуществляться лишь с согласия автора первоначального произведения. Такое произведение считается зависимым от основного. Так, у лица, составившего сборник произведений, возникает авторское право на этот сборник в целом (на подбор материала), но лица, произведения которых включены в сборник, сохраняют авторские права на свои произведения, поэтому составитель может использовать сборник (пользоваться авторским правом) лишь с их согласия.

Не получают охрану по авторскому праву объекты, которые не имеют признаков произведения, например телефонный справочник, либо не выраженные в объективной форме (например, научная идея сама по себе).

Некоторые объекты, обладающие всеми признаками произведения науки, литературы или искусства, тем не менее, не охраняются авторским правом. К ним относятся:

- произведения, на которые истек срок действия авторского права;
- произведения иностранных граждан, находящихся за пределами территории РФ, или впервые опубликованные за ее пределами, если они не охраняются на основе международных договоров РФ (т.е. на основе Всемирной конвенции об авторском праве или на основе двусторонних соглашений);
- официальные документы: законы, судебные решения и др.;

- произведения народного творчества, если авторы их неизвестны (фольклор, произведения народных художественных промыслов, народная архитектура и др.).

В Российской Федерации признаются частная, государственная, муниципальная и иные формы собственности. В соответствии с этим различаются и субъекты (носители) авторского права: частные лица, государственные и местные организации и т.д.

Субъектом авторского права автоматически, без необходимости получения какого-либо разрешения и без регистрации произведения, считается автор, создатель произведения.

В случаях, когда произведение создано творческим трудом двух или большего числа авторов, возникает соавторство. Авторское право на такое коллективное произведение принадлежит соавторам совместно, независимо от того, образует ли такое произведение одно целое или состоит из частей, которые могут быть использованы самостоятельно. В последнем случае каждый соавтор сохраняет право на созданную лично им часть коллективного произведения (если только в соглашении между ними не указано иное). Для возникновения соавторства не требуется чьего-либо разрешения или соглашения соавторов.

Авторское вознаграждение делится пропорционально объему частей произведения, принадлежащих отдельным соавторам. Если же соавторство «неделимое», то вознаграждение выплачивается по соглашению между соавторами. При отсутствии соглашения (споров между соавторами) вознаграждение распределяется поровну или в пропорции, установленной соответствующим законодательным актом. На практике при использовании коллективного произведения должен заключаться договор со всеми соавторами, а при использовании составного произведения (например, сборника статей и др.) – один договор на каждое входящее в него произведение.

Научные работники зачастую пытаются определить авторство и соавторство на научные статьи не по нормам авторского права, а по широко распространенным в научной среде морально-этическим нормам. Исходя из последних, автором обязательно должен становиться автор научной идеи, гипотезы и т.п., независимо от того, принимал ли он участие в подготовке самой статьи. Эти правила не находят отражения в нормах авторского права.

В порядке наследования авторское право может переходить наследникам по закону или завещанию. Если нет наследников ни по закону, ни по завещанию, либо наследники не приняли наследства или лишены завещателем наследства, то имущество умершего переходит к государству. В последнем случае авторское право прекращается, и произведение может использоваться свободно и без выплаты вознаграждения. Это не относится к случаям, когда в качестве наследника авторского права выступает не государство в целом, а отдельная государственная организация.

Авторское право может принадлежать юридическим лицам (организациям, предприятиям). Так, например, организациям, выпускающим периодические издания, принадлежит авторское право на них. При этом авторы отдельных произведений, включенных в эти издания, сохраняют авторские права на свои произведения.

Когда автор заключает с организацией авторский договор об использовании произведения, то он передает ей право на воспроизведение и распространение произведения. В пределах договора организация становится субъектом (носителем) авторских прав, правопреемником автора. После прекращения договора или в случае его расторжения все авторские права возвращаются к автору. Организация может приобретать авторские права по договору лишь в случаях, предусмотренных ее уставом или иным нормативным документом, и в пределах, установленных этим документом.

Если произведение создано автором в порядке выполнения служебного задания, выданного ему научной или иной организацией, то некоторые авторские права на данное произведение принадлежат этой организации: использование произведения должно осуществляться в порядке, предусмотренном в трудовом договоре организации с автором, или так, как это вытекает из его (договора) содержания.

Промышленная собственность – это (наряду с собственностью, охраняемой авторским правом) разновидность интеллектуальной собственности. Ее объектами являются: изобретения; полезные модели; товарные знаки; промышленные образцы; знаки обслуживания; фирменные наименования; указания на источник происхождения; наименования места происхождения; пресечение недобросовестной конкуренции.

Самое раннее и наиболее общее соглашение в области промышленной собственности – заключенная в 1883 г. и дополненная в 1991 г. в Мадриде Пояснительным протоколом Парижская конвенция по охране промышленной собственности – предусматривает общие принципы международной охраны промышленной собственности.

Термин «промышленная собственность» указывает на область использования перечисленных объектов. Однако этот термин не вполне точен, так как эти объекты могут использоваться в сельском хозяйстве и торговле, медицине и военной области, в сфере образования и досуга. Собственностью указанные объекты являются постольку, поскольку благодаря осуществляемой в отношении их охране право распоряжения ими принадлежит их обладателю, т.е. собственнику. Без его согласия никто не может пользоваться этими объектами, о чем свидетельствует охрана объектов промышленной собственности от посягательств со стороны третьих лиц.

Охрана прав обладателей объектов промышленной собственности осуществляется с помощью соответствующих охранных документов: патентов, свидетельств и др.

Изобретения, как правило, охраняются патентами, называемыми патентами на изобретение. Патентная форма охраны изобретений применяется почти в 140 странах мира. Патентная форма охраны придает исключительному праву патентообладателя монопольный характер, но в отличие от других форм права собственности оно ограничено во времени.

Патент предоставляет исключительное право на изобретение на территории той страны, где он выдан, либо на территории ряда стран, когда они договорились об этом. За указанными пределами изобретение уже не пользуется патентной защитой. Чтобы получить ее, необходимо подать заявку и получить патент в стране, на территории которой необходимо обеспечить охрану данного изобретения.

Объектами изобретения признаются: устройство, способ, вещество, штамм микроорганизма, культура клеток растения или животного, а также применение известного ранее устройства (способа, вещества, штамма, культуры) по новому назначению. К устройствам, как объектам изобретения, относятся изделия или совокупность конструктивных элементов, находящихся в функциональном конструктивном единстве, и обладающие названными ранее признаками изобретения.

Способ как объект изобретения – это процесс выполнения взаимосвязанных действий, необходимых для достижения поставленной технической цели, обладающий признаками изобретения.

Вещество как объект изобретения – это искусственно созданное материальное образование, являющееся совокупностью взаимосвязанных элементов и обладающее названными ранее признаками изобретения. К ним также условно отнесены высокомолекулярные соединения и продукты генной инженерии: рекомбинантные нуклеиновые кислоты, векторы и т.п., композиции (составы, смеси), продукты ядерного превращения.

Таким образом, ***изобретение*** – это техническое решение в любой области, относящееся к продукту (в частности, устройству, веществу, штамму микроорганизма, культуре клеток растений или животных) или способу (процессу осуществления действий над материальным объектом с помощью материальных средств).

Общими для патентного права различных стран являются обстоятельства, порочащие новизну изобретения и тем самым исключающие возможность выдачи патента. Это – предшествующая заявке публикация, раскрывающая сущность технического решения, и открытое его применение. Публикацией называется информация о сущности технического решения, которая становится доступной неопределенному кругу лиц (с помощью печати, радио, телевидения и др.).

Под открытым применением понимается такое использование изобретения, при котором изобретение становится известным неопределенному кругу лиц в такой степени, что специалисты могут воспроизвести его, не прибегая к дополнительным источникам информации.

Полезная модель – это техническое решение, относящееся к устройству; полезная модель предусмотрена законодательством некоторых стран – Россия, КНР, ФРГ, Япония). В США и Великобритании, например, эта категория объектов промышленной собственности не используется.

Полезные модели описываются так же, как и изобретения на устройства. Отличие полезных моделей от изобретений состоит в следующем:

а) в случае полезной модели уровень технического решения, требуемый для его признания этим объектом промышленной собственности, ниже, чем соответствующий уровень в случае изобретения;

б) максимальный срок охраны, предусмотренный законодательством для полезной модели, обычно гораздо короче, чем максимальный срок охраны, предусмотренный для изобретения.

Охранные документы на изобретение и на полезную модель имеют в одних странах одинаковое название (патент), в других, например в России, разное (соответственно патент и свидетельство).

Промышленные образцы – это объекты промышленной собственности в сфере дизайна, внешнего вида изделия. Промышленный образец представляет собой решение эстетической или декоративной стороны полезного изделия. Декоративная сторона может быть выражена в форме, структуре или цвете изделия и она должна воздействовать на зрительное восприятие.

Изделие должно быть воспроизводимо промышленными средствами. Если этот элемент отсутствует, изделие подпадает скорее под категорию произведения искусства, охраняемого авторским правом.

Для того чтобы получить охрану, промышленный образец должен быть согласно одним законодательствам новым, согласно другим – оригинальным.

Промышленные образцы обычно защищаются от неправомочного копирования или имитации

В последнее время развитие искусства дизайна приводит все к большей заинтересованности потребителя в приобретении привлекательных и оригинальных изделий. Это делает защиту внешнего вида последних актуальной задачей разработчиков, конструкторов и патентоведов.

Согласно Патентному закону Российской Федерации патент на изобретение действует в течение 20 лет; свидетельство на полезную модель – 5 лет (продлевается по ходатайству патентообладателя, но не более чем на 3 года); патент на промышленный образец – 10 лет (продлевается по ходатайству патентообладателя, но не более чем на 5 лет) с даты поступления заявки в патентное ведомство.

Товарный знак – это объект промышленной собственности в виде символа, указывающего, кто несет ответственность за предлагаемый потребителю товар. Одинаковые товары могут изготавливаться разными производителями и распространяться разными продавцами, и все они могут использовать свои товарные знаки. При выборе товаров покупатели ориентируются также и на товарные знаки, каждый из которых свидетельствует в пользу (или наоборот) доброкачественности товаров той или иной фирмы. Другими словами, товарный знак – это элемент отличия конкурирующих товаров. Поэтому товарные знаки должны четко отличаться друг от друга.

Различают словесные, графические и комбинированные товарные знаки.

Некоторые из них выступают как неотъемлемая часть товара, например особые кромки ткани, особая форма горлышка бутылки и др.

Если товарный знак связан с предоставлением услуг, его называют знаком обслуживания. Так, знаки обслуживания используются отелями, прачечными, химчистками, агентствами по продаже автомобилей и др.

Товарный знак имеет несколько функций: он служит ориентиром для покупателя при выборе товара, качеству которого он доверяет по своему прошлому опыту; он также дает возможность изготовителям товаров отличать свою, находящуюся у потребителя, продукцию от подделки; кроме того, товарный знак дает возможность органам, контролирующим качество товара, определить его владельца и т.д.

Защита товарного знака как вида собственности может осуществляться:

а) путем регистрации;

б) предоставлением права запрета на использование товарных знаков другими лицами без разрешения официальных владельцев.

Фирменные наименования - это объекты промышленной собственности, служащие для распознавания предприятий и выделения их среди других, без какой-либо связи с товарами и услугами этого предприятия. Фирменное наименование характеризует репутацию и положение предприятия в целом. Охрана фирменных наименований предусмотрена большинством национальных законодательств.

Разновидностью объектов промышленной собственности являются также указания на источники и наименования мест происхождения, которые иногда называются географическими указаниями.

Указание на источник происхождения представляет собой какое-либо наименование, выражение или знак, показывающее, что продукт или услуга произведены в той или иной стране, регионе или конкретном месте (например, «Сделано в …»). Общее правило состоит в том, что использование ложных или вводящих в заблуждение указаний на источник происхождения является незаконным.

В наименовании места происхождения входит название страны, региона или конкретного места, где произведен продукт, специфические свойства или качества которого в существенной мере или полностью определены географическими условиями, другими словами, природным и/или человеческим факторами. Использовать наименование места происхождения правомочны только те производители, чьи предприятия расположены в данной географической зоне, и только применительно к конкретным продуктам, производимым в этой зоне (например, бордо, шампанское).

Если указание на источник говорит только о том, откуда поступил продукт, то по наименованию места происхождения продукта можно судить также о его специфических свойствах и качествах, которые определяются географическими условиями района, где этот продукт произведен.

Еще одним объектом промышленной собственности является право на пресечение недобросовестной конкуренции, т.е. таких актов конкуренции, которые противоречат частной промышленной или торговой практике. Парижская конвенция определяет как недобросовестную конкуренцию следующие ее три вида:

1) все действия, ведущие к тому, что потребитель может принять предприятие, товары, промышленную или коммерческую деятельность данной фирмы за предприятие, товары, промышленную или коммерческую деятельность конкурента;

2) ложные заявления в ходе коммерческой деятельности, дискредитирующие предприятие, товары, промышленную или коммерческую деятельность конкурента;

3) использование в ходе коммерческой деятельности указаний или обозначений, которые вводят потребителя в заблуждение относительно природы, способа изготовления, характеристик, пригодности для определенных целей или количества товаров.

Еще 12 видов деятельности определяются как недобросовестная конкуренция в комментарии к Типовому закону по товарным знакам, фирменным наименованиям и актам недобросовестной конкуренции для развивающихся стран:

1) переманивание покупателей конкурентов, направленное на то, чтобы привлечь их в качестве клиентов и сохранить на будущее их признательность;

2) выяснение производственных или коммерческих тайн конкурента путем шпионажа или подкупа его служащих;

3) неправомочное использование или раскрытие ноу-хау конкурента;

4) побуждение служащих конкурента к нарушению или разрыву их связей или контрактов с нанимателем;

5) угроза конкурентам исками о нарушении патентов или товарных знаков, если это делается недобросовестно и с целью противодействия конкуренции в сфере торговли;

6) бойкотирование торговли другой фирмы или противодействие ей для недопущения конкуренции;

7) демпинг, т.е. продажа своих товаров ниже стоимости с намерением противодействовать конкуренции или подавить ее;

8) создание впечатления, что потребителю предоставляется возможность покупки на чрезвычайно выгодных условиях, когда на самом деле этого нет;

9) намеренное копирование товаров, услуг, рекламы или других аспектов коммерческой деятельности конкурента;

10) поощрение нарушений контрактов, заключенных конкурентами;

11) выпуск рекламы, в которой проводится сравнение с товарами или услугами конкурентов;

12) нарушение правовых положений, не имеющих прямого отношения к конкуренции, когда такое нарушение позволяет добиться неоправданного преимущества перед конкурентами.

7.3 Правовая охрана изобретений, полезных моделей и промышленных образцов

Изобретению предоставляется правовая охрана, если оно является новым, имеет изобретательский уровень и промышленно применимо. Изобретение является новым, если оно не известно из уровня техники. Изобретение имеет изобретательский уровень, если оно для специалиста явным образом не следует из уровня техники. Уровень техники включает любые сведения, ставшие общедоступными в мире до даты приоритета изобретения.

При установлении новизны изобретения в уровень техники включаются, при условии их более раннего приоритета, все поданные в Российской Федерации другими лицами заявки на изобретения и полезные модели (кроме отозванных), а также запатентованные в Российской Федерации изобретения и полезные модели.

Изобретение является промышленно применимым, если оно может быть использовано в промышленности, сельском хозяйстве, здравоохранении и других отраслях деятельности.

Не признается обстоятельством, препятствующим признанию патентоспособности изобретения, такое раскрытие информации, относящейся к изобретению, автором, заявителем или любым лицом, получившим от них прямо или косвенно эту информацию, при котором сведения о сущности изобретения стали общедоступными, если заявка на изобретение подана в Патентное ведомство не позднее шести месяцев с даты раскрытия информации. При этом обязанность доказывания данного факта лежит на заявителе.

Как уже было сказано ранее, объектами изобретения могут являться: устройство, способ, вещество, штамм микроорганизма, культуры клеток растений и животных, а также применение известного ранее устройства, способа, вещества, штамма по новому назначению.

Не признаются патентоспособными изобретениями:

- научные теории и математические методы;
- методы организации и управления хозяйством;
- условные обозначения, расписания, правила;
- методы выполнения умственных операций;
- алгоритмы и программы для вычислительных машин;
- проекты и схемы планировки сооружений, зданий, территорий;
- решения, касающиеся только внешнего вида изделий, направленные на удовлетворение эстетических потребностей;
- топологии интегральных микросхем;
- сорта растений и породы животных;
- решения, противоречащие общественным интересам, принципам гуманности и морали.

К полезным моделям относится конструктивное выполнение средств производства и предметов потребления, а также их составных частей. Полезной модели предоставляется правовая охрана, если она является новой и промышленно применимой. Полезная модель является новой, если совокупность ее существенных признаков не известна из уровня техники. Полезная модель является промышленно применимой, если она может быть использована в промышленности, сельском хозяйстве, здравоохранении и других отраслях деятельности.

Не признается обстоятельством, препятствующим признанию патентоспособности полезной модели, такое раскрытие информации, относящейся к полезной модели, автором, заявителем или любым лицом, получившим от них прямо или косвенно эту информацию, при котором сведения о сущности полезной модели стали общедоступными, если заявка на полезную модель подана в Патентное ведомство не позднее шести месяцев с даты раскрытия информации. При этом обязанность доказывания данного факта лежит на заявителе.

В качестве полезных моделей не охраняются:

- способы, вещества, штаммы микроорганизмов, культур клеток растений и животных, а также их применение по новому назначению;
- объекты, признаваемые непатентоспособными для изобретений.

К промышленным образцам относится художественно-конструкторское решение изделия, определяющее его внешний вид. Промышленному образцу предоставляется правовая охрана, если он является новым, оригинальным и промышленно применимым. Промышленный образец признается новым, если совокупность его существенных признаков, определяющих эстетическое и (или) эргономические особенности изделия, не известна из сведений, ставших общедоступными в мире до даты приоритета промышленного образца.

При установлении новизны промышленного образца учитывается, при условии их более раннего приоритета, все поданные в Российской Федерации другими лицами заявки на промышленные образцы (кроме отозванных), а также запатентованные в Российской Федерации промышленные образцы.

Промышленный образец признается оригинальным, если его существенные признаки обусловливают творческий характер эстетических особенностей изделия. Промышленный образец признается промышленно применимым, если он может быть многократно воспроизведен путем изготовления соответствующего изделия.

Не признается обстоятельством, препятствующим признанию патентоспособности промышленного образца, такое раскрытие информации, относящейся к промышленному образцу, автором, заявителем или любым лицом, получившим от них прямо или косвенно эту информацию, при котором сведения о сущности промышленного образца стали общедоступными, если заявка на промышленный образец подана в Патентное ведомство не позднее шести месяцев с даты раскрытия информации. При этом обязанность доказывания данного факта лежит на заявителе.

Не признаются патентоспособными промышленными образцами решения:

- обусловленные исключительно технической функцией изделия;
- объектов архитектуры (кроме малых архитектурных форм), промышленных, гидротехнических и других стационарных сооружений;
- печатной продукции как таковой;
- объектов неустойчивой формы из жидких, газообразных, сыпучих или им подобных веществ;
- изделий, противоречащих общественным интересам, принципам гуманности и морали.

7.4 Правила составления, подачи и рассмотрения заявок на изобретения и полезные модели

Правом на подачу заявки и получение патента (п.1 ст.15 и п.1 ст.8 Закона) обладает автор изобретения, работодатель или правопреемник (заявитель).

Автор изобретения – физическое лицо, творческим трудом которого оно создано, имеет право на подачу заявки и получение патента в следующих случаях:

- если изобретение не является служебным;
- если изобретение является служебным, но договором между автором и работодателем предусмотрено право автора на получение им патента, если работодатель в течение 4 месяцев с даты уведомления его автором о созданном служебном изобретении не подаст заявки, не переуступит право на подачу заявки другому лицу и не сообщит автору о сохранении изобретения в тайне.

Подтверждение права на подачу заявки каким-либо документом не требуется.

К устройствам, как объектам изобретения, относятся конструкции и изделия. К способам, как объектам изобретения, относятся процессы выполнения действий над материальным объектом с помощью материальных объектов. К веществам, как объектам изобретения, относятся:

- индивидуальные химические соединения, к которым также условно отнесены высокомолекулярные соединения и объекты генной инженерии;
- композиции (составы, смеси);
- продукты ядерного превращения.

К применению по новому назначению приравнивается первое применение известных веществ (природных и искусственно полученных) для удовлетворения общественной потребности.

Заявка должна относиться к одному изобретению или группе изобретений, связанных между собой настолько, что они образуют единый изобретательский замысел. Единство изобретения признается соблюденным, если:

- в формуле изобретения, имеющей один независимый пункт, охарактеризовано одно изобретение;

- в формуле изобретения в нескольких независимых пунктах охарактеризована группа изобретений:

а) одно из которых предназначено для получения (изготовления) другого (например, устройство или вещество и способ получения (изготовления) устройства или вещества в целом или их части);

б) одно из которых предназначено для осуществления другого (например, способ и устройство для осуществления способа в целом или одного из его действий);

в) одно из которых предназначено для использования другого (в другом) (например, способ и вещество, предназначенное для использования в способе; способ или устройство и его часть; применение устройства или вещества по новому назначению и способ с их использованием в соответствии с этим назначением; применение устройства или вещества по новому назначению и устройство или композиция, составной частью которых они являются);

г) относящихся к объектам одного вида, одинакового назначения, обеспечивающих получение одного и того же технического результата (варианты).

Состав заявки включает в себя:

- заявление;
- описание;
- формула изобретения;
- чертежи или иные материалы;
- реферат.

Документы, прилагаемые к заявке:

- документ об уплате пошлины (или документ, подтверждающий основания для освобождения от уплаты пошлины);
- документ, подтверждающий полномочия заявителя.

Рассмотрим более подробно содержание перечисленных выше документов. Каждый раздел описания несет определенную информативную нагрузку. В Международной патентной классификации (МПК) заложены два принципа: отраслевой и функциональный (см. раздел 7.6). При современном состоянии патентной культуры (защите изобретений патентами, оплате патентных пошлин и др.) заявитель будет проводить информационный поиск по патентной документации, следовательно, требуется знание основ классификации и применения МПК по следующим источникам:

- официальный бюллетень Патентного ведомства;
- описания к охранным документам;
- патентная документация США, Великобритании, Германии, Франции, Японии, Швейцарии и документация Европейского патентного ведомства;
- непатентная литература с ретроспективой не менее 5 лет

Название изобретения должно отражать существо изобретения. Например, если существо изобретения заключается в конструировании узла машины, то название изобретения – «Машина» - не целесообразно и неверно (т.к. сужается объем). Верное название изобретения – «узел … для машины …». В частности, вольтметр может быть назван «устройством для измерения напряжения». В случае использовании широких терминов в качестве названия (машина, комбайн, измельчитель) последние конкретизируются.

Область техники, к которой относится изобретение, как раздел, должен давать достаточную информацию об области техники того рода объектов, к которому относится заявленное изобретение. Приводится преимущественная область использования изобретения, т.е. четко и ясно формулируется общественная потребность, в соответствии с которой может быть использовано изобретение.

Например, если заявляется конструкция специальной шайбы, указывается:

«Изобретение относится к крепежным средствам для стопорения деталей в процессе завинчивания или затяжки». Может быть указана также одна или несколько областей техники, к которым изобретение относится.

Если заявленное изобретение – отраслевое, обязательно указывается преимущественная область использования изобретения.

Например, если заявляется «Термоэлектрический микрохолодильник» содержание раздела может быть следующим:

«Изобретение относится к холодильной технике, а именно к холодильным устройствам с малым объемом холодильной камеры, работающим на термоэлектрических элементах и предназначенных для использования в промышленности, в частности, для охлаждения приборных отсеков, в каютах пассажирских судов, на железнодорожном и автомобильном транспорте и в быту».

В разделе ***уровень техники*** указывают два-три аналога. В качестве аналога изобретения указывают средство того же назначения, известное из сведений, ставших общедоступными до даты приоритета изобретения, характеризуемое совокупностью признаков, сходной с совокупностью существенных признаков изобретения.

В качестве источников информации могут быть использованы опубликованные печатные работы, патенты и т.п. Наиболее близким к предполагаемому изобретению техническим решением из числа аналогов является прототип. Степень раскрытия признаков прототипа должен соответствовать степени раскрытия соответствующего признака у изобретения.

Например, если один из существенных признаков изобретения: «Нагревание при непрерывном перемешивании при температуре 100 – 150 °С», то при описании сходного признака прототипа недостаточно привести его характеристику в таком виде: «Нагревание при непрерывном перемешивании», так как именно разница в температуре может оказаться причиной получаемого «технического результата».

Сущность изобретения выражается в совокупности существенных признаков, достаточной для достижения обеспечиваемого изобретением технического результата. Признаки относятся к существенным, если они влияют на достижение технического результата, т.е. находятся в причинно-следственной связи с указанным результатом. Подробно раскрывается задача, на решение которой направлено изобретение, с указанием технического результата, который может быть получен при осуществлении изобретения.

Перечень фигур чертежей и иных материалов содержит сведения о представляемых наглядных материалах заявки.

Сведения, подтверждающие возможность осуществления изобретения иллюстрируются примерами конкретного осуществления результатов изобретения.

Формула изобретения, составленная по установленным правилам, краткая словесная характеристика, выражающая сущность изобретения (рисунок 7.2). Формула изобретения предназначается для определения объема правовой охраны, представляемой патентом.

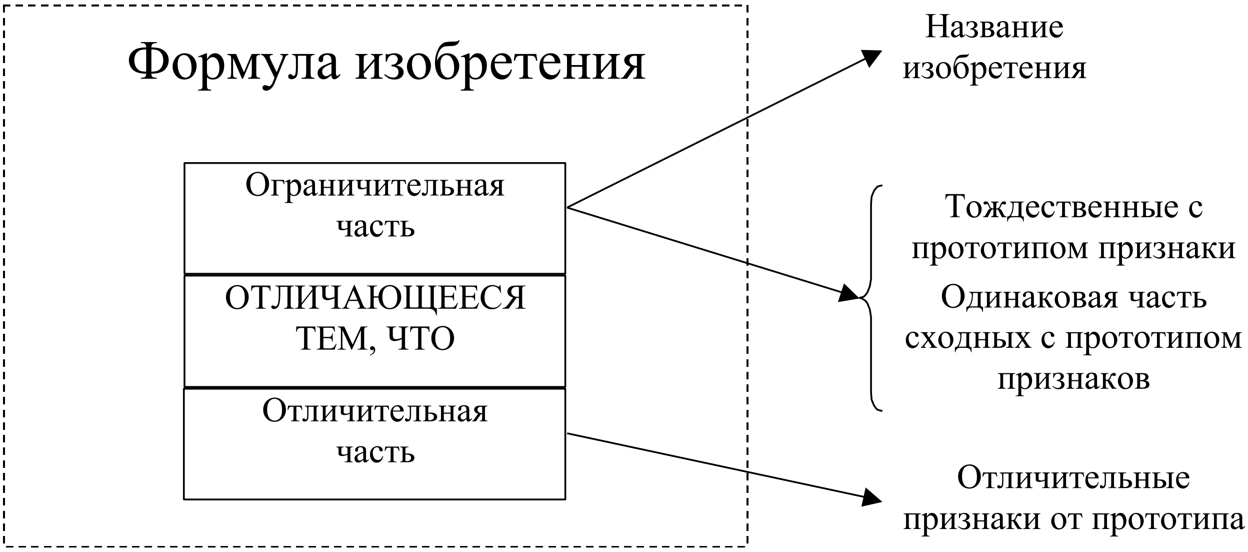


Рисунок 7.2 – Структура формулы изобретения

Формула изобретения должна быть полностью основана на описании, характеризовать изобретение понятиями, содержащимися в его описании.

Формула изобретения признается выражающей сущность изобретения, если она содержит совокупность его существенных признаков, достаточную для достижения указанного заявителем технического результата.

Признаки изобретения выражаются в формуле изобретения таким образом, чтобы обеспечить возможность их идентифицирования, то есть однозначного понимания специалистом на основании известного уровня техники их смыслового содержания.

Характеристика признака в формуле изобретения не может быть заменена отсылкой к источнику информации. Замена характеристики отсылкой к описанию заявки допускается лишь в том случае, когда без такой отсылки признак невозможно охарактеризовать, не нарушая требования следующего подпункта формулы изобретения.

Признак изобретения целесообразно характеризовать общим понятием (выражающим функцию, свойство и т.п.), охватывающим разные частные формы его реализации, если характеристики, содержащиеся в общем понятии, обеспечивают в совокупности с другими признаками получение указанного заявителем технического результата.

Признак может быть выражен в виде альтернативы при условии, что такой признак при любом допускаемом указанной альтернативой выборе в совокупности с другими признаками изобретения обеспечивает получение одного и того же технического результата.

Формула изобретения может быть однозвенной и многозвенной и включать, соответственно, один или несколько пунктов.

Однозвенная формула изобретения применяется для характеристики одного изобретения совокупностью существенных признаков, не имеющей развития или уточнения для частных случаев его выполнения или использования.

Многозвенная формула изобретения применяется для характеристики одного изобретения с развитием и(или) уточнением совокупности его признаков применительно к частным случаям выполнения (или) использования изобретения или для характеристики группы изобретений.

Многозвенная формула, характеризующая одно изобретение, имеет один независимый пункт и следующий (следующие) за ним зависимый (зависимые) пункт (пункты).

Многозвенная формула, характеризующая группу изобретений, имеет несколько независимых пунктов, каждый из которых характеризует одно из изобретений группы, при этом каждое изобретение группы может быть охарактеризовано с привлечением зависимых пунктов, подчиненных соответствующему независимому.

Пункты многозвенной формулы нумеруются арабскими цифрами последовательно, начиная с 1, в порядке их изложения.

При изложении формулы, характеризующей группу изобретений, соблюдаются следующие правила:

- независимые пункты, характеризующие отдельные изобретения, как правило, не содержат ссылок на другие пункты формулы (такая ссылка допустима лишь в случае, когда необходимо изложить данный независимый пункт без полного повторения в нем содержания другого пункта);

- зависимые пункты группируются вместе с тем независимым пунктом, который к ним относится.

Пункт формулы состоит, как правило, из ограничительной части, включающей признаки изобретения, совпадающие с признаками наиболее близкого аналога, в том числе, родовое понятие, отражающее назначение, с которого начинается изложение формулы, и отличительной части, включающей признаки, которые отличают изобретение от наиболее близкого аналога.

При составлении пункта формулы изобретения с разделением на ограничительную и отличительную части после изложения ограничительной части вводится словосочетание: «отличающийся тем, что», непосредственно после которого излагается отличительная часть.

Формула изобретения составляется без разделения ограничительную и отличительную части, в частности, если она характеризует:

- индивидуальное химическое соединение;
- штамм микроорганизма, культуры клеток растений и животных;
- применение ранее известного устройства, способа, вещества, штамма по новому назначению;
- изобретение, не имеющее аналогов.

Пункт формулы излагается в виде одного предложения.

Независимый пункт формулы изобретения должен относится только к одному изобретению. Он характеризует изобретение совокупностью его признаков, определяющей объем испрашиваемой правовой охраны, и излагается в виде логического определения объекта изобретения.

Независимый пункт формулы не признается относящимся к одному изобретению, если содержащаяся в нем совокупность признаков:

- включает выраженные в виде альтернативы признаки, не обеспечивающие получение одного и того же технического результата, либо выраженные в виде альтернативы группы признаков, каждая из которых включает несколько функционально самостоятельных признаков (узел или деталь устройства; операция способа, вещество, материал, приспособление, применяемое в способе); в том числе, когда выбор той или иной альтернативы для какого-либо из таких признаков зависит от выбора, произведенного для другого (других) признака (признаков);
- включает характеристику изобретений, относящихся к объектам разного вида, или совокупности средств, каждое из которых имеет собственное назначение, без реализации указанной совокупностью средств общего назначения.

Зависимый пункт формулы изобретений содержит развитие и/или уточнение совокупности признаков изобретения, приведенных в независимом пункте, признаками, характеризующими изобретение лишь в частных случаях его выполнения или использования.

Ограничительная часть зависимого пункта формулы изобретения состоит из родового понятия, отражающего назначение изобретения, изложенного, как правило, сокращенно по сравнению с приведенным в независимом пункте, и ссылки на независимый пункт и/или зависимый (зависимые) пункт (пункты), к которому (которым) относится данный зависимый пункт. При подчиненности зависимого пункта нескольким пунктам формулы ссылки на них указываются с использованием альтернативы.

Если для характеристики изобретения в частном случае его выполнения или использования наряду с признаками зависимого пункта необходимы лишь признаки независимого пункта, используется подчиненность этого зависимого пункта непосредственно независимому пункту. Если же для указанной характеристики необходимы признаки одного или нескольких других зависимых пунктов формулы, используется подчиненность данного зависимого пункта независимому через соответствующие зависимые пункты.

Не следует излагать зависимый пункт формулы изобретений таким образом, что при этом происходит замена или исключение признаков изобретения, охарактеризованного в предыдущем пункте формулы.

В отличие от изобретения для регистрации полезной модели не является необходимым такое условие как изобретательский уровень. Процедура оформления и подачи заявок на полезную модель и изобретение во многом схожи. Однако порядок регистрации полезной модели иной. Так, по ней не проводится экспертиза по существу, то есть не осуществляется проверка соответствия условиям патентоспособности полезной модели. Если по заявке на изобретение в течение трех лет с даты поступления в Патентное ведомство не была проведена экспертиза по существу, то она считается отозванной. На полезную модель после проведения формальной экспертизы выдается охранный документ (свидетельство), но делается это под ответственность заявителя и без гарантии его действительности. В этом и преимущество, и недостаток регистрации полезной модели. С одной стороны, это оперативное получение охранного документа, с другой – весь риск от использования представленных исключительных прав ложится на заявителя.

Эти обстоятельства представляются чрезвычайно важными, так как совершенно очевидно, что на все патентоспособные объекты, относящиеся к категории устройств и имеющие необходимый изобретательский уровень, может испрашиваться как патент на изобретение, так и свидетельство на полезную модель. Главное – как поступить заявителю в каждом конкретном случае, зависит от его интересов.

7.5 Особенности защиты интеллектуальной собственности и патентного права в различных странах

Вопросы охраны промышленной собственности регламентируются национальными законами и международными соглашениями. В патентных законах предусматриваются различные процедуры, касающиеся подачи и рассмотрения заявок на получение охранных документов на объекты промышленной собственности, их экспертизы, выдачи охранных документов, публикации о выдаче, рассмотрения патентных споров, уступки права, как на сами объекты промышленной собственности, так и на их использование, а также касающиеся других аспектов патентного права, действующего в данной стране.

Выданные патенты (здесь и далее имеются в виду и все другие виды охранных документов на объекты промышленной собственности), а также все сделки с патентами заносятся в патентный реестр.

Требования, предъявляемые в различных зарубежных странах к оформлению заявок на объекты промышленной собственности, несколько отличаются от российских требований. Например, в некоторых странах (США, Канада, Нидерланды) наряду с новизной признаком изобретения называется его полезность. В США и Великобритании отсутствует такая категория объектов промышленной собственности как полезная модель.

В ряде стран объектами патентования являются новые сорта растений (США), в других странах они не патентуются (Швеция, Норвегия, Финляндия, Италия), как и новые виды животных.

В различных странах по-разному регулируется охрана микробиологических продуктов и способов их получения. В некоторых странах возможна охрана не только способа получения микроорганизма, но и создаваемой на его основе культуры микроорганизмов, однако при условии предварительного (до подачи заявки) депонирования (представления для сохранения) этой культуры в компетентном учреждении. Передача ее на хранение в этот орган позволяет заинтересованным лицам ознакомиться с ней и в случае возникновения спора провести сопоставление с иной культурой микроорганизма, составляющей предмет другой заявки.

Если в защищенное патентом изобретение внесено усовершенствование, то на него может быть получен в ряде стран так называемый дополнительный патент. Первый же патент называется основным.

Существуют и так называемые «ввозные» или «подтвержденные» патенты. Они оформляются на основании ранее выданного иностранного патента. Заявитель, имеющий патент на свое имя, испрашивает ввозной (подтвержденный) патент и получает его в упрощенном порядке, без всякой проверки технического решения.

В разных странах законы о недобросовестной конкуренции базируются на общеконституционных принципах и специальных законах. Они могут касаться незаконного использования различных объектов промышленной собственности или служить дополнением к другим специальным законам в случаях, не предусмотренных этими законами (например, незарегистрированный товарный знак в стране, где регистрация является единственным основанием для его охраны по закону о товарных знаках, может быть защищен от незаконного использования законом о недобросовестной конкуренции). Во всяком случае, закон о недобросовестной конкуренции может обеспечить защиту даже тогда, когда другие законы о защите промышленной собственности или их разделы эту защиту не обеспечивают.

Патентное право предусматривает либо заявительную, либо авторскую системы патентования. При заявительной системе патент выдается любому первому заявителю на его имя, будь то автор, либо законный преемник автора, либо лицо, присвоившее изобретение действительного автора. При авторской системе патент может получить лишь автор либо его правопреемник, причем имя автора должно быть названо в заявочной документации и в патенте, за исключением случаев, когда сам автор и заявитель просят не указывать имя автора.

В патентных законодательствах большинства стран (США, ЕЭС и др.) провозглашен принцип, согласно которому получить патент может только действительный автор либо тот, кому автор уступил свое изобретение. Упоминание имени автора в патенте само по себе не дает ему никаких реальных прав на получение выгод от патента, если патент выдается предприятию (нанимателю). Такие изобретения обычно называются служебными. Иностранцы могут, как правило, получать патенты наравне с гражданами данной страны.

Для патентного права зарубежных стран характерны (по роли в них экспертизы) две системы выдачи патента – явочная и проверочная (исследовательская), разновидностью которой является отложенная (отсроченная) система.

При явочной системе заявка рассматривается только для того, чтобы выяснить:

1) соблюдены ли заявителем формальные требования (приложены ли к заявке все необходимые документы в установленном количестве экземпляров, а если заявка подана через доверенное лицо, имеется ли доверенность);

2) не испрашивает ли заявитель патент на объекты, которые нельзя патентовать;

3) правильно ли составлено описание, чертежи, формула изобретения и т.д., в частности, содержится ли в заявке одно предполагаемое изобретение или их группа, связанная единством замысла. Новизна же изобретения патентным ведомством специально не исследуется. Она предполагается, поскольку ее наличие обязательно и при явочной системе.

Преимущество явочной системы состоит в том, что заявитель сравнительно быстро получает патент, а общество – информацию об изобретении. Однако она имеет и ряд отрицательных сторон: патенты выдаются «на страх и риск» заявителей, определенное число патентов, выданных ошибочно, аннулируется из-за отсутствия новизны или технической значимости. Аннулирование происходит после рассмотрения дел в суде по заявлениям заинтересованных лиц. Явочная система принята во многих странах: Бельгии, Греции, Испании, Италии, в ряде стран Азии, Африки, Южной Америки и др.

Проверочная (исследовательская) система характеризуется тем, что заявка подвергается проверке с целью выяснить, соответствует ли заявляемое техническое решение условиям патентоспособности.

Проверочная система принята в Индии, Колумбии, США, Швеции и других странах (их меньше, чем стран с явочной системой). Она избавляет заявителя от дальнейших неоправданных затрат (например, при отсутствии новизны), связанных с патентным производством или аннулированием патента. Патент, выданный в стране с проверочной системой, пользуется большим доверием. Число споров по таким патентам обычно меньше, чем в странах с явочной системой. К недостаткам этой системы можно отнести более длительные сроки процедуры и возможность необоснованных отказов в выдаче патента.

Промежуточная система представляет собой переходную стадию от явочной к проверочной системе. Для нее характерно проведение частичной, неполной проверки заявки (Египет, Ливия, Тунис и др.). Отклонение заявки в этих странах возможно не только в результате формальной экспертизы, но и по материалам проверки, проведенной третьими лицами, оспаривающими заявку.

Отложенная (отсроченная) система представляет собой модификацию проверочной системы. При этой системе проверка проводится только по просьбе заявителя или другого заинтересованного лица. Заявка (обычно спустя не более 18 месяцев с даты ее подачи) подлежит обязательной публикации («выкладке»). По выложенной заявке каждый вправе подать обоснованные возражения. С момента публикации заявки изобретение получает временную охрану. Отложенная экспертиза введена в Австралии, Нидерландах, России, ФРГ, Японии и некоторых других странах. Патент выдается лишь после проведения экспертизы, если она дает положительный результат. Если просьба о ее проведении не поступает (в Нидерландах и ФРГ – в течение 7 лет, в Австралии – 5 лет, в России – 3 лет), то право на получение патента утрачивается.

7.6 Международная патентная классификация

Самые первые классификационные системы представляли собой алфавитные перечни выданных патентов. Во Франции, например, такой перечень был составлен в 1791 г. 24 марта 1791 г. в Страсбурге (Франция) было принято соглашение о Международной классификации изобретений. Нынешнее название этой системы – Международная патентная классификация. Классификация пересматривается и переиздается в виде новой редакции каждые 5 лет. С 1 января 2000 г. действует седьмая редакция МПК.

Основной целью МПК является создание эффективного поискового инструмента и обеспечение возможности классифицировать любое техническое понятие, которое относится к изобретению.

МПК состоит из разделов, классов, подклассов, групп (основных групп и подгрупп). Она охватывает все области техники, изобретения в которых охраняются патентами. Эти области техники делятся на восемь разделов:

A – Удовлетворение жизненных потребностей человека;

B – Различные технологические процессы; транспортирование;

C – Химия; металлургия;

D – Текстиль; бумага;

E – Строительство; горное дело;

F – Механика; освещение; отопление; двигатели и насосы; оружие и боеприпасы; взрывные работы;

G – Физика;

H – Электричество.

В каждом разделе имеется ряд подразделов, имеющих информативные подзаголовки, не обозначенные символами. Например, в разделе А «Удовлетворение жизненных потребностей человека» имеются подразделы «Сельское хозяйство», «Пищевые продукты, табак» и др.

Каждый раздел делится на классы, имеющие индекс и заголовок. Индекс класса состоит из индекса раздела, за которым следует двухзначная цифра. Заголовок класса раскрывает его содержание. Например, класс G05 (раздел G «Физика») называется «Управление, регулирование».

Каждый класс содержит один или несколько подклассов, также обозначаемых индексом и заголовком. Индекс каждого подкласса состоит из индекса класса, за которым следует прописная буква. Заголовок подкласса раскрывает его содержание более точно, нежели заголовок класса. Например, подкласс G05D (раздел G «Физика», класс G05 «Управление, регулирование») называется «Системы управления или регулирования неэлектрических величин».

При нумерации классов в конце каждого подраздела оставляют пропуски, которые позволяют вносить новые классы позднее при пересмотре МПК. В ряде случаев для удобства произношения не употребляются гласные для обозначения подклассов.

Каждый подкласс подразделяется на группы, которые представляют собой или основные группы, или подгруппы, обозначаемые индексом и содержащие заголовок. Индекс каждой группы состоит из индекса подкласса, за которым следуют два числа, отделенных наклонной чертой. Индекс основной группы состоит из индекса подкласса, после которого следует одно-, двух- или трехзначное число, наклонная черта и два нуля. Например, G05D23/00. Заголовок основной группы определяет тематику, являющуюся приемлемой при поиске изобретений. Группа G05D23/00 называется «Управление и регулирование температуры».

Индекс подгруппы состоит из индекса подкласса, за которым следует одно-, двух-, трехзначное число его основной группы, наклонная черта и, по меньшей мере, одна цифра, отличающаяся от «00». Например, G05D23/24.

Заголовок подгруппы определяет предметную область в пределах основной группы, являющуюся приемлемой при поиске изобретений. Заголовку предшествует одна или несколько точек, указывающих на иерархическое положение этой подгруппы. Во всех случаях подгруппа должна читаться как зависимая от текста группы, под которым она расположена со сдвигом. Например, подгруппа G05D23/24 называется «Управление и регулирование температуры с помощью термочувствительных элементов, сопротивление которых зависит от температуры».

Начиная со второй редакции МПК, введенной в действие 1 января 1975 г., была предусмотрена возможность после двойной наклонной черты проставлять индексы добавленной к основной информации, относящейся к составляющим техническую сущность элементам изобретения, классифицируемого в целом основными индексами. Например, в обозначении A21B1/02//A47J37/00 содержится информация, расшифровывающаяся как «Хлебопекарные печи, отличающиеся типом нагревательных устройств и применяемые в домашнем оборудовании для хлебопечения».

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Основы научных исследований. Учебное пособие. Под ред. В.И.Крутова. М.: Высшая школа, 1989. – 400 с.

2. 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.

3. Голдстейн М., Голдстейн И. Как мы познаем. М.: Знание, 1984. – 256 с.

4. Скурихин В.И., Шифрин В.Б., Дубровский В.В. Математическое моделирование. Киев, Техника, 1983. – 270 с.

5. Чкалова О.Н. Основы научных исследований. Киев: Вища школа, 1976. – 195 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.

9. Комментарии к Российскому патентному законодательству (справочное пособие патентоведов и изобретателей). Роспатент – НИИГПЭ. М.: 1993.

10. Бромберг Г.В. Основы патентного дела. М.: ИНИЦ, 2000. – 172 с.

11. Шестимиров А.А. Составление заявок на изобретение в Российской Федерации. Учебное пособие. М.: 1996.

12. Интеллектуальная собственность. Т.1. Авторское право и смежные права/ Составитель Р.Попова. 1997. – 560 с.