

Министерство образования и науки Российской Федерации  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ  
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ  
(ТУСУР)

**Кафедра физической электроники (ФЭ)**

**Н.В. Зариковская**

# **ОСНОВЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ**

Учебное пособие

2012

**Зариковская Н.В.**

Основы математического моделирования. Учебное пособие. – Томск: Томский государственный университет систем управления и радиоэлектроники (ТУСУР), 2012.– 244с.

Технический редактор:

**Кондратьев Е.А.**

В настоящее время большую роль при разработке новых материалов и технологий играет физическое и математическое моделирование. Математические модели позволяют выявить основные достоинства и недостатки объектов еще на ранней стадии проектирования.

Данное пособие является основой при изучении курса «Математическое моделирование в микроэлектронике» для студентов специальности 210104 «Микроэлектроника и твердотельная электроника». В нем рассматриваются вопросы реализации основных типов математических моделей полупроводниковых приборов и элементов. Пособие содержит базовый теоретический материал, а также результаты исследований, проводимых в области математического моделирования в микроэлектронике.

© Зариковская Н.В. 2012

© **Томский государственный университет систем управления и радиоэлектроники (ТУСУР),**

## СОДЕРЖАНИЕ

|  |            |
|--|------------|
| Введение .....   | 5          |
| <b>ГЛАВА 1. ОПРЕДЕЛЕНИЕ И НАЗНАЧЕНИЕ<br/>МОДЕЛИРОВАНИЯ.....</b>  | <b>8</b>   |
| 1.1 Понятие модели.....  | 8          |
| 1.2 Классификация моделей .....  | 17         |
| 1.3 Классификация математических моделей.....  | 31         |
| 1.3.1 Классификация в зависимости от сложности объекта<br>моделирования .....                          | 32         |
| 1.3.2 Классификация в зависимости от оператора модели .....  | 34         |
| 1.3.3 Классификация в зависимости от параметров модели .....   | 38         |
| 1.3.4 Классификация в зависимости от целей моделирования .....   | 44         |
| 1.3.5 Классификация в зависимости от методов исследования.....   | 46         |
| <b>ГЛАВА 2. ЭТАПЫ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКОЙ<br/>МОДЕЛИ.....</b>  | <b>49</b>  |
| 2.1 Обследование объекта моделирования .....   | 51         |
| 2.2 Концептуальная постановка задачи моделирования .....   | 55         |
| 2.3 Математическая постановка задачи моделирования .....   | 61         |
| 2.4 Выбор и обоснование выбора метода решения задачи.....  | 68         |
| 2.5 Реализация математической модели в виде программы<br>для ЭВМ .....                                 | 74         |
| 2.6 Проверка адекватности модели .....   | 78         |
| 2.7 Практическое использование построенной модели и<br>анализ результатов моделирования .....          | 82         |
| 3.1 Что такое структурная модель?.....   | 88         |
| 3.2 Способы построения структурных моделей.....  | 98         |
| <b>ГЛАВА 4. МОДЕЛИРОВАНИЕ В УСЛОВИЯХ<br/>НЕОПРЕДЕЛЕННОСТИ.....</b>                                     | <b>108</b> |
| 4.1 Причины появления неопределенностей и их виды.....   | 109        |
| 4.2 Моделирование в условиях неопределенности,<br>описываемой с позиций теории нечетких множеств ..... | 117        |
| 4.3 Моделирование в условиях стохастической<br>неопределенности .....                                  | 133        |
| 4.4 Моделирование марковских случайных процессов.....  | 158        |
| Пример 1.....  | 160        |

|  |            |
|--|------------|
| <b>ГЛАВА 5. ЛИНЕЙНЫЕ И НЕЛИНЕЙНЫЕ МОДЕЛИ.....</b>  | <b>167</b> |
| <b>5.1 О законе Гука и границах линейности .....</b>   | <b>168</b> |
| <b>5.2 Поля, сплошные среды и уравнения математической физики. Линейные уравнения и принцип суперпозиции ...</b> | <b>170</b> |
| <b>5.3 Вывод волнового уравнения из законов механики....</b>   | <b>175</b> |
| <b>5.4 Решение волнового уравнения методом Фурье .....</b>   | <b>179</b> |
| <b>5.5 О характеристиках уравнений математической физики. Решение волнового уравнения методом Даламбера.....</b> | <b>184</b> |
| <b>5.6 Уравнения Максвелла .....</b>   | <b>189</b> |
| <b>5.7 О классификации квазилинейных систем.....</b>   | <b>192</b> |
| <b>5.8 Связь непрерывного и дискретного на примерах уравнения колебаний струны и уравнения Шредингера ...</b>    | <b>195</b> |
| <b>5.9 Анализ подобия и размерности .....</b>  | <b>200</b> |
| <b>5.10 Автомодельность .....</b>  | <b>206</b> |
| <b>5.11 Самоорганизация и структуры в нелинейных средах</b>  | <b>210</b> |
| <b>5.12 О нелинейных волнах в сплошных средах .....</b>  | <b>216</b> |
| <b>5.13 Иерархические модели турбулентности и многомасштабные функциональные базисы.....</b>                     | <b>224</b> |
| <b>5.14 Вейвлеты .....</b>   | <b>234</b> |
| <b>Список использованной литературы.....</b>   | <b>242</b> |

## Введение

...Становится ясно, что в следующем веке понадобятся не только эксперты по некоторым аспектам отдельных стадий избранных процессов. Понадобятся специалисты по решению проблем... По-видимому, междисциплинарность будет в цене. А в институтах будут стараться учить не "предметам", а стилям мышления.

Г.Г.Малинецкий

Понятие “математическое моделирование” в последние два – три десятилетия является едва ли не самым распространенным в научной литературе, по крайней мере, – в естественно-научной и технической. Сегодня трудно представить себе проектную или конструкторскую организацию, не использующую в своей практике в той или иной мере математические модели. Все более распространенным и эффективным становится применение математического моделирования в научных исследованиях. Подавляющее большинство известных авторам диссертационных работ по специальностям естественно-научных и инженерно-технических направлений связано с разработкой и использованием соответствующих математических моделей. В последние 10–15 лет эта тенденция получает все более широкое распространение при подготовке дипломных работ выпускников вузов. Можно констатировать, что математическое моделирование в последние десятилетия оформилось в отдельную междисциплинарную область знаний, с присущими ей объектами, подходами и методами исследования.

В связи с этим все более актуальной становится задача целенаправленной подготовки специалистов – “матмодельеров” в вузах различного профиля, в рамках различных направлений и специальностей. Кроме того, по мнению авторов, эту подготовку, формирование соответствующего “стиля мышления” можно (а возможно – и необходимо) начинать со старших классов общеоб-

разовательной школы. Определенный положительный опыт подобной работы с учащимися специализированных физико-математических школ и классов позволяет, по крайней мере, не отвергать данную гипотезу.

Реализация образовательного процесса по подготовке специалистов “матмодельеров”, естественно, требует наличия соответствующего методического обеспечения. За последние годы издано немало прекрасных монографий, статей, научно-популярных брошюр, пособий, часть из которых цитируется в тексте. К их числу в первую очередь следует отнести работы отечественных ученых: А.А.Самарского, Н.Н.Моисеева, С.П.Курдюмова, Г.Г.Малинецкого и многих других, трудами которых математическое моделирование и превратилось в самостоятельную область знаний. В этих работах, большей частью написанных математиками, достаточно подробно и прозрачно освещены такие вопросы, как предмет, подходы, методы математического моделирования, приведено огромное количество ярких примеров математических моделей. Как правило, в работах этого направления основное внимание уделяется методам исследования собственно математических моделей, качественному анализу решений, новым эффектам в исследуемых процессах и явлениях.

Следует отметить, что в настоящее время в значительной части учебников и учебных пособий по различным дисциплинам включаются некоторые понятия, методы и примеры применения математического моделирования. Здесь обычно используется некоторый набор “базовых” моделей данной дисциплины или смежных с ней, из этих “кубиков” в дальнейшем строится модель анализируемого процесса. При этом, как правило, базовые модели принимаются как данность, не обсуждается правомерность их применения, область применимости, степень адекватности описания. “За кадром” в большинстве случаев остается и собственно процесс создания математической модели, процесс перехода от “языка природы” к “языку” математики.

Понятно, что указанные обстоятельства обусловлены сложившимися в различных областях традициями, спецификой дисциплин, личными склонностями авторов. В то же время нам представляется полезным наличие пособия, в котором более де-

тально раскрывалась бы “кухня” разработчиков математических моделей. Несмотря на то, что создание любой новой модели – процесс творческий, близкий к искусству, - существуют достаточно общие подходы, методы, “инструменты”, пригодные для различных предметных областей. Именно этому кругу вопросов, “технологии” создания математических моделей будет уделено наибольшее внимание в предлагаемом пособии. Насколько это нам удалось – судить читателю.

Пособие ориентировано в первую очередь на студентов старших курсов. Для работы с большинством разделов пособия достаточно знания вузовского курса математики (математический анализ, линейная алгебра и аналитическая геометрия, элементы теории вероятности, теории обыкновенных дифференциальных уравнений и уравнений математической физики), изучаемого обычно на младших курсах. При этом мы стремились сделать пособие “замкнутым”, то есть содержащим по возможности все необходимые понятия, определения и другие сведения.

# ГЛАВА 1. ОПРЕДЕЛЕНИЕ И НАЗНАЧЕНИЕ МОДЕЛИРОВАНИЯ

## 1.1 Понятие модели

Что значит знать? Вот, друг мой, в чем вопрос. На этот счет у нас не все в порядке.

Гете

### *Модели вокруг нас*

С процессом моделирования и различными моделями человек начинает сталкиваться с самого раннего детства. Так, еще не научившись уверенно ходить, малыш начинает играть с кубиками, сооружая из них (на первых порах – с помощью старших) различные конструкции. Его окружают различные игрушки: плюшевые, резиновые, металлические, различающиеся по размерам, форме, цвету, предназначению и т.д. При этом большинство игрушек повторяют (моделируют) отдельные свойства и форму реально существующих предметов и объектов. В этом смысле такие игрушки можно рассматривать в качестве моделей соответствующих объектов.

В школе практически все обучение построено на использовании моделей в той или иной форме. Действительно, для знакомства с основными конструкциями и правилами родного языка используются различные структурные схемы и таблицы, которые можно считать моделями, отражающими те или иные свойства языка. Процесс написания сочинения следует рассматривать как моделирование некоторого события или явления средствами родного языка. На уроках биологии, физики, химии и анатомии к плакатам и схемам (т.е. моделям) добавляются макеты (тоже модели) изучаемых реальных объектов. На уроках рисования или черчения на листе бумаги или ватмана создаются модели тех или иных объектов, выраженные изобразительным языком или более формализованным языком чертежа.

Даже такую трудно формализуемую область знания, как история, также можно рассматривать как непрерывно эволюционирующую совокупность моделей прошлого того или иного народа, государства и т.д. Устанавливая закономерности в наступлении тех или иных исторических событий (революций, войн, ускорений или застоев исторического развития), можно не только выяснить причины, приведшие к данным событиям, но и прогнозировать и даже управлять их появлением и развитием в будущем.

На протяжении всей своей жизни человек ежедневно сталкивается с моделями и сам создает новые.

Так, картину, написанную художником, можно считать моделью реальных объектов, записанной языком живописи. Такими же моделями можно считать художественное произведение и скульптуру. Даже жизненный опыт человека, его представления о мире является примером модели. Причем поведение человека определяется сформировавшейся в его сознании моделью. Психолог или учитель, изменяя параметры такой внутренней модели, могут в отдельных случаях существенно изменять поведение человека. Без преувеличения можно утверждать, что в своей осознанной жизни человек имеет дело исключительно с моделями тех или иных реальных объектов, процессов, явлений. При этом один и тот же объект воспринимается различными людьми по-разному, иногда – с точностью “до наоборот”. Как говорится в известной пословице – “на вкус и цвет...”. Это восприятие, мысленный образ объекта также является разновидностью модели последнего (так называемой *когнитивной моделью*), и существенным образом зависит от множества факторов: качества и объема знаний, особенностей мышления, эмоционального состояния конкретного человека “здесь и сейчас” и от множества других причин, зачастую не доступных рациональному осознанию. Особенно велика роль моделей и моделирования в современной науке и технике.

Все существующие в настоящее время приемы моделирования можно (условно) разделить на материальные и идеальные.

*Материальное моделирование* – это моделирование, при котором исследование объекта выполняется с использованием его материального аналога (от греч. *analogia* – соответствие, сораз-

мерность), воспроизводящего основные физические, геометрические, динамические и функциональные характеристики данного объекта. К таким моделям, например, можно отнести использование макетов в архитектуре, моделей и экспериментальных образцов при создании различных транспортных средств.

*Идеальное моделирование* отличается от материального тем, что оно основано не на материальной аналогии объекта и модели, а на аналогии идеальной, мыслимой и всегда носит теоретический характер.

Можно ли обойтись в технике без применения тех или иных видов моделей? Очевидный ответ – нет! Конечно, можно построить новый самолет “из головы” (без предварительных расчетов, чертежей, экспериментальных образцов, т.е. используя только единственную “идеальную” модель, существующую в “голове” конструктора). Однако едва ли это будет достаточно эффективная и надежная конструкция. Единственным ее достоинством можно считать ее уникальность. Ведь даже автор не сможет повторно изготовить точно такой же самолет, т.к. в результате изготовления первого экземпляра будет получен некоторый опыт, который обязательно изменит “идеальную” модель в “голове” самого конструктора.

Чем более сложным и надежным должно быть техническое изделие, тем большее число видов моделей необходимо применить на этапе его проектирования.

Как правило, сложные изделия создаются целыми коллективами разработчиков. Вся совокупность применяемых ими разнообразных моделей позволяет сформировать общую для всего коллектива “идеальную” модель разрабатываемого изделия. Реальное техническое изделие можно рассматривать как “материальную” модель (аналог) созданной авторами “идеальной” модели.

### ***Определение модели***

Понятия модели и моделирования наиболее распространены в сфере обучения, научных исследованиях, при выполнении проектно-конструкторских работ, в серийном техническом производстве. В каждой из этих областей моделирование имеет свои

особенности. Далее в этой книге моделирование будет рассматриваться главным образом применительно к научным исследованиям. Чаще всего термин “модель” используют для обозначения:

- устройства, воспроизводящего строение или действие какого-либо другого устройства (уменьшенное, увеличенное или в натуральную величину);
- аналога (чертежа, графика, плана, схемы, описания, ...) какого-либо явления, процесса или предмета.

К недостаткам термина “модель” следует отнести его многозначность. В словарях можно найти до восьми различных значений данного термина, из которых в научной литературе наиболее распространены два:

- модель как аналог реального объекта;
- модель как образец будущего изделия.

Научное познание сосредоточено на изучении предметов, явлений и процессов, существующих вне нашего сознания и называемых объектами исследования (от лат. *objectum* – предмет).

Важную роль при проведении исследований играют *гипотезы* (от греч. *hypothesis* – основание, предположение), т.е. определенные предсказания, предположительные суждения о причинно-следственных связях явлений, основанные на некотором количестве опытных данных, наблюдений, догадок. Формулирование и проверка правильности гипотез основывается, как правило, на аналогиях.

*Аналогия* – это представление о каком-либо частном сходстве двух объектов, причем такое сходство может быть как существенным, так и несущественным. Существенность сходства или различия двух объектов условна и зависит от уровня абстрагирования (от лат. *abstrahere* – отвлекать), определяемого конечной целью исследования. В данном случае уровень абстрагирования определяется набором учитываемых параметров объекта исследования. Например, при изучении механических свойств в качестве объектов исследования могут быть выделены материалы из дерева, металла, пластмассы и т.д. В свою очередь материалы из дерева можно подразделить по видам древесины на лиственные и

хвойные, лиственные – на “березу”, “тополь”, “ясень” и т.д. В данном примере степень абстрагирования снижается при добавлении учитываемых параметров. Следует заметить, что уровень абстрагирования данного объекта всегда определяется по отношению к другим объектам.

Гипотезы и аналогии, в определенной мере отражающие реальный, объективно существующий мир, должны обладать наглядностью или сводиться к удобным для исследования логическим схемам. В науке моделями называются логические схемы, упрощающие рассуждения и логические построения или позволяющие проводить эксперименты, уточняющие понимание природы явлений. Другими словами, модель – это объект-заменитель объекта-оригинала, обеспечивающий изучение некоторых интересующих исследователя свойств оригинала.

Под *моделью* (от лат. *modulus* – мера, образец, норма) понимают такой материальный или мысленно представляемый объект, который в процессе познания (изучения) замещает объект-оригинал, сохраняя некоторые важные для данного исследования типичные его черты. Процесс построения и использования модели называется моделированием.

Представленное определение является достаточно общим и может трактоваться по-разному. В частности, любое знание можно рассматривать как некоторую идеальную модель природного объекта или явления. В свою очередь, любой искусственный (т.е. созданный человеком) объект или процесс есть материальная модель, построенная на основе соответствующих знаний (идеальных моделей). В этом смысле можно говорить о трех реальностях или трех сферах (что близко к мыслям, высказанным В.И.Вернадским еще в 1922 году), в которых живет человек.

Первой реальностью является живая и неживая природа, законы развития которой не зависят от человека. Поэтому природные объекты и явления нельзя рассматривать по отношению к человеку как модели. Однако познание и использование человеком природных объектов возможно только через их модели, которые в результате изучения самих объектов также изменяются.

Объекты природы находят свое отражение во второй реальности или *ноосфере*, включающей знания, накопленные всем че-

ловечеством и практически мало зависящие от конкретного человека. Данная реальность, состоящая из идеальных моделей, зависит от эволюции человечества и изменяется в процессе познания, пополняясь новыми и изменяя старые модели. Можно сказать, что процесс познания в любой области знаний представляет собой непрерывное совершенствование существующих и построение новых моделей исследуемых объектов. Этот ряд моделей (с оптимистической точки зрения) бесконечен.

Наконец, третья реальность или *техносфера* может рассматриваться как отражение второй реальности, включает все материальные модели, созданные человеком. К составляющим техносферы следует отнести также искусственное разведение животных и растений, их селекция и (в последнее время) их клонирование. Хотя сами живые организмы являются представителями живой природы и моделями не являются, но процесс их появления управляется человеком на основании некоторых модельных представлений о данных объектах.

В контексте данных рассуждений и приведенного выше определения модели можно сделать вывод о том, что человек в своей жизни в основном занимается знакомством с уже созданными ранее моделями и созданием на их основе новых идеальных или материальных моделей. Поэтому понятие “человек моделирующий” можно считать тождественным понятию “человек разумный”.

### *Свойства моделей*

В настоящее время нет предпосылок к выделению “самых элементарных” и “неделимых” кирпичиков мироздания. Поэтому можно утверждать, что любой объект исследования является бесконечно сложным и характеризуется бесконечным числом параметров. При построении модели исследователь всегда исходит из *целей* своего исследования, учитывает только наиболее существенные для достижения поставленных целей факторы. Поэтому любая модель нетождественна объекту-оригиналу и, следовательно, *неполна*, поскольку при ее построении исследователь выделил только наиболее существенные с его точки зрения факторы. Отброшенные факторы, несмотря на свое относительно малое

влияние на поведение объекта по сравнению с факторами, выбранными в качестве существенных, все же в совокупности могут приводить к значительным различиям между объектом и его моделью. “Полная” модель, очевидно, будет полностью тождественна оригиналу. Данную мысль хорошо выразили Артуро Розенблют и Норберт Винер, заметив, что “*наилучшей моделью кота является другой кот, а еще лучше – тот же самый кот*”. С другой стороны, как отметил М.Вартофский, при моделировании должно “*исключаться какое-то бы ни было самоотнесение, ничто не может быть моделью самого себя*”.

Если результаты моделирования удовлетворяют исследователя и могут служить основой для прогнозирования поведения или свойств исследуемого объекта, то говорят, что модель *адекватна* (от лат. *adaequatus* – приравненный) объекту. При этом адекватность модели зависит от целей моделирования и принятых критериев. Учитывая заложенную при создании неполноту модели, можно утверждать, что идеально адекватная модель принципиально невозможна.

В качестве одной из характеристик модели может выступать *простота модели* (или *сложность*). Очевидно, что из двух моделей, позволяющих достичь желаемой цели и получить требуемые результаты с заданной точностью, предпочтение должно быть отдано более простой. При этом адекватность и простота модели далеко не всегда являются противоречивыми требованиями. Учитывая бесконечную сложность любого объекта исследования, можно предположить существование бесконечной последовательности его моделей, различающихся по степени полноты, адекватности и простоты.

В качестве еще одного свойства модели можно рассматривать *потенциальность модели* (от лат. *potentia* – мощь, сила) или *предсказательность* для получения новых знаний об исследуемом объекте. Данное свойство модели подчеркивается в определении Н.Н. Моисеева: “*Под моделью мы будем понимать упрощенное, если угодно, упакованное знание, несущее вполне определенную, ограниченную информацию о предмете (явлении), отражающее те или иные его свойства. Модель можно рассматривать как специальную форму кодирования информации. В отли-*

чие от обычного кодирования, когда известна вся исходная информация, и мы лишь переводим ее на другой язык, модель, какой бы язык она не использовала, кодирует и ту информацию, которую люди еще не знали. Можно сказать, что модель содержит в себе потенциальное знание, которое человек, исследуя ее, может приобрести, сделать наглядным и использовать в своих практических жизненных нуждах". Поэтому же поводу высказываются Т.Тоффли и Н.Марголус: "В науке мало пользы от моделей, которые рабски подчиняются нашим желаниям. Мы хотим иметь модели, которые дерзят нам; модели, которые имеют свой собственный ум. Мы хотим получать от моделей больше, чем в них вложили". Именно свойство потенциальности (иногда называемое богатством модели) позволяет модели выступать в качестве самостоятельного объекта исследования.

Модели в научных исследованиях, не обладающие определенной "предсказательностью", едва ли могут считаться удовлетворительными.

Известно немало случаев, когда изучение или использование моделей позволило сделать открытия. В качестве примера можно привести открытие планеты Нептун, положение которой было предсказано французским астрономом Лаверье на основании расчетов, выполненных с использованием закона всемирного тяготения (т.е. модели) и данных о движении планеты Уран. В наше время только на основании результатов теоретического моделирования открыты "черные дыры" в астрофизике и кварки в физике элементарных частиц (и те, и другие экспериментально подтверждены пока косвенно), высокотемпературный Т-слой в плазме (подтвержден экспериментально), использование которого позволяет значительно повысить коэффициент полезного действия магнитогидродинамических генераторов, которые в настоящее время рассматриваются как перспективные устройства для получения электрической энергии.

### ***Цели моделирования***

Хорошо построенная модель, как правило, доступнее, информативнее и удобнее для исследователя, нежели реальный объект.

Рассмотрим основные цели, преследуемые при моделировании в научной сфере. Самым важным и наиболее распространенным предназначением моделей является их применение при изучении и прогнозировании поведения сложных процессов и явлений. Следует учитывать, что некоторые объекты и явления вообще не могут быть изучены непосредственным образом. Недопустимы, например, эксперименты с экономикой страны или со здоровьем ее населения. Принципиально неосуществимы эксперименты с прошлым какого-то государства или народа (*“История не терпит сослагательного наклонения”*). Невозможно (по крайней мере, в настоящее время) провести эксперимент по прямому исследованию структуры звезд. Многие эксперименты неосуществимы в силу своей дороговизны или рискованности для человека и/или среды его обитания.

Как правило, в настоящее время предварительные всесторонние исследования различных моделей явления предшествуют проведению любых сложных экспериментов. Более того, эксперименты на моделях с применением ЭВМ позволяют разработать план натуральных экспериментов, выяснить требуемые характеристики измерительной аппаратуры, наметить сроки проведения наблюдений, а также оценить стоимость такого эксперимента.

Другое, не менее важное, предназначение моделей состоит в том, что с их помощью выявляются наиболее существенные факторы, формирующие те или иные свойства объекта, поскольку сама модель отражает лишь некоторые основные характеристики исходного объекта. Например, исследуя движение массивного тела в атмосфере вблизи поверхности Земли, можно выяснить, что его ускорение существенно зависит от массы и геометрической формы (в частности – от величины поперечного к направлению движения сечения тела), но не зависит от цвета поверхности.

Модель позволяет научиться правильно управлять объектом, апробируя различные варианты управления на модели этого объекта. Использовать для этого реальный объект часто бывает рискованно или просто невозможно. Например, получить первые навыки в управлении современным самолетом безопаснее, быстрее и дешевле на тренажере (т.е. модели), чем подвергать себя и дорогую машину риску.

Если свойства объекта с течением времени меняются, то особое значение приобретает задача прогнозирования состояний такого объекта под действием различных факторов. Например, при проектировании и эксплуатации любого сложного технического устройства желательно уметь прогнозировать изменение надежности функционирования как отдельных подсистем, так и всего устройства в целом.

Итак, модель нужна для того, чтобы:

- понять, как устроен конкретный объект: какова его структура, основные свойства, законы развития, саморазвития и взаимодействия с окружающей средой;
- научиться управлять объектом или процессом, определять наилучшие способы управления при заданных целях и критериях;
- прогнозировать прямые и косвенные последствия реализации заданных способов и форм воздействия на объект.

## 1.2 Классификация моделей

В действительности вообще нет никаких строго проведенных межей и граней, к великой горести всех систематиков.

А.И. Герцен

Отмеченная выше неоднозначность термина “модель”, огромное количество типов моделирования и их быстрое развитие затрудняет в настоящее время построение логически законченной, удовлетворяющей всех классификации моделей. Любая подобная классификация условна в силу того, что она отражает, с одной стороны, пристрастия авторов, а с другой – ограниченность их знаний в конечном числе областей научного познания. Данную классификацию следует рассматривать как попытку построения некоторого инструмента или модели для исследования свойств и характеристик самого процесса моделирования.

Как уже было отмечено, в настоящее время существует несколько подходов к моделированию, которые условно можно объединить в две большие группы: материальное (“овеществлен-

ное”, предметное) и идеальное (мысленное или теоретическое) моделирование. Учитывая, что идеальное моделирование является первичным по отношению к материальному (вначале в сознании человека формируется идеальная модель, а затем на ее основе строится материальная), существующие типы моделирования можно представить следующей схемой (рисунок 1.1). Знакомство с видами моделирования начнем с материального, которое, хотя и является вторичным, но более наглядно и просто для понимания.

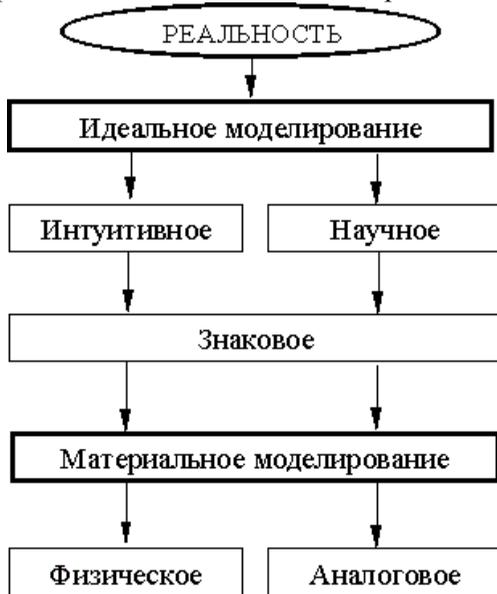


Рисунок 1.1 – Типы моделирования

### ***Материальное моделирование***

Основными разновидностями материального моделирования являются физическое и аналоговое моделирование. При этом оба вида материального моделирования основаны на свойствах геометрического или физического подобия. Две геометрические фигуры подобны, если отношение всех соответственных длин и углов одинаковы. Если известен коэффициент подобия – масштаб, то простым умножением размеров одной фигуры на величину

масштаба получаются размеры другой, ей подобной геометрической фигуры. Два явления физически подобны, если по заданным характеристикам одного можно получить характеристики другого простым пересчетом, который аналогичен переходу от одной системы единиц измерения к другой. Изучением условий подобия явлений занимается теория подобия.

*Физическое моделирование* – это такое моделирование, при котором реальному объекту ставится в соответствие его увеличенный или уменьшенный материальный аналог, допускающий исследование (как правило, в лабораторных условиях) с помощью последующего перенесения свойств изучаемых процессов и явлений с модели на объект на основе теории подобия.

К примерам физических моделей можно отнести макеты в архитектуре, модели судов в судостроении. Следует отметить, что именно с натурных моделей судов в середине XIX века моделирование стало развиваться как научная дисциплина, а сами модели – активно использоваться при проектировании новых технических устройств. Середина XIX века связана в судостроении с окончанием эпохи парусных судов и началом эпохи парового флота. Оказалось, что использование паровых машин требует принципиального изменения конструкции судов. В первую очередь это осознали строители военных кораблей. Как известно, в условиях морского сражения время жизни судна зависит главным образом от его маневренности и скорости. Для парусных судов в результате многовекового опыта были выработаны оптимальные сочетания формы корпуса и парусов. Для кораблей с паровой машиной скорость определяется мощностью последней. В тот период тепло для машин получали от сжигания угля в топках котлов. Поэтому чем выше требуемая мощность машины, тем большее количество котлов необходимо использовать и иметь на судне большой запас угля. Все это утяжеляло судно и снижало его скорость, сводя к нулевому эффекту увеличение мощности машины. Учитывая, что строительство одного крейсера занимало несколько лет, а его стоимость была весьма значительной, можно понять стремление судостроителей найти более быстрый и дешевый (по сравнению с традиционным методом проб и ошибок) способ поиска оптимальных параметров судна. Выход был най-

ден в моделировании. Протягивая в бассейнах небольшие модели будущих судов и измеряя силу сопротивления, были найдены рациональные решения, как по форме корпуса судна, так и по мощности силовой установки.

В настоящее время методы натурального моделирования находят самое широкое применение в судостроении, авиастроении, автомобилестроении, ракетостроении и других областях. Например, при разработке нового самолета большое значение имеют эксперименты с натурными моделями, испытываемыми в аэродинамической трубе. Проведенные исследования позволяют изучить особенности обтекания фюзеляжа воздушными потоками, найти наиболее рациональную форму корпуса и отдельных узлов. Натурные модели используют и при исследовании причин крупных аварий и катастроф. Активно используются натурные модели в сочетании с другими методами моделирования (например, компьютерного) при съемке кинофильмов. Так, на съемках американского фильма “Титаник” для сцен гибели корабля было использовано более десяти моделей судна.

*Аналоговое моделирование* – это моделирование, основанное на аналогии процессов и явлений, имеющих различную физическую природу, но одинаково описываемых формально (одними и теми же математическими соотношениями, логическими и структурными схемами).

В основу аналогового моделирования положено совпадение математического описания различных (в большинстве случаев – качественно) объектов. В качестве примеров аналоговых моделей можно привести электрические и механические колебания, которые с точки зрения математики описываются одинаковыми соотношениями, но относящиеся к качественно отличающимся физическим процессам. Поэтому изучение механических колебаний можно вести с помощью электрической схемы, и наоборот. При некоторых допущениях аналогичными можно считать процессы распространения тепла в теле, диффузии примесей и просачивания жидкости.

К числу интересных примеров можно отнести известную в теории упругости аналогию Прандтля. Прандтлем было показано, что уравнения для функции напряжений в задаче о кручении

стержня произвольного поперечного сечения идентичны уравнениям, определяющим прогиб нерастяжимой мембраны, натянутой на упругий контур той же формы, под действием равномерного давления. Это позволяет заменить отнюдь не простые эксперименты по определению напряжений в скручиваемом стержне простыми измерениями прогибов мембраны.

Модели физического и аналогового типа являются материальным отражением реального объекта и тесно связаны с ним своими геометрическими, физическими и прочими характеристиками. Фактически процесс исследования моделей данного типа сводится к проведению ряда натуральных экспериментов, где вместо реального объекта используется его физическая или аналоговая модель.

### ***Идеальное моделирование***

Идеальное моделирование разделяют на два основных типа: интуитивное и научное.

Интуитивное моделирование – это моделирование, основанное на интуитивном (не обоснованном с позиций формальной логики) представлении об объекте исследования, не поддающимся формализации или не нуждающимся в ней.

В качестве наиболее яркого примера интуитивной модели окружающего мира можно считать жизненный опыт любого человека. К данному типу моделирования можно отнести умения и знания, накопленные многовековым опытом и передающиеся от поколения к поколению (например, умение лечить болезни с использованием трав и приемов народной медицины). Любое эмпирическое (т.е. полученное из эксперимента или в процессе наблюдения) знание без объяснения причин и механизмов наблюдаемого явления также следует считать интуитивным.

Следует подчеркнуть чрезвычайно важную роль интуиции, интуитивных моделей в науке, без них не обходится не одно сколько-нибудь новое знание. Последнее недостижимо только методами формальной логики. В связи с этим интересным представляется следующее высказывание А. де Сент-Экзюпери: *“Теоретик верит в логику. Ему кажется, будто он презирает мечту, интуицию, поэзию. Он не замечает, что они, эти три феи, про-*

*сто переделись, чтобы обольстить его, как влюбчивого мальчишку.... Они являются ему под именем “рабочих гипотез”, “произвольных гипотез”, “аналогий”, и может ли теоретик подозревать, что, слушая их, он изменяет суровой логике и внемлет напевам муз”.*

*Научное моделирование* – это всегда логически обоснованное моделирование, использующее минимальное число предположений, принятых в качестве гипотез на основании наблюдений за объектом моделирования.

Главное отличие научного моделирования от интуитивного заключается не только в умении выполнять необходимые операции и действия по собственно моделированию, но и в знании “внутренних” механизмов, которые используются при этом. Можно сказать, что научное моделирование не только знает, как необходимо моделировать, но и знает, почему так нужно делать.

Следует заметить, что в основе любого логического рассуждения лежат гипотезы или аксиомы, принимаемые на веру и не противоречащие имеющемуся опыту или эксперименту. Поэтому можно говорить об интуитивной первооснове любого научного знания. В результате познания человек разбирается в причинах тех или иных явлений, отодвигая границы своих интуитивных представлений об окружающем мире. Учитывая бесконечность данного процесса, можно предположить, что интуитивная первооснова у любого научного знания будет присутствовать всегда. О мере интуитивности знания можно судить по количеству использованных гипотез и аксиом. В этом смысле деление моделирования на интуитивное и научное следует признать относительным.

*Знаковым* называют моделирование, использующее в качестве моделей знаковые изображения какого-либо вида: схемы, графики, чертежи, иероглифы, руны, наборы символов, включающее также совокупность законов и правил, по которым можно оперировать с выбранными знаковыми образованиями и элементами. В качестве примеров таких моделей можно назвать любой язык. Например, язык устного и письменного человеческого общения, алгоритмический язык, язык химических формул, язык живописи и язык нот для записи музыки. Моделирование с помощью мате-

матических соотношений также является примером знакового моделирования.

Говоря о научном моделировании, следует пояснить смысл таких терминов, как “модель” и “теория”. В современной научной литературе данные понятия трактуются неоднозначно, а граница между ними размыта. В методологии науки признана в настоящее время следующая трактовка:

*Модель* – это инструмент, ориентированный в первую очередь на исследование поведения и свойств конкретного объекта в целях управления этим объектом или предсказания его свойств.

*Теория* – более абстрактное, чем модель, средство, основной целью которого является объяснение поведения или свойств не конкретного объекта, а некоторого класса объектов. Можно сказать, что теория содержит конечную или даже бесконечную совокупность конкретных моделей.

Например, для описания течения ньютоновской или нелинейно-вязкой жидкости в канале заданной формы разрабатывается соответствующая модель. При создании конкретной модели используются законы и уравнения соответствующей теории, в данном случае – механики жидкостей. Можно сказать, что модель дает ответы на вопросы: “Каким образом?” и “Почему?” для конкретного объекта, а теория – для целого семейства объектов, обладающих схожими свойствами. Следует, однако, отметить, что при разработке моделей сложных процессов и явлений зачастую приходится использовать понятия и соотношения нескольких теорий, относящихся к различным разделам, дисциплинам и даже областям знаний.

Интуитивное и научное (теоретическое) моделирование ни в коей мере нельзя противопоставлять друг другу. Они хорошо дополняют друг друга, разделяя области своего применения. Привычная буквенно-цифровая (знаковая) форма представления научного знания исторически сложилась как технология “передела” готового знания. Создание же качественно нового знания, рождение принципиально новых научных идей не может быть сведено к процессу чистой дедукции, к процессу формально-логического вывода следствий из множества уже открытых, готовых фактов, гипотез и теорий, составляющих информационную базу данных

сегодняшней науки. *“Подлинной ценностью, – говорил А. Эйнштейн, – является, в сущности, только интуиция. Для меня не подлежит сомнению, что наше мышление протекает, в основном минуя символы (слова) и к тому же бессознательно”*. Не менее определенно в пользу внелогического характера научного творчества высказывался А. Пуанкаре: *“Чистая логика всегда привела бы нас только к тавтологии; она не могла бы создать ничего нового; сама по себе она не может дать начало никакой науке... Для того, чтобы создать арифметику, как и для того, чтобы создать геометрию или какую бы то ни было науку, нужно нечто другое, чем чистая логика. Для обозначения этого другого у нас нет иного слова, кроме слова “интуиция”*.

Даже в самой абстрактной области фундаментальной науки – математике – интуиция играет определяющую роль: *“Вы должны догадаться о математической теореме прежде, чем вы ее докажете; вы должны догадаться об идее доказательства, прежде чем вы его проведете в деталях  $\frac{1}{4}$ ; доказательство открывается  $\frac{1}{4}$  с помощью догадки”*. Сходные мысли высказывает Ф.Клейн: *“Извечный секрет необычайной продуктивности гения – в его умении находить новые постановки задач, интуитивно предугадывать теоремы, приводящие к новым значительным результатам и к установлению важных зависимостей. Не будь новых концепций, новых идей, математика с присущей ей строгостью логических выводов вскоре исчерпала бы себя и пришла в упадок, ибо весь материал оказался бы израсходованным. В этом смысле можно сказать, что математику движут вперед в основном те, кто отмечен даром интуиции, а не строгого доказательства”*.

Итак, интуитивное знание является генератором нового знания. Однако далеко не все догадки и идеи выдерживают последующую проверку экспериментом и методами формальной логики, свойственными научному подходу, выступающему в виде своеобразного фильтра для выделения наиболее ценных знаний.

### ***Когнитивные, концептуальные и формальные модели***

Как уже отмечалось выше, при наблюдении за объектом-оригиналом в голове исследователя формируется некий мыслен-

ный образ объекта, его идеальная модель, которую в научной литературе принято называть *когнитивной* (мысленной, способствующей познанию) моделью. Формируя такую модель, исследователь, как правило, стремится ответить на определенные конкретные вопросы, поэтому от бесконечно сложного устройства объекта отсекается все ненужное с целью получения его более компактного и лаконичного описания.

Представление когнитивной модели на естественном языке называется *содержательной* моделью.

Когнитивные модели субъективны, так как формируются умозрительно (“в голове” исследователя) на основе всех его предыдущих знаний и опыта. Получить представление о когнитивной модели, можно только описав ее в знаковой форме. Нельзя утверждать, что когнитивные и содержательные модели эквивалентны, так как когнитивная модель может содержать элементы, которые исследователь не может или не хочет сформулировать. С другой стороны, если содержательная модель сформулирована кем-то другим или является продуктом коллективного творчества, то ее интерпретация, уровень понимания, степень доверия могут существенно изменяться от человека к человеку. В естественно – научных дисциплинах и в технике содержательную модель часто называют концептуальной или технической постановкой проблемы.

По функциональному признаку и целям содержательные модели подразделяются на описательные, объяснительные и прогностические.

*Описательной* моделью можно назвать любое описание объекта.

*Объяснительные* модели позволяют ответить на вопрос, почему что-либо происходит.

Наконец, *прогностические* модели должны описывать будущее поведение объекта. Можно заметить, что прогностическая модель не обязана включать в себя объяснительную.

*Концептуальной* моделью принято называть содержательную модель, при формулировке которой используются понятия и представления предметных областей знаний, занимающихся изучением объекта моделирования.

В более широком смысле под концептуальной моделью понимают содержательную модель, базирующуюся на определенной концепции или точке зрения. Выделяют три вида концептуальных моделей: логико-семантические, структурно-функциональные и причинно-следственные.

*Логико-семантическая* модель является описанием объекта в терминах и определениях соответствующих предметных областей знаний, включающим все известные логически непротиворечивые утверждения и факты. Анализ таких моделей осуществляется средствами логики с привлечением знаний, накопленных в соответствующих предметных областях.

При построении *структурно-функциональных* моделей объект обычно рассматривается как целостная система, которую расчленяют на отдельные элементы или подсистемы. Части системы связываются структурными отношениями, описывающие подчиненность, логическую и временную последовательность решения отдельных задач. Для представления подобных моделей удобны различного рода схемы, карты и диаграммы.

*Причинно-следственные* модели часто используют для объяснения и прогнозирования поведения объекта. Данные модели ориентированы в основном на описание динамики исследуемых процессов, при этом время далеко не всегда учитывается в явном виде.

*Формальная модель* является представлением концептуальной модели с помощью одного или нескольких формальных языков (например, языков математических теорий или алгоритмических языков).

В гуманитарных науках процесс моделирования обычно заканчивается созданием концептуальной модели объекта. В естественно-научных дисциплинах, как правило, удается построить формальную модель. Таким образом, когнитивные, содержательные и формальные модели составляют три взаимосвязанных уровня моделирования.

Перечисленные выше разновидности моделей нельзя рассматривать изолированно друг от друга. “Взаимоотношения” моделей между собой представлены на рисунок 1.2.



Рисунок 1.2 – Взаимоотношения моделей между собой

Взаимовлияние уровней моделирования друг на друга связано со свойством потенциальности моделей. Создание любой модели сопряжено с появлением новых знаний об исследуемом объекте, что приводит к переоценке и уточнению концепций и взглядов на объект моделирования. Данное обстоятельство приводит, в свою очередь, к необходимости пересмотра соответствующих содержательных и когнитивных моделей, реализуя спиральное развитие всех уровней моделирования исследуемого объекта.

Следует обратить внимание на тот факт, что если значение содержательных и формальных моделей для процесса познания более или менее понятно, то роль когнитивных моделей часто недооценивается. Это связано с субъективностью этих моделей и скрытостью процесса мышления. Однако существуют объекты и процессы, для которых роль когнитивных моделей особенно велика. Например, оператор или лицо, принимающее решения, осуществляет управление объектом или процессом главным образом на основании собственных когнитивных моделей. Велика роль данного типа моделей и в социальных науках. В настоящее время изучением свойств и особенностей когнитивных моделей

занимается новая, быстро развивающаяся дисциплина – *когнитология*.

Как уже отмечалось, одним из видов знакового моделирования является математическое моделирование.

*Математическое моделирование* – это идеальное научное знаковое формальное моделирование, при котором описание объекта осуществляется на языке математики, а исследование модели проводится с использованием тех или иных математических методов.

В качестве примера математического моделирования можно назвать классическую механику точки И. Ньютона, с помощью которой можно описать движение любого материального объекта, размеры которого малы по сравнению с характерными расстояниями. Фактически все современные разделы физики посвящены построению и исследованию математических моделей различных физических объектов и явлений. Так, физики-“ядерщики” до проведения экспериментальных исследований выполняют серьезные исследования с применением математических моделей. При этом на основании результатов теоретического моделирования разрабатывается и уточняется методика натуральных экспериментов, выясняется, какие эффекты, где и когда следует ожидать, когда и что регистрировать. Такой подход позволяет значительно снизить затраты на проведение эксперимента, повысить его эффективность. Аналогичные замечания можно сделать о других современных дисциплинах.

Как правило, значительные успехи в биологии и химии в последнее время были связаны с разработкой и исследованием математических моделей для биологических систем и химических процессов. В настоящее время широким фронтом идут работы по созданию математических моделей в экологии, экономике и социологии. Нельзя переоценить использование математических моделей в медицине и промышленности. Появилась возможность на научной (т.е. логически обоснованной) основе подходить ко многим экологическим и медицинским проблемам: имплантации и замене различных органов, прогнозированию развития эпидемий, обоснованной разработке планов ликвидации последствий крупных аварий и катастроф. Очень часто методы математиче-

ского моделирования являются единственно возможными. Например, всестороннее математическое моделирование и “проигрывание” различных вариантов на ЭВМ позволило в кратчайшие сроки (1-2 недели) обоснованно спланировать и приступить к реализации плана ликвидаций последствий Чернобыльской катастрофы. Уникальные результаты были получены по проекту “Гея”, связанному с математическим моделированием последствий ядерной войны. Было выяснено, что в результате сильного запыления атмосферы возможно значительное глобальное похолодание (“ядерная зима”) и связанное с этим вымирание практически всего живого. Подобных примеров эффективного использования математических моделей можно приводить очень много. В настоящее время это один из результативных и наиболее часто применяемых методов научного исследования.

Приведенное выше определение математического моделирования нельзя считать единственно возможным. Учитывая, что сама математика подразделяется (конечно, с большей степенью условности) на “чистую” и “прикладную”, можно привести несколько вариантов подобного определения.

Пример определения математической модели, характерного для “чистой” математики, можно найти в известном справочнике Т.Корн и Г.Корн, где под математической моделью понимается *“класс абстрактных и символьных математических объектов, таких, как числа или векторы, и отношения между ними”*. Под математическим отношением понимается *“гипотетическое правило, связывающее два или более символических объекта”*. Вводится абстрактное и конструктивное определение математической модели. При абстрактном определении новая модель задается *“непротиворечивым набором правил (определяющих аксиом), вводящих операции, которыми можно пользоваться, и устанавливающих общие отношения между их результатами. Конструктивное определение вводит новую математическую модель, пользуясь уже известными математическими понятиями (например, определение сложения и умножения матриц в терминах сложения и умножения чисел)”*.

С одной стороны, недостатком приведенного определения “чистых” математиков можно считать излишний формализм,

оперирование абстрактными понятиями. С другой стороны, абстрактность понятий повышает общность определения модели, делает его применимым к моделям самых разных по природе, не похожих друг на друга объектов и явлений. Кроме того, несомненным достоинством этого определения является его строгость и логичность. Для “прикладных” математиков характерна меньшая оторванность от реальной жизни, когда математические соотношения связывают не просто абстрактные математические объекты, а вполне определенные параметры реальных физических, химических, биологических или социальных явлений или процессов. Желательно иметь определение математической модели, сохраняющее строгость и логичность определения “чистой” математики и, в то же время, пригодное для классификации существующих и создаваемых моделей, сравнения их между собой. Вариант подобного определения можно сформулировать из следующих соображений.

Любая математическая модель, предназначенная для научных исследований, позволяет по заданным исходным данным определить значения интересующих исследователя параметров моделируемого объекта или явления. Поэтому можно предположить, что суть любой подобной модели заключается в отображении некоторого заданного множества  $\Omega_X$  значений “входных” параметров  $X$  на множество значений  $\Omega_Y$  “выходных” параметров  $Y$ . Данное обстоятельство позволяет рассматривать математическую модель как некоторый математический оператор  $A$  и сформулировать следующее определение.

Под математической моделью будем понимать любой оператор  $A$ , позволяющий по соответствующим значениям входных параметров  $X$  установить выходные значения параметров  $Y$  объекта моделирования:

$$A: X \rightarrow Y, X \text{ из } \Omega_X, Y \text{ из } \Omega_Y,$$

где  $\Omega_X$  и  $\Omega_Y$  – множества допустимых значений входных и выходных параметров для моделируемого объекта. В зависимости от природы моделируемого объекта элементами множеств  $\Omega_X$  и  $\Omega_Y$  могут являться любые математические объекты (числа, векторы, тензоры, функции, множества и т.п.).

Понятие оператора в приведенном определении может трактоваться достаточно широко. Это может быть как некоторая функция, связывающая входные и выходные значения, так и отображение, представляющее символическую запись системы алгебраических, дифференциальных, интегро-дифференциальных или интегральных уравнений. Наконец, это может быть некоторый алгоритм, совокупность правил или таблиц, обеспечивающих определение выходных параметров по заданным исходным значениям.

Определение математической модели через понятие оператора является более конструктивным с точки зрения построения классификации таких моделей. С другой стороны, подобное определение включает все многообразие имеющихся в настоящее время математических моделей.

С развитием вычислительной техники большое распространение получили *информационные модели*, представляющие по существу автоматизированные справочники, реализованные с помощью систем управления базами данных. Получая на входе некоторый запрос на поиск требуемой информации, подобные модели позволяют найти всю имеющуюся в базе данных информацию по интересующему вопросу. Однако данные модели не могут генерировать новое знание, отсутствующее в базе данных. Можно сказать, что это модели с нулевым потенциалом. Так как в качестве входных или выходных параметров моделей могут выступать математические объекты, а саму процедуру поиска данных можно представить в виде некоторого оператора, то информационные модели можно считать специфической разновидностью математических моделей. В данном пособии информационные модели и особенности их построения рассматриваться не будут.

### 1.3 Классификация математических моделей

Черная королева покачала головой:  
"Вы, конечно, можете назвать это  
чушью, но я-то встречала чушь та-

кую, что в сравнении с ней эта кажется толковым словарем"

Л.Кэрролл

Бурное развитие методов математического моделирования и многообразие областей их использования привело к появлению огромного количества моделей самого разного типа. В связи с этим возникает необходимость в определенном упорядочивании, классификации существующих и появляющихся математических моделей. Учитывая большое число возможных классификационных признаков и субъективность их выбора, появление все новых классов моделей, следует отметить условность и незавершенность рассматриваемой ниже классификации.

Представляется возможным подразделить математические модели на различные классы в зависимости:

- от сложности объекта моделирования;
- от оператора модели (подмодели);
- от входных и выходных параметров;
- от способа исследования модели;
- от цели моделирования.

### ***1.3.1 Классификация в зависимости от сложности объекта моделирования***

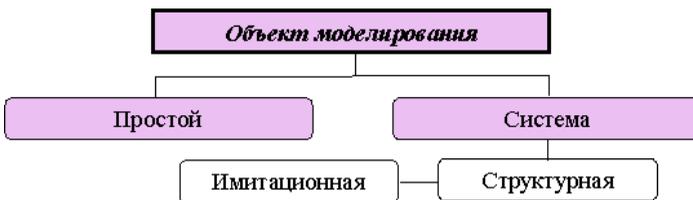


Рисунок 1.3 – Объекты моделирования

В качестве объекта моделирования может выступать как некоторое материальное тело или конструкция, так и природный, технологический или социальный процесс или явление. Все объекты моделирования можно разделить на две группы: простые и

объекты-системы (см. рисунок 1.3). В первом случае при моделировании не рассматривается внутреннее строение объекта, не выделяются составляющие его элементы или подпроцессы. В качестве примера подобного объекта можно привести материальную точку в классической механике.

Система есть совокупность взаимосвязанных элементов, в определенном смысле обособленная от окружающей среды и взаимодействующая с ней как целое.

Для сложных систем характерно наличие большого числа взаимно связанных, взаимодействующих между собой элементов. При этом связь между элементами  $A$  и  $B$  системы может отличаться от связи между элементами  $B$  и  $A$ . Если система имеет  $N$  элементов и каждый элемент связан с каждым, то общее число связей равно  $N(N-1)$ . Если все  $N$  элементов имеют по  $M$  состояний, то общее число состояний  $S$  для такой системы равно  $M^N$ . Например, для системы при  $M=2$  и  $N=3$  имеем  $S = 2^3 = 8$  состояний.

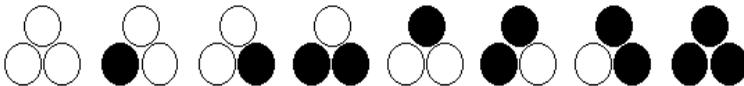


Рисунок 1.4 – Возможные состояния системы из 3-х элементов, имеющих по 2 состояния

Максимальное число связей в подобной системе равно 6. Если поведение системы описывается процессом перехода из одного состояния в другое, то общее число возможных переходов равно  $S^2$ . Для рассматриваемого примера число сценариев возможного поведения системы равно  $S = 8^2 = 64$ . Поведение системы быстро усложняется с ростом числа элементов системы. Так, для системы из 10 элементов при  $M=2$  число состояний  $S$  равно 1024, а число сценариев 1 048 576. Данное обстоятельство, с одной стороны, говорит о сложности систем и многовариантности их поведения. С другой стороны, следует ожидать наличия больших трудностей, возникающих при изучении и моделировании систем.

Конечно, деление объектов исследования на “простые” и “сложные” условно. Поскольку для любых известных процессов,

явлений, материальных тел невозможно выделить их “элементарные кирпичики”, “атомы”, то любой объект исследования можно считать бесконечно сложным. “Упрощение” его строения при разработке модели выполняется в результате отбрасывания малозначимых, несущественных для достижения поставленных на данный момент целей исследования связей между составляющими объект элементами. При изменении целей исследования или повышении требований к точности моделирования приходится, как правило, пересматривать уровень детализации объекта.

Модели объектов-систем, учитывающие свойства и поведение отдельных элементов, а также взаимосвязи между ними, называются *структурными*. Структурные динамические модели выделяют в отдельный класс *имитационных* систем; при этом рассматриваются системы, состоящие из конечного числа элементов, каждый из которых имеет конечное число состояний. Число связей между элементами также предполагается конечным. Моделирование взаимодействий элементов системы друг с другом осуществляется с помощью некоторого алгоритма, реализуемого обычно с использованием ЭВМ. Для моделирования на ЭВМ реального времени вводится понятие *системного времени*. В качестве моделей отдельных элементов могут быть использованы модели любого типа.

Как правило, взаимодействие внешней среды со сложной системой полностью проследить не удастся, что приводит к неопределенности внешних воздействий и, как следствие, неоднозначности в поведении самой системы. Наличие подобной неопределенности является характерной особенностью сложных систем, учитываемой при моделировании, например, с использованием генераторов случайных чисел.

### *1.3.2 Классификация в зависимости от оператора модели*

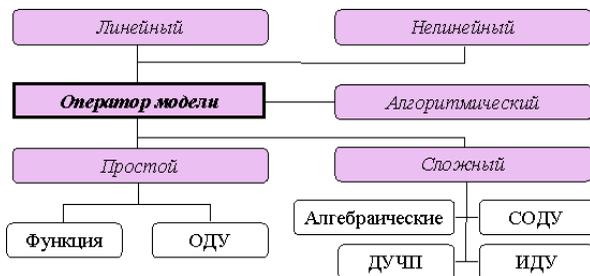


Рисунок 1.5 – Классификация по оператору модели

В зависимости от оператора математические модели можно разделить как на линейные и нелинейные, так и в соответствии с конкретным видом оператора.

Любая математическая модель, как отмечено выше, может рассматриваться как некоторый оператор  $A$ , который является алгоритмом или определяется совокупностью уравнений (алгебраических, обыкновенных дифференциальных уравнений (ОДУ), систем ОДУ (СОДУ), дифференциальных уравнений в частных производных (ДУЧП), интегро-дифференциальных уравнений (ИДУ) и других).

Если оператор обеспечивает линейную зависимость “выходных” параметров  $Y$  от значений “входных” параметров  $X$ , то математическая модель называется *линейной*. Линейные модели более просты для анализа. Например, из свойства линейности следует свойство суперпозиции решений, то есть если известны решения  $Y_1$  при  $X_1$  и  $Y_2$  при  $X_2$ , то решение для “выходных” параметров при  $X=X_1+X_2$  есть  $Y=Y_1+Y_2$ . Предельные значения  $Y$  для линейных моделей достигаются, как правило, на границах областей  $\Omega_X$  допустимых значений “входных” параметров.

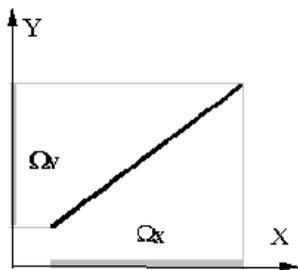


Рисунок 1.6 – Линейная математическая модель

Исторически первыми стали разрабатываться и исследоваться именно линейные математические модели. Область применения подобных моделей очень широка. Она охватывает классическую механику, электродинамику, аналитическую химию и биологию. Методы их решения, разрабатывавшиеся в течение столетий, обладают большой общностью и эффективностью.

Линейное поведение свойственно относительно “простым” объектам. Системам, как правило, присуще нелинейное многовариантное поведение. В настоящее время все чаще возникает потребность не только в повышении точности моделирования, но и создании качественно новых моделей, учитывающих нелинейность поведения реальных объектов исследования. Анализ подобных моделей намного сложнее, чем линейных, причем создание методики и общих подходов к их исследованию в настоящее время далеко от завершения. Являясь более богатым и сложным, мир нелинейных моделей представляется для современной науки более перспективным в плане открытия новых закономерностей и описания сложных явлений. Например, такие явления как солитоны и хаос нельзя с достаточной степенью адекватности описать в рамках традиционных линейных моделей. Методы исследования нелинейных моделей в настоящее время быстро прогрессируют, складываясь в новые научные направления. К таким относительно новым направлениям можно отнести, например, синергетику – науку о сложных самоорганизующихся системах.

В зависимости от вида оператора математические модели можно разделить на *простые* и *сложные*.

В случае, когда оператор модели является алгебраическим выражением, отражающим функциональную зависимость выходных параметров  $Y$  от входных  $X$ , модель будем называть *простой моделью*.

Простые модели чаще всего являются результатом обобщения и анализа экспериментальных данных, полученных в результате наблюдений за исследуемым объектом или явлением. На основании анализа таких данных выдвигается гипотеза о возможной функциональной связи входных и выходных параметров. После этого данная гипотеза проверяется на имеющемся экспериментальном материале, уточняется степень ее адекватности (т.е. степень соответствия результатов моделирования, полученных с применением данной гипотезы, имеющимся знаниям об исследуемом объекте). Если результаты проверки неудовлетворительны, то принятая гипотеза отвергается и заменяется новой. Процесс повторяется до получения желаемой степени соответствия результатов эксперимента и модели. В качестве примеров простых моделей можно привести многие законы физики (закон всемирного тяготения, закон Ома, закон Гука, закон трения Амонтона-Кулона), а также все эмпирические (т.е. полученные из опыта) алгебраические зависимости между входными и выходными параметрами. Так, в теории резания металлов очень часто используются соотношения, связывающие время и стоимость обработки детали на станке в зависимости от скорости ее вращения (скорости резания) и скорости осевого перемещения (скорости подачи) детали.

Модель, включающая системы дифференциальных и интегральных соотношений, уже не может быть отнесена к простым, так как для своего исследования требует применения довольно сложных математических методов. Однако в двух случаях она может быть сведена к простым:

1. если полученная для подобной модели система математических соотношений может быть разрешена аналитически, то результат решения может рассматриваться как простая модель;
2. если результаты вычислительных экспериментов со сложной моделью аппроксимированы некоторой алгебраической

зависимостью. В настоящее время известно достаточно большое количество подходов и методов аппроксимации (например, метод наименьших квадратов или метод планирования экспериментов).

На практике довольно часто возникают ситуации, когда удовлетворительное описание свойств и поведения объекта моделирования (как правило, сложной системы) не удается выполнить с помощью математических соотношений. Однако в большинстве случаев можно построить некоторый имитатор поведения и свойств такого объекта с помощью алгоритма, который можно считать оператором модели.

Например, если в результате наблюдения за объектом получена таблица соответствия между “входными”  $X$  и “выходными”  $Y$  значениями параметров, то определить оператор  $A$ , позволяющий получить “выход” по заданному “входу”, зачастую бывает проще с помощью алгоритма.

### *1.3.3 Классификация в зависимости от параметров модели*

В общем случае параметры, описывающие состояние и поведение объекта моделирования, разбиваются на ряд непересекающихся подмножеств:

- совокупность входных (управляемых) воздействий на объект ( $\Omega_X$ );
- совокупность воздействий внешней среды (неуправляемых) ( $\Omega_E$ );
- совокупность внутренних (собственных) параметров объекта ( $\Omega_I$ );
- совокупность выходных характеристик ( $\Omega_Y$ ).



Рисунок 1.7 – Классификация по параметрам модели

Например, при моделировании движения твердого тела в атмосфере в поле сил тяжести в качестве входных параметров могут выступать начальное положение и начальная скорость точки, принятой за полюс, и угловая скорость в момент времени  $t=0$ . Сила сопротивления и сила тяжести характеризуют воздействие внешней среды. Масса тела и его форма являются собственными параметрами тела. Координаты и скорости точек тела (при  $t>0$ ) относятся к выходным величинам. В то же время отнесение параметров к тому или иному типу зависит от постановки конкретной задачи. Например, в приведенном примере можно переформулировать задачу, сделав ее обратной к исходной: определить начальное положение и скорости (линейную скорость полюса и угловую скорость тела) по заданному положению и скоростям в момент времени  $t_1>0$ . Понятно, что при этом входные и выходные параметры меняются местами.

Следует отметить, что количество параметров всех типов в математических моделях, как правило, конечно. При этом каждый из параметров может иметь различную “математическую природу”: быть постоянной величиной или функцией, скаляром или вектором (или тензором второго, третьего и выше рангов), четким или нечетким множеством и т.д. В математических моделях, рассматриваемых в естественных науках, наиболее распространенными являются параметры, являющиеся скалярно- или тензор – значными функциями. В качестве независимых переменных (аргументов) при этом обычно выступают координаты точек трехмерного пространства и/или время (или некоторый необязывающий параметр – аналог времени).

Входные параметры  $X$ , параметры, описывающие воздействие внешней среды  $E$ , и внутренние (собственные) характеристики  $I$  объекта относятся к *независимым (экзогенным)* величинам. Выходные параметры  $Y$  являются *зависимыми (эндогенными)* величинами.

В общем случае оператор модели  $A$  преобразует экзогенные параметры в эндогенные  $A: \{X, E, I\} \rightarrow Y$ .

Введение тех или иных количественных характеристик объекта моделирования возможно при наличии некоторого эталона сравнения. Например, для характеристики размеров тела используется эталонный образец – метр. По своей природе характеристики объекта могут быть как качественными, так и количественными. Для количественной характеристики вводятся числа, выражающие отношения между данным параметром и эталоном. Кроме того, количественные значения параметра могут выражаться дискретными или непрерывными величинами. Качественные характеристики находятся, например, с помощью метода экспертных оценок. В зависимости от вида используемых множеств параметров модели могут различаться на *качественные* и *количественные, дискретные и непрерывные, а также смешанные*.

При построении моделей реальных объектов и явлений очень часто приходится сталкиваться с недостатком информации. Как правило, для любого исследуемого объекта распределение свойств, воздействия и начальное состояние известны с той или

иной степенью неопределенности. Это связано с множеством трудно учитываемых факторов, с ограниченностью числа используемых параметров модели, с конечной точностью экспериментальных измерений. При построении модели описание неопределенности параметров может быть осуществлено следующими способами:

1. *детерминированное* – значения всех параметров модели определяются детерминированными величинами (т.е. каждому параметру соответствует конкретное целое, вещественное или комплексное число или соответствующая функция). Данный способ соответствует полной определенности параметров;
2. *стохастическое* – значения всех или отдельных параметров модели определяются случайными величинами, заданными плотностями вероятности. В литературе наиболее полно исследованы случаи нормального (гауссова) и показательного распределения случайных величин;
3. *случайное* – значения всех или отдельных параметров модели устанавливаются случайными величинами, заданными оценками плотностей вероятности, полученными в результате обработки ограниченной экспериментальной выборки данных параметров. Данная форма описания тесно связана с предыдущей. Однако в данном случае получаемые результаты моделирования будут существенным образом зависеть от точности оценок моментов и плотностей вероятности случайных параметров, от постулируемых законов распределения, объема выборок;
4. *интервальное* – значения всех или отдельных параметров модели описываются интервальными величинами, заданными интервалом, образованным минимальным и максимально возможными значениями параметра;
5. *нечеткое* – значения всех или отдельных параметров модели описываются функциями принадлежности соответствующему нечеткому множеству. Такая форма используется, когда информация о параметрах модели задается экспертом на естественном языке, а следовательно, в “нечетких” (с по-

зиции математики) терминах типа “много больше пяти”, “около нуля”.

В теории игр встречается еще один вид неопределенности параметров модели, называемый *игровой неопределенностью*. В данном пособии указанный вид неопределенности не рассматривается.

Совокупность значений параметров модели в некоторый момент времени или на данной стадии называется *состоянием объекта*.

Деление параметров на стационарные и нестационарные используется для моделей, где одним из независимых аргументов может являться время или неубывающий параметр, характеризующий “направление” процесса или последовательность его стадий.

Рассмотрим течение жидкости в длинной трубе постоянного сечения. Эксперименты показывают, что на достаточно большом удалении от входа в трубу частицы жидкости движутся параллельно оси трубы.

При этом если условия на входе не изменяются и скорость течения невелика, то профиль скоростей частиц в данном сечении трубы с течением времени остается неизменным. В этом случае можно говорить о *стационарном* движении жидкости в трубе, характеристики движения которой в каждой точке пространства не зависят от времени. Как правило, стационарные модели применяются для описания различных потоков (жидкости, газа, тепла) в случае постоянства условий на входе и выходе потока. Процессы, для которых состояние объекта в каждой фиксированной точке пространства не изменяется с течением времени, называются стационарными. Для таких процессов время может быть исключено из числа независимых переменных.

Если в качестве одной из существенных независимых переменных модели необходимо использовать время (или его аналог), то модель называется *нестационарной*. Примером нестационарной модели является движение жидкости в трубе, но вытекающей из некоторого сосуда. По мере понижения уровня жидкости в сосуде давление на входе в трубу будет уменьшаться, что приведет

к изменению параметров течения жидкости в самой трубе. Обычно такие модели значительно сложнее стационарных и требуют больших временных затрат для своей реализации на ЭВМ.

Очень часто параметры, характеризующие состояние объекта исследования, изменяются при переходе от одной точки объекта к другой, образуя пространственные поля значений. В заданной исследователем системе координат подобные полевые параметры являются функциями координат.

Разделение моделей на *одномерные*, *двумерные* и *трехмерные* применимо для моделей, в число параметров которых входят координаты пространства, и связано с особенностями реализации данных моделей, равно как и с увеличением их сложности при возрастании размерности (с “проклятием размерности” по образному выражению Р. Беллмана). Как правило, увеличение размерности модели приводит к увеличению числа используемых математических соотношений. Особенно сложны в реализации трехмерные модели, требующие высокопроизводительной вычислительной техники с большим объемом оперативной и дисковой памяти. Реализация таких моделей стала возможной лишь с появлением вычислительных машин третьего поколения и потребовала создания специальных вычислительных методов и приемов. Среди характерных вычислительных трудностей, с которыми сталкиваются при создании моделей в трехмерной постановке, можно отметить необходимость хранения и решения систем уравнений большой размерности (10 тысяч уравнений и более), проблема подготовки исходной информации и ее проверка, наглядное отображение полученных результатов. При разработке модели стараются (если это возможно) понизить размерность. Однако необоснованное понижение размерности модели может существенно исказить результаты моделирования. Например, если для исследования движения брошенного мяча в вертикальной плоскости использование двумерной модели может быть оправдано, то для исследования движения бумеранга двумерную модель строить бесполезно.

### 1.3.4 Классификация в зависимости от целей моделирования



Рисунок 1.8 – Классификация по цели моделирования

Целью *дескриптивных моделей* (от лат. *descriptio* – описание) является построение законов изменения параметров модели. В качестве примера такой модели можно привести модель движения материальной точки под действием приложенных сил, использующая второй закон Ньютона. Задавая положение и скорость точки в начальный момент времени (входные параметры), массу (собственный параметр) и закон изменения прикладываемых сил (внешние воздействия), можно определить скорость и координаты материальной точки в любой момент времени (выходные параметры). Полученная модель *описывает* зависимость выходных величин от значений входных параметров. Поэтому дескриптивные модели являются реализацией описательных и объяснительных содержательных моделей на формальном уровне моделирования.

В качестве другого примера дескриптивной модели можно привести модель движения ракеты после старта с поверхности земли. В качестве параметров модели в данном случае могут выступать начальное положение и начальная скорость ракеты (входные), ее начальная масса, импульс двигателя, режим его работы (собственные параметры), закон изменения сил притяжения и сил сопротивления атмосферы (внешние воздействия). В качестве выходных величин имеем положение и скорость центра масс ракеты и ее ориентацию в пространстве в произвольный момент времени.

*Оптимизационные модели* предназначены для определения оптимальных (наилучших) с точки зрения некоторого критерия параметров моделируемого объекта или же для поиска оптималь-

ного (наилучшего) режима управления некоторым процессом. Часть параметров модели относят к *параметрам управления*, изменяя которые можно получать различные варианты наборов значений выходных величин. Как правило, данные модели строятся с использованием одной или нескольких дескриптивных моделей и включают некоторый критерий, обеспечивающий сравнение различных вариантов наборов значений выходных величин между собой с целью выбора наилучшего. На область значений входных параметров могут быть наложены ограничения в виде равенств и неравенств, связанные с особенностями рассматриваемого объекта или процесса. Целью оптимизационных моделей является поиск таких допустимых параметров управления, при которых критерий выбора достигает своего “наилучшего значения”.

Примером оптимизационной модели может служить моделирование процесса запуска ракеты с поверхности земли с целью подъема ее на заданную высоту за минимальное время при ограничениях на величину импульса двигателя, время его работы, начальную и конечную массу ракеты. Математические соотношения дескриптивной модели движения ракеты выступают в данном случае в виде ограничений типа равенств.

*Управленческие модели* применяются для принятия эффективных управленческих решений в различных областях целенаправленной деятельности человека. Следует отметить, что принятие решений в общем случае является процессом, по своей сложности сравнимым с процессом мышления в целом. Однако на практике под “принятием решений” обычно понимается выбор некоторых альтернатив из заданного их множества, а общий процесс принятия решений представляется как последовательность таких выборов альтернатив. Например, на предприятии освободилось должностное место главного инженера, и задача директора состоит в выборе из имеющегося множества кандидатов на данную должность одного, отвечающего заданным требованиям. Сложность данной задачи заключается в наличии неопределенности как по исходной информации (неполные данные о кандидатах) и характеру воздействия внешних условий (случайное: выбранный кандидат заболел или отказался; игровое: министерство против вы-

бранной кандидатуры), так и по целям (противоречивые требования к выбираемой кандидатуре: должен быть хорошим специалистом и администратором, опытен, энергичен, молод и прочее). Поэтому в отличие от оптимизационных моделей, где критерий выбора считается определенным и искомое решение устанавливается из условий его экстремальности, в управленческих моделях необходимо введение специфических критериев оптимальности, которые позволяют проводить сравнение альтернатив при различных неопределенностях задачи. Поскольку оптимальность принятого решения даже в одной и той же ситуации может пониматься по-разному, вид критерия оптимальности в управленческих моделях заранее не фиксируется. Именно в этом состоит основная особенность данных моделей. Методы формирования критериев оптимальности в зависимости от вида неопределенности рассматриваются в теории выбора и принятия решений, которая базируется на теории игр и исследовании операций.

### 1.3.5 Классификация в зависимости от методов исследования

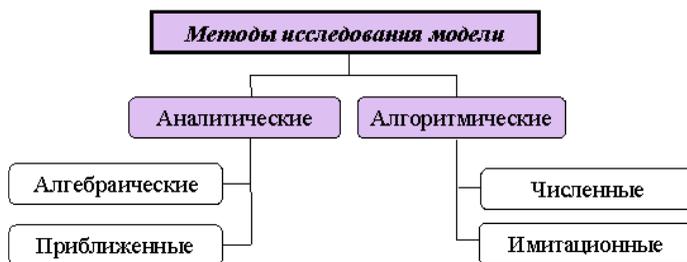


Рисунок 1.9 – Классификация по методам исследования

Метод исследования модели относят к *аналитическим*, если он позволяет получить выходные величины в виде аналитических выражений, т.е. выражений, в которых используется не более чем счетная совокупность арифметических операций и переходов к пределу по натуральным числам. Примеры *аналитических выражений*:

$$\sum_{k=0}^{\infty} \frac{a_k x^k}{x^k + 1}, \quad \lim_{n \rightarrow \infty} \left( 1 + \frac{x}{n} \right)^n.$$

Частным случаем аналитических выражений являются *алгебраические выражения*, в которых используется конечное или счетное число арифметических операций, операций возведения в целочисленную степень и извлечения корня. Примеры алгебраических выражений:

$$ax^2 + bx + c, \quad a + b\sqrt{x^3 + 4ac}.$$

На практике для оценки величины параметра моделирования приходится ограничивать количество членов совокупности арифметических операций в аналитических выражениях некоторым конечным числом. Поэтому величина оценки параметра в этом случае получается приближенной, а модели, использующие подобный прием, называются *приближенными*.

Аналитические методы исследования модели являются более ценными в том плане, что позволяет с меньшими вычислительными затратами изучить свойства объекта моделирования, применяя традиционные хорошо развитые математические методы анализа аналитических функций. Существенно, что применение аналитических методов возможно без использования ЭВМ. Кроме того, знание аналитического выражения для искомым параметров позволяет исследовать фундаментальные свойства объекта, его качественное поведение, строить новые гипотезы о его внутренней структуре. Следует отметить, что возможности аналитических методов существенно зависят от уровня развития соответствующих разделов математики.

К сожалению, существующие в настоящее время математические методы позволяют получить аналитические решения лишь для относительно несложных математических моделей в узком диапазоне значений параметров. В большинстве случаев при исследовании моделей приходится использовать алгоритмические подходы, позволяющие получить лишь приближенные значения искомым параметров.

При *численном подходе* совокупность математических соотношений модели заменяется конечномерным аналогом. Это чаще всего достигается дискретизацией исходных соотношений, то

есть переходом от функций непрерывного аргумента к функциям дискретного аргумента. После дискретизации исходной задачи выполняется построение вычислительного алгоритма, то есть последовательности арифметических и логических действий, выполняемых на ЭВМ и позволяющих за конечное число шагов получить решение дискретной задачи. Полученное решение дискретной задачи принимается за приближенное решение исходной математической задачи.

Степень приближения получаемых с помощью численного метода искомых параметров модели зависит как от *погрешностей самого метода*, связанных с заменой исходной модели ее дискретным аналогом, так и от *ошибок округления*, возникающих при выполнении любых расчетов на ЭВМ в связи с конечной точностью представления чисел в ее памяти. Основным требованием к вычислительному алгоритму является необходимость получения решения исходной задачи с заданной *точностью* за конечное число шагов.

К настоящему времени круг вопросов, связанных с разработкой и использованием численных методов, а также с построением на их основе вычислительных алгоритмов, выделился в самостоятельный быстро развивающийся и обширный раздел математики – *вычислительную математику*.

Если при численном подходе дискретизации подвергалась полученная система математических соотношений, то при *имитационном подходе* на отдельные элементы разбивается сам объект исследования. В этом случае система математических соотношений для объекта-системы в целом не записывается, а заменяется некоторым алгоритмом, моделирующим ее поведение и учитывающим взаимодействие друг с другом моделей отдельных элементов системы. В качестве моделей отдельных элементов могут быть использованы как аналитические, так и алгебраические модели.

Алгоритмические модели, использующие как численный, так и имитационный подход, не позволяют получить решения задач в аналитической форме, что затрудняет и усложняет процесс анализа результатов моделирования. Так как применение моделей данного типа возможно лишь при наличии вычислительной тех-

ники, то их эффективность зависит от мощности и быстродействия ЭВМ. Несомненным достоинством алгоритмических моделей является отсутствие принципиальных ограничений на сложность модели, что позволяет применять их для исследования систем произвольной сложности.

Использование математической модели, построенной с применением алгоритмических методов, аналогично проведению экспериментов с реальным объектом, только вместо реального эксперимента с объектом проводится *вычислительный эксперимент* с его моделью. Задаваясь конкретным набором значений исходных параметров модели, в результате вычислительного эксперимента получают конкретный набор приближенных значений искомых параметров. Для исследования поведения объекта при новом наборе исходных данных необходимо проведение нового вычислительного эксперимента.

## **ГЛАВА 2. ЭТАПЫ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ**

### ***Введение***

Процесс создания математических моделей трудоемок, длителен и связан с использованием труда различных специалистов достаточно высокого уровня, обладающих хорошей подготовкой как в предметной области, связанной с объектом моделирования, так и в области прикладной математики, современных численных методов, программирования, знающих возможности и особенности вычислительной техники. Отличительной особенностью математических моделей, создаваемых в настоящее время, является их комплексность, связанная со сложностью моделируемых объектов. Например, при моделировании процессов деформирования различных конструкций под действием приложенной нагрузки приходится учитывать не только происходящие при деформировании процессы массопереноса, но и теплоперенос, связанное с этим изменение структуры и свойств материала, влияние различных видов излучения, воздействие гравитационных и электромагнитных полей, предыстории деформирования. Кроме того, для современных моделей характерно представление объекта моде-

лирования в виде более или менее сложной системы взаимодействующих элементов. Все отмеченные выше особенности приводят к усложнению модели, к необходимости совместного использования множества теорий, применения современных вычислительных методов и вычислительной техники для получения и анализа результатов моделирования.

Наблюдаемое в настоящее время внедрение вычислительной техники во все сферы человеческой деятельности привело к повсеместному использованию математических моделей. Данное обстоятельство связано с тем, что ЭВМ – это только "железо", "умным" и полезным его делают программы, которые в большинстве случаев являются реализациями алгоритмов соответствующих математических моделей. Все это приводит к необходимости создания большого количества разнообразных математических моделей с широкими возможностями, удовлетворяющими различным, зачастую противоречивым, требованиям. Удовлетворить всем наложенным требованиям в одной модели в случае сложных объектов, как правило, невозможно. Поэтому приходится создавать целый спектр моделей одного и того же объекта (в некоторых случаях – иерархическую совокупность "вложенных" друг в друга моделей), каждая из которых наиболее эффективно решает возложенные на нее задачи моделирования. Так, модели, ориентированные на исследовательские цели, способны моделировать объект в широком диапазоне исходных параметров с удовлетворительной точностью. При этом практически нет ограничений на сложность подобной модели, а также на время, затрачиваемое на получение результатов. Исследовательские модели могут быть ориентированы как на получение количественных, так и качественных результатов. Модели, используемые в автоматизированных системах управления (АСУ), в отличие от исследовательских имеют достаточно жесткие ограничения на время, затрачиваемой на получение результатов, а также – на точность самих результатов.

Необходимость в массовом создании моделей требует разработки некоторой совокупности правил и подходов, которые позволили бы снизить затраты на создание модели и уменьшить вероятность появления трудно устранимых впоследствии оши-

бок. Подобную совокупность правил можно было бы назвать *технологией создания математических моделей*.

Процесс создания любой математической модели представляется возможным рассматривать как последовательность этапов, изображенных на рисунке 2.1.



Рисунок 2.1 – Этапы построения математической модели

## 2.1 Обследование объекта моделирования

Первоначальная формулировка проблемы является самой трудной частью, так как здесь необходимо все время использовать свои мысли, позднее взамен их может использоваться математика.

А.Эддингтон

Математические модели, особенно использующие численные методы и вычислительную технику, требуют для своего создания значительных интеллектуальных, финансовых и временных затрат. Поэтому решение о создании новой модели принима-

ется лишь в случае отсутствия иных, более простых путей решения возникших проблем (например, модификации одной из существующих моделей). Необходимость в новой модели может возникнуть в связи с проведением научных исследований (особенно – на стыке различных областей знания), выполнением проектных и конструкторских работ на производстве, созданием систем автоматического управления, планирования и контроля. Человека или организацию, заинтересованных в создании новой математической модели, для краткости будем называть *заказчиком*. После принятия решения о необходимости создания новой математической модели заказчик ищет *исполнителя* своего заказа. В качестве исполнителя, как правило, может выступать рабочая группа, включающая специалистов разного профиля: прикладных математиков, специалистов, хорошо знающих особенности объекта моделирования, программистов.

Итак, если решение о создании модели принято, и рабочая группа сформирована, то можно приступать к этапу обследования объекта моделирования. Основной целью данного этапа является подготовка содержательной постановки задачи моделирования.

Перечень сформулированных в содержательной (словесной) форме основных вопросов об объекте моделирования, интересующих заказчика, составляет *содержательную постановку задачи моделирования*.

Подготовка списка вопросов, на которые должна ответить новая модель, зачастую является самостоятельной проблемой, требующей для своего решения специалистов со специфическими знаниями и способностями. Они должны не только хорошо разбираться в предметной области моделирования, знать возможности современной вычислительной математики и техники, но и быть достаточно коммуникабельными, то есть уметь общаться с людьми. Подобных специалистов в настоящее время называют *постановщиками задач*. Источником информации для постановщика могут служить беседы с представителями заказчика, имеющаяся информация об объекте моделирования (в особенности – данные экспериментальных исследований), а также модели, разработанные ранее. На основании анализа всей собранной инфор-

мации постановщик задачи должен сформулировать такие требования к будущей модели, которые, с одной стороны, удовлетворяли бы заказчика; с другой стороны, реализация модели должна быть осуществлена в заданные сроки в рамках выделенных материальных средств. Специалисты – постановщики должны обладать способностью из большого объема слабо формализованной разнообразной информации об объекте моделирования, из различных, не четко высказанных и сформулированных пожеланий и требований заказчика к будущей модели, выделить то главное, что может быть действительно реализовано. Из перечисленных требований к постановщикам задач видно, насколько велика ответственность, возложенная на них, и насколько могут быть тяжелы ошибки и просчеты, допущенные ими. Неправильная оценка срока и стоимости реализации требуемой модели может привести к неудаче всего проекта, к напрасной потере времени и средств. Специалисты, предрасположенные к работе в качестве постановщиков задач, особенно ценятся и являются, без преувеличения, золотым фондом научных коллективов. По этому поводу Г. Биркгоф отмечает, что прикладники-математики, *"способные к глубокому общению с другими учеными и инженерами и знакомые с мощью и ограничениями цифровых машин, ... призваны стать вождями завтрашнего математического мира, но их будет крайне трудно найти и развить"*. С учетом данного высказывания, а также имея в виду конечную цель деятельности рабочей группы – построение математической модели, – представляется целесообразным рекомендовать в качестве руководителя группы именно прикладника-математика.

Этап обследования проводится членами рабочей группы под руководством постановщиков задач и включает следующие работы:

- тщательное обследование собственно объекта моделирования с целью выявления основных факторов, механизмов, определяющих его поведение, определения соответствующих параметров, позволяющих описывать моделируемый объект;

- сбор и проверка имеющихся экспериментальных данных об объектах-аналогах, проведение при необходимости дополнительных экспериментов;
- аналитический обзор литературных источников, анализ и сравнение между собой построенных ранее моделей данного объекта (или подобных рассматриваемому объекту);
- анализ и обобщение всего накопленного материала, разработка общего плана создания математической модели.

На основе собранной информации об объекте моделирования постановщики совместно с заказчиком формулируют содержательную постановку задачи моделирования, которая, как правило, не бывает окончательной и может уточняться и конкретизироваться в процессе разработки модели. Однако, с учетом изложенного выше, все последующие уточнения и изменения содержательной постановки должны носить частный, не принципиальный характер. Если объектом моделирования является технологический процесс, машина, конструкция или деталь, то содержательную постановку задачи моделирования очень часто называют технической постановкой задачи.

Весь собранный в результате обследования материал о накопленных к данному моменту знаниях об объекте, содержательная постановка задачи моделирования, дополнительные требования к реализации модели и представлению результатов оформляются в виде *технического задания на проектирование и разработку модели*.

Техническое задание является итоговым документом, заканчивающим этап обследования. Как уже отмечалось, данный этап является очень важным и ответственным. Чем более полную информацию удастся собрать об объекте на этапе обследования, тем более четко можно выполнить содержательную постановку задачи, более полно учесть накопленный опыт и знания, избежать многих сложностей на последующих этапах разработки модели. Особенно строго необходимо формулировать требования к будущей модели. Неконкретные и нечеткие требования могут серьезно затруднить процесс сдачи модели заказчику, вызвать бесконечные доработки и улучшения. В целом этап проработки технического задания может составлять до 30% времени, отпущенного

на создание всей модели, а с учетом возможного уточнения и переформулировки – и более.

Учитывая огромную важность рассматриваемого этапа, техническое задание следует подвергать внутренней (внутри организации) и внешней экспертизе независимыми экспертами, не участвующими в его разработке. Обязательным условием на этапе разработки технического задания является участие в его обсуждении всех членов рабочей группы.

### **Пример**

*Содержательная постановка задачи о баскетболисте*

Разработать математическую модель, позволяющую описать полет баскетбольного мяча, брошенного игроком в баскетбольную корзину.

Модель должна позволять:

- вычислять положение мяча в любой момент времени;
- определять точность попадания мяча в корзину после броска при различных начальных параметрах.

Исходные данные:

- масса и радиус мяча;
- начальные координаты, начальная скорость и угол броска мяча;
- координаты центра и радиус корзины.

## **2.2 Концептуальная постановка задачи моделирования**

Наилучшие гипотезы – это простые гипотезы, которые легко подтвердить экспериментально, если они верны, и легко опровергнуть с помощью надлежащим образом подобранных решающих экспериментов или наблюдений, если они неверны.

Н.Бейли

В отличие от содержательной концептуальная постановка задачи моделирования, как правило, формулируется членами ра-

бочей группы без привлечения представителей заказчика, на основании разработанного на предыдущем этапе технического задания, с использованием имеющихся знаний об объекте моделирования и требований к будущей модели.

Анализ и совместное обсуждение членами рабочей группы всей имеющейся информации об объекте моделирования позволяет сформировать *содержательную модель объекта*, являющуюся синтезом когнитивных моделей, сложившихся у каждого из членов рабочей группы. На основании содержательной модели разрабатывается уже *концептуальная* или "естественнонаучная" (физическая, химическая, биологическая и т.д.) постановка задачи моделирования, служащая основой для *концептуальной модели объекта*.

*Концептуальная постановка задачи моделирования* – это сформулированный в терминах конкретных дисциплин (физики, химии, биологии и т.д.) перечень основных вопросов, интересующих заказчика, а также совокупность гипотез относительно свойств и поведения объекта моделирования.

Наибольшие трудности при формулировке концептуальной постановки приходится преодолевать для моделей, находящихся на "стыке" различных дисциплин. Различия традиций, понятий и языков, используемых для описания одних и тех же объектов, являются очень серьезными препятствиями, возникающими при создании "междисциплинарных" моделей. Например, такие понятия как "прибыль" и "баланс" вызывают совершенно разные ассоциации у экономиста и математика-прикладника. Можно сказать, что когнитивные модели, стоящие за этими понятиями, у этих двух специалистов совершенно различны. Если экономист, говоря о прибыли и балансе, связывает с этими понятиями конкретное производство, цену и себестоимость продукции, то для математика данные понятия выглядят более формально, как результаты решения некоторых математических соотношений. При этом практически невозможно научить математика мыслить как экономиста, а экономиста – как математика. И тот, и другой способ восприятия имеет свои достоинства и недостатки. Экономист никогда не сделает ошибок, которые может натворить математик, обращаясь с параметрами модели формально, без должных зна-

ний в рассматриваемой предметной области. С другой стороны, используя формальные преобразования математических соотношений, математик может получить решения, которые очень сложно получить экономисту, пользующемуся своими подходами и методами (обычно – более простыми с точки зрения математики). Поэтому эффективность деятельности рабочей группы в большой степени зависит от способности ее членов поставить себя на место специалиста другого профиля, изучить его точку зрения (т.е. особенности его когнитивной модели) и найти некоторый компромисс, учитывающий все ценное.

Как отмечалось выше, концептуальная модель строится как некоторая идеализированная модель объекта, записанная в терминах конкретных (например, естественнонаучных) дисциплин. Для этого формулируется совокупность гипотез о поведении объекта, его взаимодействии с окружающей средой, изменении внутренних параметров. Как правило, эти гипотезы правдоподобны в том смысле, что для их обоснования могут быть приведены некоторые теоретические доводы и экспериментальные данные, основанные на собранной ранее информации об объекте. В выборе и обосновании принимаемых гипотез в значительной степени проявляется искусство, опыт и накопленные знания членов рабочей группы. На основании принятых гипотез определяется множество параметров, описывающих состояние объекта, а также перечень законов, управляющих изменением и взаимосвязью этих параметров между собой.

### **Пример**

*Концептуальная постановка задачи о баскетболисте*

Движение баскетбольного мяча может быть описано в соответствии с законами классической механики Ньютона.

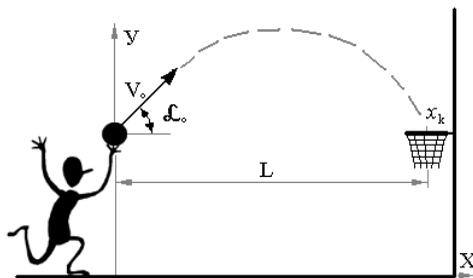


Рисунок 2.2 – К пояснению концептуальной постановки задачи о баскетболисте

Примем следующие гипотезы:

- объектом моделирования является баскетбольный мяч радиуса  $R$ ;
- мяч будем считать материальной точкой массой  $m$ , положение которой совпадает с центром масс мяча;
- движение происходит в поле сил тяжести с постоянным ускорением свободного падения  $g$  и описывается уравнениями классической механики Ньютона;
- движение мяча происходит в одной плоскости, перпендикулярной поверхности Земли и проходящей через точку броска и центр корзины;
- пренебрегаем сопротивлением воздуха и возмущениями, вызванными собственным вращением мяча вокруг центра масс.

В соответствии с изложенными гипотезами в качестве параметров движения мяча можно использовать координаты ( $x$  и  $y$ ) и скорость (ее проекции  $v_x$  и  $v_y$ ) центра масс мяча. Тогда для определения положения мяча в любой момент времени достаточно найти закон движения центра масс мяча, т.е. зависимость координат  $x$  и  $y$  и проекций вектора скорости  $v_x$  и  $v_y$  центра мяча от времени. В качестве оценки точности броска  $\Delta$  можно рассматривать величину расстояния по горизонтали (вдоль оси  $x$ ) от центра корзины до центра мяча, когда последний пересекает горизонтальную плоскость, проходящую через плоскость кольца корзины.

С учетом выше изложенного, можно сформулировать концептуальную постановку задачи о баскетболисте в следующем виде:

определить закон движения материальной точки массой  $m$  под действием силы тяжести, если известны начальные координаты точки  $x_0$  и  $y_0$ , ее начальная скорость  $v_0$  и угол бросания  $\alpha_0$ . Центр корзины имеет координаты  $x_k$  и  $y_k$ . Вычислить точность броска  $\Delta = x(t_k) - x_k$ , где  $t_k$  определяется из условий:

$$t_k > 0, v_y < 0, y(t_k) = y_k.$$

Рассмотрим особенности приведенной в примере концептуальной постановки задачи о баскетболисте. Первая из перечисленных гипотез особенно важна, так как она выделяет объект моделирования. В данном случае объект можно считать простым. Однако в качестве объекта моделирования можно рассматривать систему ИГРОК-МЯЧ-КОЛЬЦО. Требуемая для описания подобной системы модель будет уже намного сложнее, так как ИГРОК в свою очередь представляет сложную биомеханическую систему и его моделирование является далеко не тривиальной задачей. В данной ситуации выбор в качестве объекта моделирования только МЯЧА обоснован, так как именно его движение требуется исследовать, а влияние ИГРОКА можно учесть достаточно просто через начальные параметры броска. Для сложных систем выбор объекта моделирования является далеко не простой и не однозначной задачей.

Гипотеза о том, что мяч можно считать материальной точкой, широко применяется для исследования движений тел в механике. В данном случае она оправдана в силу симметрии формы мяча и малости радиуса мяча по сравнению с характерными расстояниями его перемещения. Предполагается, что мяч является шаром с одинаковой толщиной стенки.

Гипотезу о применимости в данном случае законов классической механики можно обосновать огромным экспериментальным материалом, связанным с изучением движения тел вблизи поверхности Земли со скоростями много меньше скорости света. Учитывая, что высота полета мяча лежит в пределах от 5 до 10 метров, а дальность – от 5 до 20 метров, предположение о постоянстве ускорения свободного падения также представляется

обоснованным. Если бы моделировалось движение баллистической ракеты при дальности и высоте полета более 100 км, то пришлось бы учитывать изменение ускорения свободного падения в зависимости от высоты и широты места.

Гипотеза о движении мяча в плоскости, перпендикулярной поверхности Земли, ограничивает класс рассматриваемых траекторий и значительно упрощает модель. Траектория мяча может не лежать в одной плоскости, если при броске он сильно подкручивается вокруг вертикальной оси. В этом случае скорость точек поверхности мяча относительно воздуха на различных сторонах мяча будут различны. Для точек, движущихся навстречу потоку, относительная скорость выше, а с противоположной стороны, движущихся по потоку, – ниже скорости центра масс мяча. В соответствии с законом Бернулли, давление газа на поверхность больше там, где его относительная скорость меньше. Поэтому для ситуации, изображенной на рисунке, на мяч будет действовать дополнительная сила, направленная (для данной схемы) сверху вниз. Данный эффект будет проявляться тем больше, чем больше скорость центра масс мяча и скорость его вращения. Для баскетбола характерны относительно низкие скорости полета мяча (до 10 м/с). При этом практически не используется подкрутка мяча рукой. Поэтому гипотеза о движении мяча в одной плоскости кажется оправданной. Ее использование позволяет отказаться от построения значительно более сложной трехмерной модели движения мяча.

Гипотеза об отсутствии влияния сопротивления воздуха наименее обоснована. При движении тела в газе или жидкости сила сопротивления увеличивается с ростом скорости движения. Учитывая невысокие скорости движения мяча, его правильную обтекаемую форму и малые дальности бросков, указанная гипотеза может быть использована с тщательной последующей проверкой получаемых результатов по отношению к экспериментальным результатам.

Следует отметить, что концептуальная постановка задачи моделирования в отличие от содержательной постановки использует терминологию конкретной дисциплины (в данном случае – механики). При этом моделируемый реальный объект (мяч) заме-

няется его механической моделью (материальной точкой). Фактически в рассматриваемом примере концептуальная постановка свелась к постановке классической задачи механики о движении материальной точки в поле сил тяжести. Концептуальная постановка более абстрактна по отношению к содержательной, так как материальной точке можно сопоставить произвольный материальный объект, брошенный под углом к горизонту: футбольный мяч, ядро, камень или артиллерийский снаряд.

### **2.3 Математическая постановка задачи моделирования**

Задача любого вида сводится к математической задаче.

Р.Декарт

Законченная концептуальная постановка позволяет сформулировать математическую постановку задачи моделирования, включающую совокупность различных математических соотношений, описывающих поведение и свойства объекта моделирования.

*Математическая постановка задачи моделирования* – это совокупность математических соотношений, описывающих поведение и свойства объекта моделирования.

Как было отмечено в главе 1, совокупность математических соотношений определяют вид оператора модели. Наиболее простой вид оператор модели имеет в случае системы алгебраических уравнений. Подобные модели можно назвать моделями *аппроксимационного типа*, так как для их получения часто используют различные методы аппроксимации имеющихся экспериментальных данных о поведении выходных параметров объекта моделирования в зависимости от входных параметров и воздействий внешней среды, а также от значений внутренних параметров объекта.

Однако модели подобного типа имеют ограниченную область применимости. Для создания математических моделей сложных систем и процессов, применимых для широкого класса реальных задач, как уже отмечалось выше, требуется привлечение большого объема знаний, накопленных в рассматриваемой дисциплине (а в некоторых случаях – и в смежных областях). В большинстве дисциплин (особенно – в естественно-научных) эти знания сконцентрированы в аксиомах, законах, теоремах, имеющих четкую математическую формулировку.

Следует отметить, что во многих областях знаний (механике, физике, биологии и т.д.) принято выделять законы, справедливые для всех объектов исследования данной области знаний, и соотношения, описывающие поведение отдельных объектов или их совокупностей. К числу первых в физике и механике, например, относятся уравнения баланса массы, количества движения, энергии и т.д., справедливые при определенных условиях для любых материальных тел, независимо от их конкретного строения, структуры, состояния, химического состава. Уравнения данного класса подтверждены огромным количеством экспериментов, хорошо изучены и в силу этого применяются в соответствующих математических моделях как данность. Соотношения второго класса в физике и механике называют определяющими соотношениями, или физическими уравнениями, или уравнениями состояния. Соотношения этого класса устанавливают особенности поведения материальных объектов или их совокупностей (например, жидкостей, газов, упругих или пластических сред и т.д.) при воздействиях различных внешних факторов; в качестве классических примеров определяющих соотношений можно привести закон Гука в теории упругости или уравнение Клапейрона для идеальных газов. Очевидно, определяющие соотношения должны отражать реальное атомно-молекулярное строение исследуемых материальных объектов. Соотношения второго класса гораздо менее изучены, а в ряде случаев их приходится устанавливать самому исследователю (особенно – при анализе объектов, состоящих из новых материалов). Необходимо отметить, что определяющие соотношения представляют собой основной элемент, "сердцевину" любой математической модели физико-

механических процессов. Именно ошибки в выборе или установлении определяющих соотношений приводят к количественно (а в некоторых случаях – и качественно) неверным результатам моделирования.

Совокупность математических соотношений указанных двух классов определяет оператор модели. В большинстве случаев оператор модели включает систему обыкновенных дифференциальных уравнений (ОДУ), дифференциальных уравнений в частных производных (ДУЧП) и/или интегро-дифференциальных уравнений (ИДУ). Для обеспечения корректности постановки задачи к системе ОДУ или ДУЧП добавляются начальные или граничные условия, которые, в свою очередь, могут быть алгебраическими или дифференциальными соотношениями различного порядка.

Можно выделить несколько наиболее распространенных типов задач, возникающих для систем ОДУ или ДУЧП:

- *задача Коши, или задача с начальными условиями*, в которой по заданным в начальный момент времени переменным (начальным условиям) определяются значения этих искомым переменных для любого момента времени;
- *начально-граничная, или краевая задача*, когда условия на искомую функцию выходного параметра задаются в начальный момент времени для всей пространственной области и на границе последней – в каждый момент времени (на исследуемом интервале);
- *задачи на собственные значения*, когда в формулировку задачи входят неопределенные параметры, определяемые из условия качественного изменения поведения системы (например, потеря устойчивости состояния равновесия или стационарного движения, появление периодического режима, резонанс и т.д.).

Для контроля правильности полученной системы математических соотношений следует выполнять ряд обязательных проверок [13]:

- *контроль размерностей*, включающий правило, согласно которому приравниваться и складываться могут только величины одинаковой размерности. При переходе к вычисле-

ниям данная проверка сочетается с контролем использования одной и той же системы единиц для значений всех параметров;

- *контроль порядков*, состоящий из грубой оценки сравнительных порядков складываемых друг с другом величин и исключением малозначимых параметров. Например, если для выражения  $x+y+z=0$  в результате оценки установлено, что для рассматриваемой области значений параметров модели  $|z| \ll |x|$  и  $|z| \ll |y|$ , то третьим слагаемым в исходном выражении можно пренебречь;
- *контроль характера зависимостей*: направление и скорость изменения выходных параметров модели, вытекающие из выписанных математических соотношений, должны быть такими, как это следует непосредственно из "физического" смысла изучаемой модели;
- *контроль экстремальных ситуаций*. Весьма полезно проследить за тем, какой вид принимают математические соотношения, а также результаты моделирования, если параметры модели или их комбинации приближаются к предельно допустимым для них значениям – чаще всего к нулю или бесконечности. В подобных экстремальных ситуациях модель часто упрощается, а математические соотношения приобретают более наглядный смысл и могут быть проще проверены. Например, в задачах механики деформируемого твердого тела деформация материала в исследуемой области при изотермических условиях возможна лишь при приложении нагрузок, отсутствие нагрузок должно приводить к отсутствию деформаций;
- *контроль граничных условий*, включающий проверку, что граничные условия действительно наложены, что они использованы в процессе построения искомого решения и что значения выходных параметров модели на самом деле удовлетворяют данным условиям;
- *контроль физического смысла* состоит в проверке физического или иного, в зависимости от характера задачи, смысла исходных и промежуточных соотношений, появляющихся по мере конструирования модели;

- *контроль математической замкнутости*, состоящий в проверке того, что выписанная система математических соотношений дает возможность, и притом однозначно, решить поставленную математическую задачу. Например, если задача свелась к отысканию  $n$  неизвестных из некоторой системы алгебраических или трансцендентных уравнений, то контроль замкнутости состоит в проверке того факта, что число независимых уравнений должно быть  $n$ . Если их меньше  $n$ , то надо установить недостающие уравнения, а если их больше  $n$ , то либо уравнения зависимы, либо при их составлении допущена ошибка. Однако если уравнения получаются из эксперимента или в результате наблюдений, то возможна постановка задачи, при которой число уравнений превышает  $n$ , но сами уравнения удовлетворяются лишь приближенно, а решение ищется, например, по методу наименьших квадратов. Неравенств среди условий также может быть любое число, как это бывает, например, в задачах линейного программирования.

Свойство математической замкнутости системы математических соотношений тесно связано с введенным Ж. Адамаром понятием *корректно поставленной математической задачи* [88], то есть задачи, для которой решение *существует, единственно и непрерывно* зависит от исходных данных. В данном случае решение считается непрерывным, если малому изменению исходных данных соответствует достаточно малое изменение решения. Понятие корректности задачи имеет большое значение в прикладной математике. Например, численные методы решения оправдано применять лишь к корректно поставленным задачам. Следует отметить, что далеко не все задачи, возникающие на практике, можно считать корректными (например, так называемые обратные задачи). Доказательство корректности конкретной математической задачи является достаточно сложной проблемой, решенной только для некоторого класса математически поставленных задач. В этом отношении проверка математической замкнутости является менее сложной по сравнению с проверкой корректности математической постановки. В настоящее время активно исследуются свойства некорректных задач, разрабатываются методы

их решения. Аналогично понятию корректно поставленной задачи можно ввести понятие *корректной математической модели*.

Математическая модель является *корректной*, если для нее выполняются все контрольные проверки: размерности, порядков, характера зависимостей, экстремальных ситуаций, граничных условий, физического смысла и математической замкнутости.

### Пример

*Математическая постановка задачи о баскетболисте*

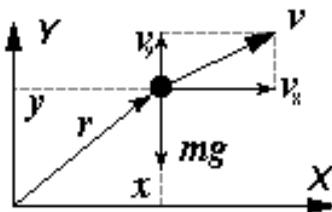


Рисунок 2.3 – К пояснению математической постановки задачи о баскетболисте

Математическую постановку задачи о баскетболисте можно представить как в векторной, так и в координатной форме.

А) *Векторная форма*

Найти зависимости от времени для векторных параметров  $r(t)$  и  $v(t)$  из решения системы обыкновенных дифференциальных уравнений:

$$m \frac{dv}{dt} = mg, \quad v = \frac{dr}{dt}, \quad (2.1)$$

при следующих начальных условиях:

$$r(0) = 0, \quad v(0) = v_0. \quad (2.2)$$

Вычислить параметр  $\Delta$  как

$$\Delta = r_x(t_k) - r_{xk}, \quad (2.3)$$

где  $t_k$  определить из следующих условий:

$$t_k > 0, \quad v(t_k) < 0, \quad y(t_k) = y_k. \quad (2.4)$$

Проецируя векторные соотношения (2.1)-(2.4) на оси координат, получим математическую постановку задачи о баскетболисте в координатной форме.

*Б) Координатная форма*

Найти зависимости  $x(t)$ ,  $y(t)$  и  $v_x(t)$  и  $v_y(t)$  из решения системы дифференциальных уравнений:

$$m \frac{dv_x}{dt} = 0, \quad v_x = \frac{dx}{dt}, \tag{2.5}$$

$$m \frac{dv_y}{dt} = -mg, \quad v_y = \frac{dy}{dt},$$

при следующих начальных условиях:

$$x(0) = x_0, \quad y(0) = y_0, \tag{2.6}$$

$$v_x(0) = v_0 \cos \alpha_0, \quad v_y(0) = v_0 \sin \alpha_0.$$

$$\text{Вычислить параметр } \Delta \text{ как } \Delta = x(t_k) - x_k, \tag{2.7}$$

где  $t_k$  определить из условий:

$$t_k > 0, \quad v(t_k) < 0, \quad y(t_k) = y_k. \tag{2.8}$$

Как можно видеть, с математической точки зрения задача о баскетболисте свелась к задаче Коши для системы ОДУ первого порядка с заданными начальными условиями. Полученная система уравнений является замкнутой, так как число независимых уравнений (4 дифференциальных и 2 алгебраических) равно числу искомых параметров задачи ( $x, y, v_x, v_y, \Delta, t_k$ ). Существование и единственность решения задачи Коши доказана математиками. Поэтому данную математическую модель можно считать корректной.

Математическая постановка задачи еще более абстрактна, чем концептуальная, так как сводит исходную задачу к чисто математической (например, к задаче Коши), методы решения которой достаточно хорошо разработаны. Умение свести исходную проблему к известному классу математических задач и обосновать правомочность такого сведения требует высокой квалификации математика-прикладника и особенно высоко ценится в исследовательских коллективах.

## 2.4 Выбор и обоснование выбора метода решения задачи

Метод решения хорош, если с самого начала мы можем предвидеть – и далее подтвердить это, – что следуя этому методу, мы достигнем цели.

Г. Лейбниц

При использовании разработанных математических моделей, как правило, требуется найти зависимость некоторых неизвестных заранее параметров объекта моделирования (например, координаты и скорость центра масс тела, точность броска), удовлетворяющих определенной системе уравнений. Таким образом, поиск решения задачи сводится к отысканию некоторых зависимостей искомых величин от исходных параметров модели. Как уже было отмечено в главе 1, все методы решения соответствующих задач, составляющих "ядро" математических моделей, можно подразделить на аналитические и алгоритмические. Следует отметить, что при использовании аналитических решений для получения результатов "в числах" также часто требуется разработка соответствующих алгоритмов, реализуемых на ЭВМ. Однако исходное решение при этом представляет собой аналитическое выражение (или их совокупность). Решения же, основанные на алгоритмических методах, принципиально не сводимы к точным аналитическим решениям соответствующей задачи.

Выбор того или иного метода исследования в значительной степени зависит от квалификации и опыта членов рабочей группы. Как уже было отмечено, аналитические методы более удобны для последующего анализа результатов, но применимы лишь для относительно простых моделей. В случае, если соответствующая математическая задача (хотя бы и в упрощенной постановке) допускает аналитическое решение, последнее, без сомнения, предпочтительнее численного. Как справедливо говорит по этому поводу А. Б. Мигдал "*... раньше, чем пользоваться счетными машинами, задачу необходимо всесторонне исследовать аналити-*

*ческими методами. Аналитические методы – "старое, но грозное оружие" – не теряют своего значения".*

Алгоритмические методы сводятся к некоторому алгоритму, реализующему вычислительный эксперимент с использованием ЭВМ. Точность моделирования в подобном эксперименте существенно зависит от использованного метода и его параметров (например, шага интегрирования). Алгоритмические методы, как правило, более трудоемки в реализации, требуют хорошего знания от членов рабочей группы методов вычислительной математики, обширной библиотеки специального программного обеспечения и мощной вычислительной техники. Современные модели на базе алгоритмических методов разрабатываются в исследовательских организациях, которые зарекомендовали себя как авторитетные научные школы в соответствующей области знания.

Как отмечалось ранее, приближенные и численные методы исследования поставленных математических задач относятся к обширному разделу современной математики – вычислительной математике. Численные методы применимы лишь для корректных математических задач, что является существенным ограничением на применение данных методов в математическом моделировании.

Общим для всех численных методов является сведение математической задачи к конечномерной. Это чаще всего достигается дискретизацией исходной задачи, то есть переходом от функции непрерывного аргумента к функциям дискретного аргумента. Например, траектория центра тяжести баскетбольного мяча ищется не как непрерывная функция времени, а как табличная (дискретная) функция координат от времени, то есть определяющая значения координат лишь для конечного числа моментов времени. Полученное решение дискретной задачи принимается за приближенное решение исходной математической задачи.

Выделяют три основных составляющих возникающей погрешности при численном решении исходной задачи:

- *неустраняемая погрешность*, связанная с неточным заданием исходных данных задачи (начальные и граничные условия, коэффициенты и правые части уравнений);

- *погрешность метода*, связанная с переходом к дискретному аналогу исходной задачи (например, заменяя производную  $y'(x)$  разностным аналогом  $(y(x+\Delta x)-y(x))/\Delta x$ , получаем погрешность дискретизации, имеющую при  $\Delta x \rightarrow 0$  порядок  $\Delta x$ );
- *ошибка округления*, связанная с конечной разрядностью чисел, представляемых в ЭВМ.

Естественным требованием для конкретного вычислительного алгоритма является согласованность в порядках величин перечисленных трех видов погрешностей.

Численный или приближенный метод реализуется всегда в виде вычислительного алгоритма. Поэтому все требования, предъявляемые к алгоритму, применимы и к вычислительному алгоритму. Прежде всего, алгоритм должен быть реализуем, то есть давать решение задачи за допустимое машинное время. Важной характеристикой алгоритма является его *точность*, то есть возможность получения решения исходной задачи с заданной точностью  $\varepsilon > 0$  за конечное число  $Q(\varepsilon)$  действий. Очевидно, чем меньше  $\varepsilon$ , тем больше затрачиваемое машинное время. Для очень малых  $\varepsilon$  время вычислений может быть недопустимо большим. Поэтому на практике добиваются некоторого компромисса между точностью и затрачиваемым машинным временем. Очевидно, что для каждой задачи, алгоритма и типа ЭВМ имеется свое характерное значение достигаемой точности.

Время работы алгоритма зависит от числа действий  $Q(\varepsilon)$  для достижения заданной точности. Для любой математической задачи, как правило, можно предложить несколько алгоритмов, позволяющих получить решение с заданной точностью, но за разное число действий  $Q(\varepsilon)$ . Алгоритмы, затрачивающие меньшее число действий для достижения одинаковой точности, будем называть более *экономичными* или более *эффективными*.

В процессе работы вычислительного алгоритма на каждом акте вычислений возникает некоторая погрешность. При этом величина погрешности может от действия к действию нарастать или не возрастать (а в некоторых случаях – даже уменьшаться). Если погрешность в процессе вычислений неограниченно нарастает, то такой алгоритм называется *неустойчивым* или *расходя-*

*щимся*. В противном случае алгоритм называется *устойчивым* или *сходящимся*.

Как уже было отмечено выше, вычислительная математика объединяет огромный пласт разнообразных быстро развивающихся численных и приближенных методов. Можно выделить следующие группы численных методов по объектам, к которым они применяются:

- интерполяция и численное дифференцирование;
- численное интегрирование;
- определение корней линейных и нелинейных уравнений;
- решение систем линейных уравнений (подразделяют на прямые и итерационные методы);
- решение систем нелинейных уравнений;
- решение задачи Коши для обыкновенных дифференциальных уравнений;
- решение краевых задач для обыкновенных дифференциальных уравнений;
- решение уравнений в частных производных;
- решение интегральных уравнений.

В данном пособии нет возможности остановиться более подробно на достоинствах и недостатках тех или иных численных методах даже одной группы. Более подробные сведения по численным методам, оценкам их точности, устойчивости и эффективности можно найти в обширной литературе по данному вопросу, а также ряд этих вопросов будет рассмотрен нами на практических занятиях.

Огромное разнообразие численных методов в значительной степени затрудняет выбор того или иного метода в каждом конкретном случае. Учитывая, что для реализации одной и той же модели можно использовать несколько альтернативных алгоритмических методов, выбор конкретного метода обусловлен тем, какой из них подходит для данной модели с точки зрения обеспечения эффективности, устойчивости и точности результатов, а также более освоен и знаком членам рабочей группы. Освоение нового метода, как правило, очень трудоемко и связано с большими временными и финансовыми затратами. Основные затраты при этом связаны с разработкой и отладкой необходимого про-

граммного обеспечения для соответствующего класса ЭВМ, обеспечивающего поддержку данного метода. Следует отметить, что вычислительная математика в определенном смысле более являет собой искусство, нежели "науку" (в понимании последней как области культуры, базирующейся на формальной логике). Очень часто эффективность применяемых методов, разработанных программ определяется нарабатываемыми годы и десятки лет интуитивными приемами, не обоснованными с математических позиций. В связи с этим эффективность одного и того же метода может весьма существенно отличаться при его применении различными исследователями.

### **Пример**

#### *Аналитическое решение задачи о баскетболисте*

Для получения решения рассмотренной выше задачи о баскетболисте в постановке (2.5)-(2.8) можно использовать как аналитические, так и численные методы. Проинтегрировав соотношения (2.5) по времени, получим

$$\begin{aligned} x(t) &= C_2 + C_1 t, & y(t) &= C_4 + C_3 t - \frac{gt^2}{2}, \\ v_x(t) &= C_1, & v_y(t) &= C_3 - gt. \end{aligned} \quad (2.9)$$

Константы интегрирования найдем из начальных условий (2.6). Тогда решение задачи можно записать следующим образом

$$\begin{aligned} x(t) &= x_0 + v_0 t \cos \alpha_0, & y(t) &= y_0 + v_0 t \sin \alpha_0 - \frac{gt^2}{2}, \\ v_x(t) &= v_0 \cos \alpha_0, & v_y(t) &= v_0 \sin \alpha_0 - gt. \end{aligned} \quad (2.10)$$

Примем для простоты, что в момент броска мяч находится в начале координат и на одном уровне с корзиной (т.е.  $x_0 = y_0 = y_k = 0$ ). Под дальностью  $L$  броска будем понимать расстояние, которое пролетит мяч от точки броска до пересечения с горизонтальной плоскостью, проходящей через кольцо корзины. Из соотношений (2.10) для координат дальность броска выразится следующим образом

$$L = \frac{v_0^2}{g} \sin 2\alpha_0 \quad (2.11)$$

Тогда точность броска с учетом (2.7) будет равна

$$\Delta = L - x_k \quad (2.12)$$

Например, при броске мяча со штрафной линии можно принять следующие исходные данные:  $x_0 = y_0 = y_k = 0$ ;  $x_k = 4,225$  м;  $v_0 = 6,44$  м/с;  $\alpha = 45^\circ$ . Тогда из (2.11) и (2.12) имеем  $L = 4,225$  м;  $\Delta = 0$  м.

### Пример

#### *Алгоритмическое решение задачи о баскетболисте*

В простейшем случае можно использовать метод Эйлера. Алгоритм решения данной задачи на псевдокоде приведен ниже.

#### Алгоритм 2.1

```

program Задача о баскетболисте.
Данные:      m, R - масса и радиус мяча;
               x0, y0 - начальные координаты мяча;
               v0, α - начальная скорость и угол броска мяча;
               xk, yk - координаты центра корзины;
               t - текущее время;
               dt - шаг по времени;
               fx, fy - силы, действующие на мяч;
               x, y, vx, vy - текущие координаты и проекции скорости мяча.
Результаты: L и Δ - дальность и точность броска.
start
  g := 9.81
  m := 0.6;           R := 0.12
  v0 := 6.44;       α := 45
  x0 := 0;          y0 := 0
  xk := 4.225;      yk := 0
  vx := v0 cos α
  vy := v0 sin α
  dt := 0.1
  t := 0;
  x := x0;          y := y0
  while ((vy > 0) or ((vy < 0) and (y >= yk)))
    t := t + dt
    fx := 0                               силы, действующие на мяч
    fy := - mg
    vx := vx + fx dt / m                 компоненты скорости
    vy := vy + fy dt / m
    x := x + dt vx                         координаты мяча
    y := y + dt vy
  end while
  L := x - x0
  Δ := x - xk
stop

```

Рисунок 2.3 – Алгоритмическое решение задачи о баскетболисте

## 2.5 Реализация математической модели в виде программы для ЭВМ

Компьютеры бесподобны: за несколько минут они могут совершить такую ошибку, которую не в состоянии сделать множество людей за многие месяцы.

Лоуренс Дж. Питер

В настоящее время математические модели, как правило, применяют в качестве основы для создания различных программных комплексов, используемых для решения разнообразных исследовательских, проектно-конструкторских и управленческих задач. Данное обстоятельство приводит к необходимости реализации модели в виде программы для ЭВМ. Процесс разработки надежного и эффективного программного обеспечения является не менее сложным, чем разработка предыдущих этапов создания математической модели. Успешное решение данного вопроса возможно лишь при уверенном владении современными алгоритмическими языками и технологиями программирования, имеющегося программного обеспечения, знании возможностей вычислительной техники, особенностей реализации на ЭВМ методов вычислительной математики, опыта решения подобных задач.

Процесс создания программного обеспечения можно разбить на ряд этапов:

- разработка технического задания на создание программного обеспечения;
- проектирование структуры программного комплекса;
- кодирование алгоритма;
- тестирование и отладка;
- сопровождение и эксплуатация.

Техническое задание на разработку программного обеспечения оформляют в виде *спецификации*. Примерная форма спецификации включает следующие семь разделов.

### 1) Название задачи

Дается краткое определение решаемой задачи, название программного комплекса, указывается система программирования для его реализации и требования к аппаратному обеспечению (компьютеру, внешним устройствам и т.д.).

### 2) Описание

Подробно излагается математическая постановка задачи, описывается применяемая математическая модель для задач вычислительного характера, метод обработки входных данных для задач не вычислительного (логического) характера и т.д.

### 3) Управление режимами работы программы

Формируются основные требования к способу взаимодействия пользователя с программой (интерфейс "пользователь-компьютер").

### 4) Входные данные

Описываются входные данные, указываются пределы, в которых они могут изменяться, значения, которые они не могут принимать, и т.д.

### 5) Выходные данные

Описываются выходные данные, указывается, в каком виде они должны быть представлены – в числовом, графическом или текстовом, приводятся сведения о точности и объеме выходных данных, способах их сохранения и т.д.

### 6) Ошибки

Перечисляются возможные ошибки пользователя при работе с программой, (например, ошибки при вводе входных данных). Указываются способы диагностики (в данном случае под диагностикой понимается выявление, обнаружение ошибок при работе программного комплекса) и защиты от этих ошибок на этапе проектирования, а также возможная реакция пользователя при совершении им ошибочных действий и реакция программного комплекса (компьютера) на эти действия.

### 7) Тестовые задачи

Приводится один или несколько тестовых примеров, на которых в простейших случаях проводится отладка и тестирование программного комплекса.

### **Пример:**

*Спецификация задачи о баскетболисте*

#### 1) Название задачи

Название программы **Basketball**

Система программирования **Delphi**

Компьютер **IBM PC Pentium**

Операционная система **Windows-9x, NT**

#### 2) Описание

Приводится математическая постановка задачи и описание метода ее решения.

#### 3) Управление режимами работы программы

Для управления режимами работы программы необходимо использовать интерфейс Windows с использованием меню, диалоговых окон, полей ввода данных, кнопок.

#### 4) Входные данные

Входными данными являются радиус и масса мяча, его начальные координаты и скорость, угол бросания, координаты корзины.

#### 5) Выходные данные

Траектория центра мяча, расчетная величина дальности и точность броска.

#### 6) Ошибки

При вводе исходных данных предусмотреть контроль:

- все вводимые значения должны быть положительны;
- угол бросания лежит в пределах от 5 до 85 градусов;
- начальная скорость мяча лежит в пределах от 0 до 30 м/с;
- горизонтальная координата центра корзины больше начальной горизонтальной координаты мяча.

При диагностировании перечисленных ошибок программа должна выдавать соответствующие сообщения, которые могут сопровождаться звуковым сигналом, и предлагать повторить ввод.

#### 7) Тестовые примеры

При  $x_0 = y_0 = y_k = 0$ ;  $x_k = 4,225$ ;  $v_0 = 6,44$ ;  $\alpha = 450$ ;

Получаем:  $L = 4,225$ ;  $\Delta = 0$ .

На этапе проектирования формируется общая структура программного комплекса. Вся программа разбивается на программные модули. Для каждого программного модуля формулируются требования по реализуемым функциям и разрабатывается алгоритм, выполняющий эти функции. Определяется схема взаимодействия программных модулей, называемая схемой потоков данных программного комплекса. Разрабатывается план и исходные данные для тестирования отдельных модулей и программного комплекса в целом.

Большинство программ, реализующих математические модели, состоят из трех основных частей:

- *препроцессор* (подготовка и проверка исходных данных модели);
- *процессор* (решение задачи, реализация вычислительного эксперимента);
- *постпроцессор* (отображение полученных результатов).

Лишь для относительно простых случаев данные три составные части могут быть оформлены в виде одной программы. При решении современных задач по моделированию поведения жидкостей, газов и твердых тел каждая из частей может включать в себя целый комплекс программ. Например, постпроцессор должен уметь представлять информацию не только в табличном, но и графическом виде (диаграммы, графики зависимости от различных параметров, отображение скалярных, векторных (тензорных) полей и т.п.). Особенно сильное развитие возможности пре- и постпроцессора получают в современных системах автоматизированного проектирования (САПР), где они в значительной степени могут сократить время на получение данных и оценку результатов моделирования.

Как правило, создание современных математических моделей в какой-либо области и доведение их до программных комплексов требует значительных временных затрат (минимум 3-5 лет). Требуется время не только на освоение методик и подходов к моделированию в исследуемой области, но и время на наработку библиотек программ по решению возникающих математических задач, по подготовке исходных данных и отображению получаемых результатов. Качественные, надежные, обладающие

дружественным интерфейсом, легко модифицируемые и хорошо сопровождаемые программные комплексы можно создавать лишь при наличии хорошо продуманной стратегии развития программного обеспечения, обеспечивающей его модульность и совместимость по входным и выходным параметрам.

Большое значение следует придавать освоению современных технологий программирования: структурной, абстрактной, объектно-ориентированной и визуальной. Назначение любой технологии – это, в первую очередь, повышение надежности программного обеспечения и увеличение производительности труда программиста. Причем, чем серьезней и объемней программный проект, тем большее значение приобретают вопросы использования современных технологий программирования. Пренебрежение данными вопросами может привести к значительным временным издержкам и снижению надежности программного комплекса.

Важнейшим фактором, определяющим надежность и малые сроки создания программного комплекса для решения определенного класса задач, является наличие развитой библиотеки совместимых между собой программных модулей. Программа получается более надежной и создается за меньшие сроки при максимальном использовании стандартных программных элементов. Для эффективной разработки программного обеспечения в области математического моделирования в первую очередь следует обратить внимание на создание следующих стандартных библиотек программ:

- приближенные и численные методы;
- средства подготовки исходных данных (препроцессоры);
- средства визуализации и представления результатов (постпроцессоры).

Разработка таких общих библиотек программ возможна лишь при стандартизации потоков передачи данных между препроцессором, процессором и постпроцессором. В простейшем случае речь может идти об унификации форматов передаваемых файлов.

## **2.6 Проверка адекватности модели**

Хороший теоретик может объяснить почти любые полученные результаты, верные или неверные, и по крайней мере может потерять массу времени на выяснение того, верны они или нет.

Р. Хемминг

Под *адекватностью* математической модели будет пониматься степень соответствия результатов, полученных по разработанной модели, данным эксперимента или тестовой задачи. Прежде чем переходить к проверке адекватности модели, необходимо убедиться в правильном комплексном функционировании всех алгоритмов и программ модели, выполнить независимое тестирование и отладку всех отдельных алгоритмов (например, используемых программных модулей, реализующих используемый численный метод).

Проверка адекватности модели преследует две цели:

1. убедиться в справедливости совокупности гипотез, сформулированных на этапах концептуальной и математической постановок. Переходить к проверке гипотез следует лишь после проверки использованных методов решения, комплексной отладки и устранения всех ошибок и конфликтов, связанных с программным обеспечением;
2. убедиться, что точность полученных результатов соответствует точности, оговоренной в техническом задании.

Проверка разработанной математической модели выполняется путем сравнения с имеющимися экспериментальными данными о реальном объекте или с результатами других, созданных ранее и хорошо себя зарекомендовавших моделей. В первом случае говорят о проверке путем сравнения с экспериментом, во втором – о сравнении с тестовой задачей.

Решение вопроса о точности моделирования зависит от требований, предъявляемых к модели, и ее назначения. При этом должна учитываться точность получения экспериментальных результатов или особенности постановок тестовых задач. Для моделей, предназначенных для выполнения оценочных и прикидочных расчетов, удовлетворительной считается точность в 10-15%.

Для моделей, предназначенных для использования в управляющих и контролирующих системах, требуемая точность может быть 1-2% и даже выше.

Как правило, различают качественное и количественное совпадение результатов сравнения. При качественном сравнении требуется лишь совпадение некоторых характерных особенностей в распределении исследуемых параметров (например, наличие экстремальных точек, положительное или отрицательное значение параметра, его возрастание или убывание и т.д.). Фактически при качественном сравнении оценивается совпадение лишь вида функции распределения параметров (убывающая или возрастающая, с одним экстремумом или с несколькими). Вопрос о количественном сравнении можно ставить лишь после удовлетворительного ответа на вопрос о качественном соответствии результатов. При количественном сравнении большое значение следует придавать точности исходных данных для моделирования и соответствующих им значений сравниваемых параметров.

Неадекватность результатов моделирования возможна, по крайней мере, по трем причинам:

а) Значения задаваемых параметров модели не соответствуют допустимой области этих параметров, определяемой принятой системой гипотез. Например, в задаче о баскетболисте гипотезу об отсутствии сопротивления воздуха можно использовать лишь при относительно малых ( $<5$  м/с) скоростях движения тела. При больших значениях начальной скорости мяча влияние силы сопротивления будет существенным.

б) Принятая система гипотез верна, но константы и параметры в использованных определяющих соотношениях установлены не точно. Например, в случае задачи о баскетболисте значение ускорения свободного падения  $g$  может быть уточнено в зависимости от широты местности, где находится баскетболист.

в) Неверна исходная совокупность гипотез.

Все три случая требуют дополнительного исследования как моделируемого объекта (с целью накопления новой дополнительной информации о его поведении), так и исследования самой модели (с целью уточнения границ ее применимости).

*Замечание.* В данном случае не анализируется влияние выбранного численного метода на точность получаемого решения, а значит, и на адекватность модели. Вопрос о сходимости алгоритма и устойчивости получаемого выбранным численным методом решения, а также накопление погрешностей, связанных с ошибками округления при использовании ЭВМ, здесь не рассматривается.

При возникновении проблем, связанных с адекватностью модели, ее корректировку следует начинать с последовательного анализа всех возможных причин, приведших к расхождению результатов моделирования с результатами эксперимента. В первую очередь следует исследовать модель и оценить степень ее адекватности при различных значениях варьируемых параметров (начальных и граничных условиях, параметров, характеризующих свойства объектов моделирования). Если модель неадекватна в интересующей исследователя области параметров, то следует попытаться уточнить значения констант и исходных параметров модели. Если же и в этом случае нет положительных результатов, то единственной возможностью улучшения модели остается изменение принятой системы гипотез. Данное решение фактически означает возвращение ко второму этапу процесса разработки модели и может повлечь не только серьезное изменение математической постановки задачи, но и изменение методов ее решения (например, переход от аналитических к численным), полной переработки программного обеспечения и нового цикла проверки модели на адекватность. Поэтому решение об изменении принятой системы гипотез должно быть всесторонне взвешено, и приниматься только в том случае, когда исчерпаны все прочие возможности по улучшению адекватности модели.

Следует предостеречь начинающих исследователей от попыток "перепрыгнуть" рассмотренный этап моделирования, от желания быстрее перейти к решению "настоящей задачи". Как показывает собственный опыт авторов, подобный образ действий приводит к огромным временным издержкам (не говоря уже о психологических). Особенно опасной является ситуация, когда при решении реальной задачи с использованием должным образом не проверенной модели получают правдоподобные резуль-

таты. Для других условий модель может дать качественно неверные результаты, но истоки ошибок разработчики будут искать уже не в модели.

## **2.7 Практическое использование построенной модели и анализ результатов моделирования**

Цель расчетов – не числа, а понимание.

Р. Хемминг

Дескриптивные модели, рассмотренные в данной главе, предназначены для описания исследуемых параметров некоторого явления или процесса, а также для изучения закономерностей изменения этих параметров. Данные модели могут использоваться

- для изучения свойств и особенностей поведения исследуемого объекта при различных сочетаниях исходных данных и при различных режимах;
- как моделирующие блоки в различных системах автоматизированного проектирования (САПР) и управления (АСУ);
- при построении оптимизационных моделей и моделей-имитаторов сложных систем и комплексов.

Модели, разрабатываемые для исследовательских целей, как правило, не доводятся до уровня программных комплексов, предназначенных для передачи сторонним пользователям. Поэтому время их существования ограничено временем выполнения исследовательских работ по соответствующему направлению. Эти модели отличает поисковый характер, применение новых вычислительных процедур и алгоритмов, неразвитый программный интерфейс. Модели и построенные на их основе программные комплексы, предназначенные для передачи сторонним пользователям или для коммерческого распространения, имеют развитый дружественный интерфейс, мощные пре- и постпроцессоры. Данные модели строятся, как правило, на апробированных и хорошо себя зарекомендовавших постановках и вычислительных процедурах. Программные комплексы имеют подробные и качественно составленные описания и руководства для пользователя,

по всем неясным вопросам фирма-производитель проводит консультации. Однако следует помнить, что такие коммерческие модели предназначены только для решения четко оговоренного класса задач, как правило, не имеют возможностей по модернизации и совершенствованию, привязаны к определенному классу вычислительных устройств и периферийному оборудованию. Так, пользователь не имеет возможности самостоятельно расширять библиотеку используемых численных методов или изменять систему исходных гипотез.

Независимо от области применения разработанной модели группа разработчиков обязана провести качественный и количественный анализ результатов моделирования. Подобный анализ может преследовать несколько целей:

- работая с моделью, разработчики становятся специалистами в области знаний, связанной с объектом моделирования. Они начинают достаточно хорошо представлять свойства объекта, предсказывать и объяснять его поведение, в некотором смысле они могут даже представлять себя в роли объекта моделирования. Поэтому всесторонний анализ результатов моделирования может позволить выполнить модификацию рассматриваемого объекта, найти его оптимальные характеристики или, по крайней мере, лучшим образом учитывать его поведение и свойства;
- качественный и количественный анализ результатов моделирования позволяет обозначить область применения модели, что особенно важно в случае использования моделей для систем автоматического управления;
- подобный анализ позволит проверить обоснованность гипотез, принятых на этапе математической постановки, оценить возможность упрощения модели с целью повышения ее эффективности при сохранении требуемой точности;
- наконец, выполненный анализ может показать, в каком направлении следует развивать модель в дальнейшем.

### **Пример**

*Анализ результатов решения задачи о баскетболисте*

Соотношения (2.10)-(2.12) представляют аналитическое решение задачи о баскетболисте. Если в зависимость для координаты  $y(t)$  подставить  $x(t)$ , то получим уравнение параболы, описывающее траекторию мяча. Приравнявая нулю левую часть в соотношении для  $y(t)$ , определяем время полета мяча на дальность  $L$ .

$$T = 2 \frac{v_0}{g} \sin \alpha_0 \quad (2.13)$$

Наивысшей точки траектории мяч достигает в момент  $T/2$ . При этом максимальная высота мяча (высота траектории) равна

$$H = \frac{v_0^2}{2g} \sin^2 \alpha_0 \quad (2.14)$$

Анализ соотношения (2.11) позволяет заключить, что максимальная дальность броска достигается при угле бросания в 45 градусов и равна

$$L_{\max} = \frac{v_0^2}{g} \quad (2.15)$$

Зададим в (2.12) точность  $\Delta=0$  и из полученного уравнения найдем начальную скорость мяча при угле бросания в 45 градусов, обеспечивающую "абсолютную" точность

$$v_{25}^2 = g x_k \quad (2.16)$$

Запишем соотношения для дальности и точности броска в безразмерном виде. Относительную дальность  $l$  броска определим как отношение дальности броска к максимальной дальности

$$l = \frac{L}{L_{\max}} = \sin 2\alpha_0 \leq 1 \quad (2.17)$$

Относительную точность  $\delta$  введем как отношение отклонения  $\Delta$  к расстоянию до центра корзины

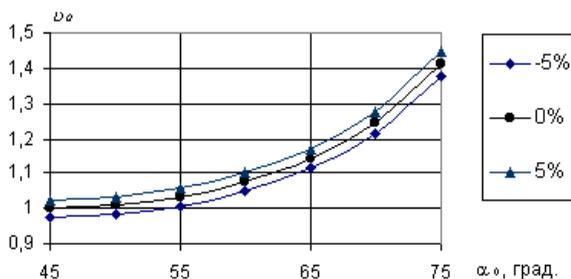
$$\delta = \frac{\Delta}{x_k} = \frac{L}{x_k} - 1 = \frac{v_0^2}{g \cdot x_k} \cdot \sin 2\alpha_0 - 1 = \left( \frac{v_0}{v_{45}} \right)^2 \cdot \sin 2\alpha_0 - 1 = v_0^2 \cdot \sin 2\alpha_0 - 1 \quad (2.18)$$

где  $v_0 = \frac{v_0}{\sqrt{g \cdot x_k}} > 0$  – относительная начальная скорость мяча.

Представление разрешающих соотношений модели в безразмерном виде более удобно для анализа. Из соотношения (2.18) видно, что заданную величину относительной точности броска

можно получить при двух значениях угла бросания, обеспечивающих *настильную* (при  $\alpha_0 < 45^\circ$ ) и *навесную* (при  $\alpha_0 > 45^\circ$ ) траектории движения мяча. При  $\alpha_0 = 45^\circ$  указанные траектории совпадают. Для обеспечения одинаковой точности для навесной и настильной траектории начальные скорости мяча должны быть одинаковы. С точки зрения попадания мяча в кольцо навесная траектория является предпочтительней по сравнению с настильной, так как даже при ударе мяча о кольцо вероятность попадания в корзину в этом случае выше. Иначе говоря, допуск на отклонение траектории мяча от центра кольца для навесной траектории больше, чем для настильной. Поэтому настильные траектории можно исключить из дальнейшего рассмотрения как менее выгодные. Для навесной траектории в первом приближении можно принять, что мяч попадает в кольцо, если отклонение траектории от центра кольца не превышает половины радиуса мяча.

Задаваясь величиной относительной точности и обозначая, из соотношения (2.18) можно получить оценку для разброса относительной начальной скорости



$$K_\alpha (1 - \delta) \leq v_0^2 \leq K(1 + \delta) \quad (2.19)$$

где  $K_\alpha = 1 / \sin 2\alpha_0 \geq 1$  – коэффициент угла бросания.

Рисунок 2.4 – Зависимость  $v_0(\alpha_0)$  при относительной точности  $\delta = 5\%$

Учитывая, что при отклонении  $\alpha_0$  от  $45^\circ$   $K_\alpha$  в (2.19) увеличивается, можно заключить, что при этом допуск на разброс значений начальной скорости при фиксированном значении  $\alpha_0$  будет

также увеличиваться. Правда, при этом во столько же раз увеличивается и сама величина относительной начальной скорости броска. Например, при  $\alpha_0 = 75^\circ$  относительная начальная скорость увеличится в  $\sqrt{2}$  раз по сравнению со значением скорости при  $\alpha_0 = 45^\circ$ . Однако и допустимая величина разброса скорости также увеличится в  $\sqrt{2}$  раз. С другой стороны, при фиксированном значении начальной скорости и точности броска, допустимый интервал разброса значений угла бросания  $\alpha_0$  максимальный (около  $10^\circ$  при  $\delta = 5\%$ ) для  $v_0 = 1$  и сужается с увеличением  $v_0$  (всего  $1,7^\circ$  при  $v_0 = 1,4$ ). Данное обстоятельство позволяет заключить, что броски под углом, близким к  $45^\circ$ , более предпочтительны, чем броски при больших значениях угла.

В заключении оценим параметры броска для реальных размеров, оговоренных в международных правилах для баскетбола. В соответствии с принятыми в баскетболе правилами длина окружности мяча может изменяться от 0,749 до 0,780 м, что соответствует изменению радиуса мяча от 0,119 до 0,124 м. Масса мяча должна лежать в пределах от 0,567 до 0,650 кг. Примем для определенности радиус мяча равным  $R=0,12$  м, а массу мяча  $m=0,6$  кг. Внутренний радиус кольца корзины равен  $r_k=0,225$  м. В соответствии с высказанными ранее допущениями об условиях попадания мяча в корзину допустимое отклонение траектории мяча от центра корзины составит  $\Delta_{\max}=r_k-R/2=0,165$  м. Расстояние от линии штрафных бросков до центра кольца корзины равно 4,225 м. Принимая данную величину в качестве  $x_k$ , можно определить значение предельной относительной точности броска  $\delta_{\max}=\Delta_{\max}/x_k=0,039$  или 3,9%. Используя соотношение (2.19) и найденную величину предельной относительной точности броска, можно оценить начальную скорость мяча при штрафном броске и допустимый интервал ее изменения при угле бросания в  $45^\circ$ :  $v_0=6,44\pm 0,25$  м/с.

Интересно также выполнить оценку параметров "трехочкового" броска, который выполняется из-за линии, расположенной на расстоянии 6,25 м от центра кольца корзины. Предельная относительная точность броска в этом случае не более 2,64%. Ве-

личина начальной скорости при угле бросания в  $45^\circ$  равна соответственно  $v_0=7,83\pm 0,21$  м/с.

В таблице 2.1 для угла бросания  $45^\circ$  представлены параметры броска в зависимости от расстояния до кольца. Из приведенных результатов наглядно видно, что при увеличении дальности броска на 10 м начальная скорость броска возрастает более чем в 3 раза. При этом допустимое отклонение в значении начальной скорости уменьшается более чем в 3 раза.

Таблица 2.1 – Параметры броска в зависимости от расстояния до кольца

|                     |      |      |      |      |      |       |
|---------------------|------|------|------|------|------|-------|
| $x_k$ , м           | 1    | 3    | 5    | 7    | 9    | 11    |
| $\delta$ , %        | 16,5 | 5,5  | 3,3  | 2,3  | 1,8  | 1,5   |
| $v_0$ , м/с         | 3,13 | 5,42 | 7,00 | 8,29 | 9,40 | 10,39 |
| $\pm\Delta v$ , м/с | 0,52 | 0,30 | 0,23 | 0,20 | 0,17 | 0,16  |

## ГЛАВА 3. СТРУКТУРНЫЕ МОДЕЛИ

### *Введение*

Как было отмечено во 2-ой главе, одним из важных этапов математического моделирования является переход от содержательной (технической) и концептуальной постановки задачи к математической, т.е. переход с технического языка (или – "языка природы") описания исследуемого объекта к языку математических формул и уравнений. Однако на практике такой переход бывает совсем не просто осуществить. В некоторых случаях это связано со сложностью самого исследуемого объекта, в других – с отсутствием необходимой информации об объекте или соответствующего математического аппарата. Поэтому часто возникает необходимость "разбить" поставленную задачу на несколько более простых подзадач, имеющих известные решения, или которые можно решить с помощью апробированных методов. Для этой цели удобно использовать методы структурного моделирования, позволяющие еще на стадии постановки упростить решаемую задачу путем исследования внутренней структуры рассматриваемого объекта, изучения свойств отдельных элементов объекта и связей между ними.

### **3.1 Что такое структурная модель?**

Системой является все, что мы хотим различать как систему.

Б.Гейне

Очень часто для достижения практических целей возникает необходимость рассматривать исследуемый объект не как нечто целое, а как совокупность отдельных элементов, связанных (взаимодействующих) между собой некоторым образом. В этом случае исследуемый объект удобно представить в виде *системы*, а при его моделировании использовать методы *системного анализа*.

Напомним основные понятия системного анализа, которые будут использоваться в дальнейшем. Одним из основополагаю-

щих понятий системного анализа является понятие искусственной системы, которую определим следующим образом:

*Система* есть совокупность взаимосвязанных элементов, выделенная из среды и взаимодействующая с окружающей средой как целое для достижения поставленной цели.

Следует отметить, что важным признаком для выделения системы из среды является возможность определения взаимодействия этой системы с окружением независимо от поведения ее отдельных элементов (именно это подразумевается под словами "взаимодействующая: как целое").

Для описания систем в системном анализе рассматриваются четыре основных модели. Если внутреннее строение системы неизвестно (или не интересует исследователя), то применяется модель "черного ящика". В данной модели системы отсутствуют (или не используются в явной форме) сведения о внутреннем содержании "ящика" (поэтому он и называется "черным"), а только задаются входные и выходные связи со средой. Обычно это сводится к заданию двух множеств входных и выходных параметров, но никаких соотношений между ними не задается. Примером модели "черного ящика" может служить экспериментальное исследование некоторого сложного объекта, когда экспериментатор, изменяя входные параметры объекта, получает на выходе различные его характеристики.

Очевидно, что исследование внутреннего устройства системы невозможно с помощью модели "черного ящика". Для этого необходимы более развитые модели. Одной из таких моделей является модель *состава системы*, описывающая, из каких элементов и подсистем состоит данная система. При этом *элементами* системы называются те части системы, которые полагаются неделимыми, а части системы, состоящие более чем из одного элемента, называются *подсистемами*. Например, если в качестве системы рассмотреть автомобиль, то ее подсистемой можно считать систему управления, а ее элементами – руль, педали и т.д.

Сложность построения модели состава системы состоит в ее неоднозначности. Это связано со следующими причинами. Во-первых, понятие "элементарности" можно определить по-разному. Во-вторых, модель состава (как и любая другая модель)

является целевой и для отличающихся целей один и тот же объект может потребовать различного разбиения на части. В третьих, всякое разбиение целого на части является относительным. Например, тормозную подсистему автомобиля можно отнести как к ходовой части, так и к подсистеме управления.

В большинстве случаев модели состава системы оказывается недостаточно для ее описания. Мало знать состав системы, кроме этого необходимо установить связи между отдельными элементами, которые называются *отношениями*. Совокупность необходимых и достаточных для достижения цели отношений между элементами называется *моделью структуры системы*. Основной сложностью при описании структуры (списка отношений) является обоснование конечного числа связей, которые являются наиболее существенными по отношению к рассматриваемой цели.

Например, при моделировании механической системы, движущейся в околоземном пространстве, обычно не учитываются силы взаимного притяжения отдельных материальных точек (элементов), но учитывается сила притяжения их к Земле (отношения).

Следует отметить, что структура системы является абстрактной моделью, так как рассматривает только связи (отношения) между элементами, но не рассматривает сами элементы (понятно, что на практике говорить об отношениях без элементов просто не имеет смысла). Однако в некоторых случаях модель структуры теоретически может быть исследована отдельно, если, например, отношения заданы в виде математических формул или уравнений.

Теперь, имея три формальные модели системы: "черного ящика", состава и структуры, – и объединив их, можно получить еще одну модель, которую называют *структурной схемой* системы или моделью "белого ящика". Данная модель включает все элементы системы, все связи между элементами внутри системы и связи системы (или ее отдельных элементов) с окружающей средой (входы и выходы системы), как изображено на рисунок 3.1.

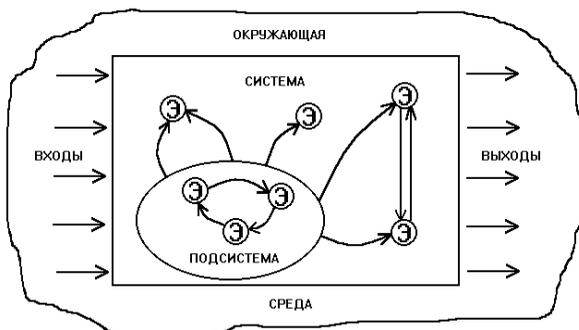


Рисунок 3.1 – Структурная схема системы

Следует отметить, что структурная схема системы является формальной моделью, отделенной от содержательного наполнения. Это позволяет рассматривать структурную схему как особый математический объект и исследовать его свойства. Такой объект называется *графом*. Он состоит из обозначений элементов произвольной природы, называемых *вершинами*, и обозначений связей между ними, называемых *ребрами*.

На рисунке 3.2 приведен пример графа, у которого вершины обозначены в виде кружочков, а ребра – в виде линий. Стрелки указывают на несимметричность некоторых связей. Такой граф называется *ориентированным*. Каждая пара вершин может быть соединена любым количеством ребер. Чтобы ввести другие различия между ребрами – кроме несимметричности, – им приписывают различные веса. Такие графы называются *взвешенными*. В качестве весов могут выступать различные характеристики сети, например, длина ребра, число каналов электросети, тип покрытия в сети автомобильных дорог и так далее. В настоящее время для графов разработана целая теория, имеющая многочисленные приложения. Графы могут изображать любые структуры. Некоторые графы получили специальные названия: линейные, древовидные (иерархические), сетевые, матричные и т.д. Пример сетевого графа приведен на рисунке 3.3.

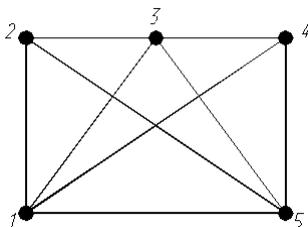
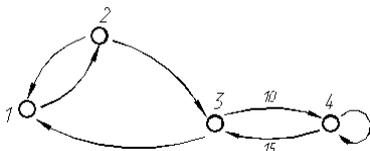


Рисунок 3.2 – Пример графа    Рисунок 3.3 – Сетевой граф

Как отмечено выше, граф (структурная схема) является формальной моделью, которую необходимо наполнить конкретным содержанием. Только после этого структурная схема становится *структурной моделью* исследуемого объекта. Например, если в качестве вершин графа, изображенного на рисунке 3.3, считать цилиндрические шарниры, а ребер – прямолинейные стержни, то получим структурную модель стержневой конструкции, широко применяемую при моделировании в строительной механике. Если же в качестве вершин рассматривать узлы связи, а за ребра принять линии связи, то получим структурную модель сети электро-связи.

Таким образом, можно дать следующее определение структурной модели.

*Структурная модель системы* – это совокупность конкретных элементов данной системы, необходимых и достаточных отношений между этими элементами и связей между системой и окружающей средой.

Рассмотрим пример построения структурной модели, поясняющий данное определение. Пусть требуется построить структурную модель абсолютно твердого тела, совершающего поступательное движение под действием приложенной силы. Напомним, что абсолютно твердое тело при движении не изменяет форму и размеры. Поэтому такое тело можно представить как совокупность материальных точек (элементов), соединенных прямолинейными невесомыми недеформируемыми стержнями (рисунок 3.4).

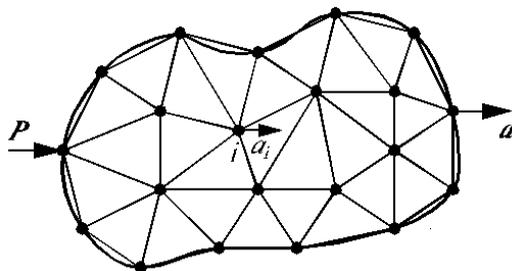


Рисунок 3.4 – Структурная модель абсолютно твердого тела при поступательном движении

Сила  $\mathbf{P}$  выступает в качестве воздействия внешней среды на тело, а откликом на это воздействие (выходным параметром) служит ускорение тела  $\mathbf{a}$  (или любой его точки вследствие их равенства при поступательном движении). Согласно второму закону Ньютона

$$\mathbf{a} = (m_T)^{-1} \mathbf{P}, \quad (3.1)$$

где  $m_T$  – масса тела.

Таким образом, для того, чтобы данная структурная модель правильно описывала поступательное движение тела, необходимо и достаточно, чтобы выполнялось следующее условие

$$\sum_{i=1}^N m_i = m_T, \quad (3.2)$$

где  $m_i$  – масса  $i$ -ого элемента,  $N$  – число элементов.

Из последнего соотношения вытекают условия на отношения между элементами структурной модели, из которых следует, что количество элементов и их распределение внутри тела не имеют значения. Другое дело, если необходимо описать вращательное движение твердого тела вокруг некоторой заданной оси под действием приложенного момента сил. В этом случае распределение элементов внутри объема тела должно быть таким, чтобы выполнялось условие равенства моментов инерции реального тела и его структурной модели относительно заданной оси

$$I_z = \sum_{i=1}^N m_i r_i^2, \quad (3.3)$$

где  $I_z$  – момент инерции тела относительно оси  $Z$ ,  $r_i$  – расстояние  $i$ -ой точки до оси  $Z$ .

Из этого равенства вытекают необходимые и достаточные условия на отношения между элементами данной структурной модели (длины и ориентации стержней, массы материальных точек), что хорошо видно из рисунка 3.5.

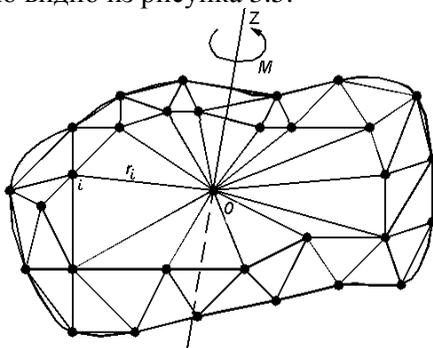


Рисунок 3.5 – Структурная модель вращающегося тела

Еще больше усложнится структурная модель в случае описания движения деформируемого (например, упругого) тела. Тогда вместо недеформируемых стержней в качестве связей между элементами (материальными точками) могут выступать пружинки с различными упругими свойствами (рисунок 3.6).

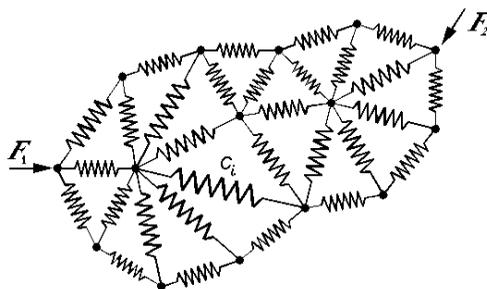


Рисунок 3.6 – Структурная модель упругого тела

В этой структурной модели для получения необходимых и достаточных отношений между элементами кроме распределения масс по объему тела необходимо учесть распределение жестко-

стей  $c_j$  пружинок, чтобы совокупность последних описывала упругие свойства реального тела.

Из этого примера видно, что структурное моделирование позволяет описывать поведение довольно сложных систем. При этом, чем сложнее система, тем структурное моделирование становится все более эффективным (а в некоторых случаях – просто необходимым). Например, при описании поступательного движения твердого тела структурная модель, в принципе, не нужна, так как в этом случае тело можно представить в виде одной материальной точки, масса которой равна массе тела. В других рассмотренных выше примерах структурная модель помогает описать поведение достаточно сложных механических систем с помощью взаимодействующих простейших элементов. Для некоторых механических систем структурное моделирование является едва ли не единственным способом описания их поведения. Это относится, например, к моделированию поведения структурно-неоднородных материалов (композитов, полимеров, керамик и т.д.).

Следует отметить, что той информации, которая содержится в структурной схеме системы, очень часто бывает недостаточно для исследования. Поэтому графы надо рассматривать как вспомогательный инструмент при моделировании. Главным же при структурном моделировании является установление конкретных функциональных связей между входными, внутренними и выходными параметрами. Поэтому без достаточно широкого арсенала методов математического моделирования здесь также не обойтись. В большинстве случаев приходится составлять уравнения, описывающие поведение каждого элемента структуры с учетом взаимодействия всех элементов. Это приводит к необходимости решения систем уравнений с большим количеством переменных. Однако с появлением современной вычислительной техники такие подходы к моделированию сложных систем становятся все более распространенными.

В заключение данного параграфа кратко остановимся на классификации структурных моделей и примерах их применения. Структурные модели бывают 4-х видов: пространственные, временные, физические и иерархические. *Пространственные*

*структуры* обычно используют для описания геометрии исследуемого объекта и расположения в пространстве его отдельных элементов. Такие структуры хорошо описываются с помощью сетевых и матричных графов, вершины которых указывают места расположения элементов, а ребра – расстояния между ними или другие условия соединения. Примером пространственной структурной модели является структурная схема телефонной сети некоторого населенного пункта, где вершинами являются узлы связи, а ребрами – линии связи с указанием, например, числа каналов связи. Другим примером пространственной структуры может служить конечно-элементная сетка, с помощью которой описывается геометрия исследуемого объекта при его, например, прочностном расчете методом конечных элементов.

*Временные структурные модели* широко используются в сетевом и календарном планировании, а также в теории массового обслуживания. Во временных структурах в качестве элементов выступают этапы происходящего процесса или состояния системы в некоторый момент времени. Отношениями здесь служат условия перехода от одного этапа к другому или из одного состояния системы в другое. Например, на производстве широко применяют так называемые сетевые графики (технологические карты). Они представляют собой графы, вершинами которых служат необходимые производственные операции, а с помощью ребер указывается последовательность и длительность этих операций. При моделировании систем массового обслуживания удобно применять структурную схему "гибели и размножения", которая представляет собой линейный граф (последовательный набор состояний системы, вытянутой в одну цепочку). Считается, что система обслуживает случайный поток заявок, поступающих в данную систему. Тогда отношениями между элементами системы (ее состояниями в различные моменты времени) служат условия поступления новой заявки, которые можно характеризовать, например, интенсивностью соответствующих случайных потоков событий.

*Физические структурные модели* применяются для описания сложных физических свойств исследуемого объекта с помощью простых структурных элементов. Пример моделирования упру-

гих свойств тела был приведен выше. Другие примеры применения физического структурного моделирования будут приведены в конце данной главы.

Для моделирования управляемых систем широко применяются *иерархические структурные схемы*, предполагающие наличие нескольких уровней обработки информации и принятия решений. Основная задача иерархической структуры – распределение функций обработки информации и принятия решений между отдельными элементами. На рисунке 3.7 приведен пример двухуровневой веерной иерархической структуры управления системой. В данной системе существует один привилегированный элемент, который имеет возможность управлять остальными элементами. Этот привилегированный элемент обычно называется *центром* (Ц), а остальные элементы – *производителями* ( $\Pi_i$ ,  $i=1 \div n$ ).

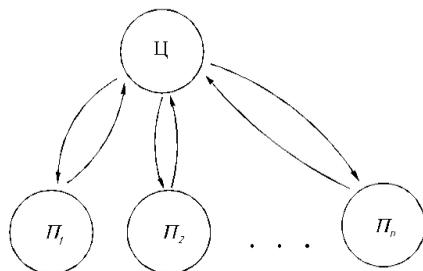


Рисунок 3.7 – Двухступенчатая веерная структурная схема

Отношениями в данной модели служат условия обмена информацией, денежными и материальными ресурсами между центром и производителями. Следует отметить, что в приведенной структурной схеме отношения между производителями (горизонтальные связи) отсутствуют. Неравноправие элементов системы проявляется в том, что центр назначает правила формирования воздействий на производителей, и тем самым имеет возможность направлять в нужное для него русло действия нижних элементов. Данную схему несложно обобщить до многоуровневой веерной структуры, которая широко используется в экономике.

В заключение данного параграфа можно отметить следующее. Как было показано, структурные модели нашли широкое

применение в различных областях целенаправленной деятельности человека. В некоторых случаях они помогают построить модель исследуемой системы или процесса, а иногда – являются единственно эффективным инструментом при моделировании.

### **3.2 Способы построения структурных моделей**

Не упрощайте жизнь, она и так проста: В ней всюду и всему отведены места. Не усложняйте жизнь, она и так сложна: Чтоб разобраться в ней, вторая жизнь нужна.

В. Кривилев

При структурном моделировании широко применяются методы анализа и синтеза. С помощью методов анализа производится разделение рассматриваемого объекта на части и исследование каждой из этих частей в отдельности. Методы синтеза, наоборот, служат для соединения частей в целое. Следует отметить, что при структурном моделировании методы анализа и синтеза необходимо применять совместно. Важно не просто разбить целое на отдельные элементы, но и соединить эти элементы таким образом, чтобы они снова образовали единое целое. С этой точки зрения синтез является завершающим этапом анализа, так как только после этого этапа можно объяснить целое через его части – в виде структуры целого. Это связано с тем, что при анализе теряются важные свойства объекта как целого (разобранный автомобиль не поедет), так и отдельных его элементов (оторванный руль "не рулит"). Поэтому, как отмечал один из ведущих специалистов по системному анализу Р. Акофф, результатом анализа является лишь вскрытие структуры системы, знание о том, как система работает, но не понимание того, почему и зачем она это делает. Только после синтеза можно объяснить поведение системы, рассматривая каждый элемент и его роль через призму всей системы.

Таким образом, анализ и синтез нельзя рассматривать как отдельные методы. Они дополняют друг друга и при структурном

моделировании должны применяться совместно. Только в этом случае построенная структурная модель будет отражать основные свойства исследуемого объекта согласно поставленным целям.

Рассмотрим теперь методы анализа и синтеза более подробно. Начнем с методов анализа, которые нашли широкое применение в науке и практике. В математике давно с успехом применяются такие аналитические методы, как разложение функций в ряды, спектральный анализ, дифференциальное и интегральное исчисление; в физике – методы молекулярной динамики; на производстве – конвейерная технология изготовления.

Как было отмечено выше, основной операцией при анализе является разделение целого на части. В дальнейшем эту операцию будем называть *декомпозицией*, и понимать под ней метод разложения системы на отдельные элементы. В результате декомпозиции исходная система распадается на подсистемы, задача – на подзадачи и т.д. При необходимости операция декомпозиции может повторяться несколько раз, что приводит к древовидным структурам системы. Основной проблемой при декомпозиции является ее неоднозначность. Понятно, что одну и ту же систему можно разбить на различные подсистемы в зависимости как от опыта исследователя, так и от применяемой методики анализа. Поэтому в системном анализе существуют специальные критерии для обоснования процесса декомпозиции. Одним из таких критериев являются полнота декомпозиции. В свою очередь, полнота декомпозиции связана с полнотой модели системы, взятой в качестве исходной при декомпозиции. Как было указано в предыдущем параграфе, в системном анализе рассматривается только четыре формальные модели системы: "черного ящика", состава, структуры и "белого ящика". Очевидно, что для декомпозиции подходят только последние три модели, которые позволяют рассматривать систему через взаимодействие ее отдельных элементов. Поэтому одной из проблем системного анализа является накопление наборов полных формальных моделей для различных исследуемых систем, которые в теории искусственного интеллекта получили название *фреймов*. Известны фреймы для некоторых информационных, организационных и социальных систем. Одна-

ко построение и обоснование фреймов для произвольной системы остается сложной задачей для исследователя, от решения которой во многом зависит успех декомпозиции. Рассмотрим, например, педагогический процесс в высшем учебном заведении (вузе). В качестве фрейма при декомпозиции в данном случае может быть взята формальная модель деятельности человека, предложенная еще К. Марксом в "Капитале" для анализа процесса труда (рисунок 3.8). В качестве элементов здесь выделены субъект деятельности, объект, на который направлена деятельность, и средства, используемые в процессе деятельности, а также все возможные связи между ними и окружающей средой. Используя эту формальную модель, можно построить модель педагогического процесса в вузе, вариант которой изображен на рисунок 3.9. Здесь в качестве субъекта выступает преподаватель вуза, объекта – студенты, а средств – методические, информационные и технические средства обучения (учебные программы изучаемых предметов, методические и учебные пособия, лабораторная база и т.д.). Окружающая среда описывается с помощью трех элементов: школа, вуз и министерство, которые оказывают существенное влияние на организацию учебного процесса (понятно, что число элементов, входящих в приведенную структурную модель, может быть гораздо больше).

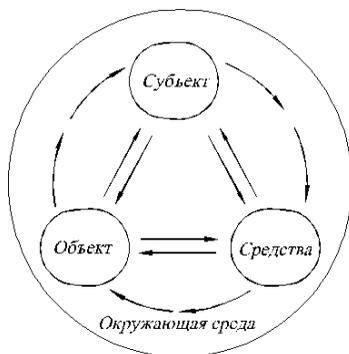


Рисунок 3.8 – Полная формальная модель деятельности человека



Рисунок 3.9 – Формальная модель педагогического процесса в вузе

Однако при декомпозиции необходимо учитывать и другой критерий – простоты, который требует сокращения размеров древовидной структуры. Таким образом, при декомпозиции должен быть принят некий компромисс между полнотой и простотой, который может быть достигнут в том случае, если в структурную модель включаются только элементы, существенные по отношению к цели анализа.

Необходимое число уровней декомпозиции (уровней древовидной структуры) выбирается из следующих соображений. Декомпозиция по каждой из ветвей древовидной структуры ведется до тех пор, пока не приведет к получению элементов системы, не требующих дальнейшего разложения. Такие составляющие называются элементарными. Необходимо отметить, что понятие элементарности составляющей системы должно быть конкретизировано в каждом рассматриваемом случае отдельно. При этом могут быть использованы как формализованные (с помощью критериев), так и неформализованные (с помощью экспертов) способы. Например, в некоторых случаях многомерную задачу механики сплошной среды удастся разложить на последовательность одномерных задач, имеющих простое (аналитическое) решение. Тогда

каждую из этих одномерных задач можно считать элементарной частью исходной системы.

Часть системы, которую нельзя считать элементарной, подлжит дальнейшей декомпозиции. При этом могут использоваться различные фреймы. Если исследователь "перебрал" все фреймы, но не достиг элементарности на какой-то ветви древовидной структуры, то вводятся новые элементы в модель, взятой в качестве основания, и декомпозиция продолжается по ним. Подробная блок-схема алгоритма декомпозиции приведена в работе.

Следует отметить, что в результате декомпозиции будет реализован только первый этап структурного моделирования, а именно – этап анализа. После этого этапа удастся разделить исследуемую систему на отдельные элементы (или исходную задачу – на более простые подзадачи). Однако, как было отмечено в начале данного раздела, на поведение каждого элемента необходимо смотреть с точки зрения целей всей системы. Другими словами, полученная совокупность элементов кроме внешней целостности (т.е. определенной обособленности от окружающей среды) должна обладать и внутренней целостностью. Внешняя целостность хорошо описывается моделью "черного ящика", а внутренняя целостность связана с моделью структуры системы, то есть установлением отношений между элементами. Для этого используется операция *агрегирования* – объединение нескольких элементов в единое целое. Результатом агрегирования является система, которую называют *агрегатом*. Необходимо отметить, что свойства агрегата не являются только совокупностью свойств его отдельных элементов. Агрегат может обладать такими свойствами, которых нет ни у одного из его элементов, взятых в отдельности. Другими словами, объединение элементов в систему влечет появление нового качества, которое не могло появиться без этого объединения. Такое "внезапное" появление новых качеств у агрегата получило название *эмерджентности* (от англ. *emergent* – внезапно возникающий). Следует отметить, что новые свойства возникают благодаря конкретным связям между элементами. Другие связи могут дать другие новые свойства агрегата.

Хорошей иллюстрацией свойства эмерджентности является пример, предложенный М. Арбибом. Пусть имеется некоторый цифровой автомат  $S$ , увеличивающий на 1 любое целое число, поступающее на его вход. При последовательном соединении двух автоматов в цепочку это свойство не изменяется. Если же соединить два таких автомата последовательно в кольцо (рис. 4.11), то в полученном агрегате обнаружится новое свойство: он генерирует возрастающие последовательности на выходах  $A$  и  $B$ , причем одна последовательность состоит из четных, а другая – из нечетных чисел. Другим ярким подтверждением свойства эмерджентности может служить пример из материаловедения. Известно, что тип кристаллической решетки (способ соединения атомов) определяет твердость материала. При этом твердость получаемого агрегата, состоящего из одинаковых элементов, может различаться в десятки раз (графит и алмаз).

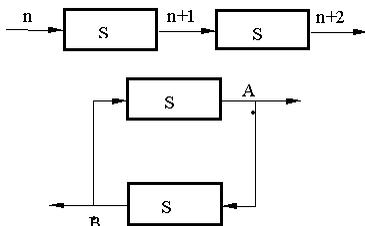


Рисунок 3.11 – Пример вычислительного агрегата

Возникновение качественно новых свойств при агрегировании есть частное, но яркое проявление одного из законов диалектики – закона перехода количества в качество. При этом считается, что чем больше свойства агрегата отличаются от свойств его элементов, тем выше организованность системы. Кибернетик У. Эшби доказал, что у системы тем больше возможностей в выборе поведения, чем сильнее степень согласованности поведения ее элементов. Высшая степень проявления согласованности поведения элементов системы – самоорганизация системы, изучением которой занимается относительно молодая междисциплинарная область знаний – *синергетика* (от греч. *synergos* – вместе действующий).

Таким образом, как следует из выше изложенного, при агрегировании большое значение имеет установление связей между элементами, то есть выбор модели структуры.

Значит, в самом общем виде агрегирование можно определить как установление отношений на заданном множестве элементов. Такое установление отношений может быть проведено различными способами: построением математических зависимостей, структурированием, статистической обработкой, классификацией и т.п. В результате получают различные агрегаты, основными из которых являются следующие: *конфигуратор*, *классификатор*, *оператор*, *статистик* и *структура*. Рассмотрим эти агрегаты более подробно.

*Конфигуратором* называется такой агрегат, который состоит из качественно различных языков описания системы и обладает тем свойством, что число этих языков минимально, но необходимо для выполнения заданной цели. Следует отметить, что конфигуратор является содержательной моделью высшего возможного уровня. Перечислив языки, на которых будет вестись описание системы, мы тем самым определяем тип системы и ее основные свойства.

Например, в радиотехнике для описания одного и того же прибора используется следующий конфигуратор: блок-схема, принципиальная схема и монтажная схема. Этот конфигуратор полностью описывает рабочие характеристики прибора. Однако, если кроме цели производства радиоаппаратуры, ставится цель ее сбыта, то в конфигуратор необходимо добавить язык рекламы (маркетинг, дизайн, цена и т.п.).

В инженерной графике для описания поверхности любого трехмерного тела в качестве конфигуратора используются совокупность трех ортогональных проекций. Число их нельзя уменьшить и нецелесообразно увеличивать.

А какой конфигуратор применяется при математическом моделировании? Выбор языка зависит от вида модели. Понятно, что основным языком для математической модели является язык математических формул. Однако, как было отмечено в первых двух главах данного учебного пособия, важными этапами математического моделирования являются содержательная и концептуальная

постановки задачи. На этапе содержательной постановки осуществляется словесная постановка задачи на том языке, на котором она формулируется заказчиком. На этапе концептуальной постановки выполняется запись задачи на языке тех областей знаний, которые используются при моделировании рассматриваемого объекта. Поэтому конфигуратором в данном случае можно считать содержательную, концептуальную и математическую постановку задачи.

В качестве *классификатора* выступает агрегат, устанавливающий отношения эквивалентности между элементами системы, то есть описывает условия образования классов.

Говорят, что на множестве  $A$  определено отношение  $R$ , если по некоторому правилу составлены упорядоченные пары элементов, находящихся в отношении. При этом пишут  $a R b$ ,  $a, b \in A$ .

Если отношение  $R$  на  $A$  удовлетворяет следующим аксиомам:

1.  $a R a$
2.  $a R b \implies b R a$
3.  $a R b, b R c \implies a R c$ ,

то оно называется отношением эквивалентности. Отношение эквивалентности разбивает множество элементов системы на классы.

Следующим типом агрегата является *оператор*, который ставит в соответствие некоторому набору отдельных элементов один элемент. Одним из наиболее распространенных в математическом моделировании видов оператора является *функция*. Этот вид оператора возникает, если агрегируемые элементы измеряются в числовых шкалах. Тогда появляется возможность задать отношение на множестве элементов в виде числовой функции многих переменных  $f$ , которая и является агрегатом, то есть

$$f: R^n \rightarrow R, \quad (3.4)$$

где  $R^n$  –  $n$ -мерное евклидово пространство,  $R$  – вещественная ось.

Здесь элементом отображаемого пространства  $R^n$  является  $n$ -мерный вектор переменных системы  $x=(x_1, x_2, \dots, x_n)$ , характеризующий ее поведение. Приведенный вид функции является одним из простейших. В общем случае, области определения и значений функции, могут относиться к более сложным простран-

вам и множествам. Конкретное задание функции  $f(x)$  связано с построением математической модели рассматриваемой системы. Поэтому на выбор функции накладываются ограничения, вытекающие из содержательной постановки задачи, то есть этот выбор не является свободным. В тех же (достаточно редких) случаях, когда оператор-функция является вполне адекватной математической моделью всей системы, свобода выбора функции, агрегирующей набор внутренних переменных, вообще отсутствует. Такой случай имеет место, например, когда закономерности природы удается с достаточной степенью адекватности отобразить безразмерными степенными одночленами физических размерных величин. В этом случае можно утверждать, что если удалось построить безразмерный степенной одночлен из размерных физических величин, образующих конфигурацию данного явления, то установлен физический закон данного явления. Это легко можно показать на примере второго закона Ньютона, описывающего поступательное движение твердого тела. Безразмерный одночлен здесь имеет вид

$$m\tau a/F = 1, \quad (3.5)$$

что подтверждает правомерность полученного закона.

К сожалению, построить функциональную зависимость, адекватно описывающую поведение сложной системы, очень трудно, а в некоторых случаях – практически невозможно. Гораздо проще установить функциональные зависимости между отдельными элементами системы. В этом случае оператор будет представлять собой некоторую (часто-нелинейную) систему уравнений. Во многих случаях внутренними переменными системы являются не числа, а функции одного или нескольких аргументов. Тогда выходными параметрами могут выступать также функции или функционалы. Например, для динамических систем, описывающих процессы и явления, изменяющиеся во времени, связь между внутренними параметрами  $x(t)$  и выходными параметрами системы  $y(t)$  в операторной форме имеют вид

$$y(t)=f(x(t)), t \in [0, T], \quad (3.6)$$

где оператор  $f$  обычно представляет собой систему дифференциальных уравнений,  $T$  – время протекания процесса. Пример

построения подобного оператора для механической системы, состоящей из нескольких тел, будет приведен в следующем параграфе.

При математическом моделировании сложных систем построить оператор  $f$  бывает совсем не просто. Это связано со многими причинами. Основной из этих причин можно считать недостаток информации о характере и механизмах взаимодействия между отдельными элементами системы. Например, эти взаимодействия могут носить случайный характер, закон которого нам не известен. В этом случае говорят, что моделирование ведется в условиях неопределенности, а оператор  $f$  может быть определен только с точностью до конечного числа параметров  $q=(q_1, q_2, \dots, q_n)$

$$y(t) = f(x(t), q), t \in [0, T]. \quad (3.7)$$

Обычно считается, что параметры  $q$  носят случайный характер или могут быть определены в ходе самого моделирования с помощью методов идентификации. В этом случае используются системы с обратной связью. Более подробно о способах моделирования в условиях неопределенности будет сказано в следующей главе. В тех случаях, когда оператор задается с помощью алгоритма, реализующего некоторый набор правил, агрегат будем называть *имитатором*.

Отдельно при агрегировании рассматривается случай, когда все параметры, описывающие поведение элементов системы, являются случайными величинами. Тогда вводится понятие *агрегата-статистика*, определяющего отношения на множестве случайных параметров системы. Для его построения используются функции выборочных значений случайных величин, в качестве которых широко используются функции распределения вероятностей или плотности распределения вероятностей случайных событий. На практике используются достаточные и оптимальные статистики. Достаточными статистиками называются такие агрегаты, которые извлекают всю полезную информацию об интересующем нас параметре из совокупности наблюдений. Например, для систем, результат деятельности которых нас интересует только в среднем, достаточным статистиком может служить математическое ожидание выходной случайной величины. Однако на

практике достаточные статистики применяются редко, так как при таком агрегировании потери информации неизбежны. Поэтому чаще применяются оптимальные статистики – такие агрегаты, которые позволяют свести потери информации к минимуму. Например, совокупность математического ожидания и дисперсии случайной величины является оптимальным статистиком для многих технических систем. Более подробно способы описания случайных величин при моделировании будут рассмотрены в следующей главе.

Последним из рассматриваемых видов агрегатов, но не последним по частоте применения при структурном моделировании, является структура системы, то есть агрегат, устанавливающий типы связей между отдельными элементами системы. Наиболее широко подобный вид агрегирования применяется при моделировании технических, информационных и организационных систем. Например, в материаловедении в качестве структуры материала используются разные модели кристаллических решеток, устанавливающие типы связи между атомами. В информационных системах применяется структура в виде первичной сети, указывающей направление и интенсивность передачи информации. В организационных системах структура описывает иерархию в процессе принятия решений и ответственность за принятые решения (распределение власти и ответственности).

## **ГЛАВА 4. МОДЕЛИРОВАНИЕ В УСЛОВИЯХ НЕ-ОПРЕДЕЛЕННОСТИ**

### ***Введение***

Нет ничего более противного разуму и природе, чем случайность.

Цицерон

Развитие современных технологий во всех сферах деятельности человека и разработка соответствующих моделей приводят к необходимости учета максимально возможного количества информации об исследуемом объекте. При этом вопросы моделиро-

вания сложных процессов и явлений чаще всего формулируются и обсуждаются на профессиональном языке (искусственном или подмножестве естественного), отражающем специфику области исследования. Следствием этого является использование в процессе моделирования качественных элементов: описаний, понятий и отношений с неопределенными или нечеткими границами, высказываний с многозначной шкалой истинности и тому подобное.

В тех случаях, когда информация существенно неопределенная, задание строгих границ "волевым" порядком или искусственное введение однозначности означает не что иное, как огрубление исходных данных, и может приводить к получению пусть четкого, но неверного результата. Исследование и учет однозначности (определенности) или неоднозначности (неопределенности) всех параметров и отношений, описывающих исследуемое явление, представляет собой необходимый и важнейший элемент математического моделирования.

#### **4.1 Причины появления неопределенностей и их виды**

- ...- Все неверно, – сказала Гусеница.
  - Да, не совсем верно, – робко согласилась Алиса, – Некоторые слова не те.
  - Все не так, от самого начала и до самого конца, – строго проговорила Гусеница.
- Льюис Кэррол

Известные закономерности, описывающие процессы и явления объективного мира, можно условно разделить на две группы: *однозначно определенные (детерминированные)* и *находящиеся в условиях неопределенности*.

К первой группе относят те, которые по заданным с определенной точностью характеристикам воздействий позволяют установить вполне определенный (детерминированный) отклик (реакцию) исследуемого объекта. Например, материальная точка падает с некоторой высоты. При заданной точности определения

начальных условий и действующих на точку внешних сил можно однозначно, с определенной точностью установить ее скорость при соприкосновении с Землей, время полета и так далее. С математической точки зрения эти закономерности описываются на основе аксиом традиционной математикой с использованием вполне определенных величин.

Вторая группа закономерностей описывает *случайные события* (такие, которые при заданном комплексе условий могут протекать по-разному при одних и тех же условиях). Например, при бросании игрального кубика нельзя заранее однозначно сказать, какая цифра выпадет. Если попытаться учесть природу этих закономерностей как явлений, находящихся в условиях неопределенности, то нужно иметь в виду, что описание этой неопределенности может быть разным в зависимости от количества и качества имеющейся информации.

Часто граница, отделяющая случайное событие от неслучайного, очень размытая. Одна из концепций случайности (которая преобладала до начала XX столетия) состояла в том, что если при описании исследуемого объекта предусмотреть все связанные с ним "детали", то никакой случайности не будет. Однако в настоящее время принято придерживаться другой концепции. Вернемся к примеру с падением материальной точки. При полете последней необходимо учитывать температуру окружающей среды, скорость ветра, положение относительно поверхности Земли и другие факторы, которые имеют неоднозначный характер и могут, в свою очередь, влиять друг на друга. Поэтому в "чистом виде" однозначно определенных процессов (явлений), наверное, нет, то есть при описании достаточно сложных процессов закономерности всегда носят стохастический характер.

При решении задач математического моделирования (задачи проектирования, описания различных технологических процессов, выбора оптимальных параметров и так далее) уже на стадии концептуальной постановки необходимо задуматься над тем, насколько однозначно определены параметры, в терминах которых осуществляется математическое описание объекта моделирования. На этом этапе необходимо определить для каждого параметра, можно ли считать его однозначно определенным или ему при-

суца некоторая неопределенность. Причем неопределенными могут быть не только параметры, но и связи между ними. Неопределенность понимается в том смысле, что соответствующие характеристики рассматриваемой системы находятся в условиях приближения и неполноты информации. Эта неопределенность может быть связана, с одной стороны, с тем, что параметры могут изменяться случайным образом, а с другой стороны, с тем, что они могут адаптироваться (изменяться, устанавливаться) в процессе функционирования исследуемой системы.

К наиболее значимым причинам появления неопределенности можно отнести следующие:

- показатели системы практически всегда зависят от большого количества различных факторов, причем часть из них может быть даже неизвестна исследователю;
- при построении модели обычно ограничиваются отбором наиболее существенных (по мнению субъекта или в силу объективных обстоятельств) переменных, что, конечно, приводит к огрублению модели;
- математические погрешности, возникающие при линейаризации модели или использовании разложения в ряд при ограничении на число членов ряда; ошибки измерений и погрешности при проведении эксперимента; и тому подобное.

В общем случае все причины возникновения неопределенности можно разбить на две основные группы: *субъективные* и *объективные*. Субъективные причины обусловлены некоторыми частными, нерегулярно повторяющимися явлениями, поэтому их достаточно сложно учесть при решении прикладных задач. Объективные причины чаще всего связаны с физическими особенностями исследуемого явления.

Например, если рассматривать задачу исследования некоторого технологического процесса, то к субъективным причинам можно отнести квалификацию работников, проводящих и регламентирующих исследуемый процесс, их навыки, реакцию, время адаптации и тому подобное. К объективным причинам появления неопределенности для такого типа задач можно отнести:

- физико-механические свойства поставляемых материалов (в частности, предел текучести, модуль Юнга, коэффициенты

теплопроводности, теплоемкости, теплоотдачи и тому подобное);

- анизотропию свойств;
- поля остаточных напряжений;
- геометрические характеристики заготовок (форма и размеры);
- характер износа инструмента и так далее.

В свою очередь, каждая из указанных объективных причин появления неопределенности может быть обусловлена целым рядом предпосылок. Так, например, неоднородность свойств материала определяется, с одной стороны, как неоднородность по объему, обусловленная особенностями технологического процесса (отливки, прокат, армированные и порошковые композиты и прочие), с другой – как неоднородность партий поставляемых заготовок. При решении прикладных задач для устранения неопределенностей обычно вводится предположение о принятии в качестве физико-механических характеристик некоторых предельных или средних значений (из возможных диапазонов). На наш взгляд, подобное предположение является весьма спорным в силу нелинейности исследуемых процессов и сложного характера взаимодействия отдельных частей объекта между собой. В силу этого возникает необходимость учета распределения соответствующей неопределенной величины.

В зависимости от полноты описания неопределенность можно разбить на три основные группы: *неизвестность, недостоверность и неоднозначность*.

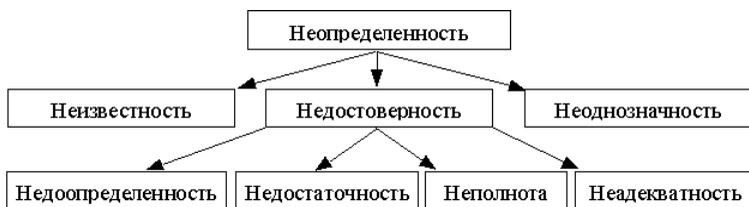


Рисунок 4.1 – Виды описания неопределенности

Рассмотрим группы описания неопределенности более подробно.

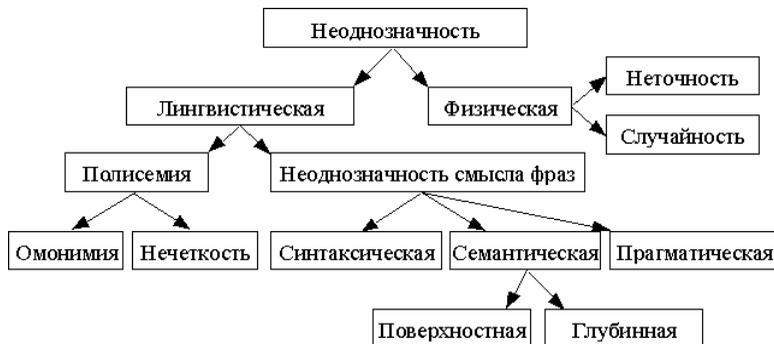
*Неизвестность* – это начальная стадия описания неопределенности, при которой информация полностью отсутствует.

*Недостоверность* – это вторая стадия описания неопределенности, которая для различных стадий сбора информации может классифицироваться как *неполнота*, *недостаточность*, *недоопределенность* и *неадекватность*. *Неполнота* характеризуется тем, что собрана не вся возможная информация; *недостаточность* – собрана не вся необходимая информация.

*Недоопределенность* – для некоторых элементов определены не их точные описания, а лишь множества, которым эти описания принадлежат; *неадекватность* – ряд элементов исследуемого объекта описан по аналогии с уже имеющимися описаниями подобных элементов, то есть имеет место так называемое "замещающее" описание, которое целям исследования не всегда удовлетворяет.

Дальнейший анализ неопределенности, учет новых факторов, определяющих исследуемое явление, может привести либо к устранению неопределенности (все элементы описаны однозначно), либо к *неоднозначности*.

*Неоднозначность* – это конечная (по полноте возможного описания) степень неопределенности, когда вся возможная информация собрана, но полностью определенное описание не получилось.



## Рисунок 4.2 – Причины возникновения неоднозначности

Причины возникновения неоднозначности могут быть *лингвистические* и *физические* (рисунок 4.2).

*Физическая неопределенность* может быть связана либо с наличием нескольких возможностей, каждая из которых *случайным* образом может стать реальностью, либо с *неточностью* вычислений или измерений. Таким образом, физическая неопределенность связана или с *физической сущностью* исследуемого явления, или с его *измеряемыми* проявлениями.

*Лингвистическая неопределенность* связана с использованием некоторого естественного языка. Она порождается, с одной стороны, множественностью значений слов (понятий и отношений) – *полисемией* (греч. *polysēma* многозначность), а с другой – *неоднозначностью смысла фраз*.

Можно выделить два вида *полисемии*: *омонимию* и *нечеткость*.

*Омонимия* (греч. *homonymia* – одноименность) характеризуется тем, что одним и тем же словом можно характеризовать различные физические объекты. Например: коса – это вид побережья, инструмент или прическа.

Если же объекты описания сходны по сути, но описывают некоторое множество понятий, то ситуацию относят к *нечеткости*. Например, понятие *несколько шагов*. Это может быть два шага, три шага, четыре шага и тому подобное

Рассматривая источники неоднозначности смысла фраз, можно выделить *синтаксическую*, *семантическую* и *прагматическую* неоднозначность.

При *синтаксической* неопределенности уточнение синтаксиса позволяет понять смысл фразы. Пример: "*казнить нельзя помиловать*" – "*казнить, нельзя помиловать*" или "*казнить нельзя, помиловать*".

*Семантическая* неопределенность бывает поверхностная и глубинная. В первом случае отдельные слова понятны, но неясен смысл фразы ("*голубые зеленые мысли яростно спят*"), во втором случае непонятны и все отдельные слова ("*глокая куздра штеко будланула бокра и курдячит бокренка*").

*Прагматическая* неопределенность связана с совместным проявлением синтаксической и семантической неопределенностей. Она чаще всего проявляется при работе с неизвестными объектами и, возможно, в неизвестной (например, языковой) среде.

Как уже отмечалось, на стадии концептуальной постановки задачи необходим детальный анализ степени неопределенности всех характеристик системы и связей между ними. При переходе к математической постановке задачи перед исследователем встает непростой вопрос: каким типом переменных описать те или иные параметры и как описать связь между этими параметрами?

Если цели исследования предполагают однозначное описание явления (процесса) или если параметры системы и связи между ними определены единственно возможным образом, то в этом случае применяется четкое описание, то есть все характеристики считаются детерминированными и связи между соответствующими четкими переменными – однозначными. В противном случае в зависимости от целей исследования и требуемой полноты описания можно использовать различные математические подходы представления неопределенностей. Отметим, что усложнение модели (например, иерархичность) также может привести к необходимости использования других типов описания переменных, характеризующих исследуемое явление.

Математически неопределенность может быть описана *стохастически, статистически, с позиций теории нечетких множеств*, а также *интервально* (рисунок 4.3).

Отмеченные формы описания неопределенности перечислены по возрастанию степени неопределенности. Рассмотрим физический смысл этих неопределенностей.

*Стохастическое описание.* Это описание используется тогда, когда неопределенные параметры имеют вероятностный (случайный) характер. Причем в этом случае необходимо, чтобы был определен закон распределения этих случайных параметров. Стохастическим описанием занимается теория вероятностей и теория случайных процессов.



Рисунок 4.3 – Формы описания неопределенности

*Статистическое описание* является, по существу, частным случаем стохастического описания. Эту форму описания применяют тогда, когда заданы только выборочные оценки каких-то характеристик случайной величины или наборы значений некоторых случайных параметров. Статистическим описанием занимается математическая статистика.

Описание с позиций *нечетких множеств*. В этом случае неопределенный параметр задается некоторым множеством возможных его значений, характеризующихся той или иной степенью принадлежности (с помощью, так называемой *функции принадлежности*) объекту, описываемому этим нечетким множеством. Функция принадлежности может принимать значения от "1" (полная принадлежность) до "0" (полная непринадлежность). Интерпретацией функции принадлежности является *субъективная мера* того, насколько полно элемент (параметр) соответствует понятию, смысл которого описывается нечетким множеством. Этим описанием занимается теория нечетких множеств.

*Интервальное описание*. Это описание можно использовать, когда неопределенные параметры заданы только диапазонами возможных значений (верхней и нижней границей), причем параметр может принимать любое значение внутри интервала и ему нельзя приписать никакой вероятностной меры. Интервальным описанием занимается интервальная математика.

Следует отметить, что зависимость математического подхода к описанию переменных от полноты имеющейся информации весьма условна. Так, например, при численной реализации тех или иных алгоритмов моделей на ЭВМ даже для детерминированных переменных (неявным образом) используется аппарат интервальных вычислений, т.к. расчеты на ЭВМ ведутся с интервальными величинами. Поэтому считать, что интервальное описание переменных "менее определенное", чем стохастическое,

наверное, нельзя. Однако с этим утверждением можно согласиться, если интервальное описание вводится уже на стадии концептуальной и математической постановок задач.

Многообразие форм описания неопределенностей приводит к различным особенностям постановки и решения соответствующих задач.

## **4.2 Моделирование в условиях неопределенности, описываемой с позиций теории нечетких множеств**

...Вряд ли можно считать, что мозг в сравнении с современными вычислительными машинами не имеет определенных преимуществ... Главное из этих преимуществ, – по-видимому, способность мозга оперировать с нечетко очерченными понятиями.

Н.Винер

При решении сложных технических, экономических, технологических, социальных и др. задач мы сталкиваемся с тем, что чем сложнее система, тем менее мы способны дать точные и в то же время имеющие практическое значение суждения о ее поведении. Такая ситуация определяется термином "*принцип несовместимости*". Следствие из этого принципа кратко можно выразить так: "*Чем глубже мы анализируем реальную задачу, тем неопределеннее становится ее решение*". Именно в этом смысле точный количественный анализ поведения сложных систем для практического исследования реальных задач, по-видимому, недостаточен. В некоторых работах предлагается подход, который опирается на предпосылку о том, что элементами исследования являются не числа, а некоторые нечеткие множества, для которых переход от "*принадлежности к классу*" к "*непринадлежности*" не скачкообразен, а непрерывен. В основе такого подхода лежит не традиционная двузначная или даже многозначная логика, а логи-

ка с нечеткой истинностью, нечеткими связями и нечеткими правилами вывода.

Этот подход имеет три отличительные черты:

1. в нем используются так называемые "лингвистические" переменные вместо числовых переменных или в дополнение к ним;
2. простые отношения между переменными описываются с помощью нечетких высказываний;
3. сложные отношения описываются нечеткими алгоритмами.

Отметим, что с математической точки зрения предложенный подход как метод описания неопределенности лежит между описанием с позиций *теории вероятностей* и *математической статистики*. В этом случае параметры системы, имеющие вероятностный, случайный характер, описываются некоторым распределением, и описанием с позиций *интервальной математики*, при котором характеристики задаются диапазонами возможных значений (верхними и нижними границами).

Подобный тип задач чаще всего имеет место в том случае, когда концептуальная постановка задачи сформулирована в виде некоторого неопределенного высказывания типа "*Если A, то B*" ( $A \Rightarrow B$ ), в котором  $A$  и  $B$  можно описать нечеткими множествами.

Прежде чем перейти к подробному обсуждению предлагаемого подхода, приведем некоторые основные положения.

*Нечеткое множество* – это математическая модель класса с нечеткими или, иначе говоря, размытыми границами.

В этом понятии учитывается возможность постепенного перехода от принадлежности к непринадлежности элемента множеству. Иными словами, элемент может иметь степень принадлежности множеству между полной принадлежностью ("1") и полной непринадлежностью ("0"). Если степень принадлежности обозначить через  $\mu$ , то  $\mu \in [0,1]$ .

Введем некоторые основные понятия и определения. С математической точки зрения нечеткое множество  $A$  можно определить следующим образом: *нечетким множеством*  $A$  в  $U$  называется совокупность пар вида  $(u, \mu_A(u))$ , где  $u \in U$ , а  $\mu_A(u)$  – функция принадлежности элементов нечеткого множества  $A$ ,  $\mu_A :$

$U \rightarrow [0,1]$ . Здесь  $U$  – некоторое множество (в обычном смысле) элементов, которое называется универсальным множеством. (Понятие "множества" считается первичным и не определяется. По существу, множество – это совокупность элементов любого вида). Для каждого элемента  $u \in U$  функция принадлежности определяет степень его принадлежности той совокупности элементов, которая формализуется данным нечетким множеством. Математически нечеткое множество определяется следующим образом

$$R = \bigcup_{(x,y) \in X \times Y} \mu_R(x,y) / (x,y).$$

Знак "/" в данном случае используется для отделения значения функции принадлежности от соответствующе элемента носителя нечеткого множества.

Например, пусть универсальное множество и совокупность функций принадлежности определены выражениями:  $U=(a, b, c, d, e, f)$ ;  $M=(0; 0,5; 1)$ .

В этом случае  $M$  – совокупность возможных значений функции принадлежности. При этом одна из возможных форм записи нечеткого множества  $A$  может быть представлена в виде:

$$A=(0,0/a; 1,0/b; 0,5/c; 0,0/d; 0,5/e; 0,0/f).$$

Рассмотрим более подробно, что понимается под функцией принадлежности. Спектр мнений по этому вопросу чрезвычайно широк. Если под функцией принадлежности понимать "меру благоприятствия" соответствующих элементов тому понятию, которое формализуется данным нечетким множеством, то в этом случае функция принадлежности отождествляется с понятием *вероятности* (определение понятия вероятности дано ниже).

Предполагается, что функция принадлежности – это некоторое "*невероятностное субъективное измерение неточности*", что она отлична от плотности вероятности и от функции распределения вероятности. Иногда под функцией принадлежности понимают возможность или полезность того или иного события.

В данной работе, под значением функции принадлежности  $\mu_A(x)$  нечеткого множества  $A$  для любого  $x \in X$  будем понимать вероятность того, что *лицо принимающее решение (ЛПР)*, отнесет элемент  $x$  к множеству  $A$ . В случае, когда  $A$  – некоторое понятие естественного языка, а  $x$  – элемент множества объектов, обозна-

чаемых понятием  $A$ ,  $\mu_A(x)$  есть вероятность того, что ЛПР использует  $A$  в качестве имени объекта. Следует отметить, что:

- приведенная интерпретация, которую будем называть вероятностной, не исключает других (в том числе невероятных);
- элемент  $x$ , как следует из определения, уже предъявлен ЛПР, а ЛПР и решает задачу отнесения элемента к нечеткому множеству  $A$ .

В качестве конкретного примера применения аппарата теории нечетких множеств для математического моделирования некоего явления рассматривается следующая задача.

### Пример

Пусть справедливым считается следующее высказывание: "Если дорога скользкая – езда опасная, в противном случае – не опасная". Необходимо определить, в каком случае езда будет более опасной: если дорога *не очень скользкая* или дорога *очень не скользкая*?

Отметим, что в качестве примера выбрано простейшее утверждение, которое легко формализуется, ответ на него почти очевиден. В реальных задачах операций может быть десятки и сотни и ответы на поставленные вопросы не так тривиальны.

Для математической постановки представленной задачи необходимо ввести ряд определений и некоторые простейшие операции с нечеткими множествами, что и делается ниже.

Носителем нечеткого множества  $A$  ( $SuppA$  или  $S(A)$ ) называется множество (в обычном смысле), определяемое как  $SuppA = \{u/u \in U, \mu_A(u) > 0\}$ .

Нечеткое отношение  $R: X \rightarrow Y$  представляет бинарное отношение нечетких множеств  $X$  и  $Y$ ;  $R$  следующим образом описывается с помощью функции принадлежности двух переменных:

$$R = \bigcup_{(x,y) \in X \times Y} \mu_R(x, y) / (x, y),$$

где функция принадлежности двух переменных в зависимости от постановки задачи показывает предпочтение или сходство элементов первого и второго нечетких множеств.

Сравнивая определение нечеткого множества с определением нечеткого отношения, можно видеть, что нечеткое отношение – это нечеткое множество с векторной базовой переменной.

В зависимости от того, для чего используются бинарные нечеткие отношения, вводятся нечеткие отношения *сходства* и нечеткие отношения *предпочтения*.

Для примера рассмотрим бинарное отношение сходства. Предположим, что

$$X = \{\text{яблоко, груша}\}, Y = \{\text{айва, апельсин}\}.$$

Сходство будем оценивать по степени сладости зрелых фруктов (функция принадлежности в данном примере выбирается субъективно; в данном случае полагается, что по "степени сладости" груша и апельсин наиболее близки, а груша и айва – наименее близки друг другу).

Бинарное нечеткое отношение *сходства* между элементами множеств  $X$  и  $Y$  можно записать в виде:

$$\text{сходство} = \{0,8/(\text{яблоко, айва}); 0,6/(\text{яблоко, апельсин}); 0,2/(\text{груша, айва}); 0,9/(\text{груша, апельсин})\}.$$

Для удобства записи в теории нечетких множеств нечеткие отношения обычно представляются в виде так называемой *матрицы отношений*; в данном случае имеем:

$$R = \begin{bmatrix} 0,8 & 0,6 \\ 0,2 & 0,9 \end{bmatrix},$$

где элемент  $R_{ij}$  равен значению функции  $\mu_R(x, y)$  для  $i$ -го элемента  $X$  и  $j$ -го элемента  $Y$ .

В реальных задачах часто приходится иметь дело со следующей ситуацией. Имеются бинарные нечеткие отношения между множествами  $X \rightarrow Y$  и  $Y \rightarrow Z$ . Необходимо установить бинарное нечеткое отношение между множествами  $X \rightarrow Z$ . Данная ситуация требует введения операции произведения отношений. Пусть  $R$  – отношение  $X \rightarrow Y$ , а  $S$  – отношение  $Y \rightarrow Z$ , тогда отношение  $X \rightarrow Z$  определится произведением  $R \circ S$ , которое в теории нечетких множеств определяется как *максиминное произведение* следующего вида:

$$R \circ S = \bigcup_{x, z \in (X, Z)} \max(\min(\mu_R(x, y), \mu_S(y, z))) / (x, z).$$

При выполнении обычного произведения матриц элемент матрицы-произведения, стоящий в  $i$ -й строке и  $k$ -м столбце, равен сумме произведений соответственных элементов  $i$ -й строки первой матрицы и  $k$ -го столбца второй матрицы. По существу, максиминное произведение определяется как обычное произведение матриц [33], где вместо операции умножения вводится  $\min$ , а вместо операции сложения –  $\max$ .

Пример вычисления максиминного произведения

Пусть

$$R = \begin{bmatrix} 0,3 & 0,8 \\ 0,6 & 0,9 \end{bmatrix}, S = \begin{bmatrix} 0,5 & 0,9 \\ 0,4 & 1,0 \end{bmatrix}, \text{ тогда } R \circ S = \begin{bmatrix} 0,3 & 0,8 \\ 0,6 & 0,9 \end{bmatrix} \circ \begin{bmatrix} 0,5 & 0,9 \\ 0,4 & 1,0 \end{bmatrix} =$$

$$\begin{bmatrix} \max(\min(0,3;0,5); \min(0,8;0,4)) & \max(\min(0,3;0,9); \min(0,8;1,0)) \\ \max(\min(0,6;0,5); \min(0,9;0,4)) & \max(\min(0,6;0,9); \min(0,9;1,0)) \end{bmatrix}$$

*Нечеткой (лингвистической) переменной* называется совокупность (кортеж) вида  $(X, U, \tilde{X})$ ,

где  $X$  – наименование нечеткой переменной;  $U = \{u\}$  – область ее определения (обычное множество);  $\tilde{X} = U, \mu_R(u) / u$  – нечеткое множество на  $U$ , описывающее числовые значения нечеткой переменной  $X$ .

Если обратить внимание на структуру наименования лингвистической переменной, то можно отметить, что в общем случае это составной термин, представляющий сочетание некоторых элементарных терминов. Эти элементарные термины можно разбить на четыре основных категории:

- первичные термины, которые являются символами специальных нечетких подмножеств, например, молодой, старый и так далее;
- отрицание **НЕ** и союзы **И**, **ИЛИ**;
- неопределенности типа: очень, слабо, более или менее и т.д.;
- маркеры (чаще всего это вводные слова).

*Пример выделения основных категорий.*

Пусть нечеткое множество описывается составным термином: "по мнению окружающих, это был не очень сильный и совсем не высокий человек.

- *первичные термины* – сильный, высокий (человек);
- *отрицание НЕ* и *союзы И*;
- *неопределенности* – очень, совсем;
- *маркеры* – по мнению окружающих.

Отрицание **НЕ**, союзы **И**, **ИЛИ**, неопределенности типа *очень, весьма, больше, меньше* и другие термины, которые входят в определение значений лингвистических переменных, могут рассматриваться как символы различных операций, определенных на нечетких подмножествах  $U$ . Рассмотрим наиболее существенные из этих операций.

Пусть  $A$  и  $B$  – нечеткие множества;  $S(A)$ ,  $S(B)$  – их носители. Обычно вводятся два набора определений основных операций над нечеткими множествами: *максиминный* ( $mm$ ) и *вероятностный* ( $p$ ). Объединением нечетких множеств  $A$  и  $B$  в  $U$  называется нечеткое множество  $A \cup B$  с функцией принадлежности вида:

$$\mu_{A \cup B}(u) = \begin{cases} \max(\mu_A(u), \mu_B(u)), & u \in U & -(mm) \\ \mu_A(u) + \mu_B(u) - \mu_A(u) \cdot \mu_B(u), & u \in U & -(p) \end{cases}$$

Объединение соответствует союзу **ИЛИ**. Таким образом, если  $X$  и  $Y$  – символы нечетких множеств, то  $X \cup Y = (X \text{ или } Y)$ .

Пример объединения.

Пусть  $A=(0,2/1; 0,5/2; 1,0/3)$ ,  $B=(0,8/1; 0,4/2; 0,5/3; 0,8/4)$ .

Тогда  $A \cup B=(0,8/1; 0,5/2; 1,0/3; 0,8/4)$  ( $mm$ )

$A \cup B=(0,84/1; 0,7/2; 1,0/3; 0,8/4)$  ( $p$ ).

Пересечением нечетких множеств  $A$  и  $B$  в  $U$  называется нечеткое множество  $A \cap B$  с функцией принадлежности вида:

$$\mu_{A \cap B}(u) = \begin{cases} \min(\mu_A(u), \mu_B(u)), & u \in U & -(mm) \\ \mu_A(u) \cdot \mu_B(u), & u \in U & -(p) \end{cases}$$

Пересечение соответствует союзу **И**. Таким образом,  $X \cap Y = (X \text{ и } Y)$ .

Пример пересечения.

Пусть  $A=(0,2/1; 0,5/2; 1,0/3)$ ,  $B=(0,8/1; 0,4/2; 0,5/3; 0,8/4)$ .

Тогда  $A \cap B=(0,2/1; 0,4/2; 0,0/3; 0,0/4)$  (*mm*)

$A \cap B=(0,16/1; 0,2/2; 0,5/3; 0,0/4)$  (*p*).

*Дополнением* нечеткого множества  $A$  называется нечеткое множество  $\bar{A}$  с функцией принадлежности:

$$\mu_{\bar{A}}(u) = 1 - \mu_A(u), \quad u \in U \quad (mm, p).$$

Определение дополнения соответствует отрицанию **НЕ**, то есть

$$\bar{X} = (\text{не } X) = \bigcup (1 - \mu_X(x)) / x.$$

Пример дополнения.

Пусть  $A=(0,2/1; 0,5/2; 1,0/3)$ .

Тогда  $\bar{A}=(0,8/1; 0,5/2; 0,0/3)$ .

*Декартово произведение*  $A_1 \times A_2 \times \dots \times A_n$  нечетких множеств  $A_i$  в  $U_i, i = \overline{1, n}$ , определяется как нечеткое множество  $A$  в декартовом произведении  $U = U_1 \times U_2 \times \dots \times U_n$  с функцией принадлежности вида:

$$\mu_A(u) = \min \{ \mu_{A_1}(u_1), \dots, \mu_{A_n}(u_n) \}, \quad u = (u_1, \dots, u_n) \in U.$$

*Обычным множеством*  $\alpha$ -уровня нечеткого множества  $A$  называется

$$S_\alpha = \{ u : u \in U, \mu_A(u) \geq \alpha \}, \quad \text{где } \alpha \in [0, 1].$$

Для определения арифметических операций  $\otimes = \{ +, -, \cdot, / \}$  в работе был сформулирован так называемый *принцип обобщения* Л.А.Заде:

Пусть  $A$  и  $B$  – два нечетких множества. Тогда нечеткое множество  $D = A \otimes B$  определяется функцией принадлежности  $\mu_D(u) = \Theta[\mu_A(u), \mu_B(u)]$ , где

$$\mu_D(u) = \Theta[\mu_A(u), \mu_B(u)] = \begin{cases} \sup_{a \otimes b = u, a \in S_A, b \in S_B} (\mu_A(u), \mu_B(u)), & (mm) \\ \sup_{a \otimes b = u, a \in S_A, b \in S_B} (\mu_A(u), \mu_B(u)), & (p) \end{cases}$$

Теперь арифметические операции  $\otimes = \{+, -, \cdot, / \}$  можно определить следующим образом:  $A \otimes B = \bigcup \mu_D(u) / (a \otimes b)$ , если операция  $(a \otimes b)$  определена.

Пример операций над множествами.

Пусть  $A=(0,2/1; 0,5/2; 1,0/3)$ ,  $B=(0,8/1; 0,4/2; 0,5/3; 0,8/4)$ .

Тогда

$$A + B = a \otimes b = u, a \in S_A, b \in S_B \begin{aligned} & \sup (0,2/2; 0,2/3; 0,2/4; 0,2/5; \\ & 0,5/3; 0,4/4; 0,5/5; 0,5/6; 0,8/4; \\ & 0,4/5; 0,5/6; 0,8/7) = \\ & = (0,2/2; 0,5/3; 0,8/4; 0,5/5; 0,5/6; \\ & 0,8/7) \text{ (mm)} \end{aligned}$$

$$A \cdot B = a \otimes b = u, a \in S_A, b \in S_B \begin{aligned} & \sup (0,16/2; 0,08/3; 0,1/4; 0,16/5; \\ & 0,4/3; 0,2/4; 0,25/5; 0,4/6; 0,8/4; \\ & 0,4/5; 0,5/6; 0,8/7) = \\ & = (0,16/2; 0,4/3; 0,8/4; 0,4/5; \\ & 0,5/6; 0,8/7) \text{ (p)} \end{aligned}$$

*Степенью* нечеткого множества  $A$  называется нечеткое множество  $A_a$  с функцией принадлежности

$$\mu_{A_a}(u) = (\mu_A(u))^a, u \in U, a > 0.$$

При  $a=2$  получаем операцию *концентрирования* (**CON**):  $\text{CON}(A)=A^2$ .

В результате применения этой операции к множеству  $A$  снижается степень нечеткости описания, причем для элементов с высокой степенью принадлежности это уменьшение относительно мало, а для элементов с малой степенью принадлежности – относительно велико.

При  $a=0,5$  получаем операцию *растяжения* (**DIL**):  $\text{DIL}(A)=A^{0,5}$ .

Эта операция увеличивает степень нечеткости исходного нечеткого множества.

Неопределенность удобно определить через некоторые основные операции (особенно операции *степень*, **CON**, **DIL**). Покажем, как это можно сделать для неопределенности *очень*. Ана-

логичным образом можно установить неопределенности *больше, меньше, много, слабо, вроде, вполне* и другие.

В обычном использовании неопределенность *очень* не имеет четко определенного значения. Она действует как усилитель, генерируя подмножества того множества, к которому она применяется. Аналогичным образом действует операция концентрирования, поэтому *очень и*, где *и* – некоторый термин, может быть определено как квадрат *и*, то есть:

$$\text{очень } i = i^2 = \bigcup \mu_U^2(i) / i$$

Например, если *и=маленький возраст*=(1,0/1; 0,8/2; 0,6/3; 0,4/4; 0,2/5),

тогда *очень маленький возраст*=(1,0/1; 0,64/2; 0,36/3; 0,16/4; 0,04/5).

Рассматриваемый как оператор, *очень* может действовать сам на себя.

Так, например, *очень очень и*=(*очень и*)<sup>2</sup> = *i*<sup>4</sup>.

Заметим, что порядок следования элементарных терминов в составном термине существенно влияет на результат. Так, например,

$$i = \text{очень не точно} = \overline{(\text{точно})}^2 \text{ и}$$

$$i = \text{не очень точно} = \overline{(\text{точно})}^2 \text{ не одно и то же.}$$

С другой стороны *не очень точно* может быть записано по-разному, хотя результат будет один и тот же:

$$i = \text{не очень точно} = \overline{\overline{\text{очень точно}}} = \overline{(\text{точно})}^2$$

Вернемся к нашему примеру о скользкой дороге.

Высказывание "*если А, то В, иначе С*" (где А,В,С нечеткие подмножества, при этом А обязательно из U, а В и С могут быть определены как на U, так и на V в зависимости от формулировки задачи; напомним, что U и V – универсальные множества, на которых определены нечеткие множества А,В,С) в терминах декартова произведения можно определить следующим образом:

$$\text{если } A, \text{ тогда } B, \text{ иначе } C \stackrel{\text{def}}{=} A \times B \cup \bar{A} \times C.$$

По существу мы получили некоторое неопределенное отношение  $R$  ( $R: U \Rightarrow V$ ). Далее (по условию задачи) мы хотим определить значения некоторых подмножеств из  $V$ , которые определяются заданными подмножествами из  $U$ .

Для решения этой и подобного типа задач в работе сформулировано *составное правило вывода*. Это правило имеет следующий вид.

Если  $R$  – неопределенное отношение  $U \Rightarrow V$  и  $x$  – неопределенное подмножество  $U$ , тогда неопределенное подмножество  $y \in V$ , которое индуцируется подмножеством  $x$ , дается композицией  $x$  и  $R$ , т.е.  $y = x \circ R$ , где " $\circ$ " – максиминное произведение.

В рассматриваемом примере необходимо определить

$x_1$  – "дорога не очень скользкая",

$x_2$  – "дорога очень не скользкая",

и, решив соответствующие задачи вывода, получим подмножества  $y_1$  и  $y_2$ .

Для сравнения полученных нечетких множеств строится некоторая четкая функция  $H(y_1, y_2)$  от нечетких аргументов, которая называется индексом ранжирования. Значения индекса для конкретной пары нечетких чисел (являющихся, по существу, нечеткими множествами) дает основание решить вопрос о том, какое из двух нечетких чисел больше (или – с какой степенью больше). Простейший индекс ранжирования имеет вид:

$$H_5^1(A, B) = \sup_{a \geq b} \min(\mu_A(a), \mu_B(b)),$$

где **sup** – обозначение точной верхней границы множества.

При этом если  $H_5^1(A, B) > H_5^1(B, A)$ , то  $A > B$ . Отметим, что приведенный индекс ранжирования в качестве наибольшего выбирает то нечеткое число, пик функции, принадлежности которого соответствует большему значению носителя. Данный индекс ранжирования является детерминированным, так как для сравнения нечетких чисел используются однозначно определенные представители этих нечетких чисел. Недостатком детерминированных индексов ранжирования является то, что они не учитывают вид и форму функции принадлежности сравниваемых нечетких чисел. Чтобы избежать такого ограничения, в теории не-

четких множеств введены так называемые интегральные индексы ранжирования. Рассмотрим один из них:

$$H_2(A, B) = H_+(A) - H_+(B), \quad H_+(A) = \sum_{\Delta\alpha} M(A_\alpha) \cdot \Delta\alpha,$$

где  $A_\alpha$  –  $\alpha$ -уровневое подмножество нечеткого множества  $A$ ;

$$M(A_\alpha) = \frac{\inf_{a \in A_\alpha} a + \sup_{a \in A_\alpha} a}{2}, \quad \text{где } \mathbf{inf} \text{ – точная нижняя граница}$$

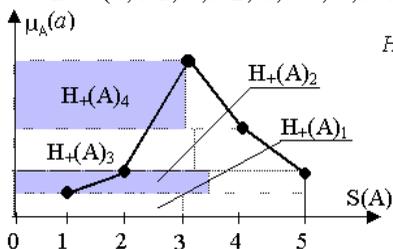
множества.

При этом если  $H_2(A, B) \geq 0 \Rightarrow A \geq B$ . Отметим, что интегральные индексы ранжирования дают более точный результат решения, чем детерминированные.

Пример вычисления индексов ранжирования.

Пусть  $A = (0,1/1; 0,2/2; 0,8/3; 0,4/4; 0,2/5)$ ;

$B = (0,4/1; 0,5/2; 0,9/3; 0,5/4; 0,1/5)$ .



$$\begin{aligned} H_+(A) &= H_+(A)_1 + H_+(A)_2 + \\ &+ H_+(A)_3 + H_+(A)_4 = \\ &= \frac{1+5}{2} \cdot 0,1 + \frac{2+5}{2} \cdot 0,1 + \\ &+ \frac{2,33+4}{2} \cdot 0,2 + \frac{3+3}{2} \cdot 0,4 \approx 2,48; \end{aligned}$$

$$H_+(B) = \frac{1+5}{2} \cdot 0,1 + \frac{2+4}{2} \cdot 0,4 + \frac{3+3}{2} \cdot 0,4 = 2,7;$$

$$H_2(A, B) = H_+(A) - H_+(B) = 2,48 - 2,7 = -0,22 < 0 \Rightarrow A < B.$$

Введем нечеткие множества *скользящая дорога* и *опасная езда*.

Первое нечеткое множество можно определить через коэффициент трения скольжения. Если использовать коэффициенты 0,01; 0,02; 0,03; 0,04; 0,05; 0,06, то можно записать:

$A = \text{скользящая дорога} = (0,8/0,01; 1,0/0,02; 0,6/0,03; 0,4/0,04; 0,2/0,05; 0,1/0,06)$ .

Для определения степени опасности можно ввести коэффициенты опасности I; II; III; IV; V. Например I – это легкие ушибы, садины; II – это переломы конечностей и т.д. При этом

$B = \text{опасная езда} = (0,1/I; 0,3/II; 0,5/III; 0,7/IV; 0,9/V)$ .

Теперь можно определить все необходимые для решения задачи нечеткие множества:

$$\text{не скользкая дорога} = \bar{A} = (0,2/0,01; 0,0/0,02; 0,4/0,03; 0,6/0,04; 0,8/0,05; 0,9/0,06);$$

$$\text{не опасная езда} = \bar{B} = (0,9/I; 0,7/II; 0,5/III; 0,3/IV; 0,1/V);$$

$$X_1 = \text{не очень скользкая дорога} = \bar{B}^2 = (0,36/0,01; 0,0/0,02; 0,64/0,03; 0,84/0,04; 0,96/0,05; 0,99/0,06);$$

$$X_2 = \text{очень не скользкая дорога} = \bar{B}^2 = (0,04/0,01; 0,0/0,02; 0,16/0,03; 0,36/0,04; 0,64/0,05; 0,81/0,06);$$

Определим нечеткое отношение  $R$ :  $R = A \times B \cup \bar{A} \times \bar{B}$

$$A \times B = \begin{bmatrix} 0,1 & 0,3 & 0,5 & 0,7 & 0,8 \\ 0,1 & 0,3 & 0,5 & 0,7 & 0,9 \\ 0,1 & 0,3 & 0,5 & 0,6 & 0,6 \\ 0,1 & 0,3 & 0,4 & 0,4 & 0,4 \\ 0,1 & 0,2 & 0,2 & 0,2 & 0,2 \\ 0,1 & 0,1 & 0,1 & 0,1 & 0,1 \end{bmatrix}, \bar{A} \times \bar{B} = \begin{bmatrix} 0,2 & 0,2 & 0,2 & 0,2 & 0,1 \\ 0,0 & 0,0 & 0,0 & 0,0 & 0,0 \\ 0,4 & 0,4 & 0,4 & 0,3 & 0,1 \\ 0,6 & 0,6 & 0,5 & 0,3 & 0,1 \\ 0,8 & 0,7 & 0,5 & 0,3 & 0,1 \\ 0,9 & 0,7 & 0,5 & 0,3 & 0,1 \end{bmatrix}$$

$$R = A \times B \cup \bar{A} \times \bar{B} = \begin{bmatrix} 0,1 & 0,3 & 0,5 & 0,7 & 0,8 \\ 0,1 & 0,3 & 0,5 & 0,7 & 0,9 \\ 0,4 & 0,4 & 0,5 & 0,6 & 0,6 \\ 0,6 & 0,6 & 0,5 & 0,4 & 0,4 \\ 0,8 & 0,7 & 0,5 & 0,3 & 0,2 \\ 0,9 & 0,7 & 0,5 & 0,3 & 0,1 \end{bmatrix}$$

Окончательно получим:

$$Y_1 = X_1 \circ R = \left( \frac{0,36}{0,01}; \frac{0,0}{0,02}; \frac{0,64}{0,03}; \frac{0,84}{0,04}; \frac{0,96}{0,05}; \frac{0,99}{0,06} \right) \circ \begin{bmatrix} 0,1 & 0,3 & 0,5 & 0,7 & 0,8 \\ 0,1 & 0,3 & 0,5 & 0,7 & 0,9 \\ 0,4 & 0,4 & 0,5 & 0,6 & 0,6 \\ 0,6 & 0,6 & 0,5 & 0,4 & 0,4 \\ 0,8 & 0,7 & 0,5 & 0,3 & 0,2 \\ 0,9 & 0,7 & 0,5 & 0,3 & 0,1 \end{bmatrix}$$

$$= (0,9/I; 0,7/II; 0,5/III; 0,6/IV; 0,6/V);$$

$$Y_2 = X_2 \circ R = \left( \frac{0,04}{0,01}; \frac{0,0}{0,02}; \frac{0,16}{0,03}; \frac{0,36}{0,04}; \frac{0,64}{0,05}; \frac{0,81}{0,06} \right) \circ \begin{bmatrix} 0,1 & 0,3 & 0,5 & 0,7 & 0,8 \\ 0,1 & 0,3 & 0,5 & 0,7 & 0,9 \\ 0,4 & 0,4 & 0,5 & 0,6 & 0,6 \\ 0,6 & 0,6 & 0,5 & 0,4 & 0,4 \\ 0,8 & 0,7 & 0,5 & 0,3 & 0,2 \\ 0,9 & 0,7 & 0,5 & 0,3 & 0,1 \end{bmatrix}$$

$$= (0,81/I; 0,7/II; 0,5/III; 0,36/IV; 0,36/V).$$

Сравним полученные результаты для  $Y_1$  и  $Y_2$  между собой, для чего воспользуемся индексом ранжирования  $H_5^1(Y_1, Y_2)$ :

$$\begin{aligned} H_5^1(Y_1, Y_2) &= \sup(\min(0,9_I; 0,81_I), \min(0,7_{II}; 0,81_I; 0,7_{II}), \\ &\min(0,5_{III}; 0,81_I; 0,7_{II}; 0,5_{III}), \\ &\min(0,6_{IV}; 0,81_I; 0,7_{II}; 0,5_{III}; 0,36_{IV}), \min(0,6_V; 0,81_I; 0,7_{II}; 0,5_{III}; \\ &0,36_{IV}; 0,36_V)) = \\ &= \sup(0,81; 0,7; 0,5; 0,36; 0,36) = 0,81. \end{aligned}$$

$$\begin{aligned} H_5^1(Y_1, Y_2) &= \sup(\min(0,81_I; 0,9_I), \min(0,7_{II}; 0,9_I; 0,7_{II}), \min(0,5_{III}; \\ &0,9_I; 0,7_{II}; 0,5_{III}), \\ &\min(0,36_{IV}; 0,9_I; 0,7_{II}; 0,5_{III}; 0,6_{IV}), \min(0,36_V; 0,9_I; 0,7_{II}; 0,5_{III}; \\ &0,6_{IV}; 0,6_V)) = \\ &= \sup(0,81; 0,7; 0,5; 0,36; 0,36) = 0,81. \end{aligned}$$

Использование детерминированного индекса ранжирования не позволило сделать вывод о том, какая дорога более опасная. Уточним решения, используя для сравнения интегральный индекс ранжирования. При этом для оценки степеней опасности введем следующие числовые значения: I – 1, II – 2, III – 3, IV – 4, V – 5 (это необходимо сделать, чтобы как-то численно оценивать элементы носителя. Можно было бы пойти по другому пути – "опасность" езды оценивать по степени ущерба для здоровья человека, например, I – 5%, II – 10% и т.д.).

$$H_+(Y_1) = \frac{1+5}{2} \cdot 0,6 + \frac{1+2}{2} \cdot 0,2 + \frac{1+1}{2} \cdot 0,2 = 2,3;$$

$$H_+(Y_2) = \frac{1+5}{2} \cdot 0,36 + \frac{1+3}{2} \cdot 0,14 + \frac{1+2}{2} \cdot 0,2 + \frac{1+1}{2} \cdot 0,11 = 1,77.$$

Таким образом, *не очень скользкая дорога* более опасна, чем *очень не скользкая дорога*. Отметим, что интегральный индекс ранжирования дает более точный результат по сравнению с де-

терминированным индексом, так как учитывает весь спектр распределений нечетких множеств. Приведенный пример является достаточно простым, имеющим тривиальный ответ, и служит только иллюстрацией применения теории нечетких множеств. Примеры использования этой теории для разработки математических моделей сложных процессов и явлений читатель может найти в работах.

Как уже отмечалось, подход с позиций теории нечетких множеств к математическому моделированию явлений или процессов наиболее предпочтителен в тех областях знаний, в которых трудно или даже невозможно получить точные количественные оценки. К таким разделам можно отнести экологию, метеорологию, педагогику, психологию, социологию и тому подобное.

### Пример

Для задачи о баскетболисте, постановка которой рассмотрена во второй главе, применение подхода с позиций теории нечетких множеств может помочь в обосновании используемых приемов бросания мяча, уточнении методики подготовки баскетболиста и тому подобное. Рассмотрим одну из возможных задач – вопрос об оценке точности попадания мяча в кольцо при броске со штрафной линии. Пусть известно, что если начальная скорость  $V_0$  мяча будет *средняя*, то точность попадания будет *высокой*, в противном случае – *не высокой*. Оценим, в каком случае точность попадания будет выше: при начальной скорости *больше* средней или при начальной скорости *меньше* средней.

Для решения этой задачи необходимо ввести неопределенности типа *больше* (средней) и *меньше* (средней).

Функцию принадлежности неопределенности *больше* (средней) определим следующим образом:

$$\mu_A(u) = \begin{cases} 0, & u < \max_{u \in S(A)} u, \\ 1 - \mu_A(u), & u \geq \max_{u \in S(A)} u. \end{cases}$$

Тогда функция принадлежности неопределенности *меньше* (средней) определим так:

$$\mu_A(u) = \begin{cases} 1 - \mu_A(u), & u < \max_{u \in S(A)} u, \\ 0,2, & u \geq \max_{u \in S(A)} u. \end{cases}$$

Введем нечеткие множества *средняя скорость* и *высокая точность*.

Пусть

$$A = \text{средняя скорость} = (0,1/5,5; 0,4/6; 0,8/6,5; 0,6/7; 0,3/7,5)$$

(элементы носителя числового нечеткого множества  $A$  измеряются в м/сек).

$$B = \text{высокая точность} = (0,2/20; 0,3/40; 0,5/60; 0,7/80; 1/100)$$

(элементы носителя числового нечеткого множества  $B$  измеряются в %).

Теперь можно определить все необходимые для решения задачи нечеткие множества:

$$\bar{A} = (0,9/5,5; 0,6/6; 0,2/6,5; 0,4/7; 0,7/7,5);$$

$$\bar{B} = (0,8/20; 0,7/40; 0,5/60; 0,3/80; 0/100);$$

$X_1 = \text{скорость больше средней} = (0,2/5,5; 0,2/6; 0,2/6,5; 0,4/7; 0,7/7,5);$

$X_2 = \text{скорость меньше средней} = (0,9/5,5; 0,6/6; 0,2/6,5; 0,2/7; 0,2/7,5).$

Определим нечеткое отношение  $R: R = A \times B \cup \bar{A} \times \bar{B}$

$$R = \begin{bmatrix} 0,8 & 0,7 & 0,5 & 0,3 & 0,1 \\ 0,6 & 0,6 & 0,5 & 0,4 & 0,4 \\ 0,2 & 0,3 & 0,5 & 0,7 & 0,8 \\ 0,4 & 0,4 & 0,5 & 0,6 & 0,6 \\ 0,7 & 0,7 & 0,5 & 0,3 & 0,3 \end{bmatrix}$$

Окончательно получим:

$$Y_1 = X_1 \circ R = (0,7/20; 0,7/40; 0,5/60; 0,4/80; 0,4/100);$$

$$Y_2 = X_2 \circ R = (0,8/20; 0,7/40; 0,5/60; 0,4/80; 0,4/100).$$

Сравним полученные результаты для  $Y_1$  и  $Y_2$  между собой, для чего воспользуемся индексом ранжирования :

$$H_s^1(Y_1, Y_2) = 0,7;$$

$$H_s^1(Y_2, Y_1) = 0,7.$$

Использование детерминированного индекса ранжирования не позволило сделать однозначный вывод. Уточним решения, используя для сравнения интегральный индекс ранжирования:

$$H_+(Y_1) = 36; H_+(Y_2) = 38.$$

Таким образом, точность попадания будет выше, если начальная скорость *меньше* средней по сравнению с броском с начальной скоростью *больше* средней.

### 4.3 Моделирование в условиях стохастической неопределенности

...Истинной логикой для этого мира является исчисление вероятностей, занимающееся нахождением величин вероятностей, которые учитывает или должен учитывать любой здравомыслящий человек.

Дж. Кларк Максвелл

При математическом описании модели любого сколь угодно сложного процесса или явления всегда возникает задача учета случайности. Сейчас уже трудно установить, кто первый поставил вопрос, пусть и в несовершенной форме, о возможности количественного измерения возможности появления случайного события. В течение долгого периода исследователи ограничивались рассмотрением разного рода игр, особенно игр в кости, поскольку их изучение позволяет ограничиваться простыми и прозрачными математическими моделями. Однако следует заметить, что многие отлично понимали то, что позднее (в 1657 году) было прекрасно сформулировано Х.Гюйгенсом: "*... Я полагаю, что при внимательном изучении предмета читатель заметит, что имеет дело не только с игрой, но здесь закладываются основы очень интересной и глубокой теории*". И эта теория, изучающая количественные закономерности случайных явлений в однородных массовых процессах, была создана и названа *теорией вероятностей*.

За последние десятилетия неизмеримо выросла роль, которую играет теория вероятностей в современном естествознании. После того, как молекулярные представления о строении вещества получили всеобщее признание, стало неизбежным широкое использование теории вероятностей в физике и в химии. Заметим, что с точки зрения молекулярной физики каждое вещество состоит из огромного числа малых частиц, находящихся в непрерывном движении, и в процессе этого движения взаимодействующих друг с другом. Естественно, что при таких условиях обычные для физических теорий методы математических исследований становятся бессмысленными.

Теория вероятностей развилась из потребностей практики: в абстрактной форме она отражает закономерности, присущие случайным событиям массового характера. Эти закономерности играют исключительно важную роль в физике, химии, различных технических дисциплинах, экономике, военном деле и так далее.

До недавнего времени теория вероятностей представляла собой еще не сложившуюся математическую науку, в которой основные понятия были недостаточно корректно определены. Сказанное относится к так называемой классической теории вероятностей. В 1900 году на II Международном математическом конгрессе в Париже немецкий ученый Д.Гильберт сделал доклад, в котором указал 23 важнейшие проблемы, стоящие перед математикой. Одна из них – аксиоматическое построение теории вероятностей. В настоящее время эта проблема успешно решена. Наиболее удачный подход к построению аксиоматической теории вероятностей предложил А.Н.Колмогоров. Этот подход тесно связывает теорию вероятностей с современной теорией меры, а также теорией множеств.

Иллюстрацию моделирования в условиях стохастической неопределенности проведем на примере задачи о баскетболисте, постановка которой рассмотрена в главе 2. Там же было приведено аналитическое решение для дальности и точности броска баскетболиста в зависимости от начальной скорости  $V_0$  мяча, угла бросания  $\alpha_0$  и расстояния  $x_k$  до кольца в момент броска.

В реальной ситуации баскетболист не может выполнить хотя бы два полностью одинаковых броска, то есть с одинаковыми

начальными параметрами  $V_0$  и  $\alpha_0$ . Это связано с субъективными и объективными причинами. В частности, дрогнула рука, изменилось положение тела, произошел вдох или выдох, сменилось эмоциональное состояние и т.п. (субъективные); изменилось направление движения воздуха, температура окружающей среды и т.д. (объективные причины). Поэтому для более полного отражения реальной действительности при моделировании необходимо учесть весь спектр возможных разбросов параметров. Если выполняется серия бросков одним мячом, то неопределенными будут являться только начальные параметры бросания. Данные величины могут быть представлены как интервальные, статистические, стохастические параметры или характеристики, описываемые нечеткими множествами. С наименьшей степенью неопределенности влияние изменения этих параметров можно учесть с позиций стохастической неопределенности. При этом необходимо знать законы распределения соответствующих случайных величин.

Введем некоторые основные определения.

*Опыт* – это осуществление какого-либо комплекса условий, который может быть воспроизведен сколь угодно большое число раз.

Под *событием* будем понимать результат опыта или наблюдения.

События могут быть *элементарными* (неразложимыми) и *составными* (разложимыми).

*Элементарное событие* – это событие, которое происходит в результате единичного опыта.

*Составное событие* – это совокупность элементарных событий.

### **Пример событий**

Игральный кубик подбрасывается два раза. Пусть составное событие определено следующим образом: "сумма выпавших цифр равна 6". Тогда элементарные события будут: "5+1", "4+2", "3+3", "2+4" и "1+5". Любые другие сочетания не относятся к рассматриваемому составному событию.

*Генеральной совокупностью* называют совокупность событий, которые могут быть реализованными в результате бесконечного числа однотипных опытов.

*Выборочной совокупностью* или просто *выборкой* называют совокупность случайно отобранных событий из генеральной совокупности.

*Объемом совокупности* называют число событий этой совокупности.

*Случайной величиной* будем называть величину, которая в результате опыта (наблюдения, испытания) может принять одно из возможных значений, но какое именно – неизвестно.

*Вероятность* некоторого события – это мера его "благоприятствия".

События будем называть *равновозможными*, если мера их "благоприятствия" одинакова. В реальных условиях, когда количество опытов *конечно*, мера "благоприятствия" определяется не вероятностью, а *частотью*. Пусть событие  $A$  наблюдалось в  $m$  опытах из  $n$  опытов (испытаний). Тогда частота события  $A$  ( $W(A)$ ) определяется формулой:

$$W(A) = m/n. \quad (4.1)$$

Когда  $n$  достаточно велико, то "работает" одна из предельных теорем (закон больших чисел – теорема Бернулли), и может быть записано приближенное равенство:

$$P(A) \cong W(A), \quad (4.2)$$

где  $P(A)$  – вероятность события  $A$ .

Отметим, что по принятому в теории вероятности соглашению вероятность произвольного события полагается изменяющейся от 0 до 1. При этом нулевая вероятность соответствует *невозможному событию (которое никогда произойти не может)*, а единичная – *достоверному событию (которое обязательно произойдет)*.

Изложенный в данном разделе подход к описанию стохастической неопределенности базируется на классической теории вероятностей. В настоящее время общепринятым является аксиоматическое построение теории вероятностей, базирующееся на аксиомах А.Н.Колмогорова.

Сформулируем основные положения этого подхода. Для удобства чтения сформулируем основные положения этого подхода на примере стохастического эксперимента.

Пусть  $\Omega$  – пространство элементарных событий любой природы (по существу,  $\Omega$  – это достоверное событие, а  $\emptyset$  (пустое множество) – это невозможное событие). Предположим, что в  $\Omega$  выделена система подмножеств  $U$ , удовлетворяющая следующим условиям:

**A1.**  $\Omega \in U$ ;

**A2.** Если  $A \in U$ , то  $\bar{A} = \Omega \setminus A \in U$ ;

**A3.** Из того, что  $A_i \in U, i=1,2,\dots,k$  следует, что

$$\bigcup_{i=1}^k A_i \in U \text{ и } \bigcap_{i=1}^k A_i \in U .$$

Система подмножеств  $U$ , удовлетворяющая (A1-A3), называется алгеброй событий. Алгебру  $U$  называют также *кольцом*, так как в  $U$  определены только две операции (сложение и умножение), не выводящие из  $U$ . Алгебра  $U$  является *кольцом с единицей*, т.к.  $\Omega \in U$  и  $A\Omega = \Omega A = A$  для любого  $A \in U$ .

Если к A1-A3 добавить еще одну аксиому:

**A4.** Из того, что  $A_i \in U, i=1,2,\dots$  следует, что

$$\bigcup_{i=1}^{\infty} A_i \in U \text{ и } \bigcap_{i=1}^{\infty} A_i \in U , \text{ то система подмножеств } U, \text{ удовле-}$$

творяющая (A1-A4), называется  $s$ -алгеброй событий.

Таким образом, алгебра есть класс множеств, замкнутый относительно конечного числа операций дополнения, объединения и пересечения;  $s$ -алгебра есть класс множеств, замкнутый относительно счетного числа этих операций.

Если задано множество  $\Omega$  и какая-нибудь алгебра или  $s$ -алгебра  $U$  его подмножеств, то говорят, что задано *измеримое пространство*  $(\Omega, U)$ .

Осталось ввести понятие вероятности события.

Вероятность на  $(\Omega, U)$  есть числовая функция, определенная на множествах из  $U$  и обладающая следующими свойствами:

**P1.**  $p(A) \geq 0$  для любого  $A \in U$ ;

**P2.**  $p(\Omega) = 1$ ;

**P3.** Аксиома конечной аддитивности. Если последовательность событий  $A_i \in U$ ,  $i=1,2,\dots,k$  такова, что

$$A_i A_j = \emptyset \text{ при } i \neq j, \bigcup_{i=1}^k A_i \in U, \text{ то } p\left(\sum_{i=1}^k A_i\right) = \sum_{i=1}^k p(A_i);$$

**P4.** Аксиома счетной аддитивности. Если последовательность событий  $A_i \in U$ ,  $i=1,2,\dots$  такова, что

$$A_i A_j = \emptyset \text{ при } i \neq j, \bigcup_{i=1}^{\infty} A_i \in U, \text{ то } p\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} p(A_i).$$

Тройка  $(\Omega, U, P)$  называется вероятностным пространством.

Отметим, что аксиоматика Колмогорова является непротиворечивой. Она реализуется на объектах, где справедлива классическая теория вероятностей.

Аксиоматика Колмогорова не является полной, т.к. не указывает способ, как приписывать событиям вероятность. Однако это скорее не недостаток, а достоинство. Дискуссии о том, что следует понимать под вероятностью, связаны с желанием увязать определение вероятности с ее "физической" природой. Ввиду сложности последней, такие попытки всегда наталкивались на существенные трудности не только математического, но и философского характера. Появление аксиоматики Колмогорова как бы отделяет математическую сторону вопроса от всего остального. При таком подходе "физическая трактовка" понятия вероятности появляется уже в виде теоремы (усиленный закон больших чисел), в силу которой частота появления некоторого события при неограниченном повторении независимых испытаний сближается (в строго определенном смысле) с вероятностью этого события.

Таким образом, можно сказать, что при аксиоматическом подходе построение вероятностного пространства  $(\Omega, U, P)$  является основным этапом в создании математической модели того или иного эксперимента. При этом произвол хотя и существует, но он по существу всегда определяется постановкой конкретных задач.

Рассмотренная выше аксиоматика Колмогорова оказалась исключительно плодотворной для развития новых разделов тео-

рии вероятностей, в частности – теории случайных процессов, примеры которых будут приведены в следующем параграфе.

Полностью охарактеризовать случайную величину можно законом распределения. Необходимо различать случайные величины (с.в.), принимающие лишь отдельные, изолированные возможные значения (дискретные случайные величины), и случайные величины, возможные значения которых, сплошным образом заполняют некоторый промежуток (непрерывные случайные величины).

*Законом распределения* случайной величины называют любое правило (таблицу, функцию), позволяющее находить вероятности всевозможных событий, связанных со случайной величиной (например, вероятность того, что она примет какое-то значение или попадет на какой-то интервал). Если с.в.  $X$  имеет данный закон распределения, то про нее говорят, что она "распределена" по этому закону (или "подчинена" этому закону распределения).

*Дискретной* называют случайную величину, которая принимает отдельные, изолированные возможные значения с определенными вероятностями. Число возможных значений дискретной случайной величины может быть конечным или счетным.

*Непрерывной* называют случайную величину, которая может принимать все значения из некоторого конечного или бесконечного промежутка. Очевидно, что число возможных значений непрерывной случайной величины бесконечно.

Рассмотрим некоторые простейшие характеристики случайной величины. К таким характеристикам относятся *математическое ожидание, дисперсия, мода и медиана случайной величины*.

*Математическим ожиданием* дискретной случайной величины называют сумму произведений всех ее возможных значений на их вероятности.

Математическое ожидание является средневзвешенным значением случайной величины, в которое каждое значение входит с "весом", равным соответствующей вероятности. Таким образом, математическое ожидание представляет собой "центр тяжести" системы "материальных точек", координаты которых суть всевозможные значения случайной величины, а "массы" равны вероятностям этих значений.

Пусть случайная величина  $X$  принимает только значения  $x_1, x_2, \dots, x_n$ , вероятности которых соответственно равны  $p_1, p_2, \dots, p_n$ . Тогда математическое ожидание  $M(X)$  определится формулой

$$M(X) = \sum_{i=1}^n x_i \cdot p_i \quad (4.3)$$

### **Пример 1**

Найти математическое ожидание случайной величины, заданной таблицей значений:

|     |     |     |     |
|-----|-----|-----|-----|
| $X$ | 3   | 5   | 2   |
| $p$ | 0,1 | 0,6 | 0,3 |

*Решение*

$$M(X) = 3 \times 0,1 + 5 \times 0,6 + 2 \times 0,3 = 3,9.$$

Математическое ожидание – важная, но не единственная характеристика "усредненного" значения случайной величины. "Усредненное" значение может характеризоваться также *модой* и *медианой*.

*Модой* случайной величины называется ее наиболее вероятное значение (то, для которого вероятность достигает максимума).

*Медианой* случайной величины называется такое ее значение, для которого  $P(X < x) \gg P(X > x)$ .

### **Пример 2**

Найти математическое ожидание, моду и медиану случайной величины, заданной таблицей значений:

|     |     |     |     |      |      |
|-----|-----|-----|-----|------|------|
| $X$ | 0   | 1   | 2   | 3    | 4    |
| $P$ | 0,2 | 0,3 | 0,4 | 0,05 | 0,05 |

*Решение*

Математическое ожидание:

$$M(X) = 0 \times 0,2 + 1 \times 0,3 + 2 \times 0,4 + 3 \times 0,05 + 4 \times 0,05 = 1,45.$$

Мода:  $M_x(X) = 2$ .

Медиана:  $X_m = 1$ .

Математическое ожидание, мода и медиана случайной величины по разному характеризуют усредненное значение случайной величины. Эти различия становятся понятными из следующего примера. Если некоторая фирма хочет оценить необходимое количество денег для выдачи зарплаты, ее интересует средняя зарплата сотрудника, то есть математическое ожидание зарплаты сотрудника, рассматриваемой как случайная величина. При устройстве на работу будущего сотрудника должна интересоваться не средняя зарплата, а ее мода – заработная плата, получаемая большинством сотрудников. Сопоставление собственной заработной платы с медианой заработной платы говорит сотруднику о том, принадлежит ли он к хорошо или плохо оплачиваемой части сослуживцев.

На практике часто требуется оценить среднее отклонение возможных значений случайной величины от ее среднего значения. Характеристикой такого отклонения служит *дисперсия случайной величины*.

*Дисперсией* дискретной случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X) = M(X - M(X))^2. \quad (4.4)$$

### **Пример 3**

Найти дисперсию случайной величины, заданной в примере 1.

*Решение*

$$M(X) = 3,9;$$

$$\begin{aligned} D(X) &= (3 - 3,9)^2 \times 0,1 + (5 - 3,9)^2 \times 0,6 + (2 - 3,9)^2 \times 0,3 = \\ &= 0,081 + 0,726 + 1,083 = 1,89 \end{aligned}$$

Использование дисперсии не всегда бывает удобно, так как ее размерность равна квадрату размерности случайной величины. Поэтому для оценки отклонения возможных значений случайной величины вокруг ее среднего значения часто используют *среднее квадратическое отклонение*.

*Средним квадратическим отклонением* случайной величины  $X$  называется квадратный корень из дисперсии:

$$\sigma(X) = \sqrt{D(X)} \quad (4.5)$$

Для примера 5.3.3:  $\sigma(X) = \sqrt{1,89} \approx 1,375$ .

Для случайных величин в большинстве случаев невозможно составить перечень всех возможных значений. Поэтому необходимо ввести новые числовые характеристики случайных величин. В качестве таких характеристик чаще всего используют *функцию распределения и плотность распределения вероятностей (непрерывной случайной величины)*.

*Функцией распределения* случайной величины  $X$  называют функцию  $F_X(x)$ , представляющую собой вероятность того, что случайная величина  $X$  в результате испытания примет значение, меньшее  $x$ , т.е.

$$F_X(x) = P(X < x) \quad (4.6)$$

Геометрически это равенство означает, что  $F_X(x)$  есть вероятность того, что случайная величина примет значение, которое на числовой оси изображается любой точкой, лежащей левее точки  $x$ .

Пусть известно статистическое распределение частотей количественного признака  $X$ . В этом случае удобно пользоваться *эмпирической функцией распределения*.

*Эмпирической функцией распределения* (функцией распределения выборки) называют функцию  $F_X^*(x)$ , определяющую частоту события, заключающегося в том, что случайная величина  $X$  в результате испытания примет значение, меньшее  $x$ , т.е.

$$F_X^*(x) = W(X < x) \quad (4.7)$$

Таким образом, эмпирическая функция распределения выборки служит для оценки теоретической функции распределения генеральной совокупности.

Непрерывные случайные величины можно исследовать с помощью еще одной функции – *плотности распределения вероятностей*.

*Плотность распределения*  $f_X(x)$  вероятностей непрерывной случайной величины  $X$  представляет собой первую производную

от функции распределения  $F_X(x)$  и характеризует скорость изменения функции распределения этой случайной величины.

Для наглядности представления *выборки* строят различные графики статистического распределения, в частности – *полигон* и *гистограмму*.

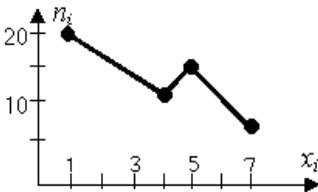
*Полигоном частот* называют ломанную, отрезки которой соединяют точки  $(x_1; n_1), (x_2; n_2), \dots, (x_k; n_k)$ , где  $n_i, i=1, \dots, k$  – число наблюдений (*частоты*), при которых наблюдалось значение признака, равное  $x_i$ . Для построения полигона частот на оси абсцисс откладывают варианты  $x_i$ , а на оси ординат соответствующие им частоты  $n_i$ . Отметим, что *сумма всех частот равна объему* выборки. Точки  $(x_i; n_i)$  соединяют отрезками прямых и получают *полигон частот*.

#### **Пример 4**

Построить полигон частот для следующего распределения:

|       |    |    |    |   |
|-------|----|----|----|---|
| $x_i$ | 1  | 4  | 5  | 7 |
| $n_i$ | 20 | 10 | 14 | 6 |

*Решение*



Если вместо  $n_i$  выбирать частоты  $w_i$ , то получим *полигон относительных частот*.

В случае *непрерывного признака* целесообразно строить *гистограмму*, для чего интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной  $h$  и находят для каждого частичного интервала  $n_i$  – сумму частот вариантов, попавших в  $i$ -й интервал.

*Гистограммой частот* называют ступенчатую функцию, состоящую из прямоугольников, основаниями которой служат частичные интервалы длиной  $h$ , а высоты равны отношению  $n_i/h$  (плотность частоты). Для построения гистограммы частот на оси

абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии  $n_i/h$ .

Площадь  $i$ -ого частичного прямоугольника равна  $h * n_i/h = n_i$  – сумме частот  $i$ -ого интервала; следовательно, *площадь гистограммы частот* равна сумме всех частот, т.е. *объему выборки*.

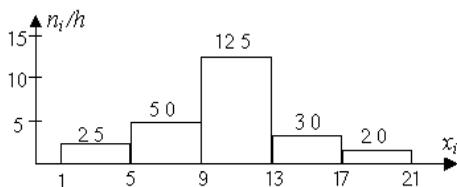
Если в качестве высот в *гистограмме частот* выбрать отношение  $w_i/h$  (плотность частоты), то получим *гистограмму относительных частот*. Отметим, что *площадь гистограммы относительных частот* равна сумме всех *относительных частот*, т.е. единице.

### **Пример 5**

Построить гистограмму частот по данному распределению выборки объема  $n = 100$ :

| Номер интервала | Частичный интервал $x_i - x_{i+1}$ | Сумма частот $n_i$ | Плотность частоты $n_i/h_i$ |
|-----------------|------------------------------------|--------------------|-----------------------------|
| 1               | 1-5                                | 10                 | 2,5                         |
| 2               | 5-9                                | 20                 | 5,0                         |
| 3               | 9-13                               | 50                 | 12,5                        |
| 4               | 13-17                              | 12                 | 3,0                         |
| 5               | 17-21                              | 8                  | 2,0                         |

*Решение*



При решении конкретных задач встречаются как дискретные, так и непрерывные случайные величины. В качестве примеров *распределений* рассмотрим одно дискретное (пуассоновское) и три непрерывных (равномерное, показательное и распределение Гаусса) распределения случайных величин.

*Распределение Пуассона.*

К распределению Пуассона приводит следующая постановка. Найти вероятность того, что при большом числе повторных ис-

пытаний, в каждом из которых вероятность появления события очень мала и равна  $p$  (возможные значения события:  $0, 1, 2, \dots, k, \dots$ ), событие произойдет ровно  $k$  раз.

Закон распределения вероятностей:

$$P_\lambda(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda), \quad k = 0, 1, 2, \dots \quad (4.8)$$

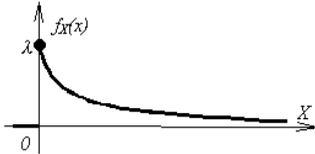
где  $k!$  (читается "ка-факториал") представляет собой произведение всех натуральных чисел от 1 до  $k$  включительно ( $k! = 1 \times 2 \times 3 \times \dots \times k$ ).

Математическое ожидание и дисперсия имеют вид:

$$M(X) = D(X) = \lambda.$$

*Показательное распределение.*

*Показательным (экспоненциальным)* называют распределение вероятностей непрерывной случайной величины  $X$ , которое описывается плотностью:



$$f_x(x) = \begin{cases} 0, & \text{при } x < 0, \\ \lambda \exp(-\lambda x), & \text{при } x \geq 0. \end{cases} \quad (4.9)$$

где  $\lambda > 0$  – константа.

Математическое ожидание и дисперсия в этом случае равны:

$$M(X) = D(X) = 1/\lambda.$$

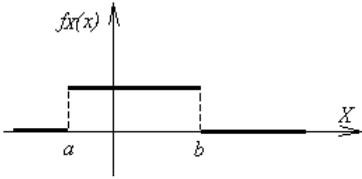
Вероятность попадания в интервал  $(a, b)$  непрерывной случайной величины  $X$ , распределенной по показательному закону, равна

$$P(a < X < b) = \exp(-\lambda a) - \exp(-\lambda b).$$

Показательное распределение широко применяется в теории случайных процессов. Например, отказы техники, появление бракованных изделий на конвейере, поступление больных в стационар и т.п. образует поток повторяющихся однотипных событий. Интервалы времени между соседними событиями в подобном потоке можно считать случайными величинами, которые при некоторых допущениях хорошо описываются показательным распределением.

*Равномерное распределение на  $[a, b]$ .*

Распределение вероятностей называется *равномерным*, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение.



Плотность равномерного распреде-

ления имеет вид:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases} \quad (4.10)$$

Распределение в интервале  $(0,1)$  будем называть *стандартным равномерным распределением*.

Математическое ожидание и дисперсия в этом случае определяются как

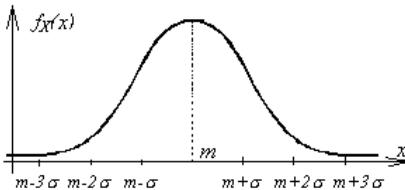
$$M(X) = \frac{a+b}{2}, \quad D(X) = \frac{(b-a)^2}{12}.$$

Вероятность попадания в интервал  $(a, b)$  непрерывной случайной величины  $X$ , распределенной по равномерному закону (при  $a \in [a, b]$ ,  $b \in [a, b]$ )

$$P(\alpha < X < \beta) = \frac{\beta - \alpha}{b - a}.$$

*Нормальное распределение (распределение Гаусса).*

*Нормальным* называется распределение случайной величины, которое полностью определяется математическим ожиданием  $M(X)=m$  и средним квадратическим отклонением  $s(X)=s$  и описывается плотностью:



$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right). \quad (4.11)$$

*Правило трех сигм:* вероятность попадания реализаций нормально распределенной случайной величины в интервал  $m \pm 3s$  составляет более 99,7%.

Нормальное распределение с параметрами  $m = 0$  и  $s = 1$  называется *стандартным нормальным распределением*.

Вероятность попадания в интервал  $(a, b)$  непрерывной случайной величины  $X$ , распределенной по нормальному закону

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right),$$

где  $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x \exp(-t^2) dt$  – функция Лапласа (интеграл ве-

роятностей, его значения приведены во всех учебных пособиях по теории вероятностей).

Нормальное распределение играет очень важную роль в теории вероятности и широко применяется в математической статистике. Например, любой количественный параметр, полученный в результате измерения, следует рассматривать как случайную величину. На конкретное значение параметра при данном измерении влияет независимо друг от друга множество трудно учитываемых или даже неизвестных факторов: положение и состояние измерительного инструмента, состояние измеряемого объекта, флуктуации параметров окружающей среды (температура, влажность, плотность, скорость движения и т.п.) и т.д. Все это приводит к отличию значений измеряемого параметра в различных измерениях. Оказалось, что ошибка измерения ведет себя как случайная величина, подчиняющаяся нормальному закону распределения.

Роль нормального распределения важна не только при проведении измерений. При сложении случайных величин, распределенных по произвольному закону, полученные суммы представляют собой случайные величины, распределенные по закону, близкому к нормальному. То есть последний играет роль некоторого предельного распределения, к которому стремятся распределения сумм произвольных случайных величин. Например, сумма 12 равномерно распределенных чисел с достаточной для

практики точностью позволяет получить нормально распределенную случайную величину.

Моделирование в условиях стохастической неопределенности широко используется в таких разделах теории вероятностей как дисперсионный анализ, регрессионный анализ и ковариационный анализ. По существу, математическая модель исследуемого явления в рамках этих разделов имеет один и тот же формализм, который можно представить следующим образом.

Пусть имеется  $n$  наблюдений или измерений. Эти наблюдения рассматриваются как  $n$  случайных величин  $y_i$ ,  $i=1,2,3,\dots,n$ , которые являются линейными комбинациями известных факторов  $x_{ij}$ ,  $i=1,2,3,\dots,n$ ;  $j=1,2,3,\dots,k$  с  $k$  неизвестными постоянными  $b_j$ ,  $j=1,2,3,\dots,k$  и ошибок  $u_i$ ,  $i=1,2,3,\dots,n$ :

$$y_i = x_{i1}b_1 + x_{i2}b_2 + x_{i3}b_3 + \dots + x_{ik}b_k + u_i, \quad i=1,2,3,\dots,n.$$

Целью *дисперсионного анализа* является получение выводов относительно ошибок  $u_i$ ,  $i=1,2,3,\dots,n$  и некоторых  $\{b_j\}$ , причем выводов, остающихся справедливыми независимо от значений других  $\{b_j\}$ , "исключить" которые было бы более желательно, чем "оценить". При этом  $x_{ij}$ ,  $i=1,2,3,\dots,n$ ;  $j=1,2,3,\dots,k$  являются целыми числами, равными обычно 0 или 1 (качественными "переменными-указателями" [106]). Если  $\{x_{ij}\}$  не являются "переменными-указателями", а пробегают непрерывные множества значений, то мы получим *регрессионный анализ*. В случае, когда среди  $\{x_{ij}\}$  есть переменные двух видов, мы получаем *ковариационный анализ*. Природа неизвестных параметров  $\{b_j\}$  может быть различна. Они могут быть или неизвестными постоянными, или случайными величинами, или часть из них случайные величины, а часть – постоянные.

Другими словами можно сказать, что отличие трех видов анализа заключается в том, что в дисперсионном анализе все факторы исследуются *качественно*, в регрессионном анализе все факторы являются количественными и исследуются *количественно*, тогда как в ковариационном анализе одна часть рассматриваемых факторов исследуется *качественно*, а другая часть – *количественно*.

Как было отмечено в главе 1, при реализации математических моделей могут использоваться как аналитические, так и ал-

горитмические методы исследования полученных математических соотношений. При применении аналитических методов выходные переменные модели могут быть представлены в виде аналитических выражений, для которых могут быть применены универсальные преобразования случайной величины. Однако подобное преобразование требует глубокого знания специальных разделов теории вероятности и применимы в относительно простых случаях. Для сложных аналитических выражений, а также для алгоритмических методов анализа модели при оценке закона распределения случайной величины, зависящей от ряда случайных параметров, используют *метод Монте-Карло*. Данный метод можно рассматривать как метод приближенного моделирования случайных величин с целью оценки характеристик их распределений. Идея моделирования случайных явлений, как известно, очень стара и, по мнению некоторых авторов, восходит ко времени Древнего Вавилона и Ветхого Завета. Современная формулировка метода статистического моделирования и само название – метод Монте-Карло была предложена Дж. Нейманом совместно с С. М. Уламом в 40-х годах нынешнего столетия.

Под моделированием случайной величины в методе Монте-Карло подразумевается ее конструирование при помощи некоторого случайного механизма. После получения ряда реализаций интересующей нас случайной величины решается типичная задача математической статистики – оценка различных характеристик (математического ожидания, дисперсии и т.д.). Применение метода Монте-Карло возможно лишь с использованием ЭВМ.

Узловой проблемой для метода Монте-Карло является моделирование случайных чисел  $r$ , равномерно распределенных в интервале  $(0,1)$ . Существующие таблицы случайных чисел неудобно использовать для решения "больших" задач на ЭВМ, тем более что еще никем не было замечено преимущество таблиц перед так называемыми "псевдослучайными числами". Уже давно разрабатываются "физические датчики" случайных чисел, когда с ЭВМ соединяется, например, счетчик числа  $\alpha$ -частиц, вылетающих из некоторого радиоактивного источника за фиксированный промежуток времени. Они применяются сравнительно редко по двум причинам:

- невозможно воспроизвести выборочную последовательность для повторения расчетов;
- практически нельзя гарантировать устойчивую удовлетворительную постоянную работу датчика.

Как правило, для решения задач методом Монте-Карло, используют рекуррентные формулы, производящие регулярную последовательность чисел, являющейся для внешнего наблюдателя случайной и удовлетворяющей основным требованиям, накладываемым на "настоящие" случайные числа. Подобную последовательность принято называть последовательностью псевдослучайных чисел. Для задания такой последовательности достаточно задать ее начальное число. Многочисленные методики получения реализаций псевдослучайных чисел можно найти в литературе. Простейший метод получения таких последовательностей заключается в следующем. Выбирается рекуррентная последовательность

$$r_0 \leq c, r_k = (m \times r_{k-1} + n) \bmod c, \quad (4.12)$$

где  $\bmod$  – операция вычисления остатка при целочисленном делении.

Соотношение (4.12) при заданных целых константах  $m$ ,  $n$ ,  $c$  позволяет получать последовательность из  $c$  целых чисел. Например, пусть  $c=8$ ,  $m=5$ ,  $n=3$ . Примем  $r_0=1$ . Тогда получаем следующий бесконечно повторяющийся ряд значений:

1 0 3 2 5 4 7 6 1 ...

Данная последовательность имеет период, равный 8, и каждое значение от 0 до 7 на периоде появляется ровно один раз. Цикл из 8 чисел слишком мал. В литературе предлагаются константы для генерации перестановки чисел от 0 до 65525:

$$m = 25173; n = 13849; c = 65536.$$

Подбор констант для получения псевдоравномерной последовательности – нетривиальная задача, требующая проверки гипотез о равномерности, независимости и однородности получаемых реализаций.

При реализации на ЭВМ значение  $r_0$  генерируют с использованием таймера машины. В этом случае для различных моментов времени запуска программы получают различные фрагменты

псевдослучайной последовательности, смещенные друг относительно друга. Для получения одной и той же последовательности достаточно в качестве  $r_0$  выбрать одно и тоже число, например, принять  $r_0=0$ .

Соотношение (4.12) позволяет получить последовательность целых чисел, распределенных равномерно в интервале от 0 до  $(c-1)$ . Для получения числа, равномерно распределенного в интервале  $(0,1)$ , необходимо  $r_k$  разделить на величину периода:

$$R_k(0,1) = r_k / c. \quad (4.13)$$

Тогда для произвольного интервала  $(a,b)$  получим следующее соотношение

$$R_k(a,b) = a + (b-a) \times R_k(0,1). \quad (4.14)$$

Для получения псевдонормальной последовательности можно воспользоваться следствием из основной теоремы теории вероятности, согласно которому при сложении случайных чисел полученные суммы представляют собой последовательность, распределение которой близко к нормальному закону. Как уже отмечалось выше, при проведении практических расчетов для получения  $k$ -ого значения последовательности стандартного нормального распределения достаточно использовать сумму из 12 равномерно распределенных чисел:

$$N_k(0,1) = \sum_{i=1}^{12} R_i(0,1) - 6 \quad (4.15)$$

Для нормального распределения со средним значением  $m$  и среднеквадратическим отклонением  $\sigma$  случайное число получаем, используя (4.15):

$$N_k(m, \sigma) = m + \sigma \cdot N_k(0,1) \quad (4.16)$$

Для получения псевдопоказательной последовательности можно воспользоваться преобразованием значения псевдоравномерного числа:

$$E_k(\lambda) = -\frac{\text{Ln}[R_k(0,1)]}{\lambda} \quad (4.17)$$

Иллюстрацию особенностей анализа моделей в условиях неопределенности проведем на примере задачи о баскетболисте, постановка которой приведена в главе 2.

### Задача о баскетболисте

Несколько изменим содержательную постановку задачи о баскетболисте, добавив вопрос об оценке вероятности попадания мяча в кольцо при броске со штрафной линии. Пусть начальная скорость  $V_0$  мяча распределена по нормальному закону со средним значением  $M(V_0)=6,44$  м/с и средним квадратическим отклонением  $s(V_0)=0,25$  м/с. В соответствии с правилом "трех сигм" практически все значения  $V_0$  в этом случае (более 99,7%) лежат в интервале  $M(V_0) \pm 3s(V_0)$ , т.е.  $V_0 \in [5,69; 7,19]$  м/с. Величину угла бросания  $\alpha_0$  примем также нормально распределенной с параметрами  $M(\alpha_0)=45^\circ$  и  $s(\alpha_0)=5^\circ$ .

Для анализа вероятности попадания воспользуемся методом Монте-Карло. Проведем первоначально 16 вычислительных экспериментов. Используя (4.16), получим последовательность из 16 значений  $V_0$  и  $\alpha_0$ . Затем по соотношению (2.12) вычислим соответствующие каждому значению скорости и угла бросания значение точности броска (таблица 4.1). В главе 2 показано, что попадание мяча в корзину соответствует отклонению траектории центра мяча от центра кольца не более чем на 0,165 м. Из 16 бросков данному условию удовлетворяют только 7, т.е. отношение числа попаданий к общему количеству бросков или *частость* попаданий равна 0,4375 или 43,75%. Однако выборка из 16 бросков не является достаточно *представительной*, так как в последующих 100 сериях по 16 бросков число попаданий изменялось от 1 до 11 попаданий в серии, т.е. частость попаданий колебалась от 0,0625 до 0,6875.

Таблица 4.1.

|            |       |       |       |       |       |       |       |       |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|
| №          | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     |
| $V_o$      | 6,33  | 6,15  | 6,34  | 6,89  | 6,61  | 6,58  | 6,38  | 6,46  |
| $\alpha_0$ | 49,94 | 47,71 | 36,75 | 47,42 | 43,26 | 35,55 | 49,14 | 44,52 |
| $\Delta_i$ | -0,21 | -0,39 | -0,30 | 0,60  | 0,23  | -0,05 | -0,12 | 0,03  |
| №          | 9     | 10    | 11    | 12    | 13    | 14    | 15    | 16    |
| $V_o$      | 6,51  | 6,62  | 6,34  | 6,48  | 6,48  | 6,65  | 6,64  | 6,76  |
| $\alpha_0$ | 51,13 | 51,93 | 49,24 | 51,80 | 43,07 | 40,83 | 42,98 | 44,64 |
| $\Delta_i$ | -0,01 | 0,11  | -0,17 | -0,07 | 0,05  | 0,24  | 0,25  | 0,43  |

Границы изменения относительной частоты попаданий, вычисленные по 100 сериям бросков, в зависимости от объема выборки показаны на рис. 5.4. Можно видеть, что с увеличением объема выборки разброс значений оценки вероятности попадания уменьшается, стремясь к некоторому предельному значению в районе 0,36. Таким образом, вероятность попадания со штрафной линии при принятых исходных данных составляет 36%, т.е. в среднем только каждый 3-ий бросок попадает в цель.

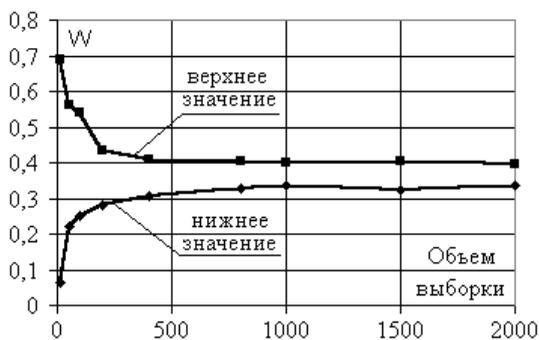


Рисунок 4.4 – Изменение предельных значений относительной частоты попаданий в зависимости от объема выборки

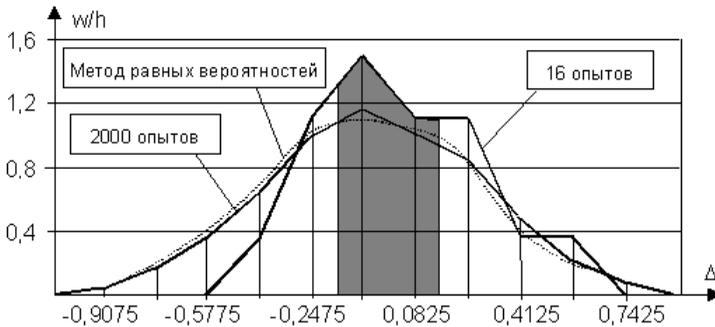


Рисунок 4.5 – Гистограмма относительных частот для точности броска

На рисунке 4.5 приведены гистограммы относительных частот для точности броска, построенные как по выборке из таблицы 4.1, так и по выборке из 2000 опытов. Для наглядности отрезками прямых линий соединены значения относительных частот в средних точках интервалов. Длина  $h$  интервалов по оси абсцисс была равной 0,165 м. Анализ приведенных гистограмм позволяет отметить следующее:

- увеличение объема выборки приводит к построению более сглаженных и более достоверных графиков;
- максимум плотности частоты (значение моды) наблюдается при  $\Delta = -0,0825$ ;
- значение медианы равно  $-0,14$ , т.е. недолет мяча более вероятен, чем перелет.

Попадание мяча в корзину происходит, если  $\Delta \in [-0,165; 0,165]$ . Вероятность попадания можно оценить по гистограмме относительных частот, вычислив площадь фигуры, заштрихованной на рисунке 4.5. Для выборки из 16 опытов значение вероятности попаданий равно 0,43 или 43%. Для выборки из 2000 опытов – 0,348 или 34,8%. Данный результат близок к полученному выше.

К существенным недостаткам метода Монте-Карло можно отнести необходимость в проведении большого числа вычислительных экспериментов для получения надежных оценок случайных переменных модели. Например, при нормальном законе рас-

пределения случайных параметров примерное число вычислений исследуемых переменных (необходимое число опытов) при построении гистограммы относительных частот можно определить из оценки отклонения относительной частоты от постоянного значения вероятности:

$$P\left(\left|\frac{m}{N} - p\right| \leq \varepsilon\right) = 2\Phi\left(\frac{\varepsilon\sqrt{N}}{\sqrt{p(1-p)}}\right) \quad (4.18)$$

где  $\Phi(\cdot)$  – функция Лапласа;  $\varepsilon$  – величина отклонения ( $\varepsilon$  – полоска);  $p$  – оцениваемая величина вероятности.

Величина отклонения  $\varepsilon$  определяет точность построения гистограммы. Например, при  $\varepsilon = 0,05$  оцениваемая величина  $p$  лежит в интервале, составляющем  $\pm 5\%$  от значения относительной частоты. Если в результате вычисления по (5.18) значение вероятности получилось равным 0,91, то это означает, что истинное значение  $p$  лежит в  $\varepsilon$ -полоске с вероятностью 91%. Задаваясь вероятностью  $P(\varepsilon)$  попадания в  $\varepsilon$ -полоску и величиной точности  $\varepsilon$ , из соотношения (5.18) можно получить оценку необходимого числа опытов

$$N = \left[ \frac{\Phi^{-1}(P(\varepsilon)/2)}{\varepsilon} \right]^2 p(1-p).$$

Из анализа данного соотношения можно заключить, что наибольшее количество опытов необходимо при  $p=0,5$  и оно обратно пропорционально квадрату точности  $\varepsilon$ . Например, при  $\varepsilon=0,05$  и  $P(\varepsilon)=0,95$  необходимо провести 384 опыта, а с точностью  $\varepsilon=0,01$  – 9600 опытов.

В случае, когда число случайных параметров модели невелико, для построения гистограммы относительных частот можно предложить более эффективный приближенный метод, который можно было бы назвать *методом равных вероятностей*. Поясним суть данного метода.

Пусть известна функция  $Y=f(X)$  и задан закон распределения случайной величины  $X$ . Требуется построить гистограмму для величины  $Y$ . Разобьем область значений  $X$  на ряд интервалов таким образом, чтобы в каждой из них функция  $f(X)$  была монотонна. Так как функция на интервале монотонна, то вероятность по-

падания значений  $X$  в интервал  $[X_k, X_{k+1}]$  равна вероятности попадания  $Y$  в соответствующий интервал  $[Y_k, Y_{k+1}]$ :

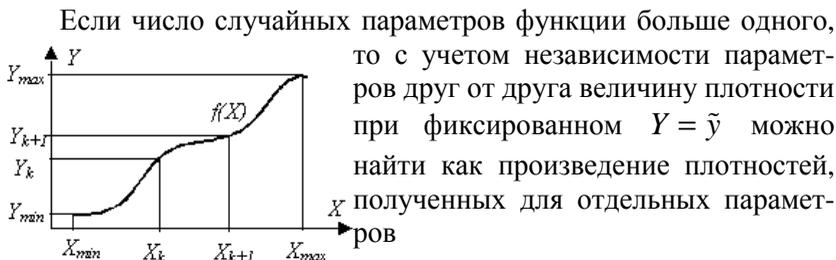
$$P(X_k \leq x \leq X_{k+1}) = P(Y_k \leq y \leq Y_{k+1}) = P(f(X_k) \leq x \leq f(X_{k+1})).$$

Тогда среднюю плотность величины  $Y$  на интервале  $[Y_k, Y_{k+1}]$  можно вычислить следующим образом

$$p(Y_k \leq Y \leq Y_{k+1}) = \frac{1}{|Y_{k+1} - Y_k|} \int_{X_k}^{X_{k+1}} p(X) dx \quad (4.19)$$

где  $p(X)$  – плотность распределения величины  $X$ .

Теперь необходимо заменить непрерывные случайные величины их дискретными аналогами на интервалах. Точность метода повышается с увеличением числа интервалов разбиения области значений величины  $X$ . Если функция на области значений  $X$  не монотонна, то с учетом *несовместности*  $X$  (т.е. невозможности одновременной реализации двух различных значений  $X$ ) значение плотности распределения для фиксированного  $Y = \tilde{y}$  определяется суммированием плотностей по всем интервалам для данного  $\tilde{y}$ .



$$p(Y = \tilde{y}) = \prod_{k=1}^M p_{x_k}(Y = \tilde{y}) \quad (4.20)$$

где  $M$  – число случайных параметров.

Возвращаясь к примеру о баскетболисте, построим гистограмму относительных частот с применением метода равных вероятностей. В модели использовано два случайных параметра  $V_0$  и  $a_0$ , имеющих заданное нормальное распределение. Область возможных значений для  $V_0$  и  $a_0$  выберем в соответствии с правилом "трех сигм". Разобьем выбранные области на 12 интервалов по 0,5s. Вероятности попаданий значений случайной величины  $X$  в

заданные интервалы для стандартного нормального распределения можно вычислить с помощью функции Лапласа (см. табл. 5.2).

Таблица 4.2.

|     |            |            |            |            |            |            |
|-----|------------|------------|------------|------------|------------|------------|
| $X$ | [0,0; 0,5] | [0,5; 1,0] | [1,0; 1,5] | [1,5; 2,0] | [2,0; 2,5] | [2,5; 3,0] |
| $P$ | 0,1915     | 0,1498     | 0,0919     | 0,0440     | 0,0166     | 0,0049     |

Используя данные, приведенные в таблице, а также соотношения (4.19) и (4.20), можно получить распределение относительной частоты. Полученная при этом гистограмма показана на рисунке 4.5 пунктирной линией. Как можно видеть, она практически совпадает с гистограммой для выборки, построенной по 2000 опытов, являясь более сглаженной. Вероятность попадания мяча, вычисленная по полученной гистограмме, составляет 35,5%.

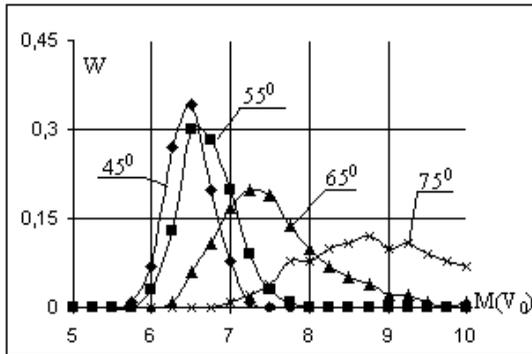


Рисунок 4.6 – Влияние  $M(V_0)$  и  $M(\alpha_0)$  на относительную частоту попаданий

В проведенном в главе 2 анализе показано, что существует бесконечно много пар значений начальных параметров броска, обеспечивающих попадание мяча в корзину. Интересно посмотреть, насколько те или иные сочетания  $V_0$  и  $\alpha_0$  предпочтительны с точки зрения попадания мяча в корзину. Пусть среднее квадратическое отклонение (степень разброса) начальных параметров ос-

тается неизменной и равной  $s(V_0)=0,25\text{ м/с}$  и  $s(\alpha_0)=5^\circ$ . Исследуем, как изменяется вероятность попадания мяча при изменении математического ожидания исходных параметров:  $5 \leq M(V_0) \leq 10 \text{ м/с}$ ;  $45^\circ \leq M(\alpha_0) \leq 75^\circ$ .

Полученные результаты для относительной частоты попаданий, представленные на рис.5.6, показывают, что с увеличением математического ожидания угла бросания относительная частота попаданий снижается. Данные результаты позволяют сделать вывод, что различные сочетания  $V_0$  и  $\alpha_0$  не являются равнозначными. Вероятность попадания при броске под углом  $75^\circ$  более чем в два раза ниже, чем при броске под углом в  $45^\circ$ , т.е. бросок под углом в  $45^\circ$  является наиболее предпочтительным.

#### 4.4 Моделирование марковских случайных процессов

Человек, имеющий одни часы, твердо знает, который час. Человек, имеющий несколько часов, ни в чем не уверен.

Закон Сегала

Любые данные, полученные в результате временных наблюдений реального физического явления, можно классифицировать как *процесс*. При этом если процесс можно описать явными математическими формулами в терминах детерминированных переменных, то это будет *детерминированный процесс*. В противном случае, если любое наблюдение дает только один вариант из множества возможных, получим *случайный процесс*. Таким образом, *случайный процесс* – это случайная величина, зависящая от времени.

Введем некоторые основные определения.

*Выборочной функцией (реализацией, траекторией)* случайного процесса будем называть конкретную реализацию описываемого случайного явления. Тогда *случайный процесс* – это совокупность выборочных функций, которые может дать исследуемое случайное явление.

*Сечением* случайного процесса будем называть случайную величину, характеризующую исследуемое случайное явление в некоторый момент времени, то есть это временной срез *случайного процесса*.

Обычно простую математическую модель, позволяющую получить решение в явном (аналитическом) виде, найти не удается. Однако в некоторых особых случаях такую математическую модель удастся построить. Подобная ситуация, например, имеет место в случае, когда исследуемая операция представляет собой (точно или приближенно) так называемый *марковский случайный процесс*.

Одним из основных признаков, по которым классифицируют случайные процессы, является отношение зависимости между сечениями случайного процесса.

*Случайный процесс* называется *марковским*, если для любого момента времени  $t_0$  вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент  $t_0$  и не зависят от того, когда и как *система* пришла в это состояние.

Приведем пример простейшего марковского процесса. Пусть система – это счетчик Гейгера, на который время от времени попадают космические частицы; состояние системы в момент времени  $t$  характеризуется показанием счетчика – числом частиц, пришедших до данного момента. Пусть в момент  $t_0$  индикатор счетчика показывает  $S_0$ . Вероятность того, что в момент  $t > t_0$  индикатор покажет то или другое число частиц  $S_1$  (или не менее  $S_1$ ), разумеется, зависит от  $S_0$ , но не зависит от того, в какие именно моменты приходили частицы до момента  $t_0$ .

Для марковского процесса будущее зависит от прошлого только через настоящее.

В реальных условиях, в сущности, любой процесс можно рассматривать как марковский, если все параметры из "прошлого", от которых зависит "будущее", включить в "настоящее".

Хотя на практике марковские процессы в чистом виде обычно не встречаются, нередко приходится иметь дело с процессами, для которых влиянием "предыстории" можно пренебречь. При изучении таких процессов можно применять марковские модели.

Наиболее простыми марковскими процессами являются марковские случайные процессы с дискретными состояниями и дискретным или непрерывным временем.

Марковский процесс называется *процессом с дискретными состояниями*, если его возможные состояния  $S_1, S_2, S_3, \dots$  можно заранее перечислить (перенумеровать), и переход системы из состояния в состояние переходит "скачком", практически мгновенно. Например, пылесос может находиться в трех состояниях: включенном, выключенном и в состоянии поломки. Процесс называется *процессом с непрерывным временем*, если моменты возможных переходов из состояния в состояние не фиксированы заранее, а неопределенны, случайны, если переход может осуществиться, в принципе, в любой момент. Если процесс марковский и моменты возможных переходов из состояния в состояние фиксированы, то этот *марковский процесс с дискретным временем* называется *марковской цепью*.

Мы будем рассматривать только марковские процессы с дискретными состояниями и непрерывным временем.

### **Пример 1**

Техническое устройство  $S$  состоит из двух узлов, каждый из которых в случайный момент времени может выйти из строя (отказ), после чего мгновенно начинается ремонт узла, тоже продолжающийся заранее неизвестное, случайное время. Графическое изображение возможных состояний системы и порядок переходов называется *графом состояний системы* (рис. 5.7). Возможные состояния системы:  $S_0$  – оба узла исправны;  $S_1$  – первый узел ремонтируется, второй исправен;  $S_2$  – второй узел ремонтируется, первый исправен;  $S_3$  – оба узла ремонтируются. Переходы системы  $S$  из состояния в состояние происходят практически мгновенно, в случайные моменты выхода из строя того или другого узла или окончания ремонта.

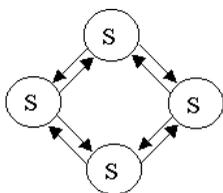


Рисунок 4.7 – Граф состояний

Чтобы построить математическую модель марковского процесса с дискретными состояниями и непрерывным временем необходимо ввести понятие *потока событий*.

*Потоком событий* называется последовательность однотипных ситуаций, следующих одна за другой в какие-то случайные моменты времени.

Например:

- поток отказов вычислительной техники;
- появление бракованных изделий.

Ситуации, образующие поток, сами вероятностью не обладают; вероятностями обладают другие, основанные на них события. Например: исходная ситуация – попадание на участок времени  $\Delta t$  некоторого события. Случайное событие – попадание на участок времени  $\Delta t$  ровно двух событий.

Введем некоторые характеристики потока событий. Важнейшей характеристикой потока событий является его *интенсивность*  $\lambda$  – среднее число событий, приходящихся на единицу времени. Интенсивность потока может быть как постоянной ( $\lambda = \text{const}$ ), так и зависящей от времени  $t$ .

Поток событий называется *регулярным*, если события следуют одно за другим через определенные равные промежутки времени.

Поток событий называется *стационарным*, если его вероятностные характеристики не зависят от времени. В частности, интенсивность  $\lambda$  стационарного потока должна быть постоянной.

Поток событий называется *поток без последовательности*, если для любых двух непересекающихся участков времени  $\tau_1$  и  $\tau_2$  число событий, попадающих на один из них, не зависит от того, сколько событий попало на другой. Например, в приведенном

выше примере со счетчиком Гейгера поток поступления частиц можно считать потоком без последействия.

Поток событий называется *ординарным*, если события в нем появляются поодиночке, а не группами по несколько событий сразу. Например, поток событий – прохождение автомобилей мимо наблюдателя по дороге с однорядным движением. Поток поступления вагонов на железнодорожную станцию не является ординарным, так как вагоны прибывают на станцию составами.

Поток событий называется *простейшим (или стационарным пуассоновским)*, если он обладает сразу тремя свойствами: *стационарен, ординарен и не имеет последействия*.

Простейший поток играет среди других потоков особую роль. Это связано с тем, что при наложении (суперпозиции) достаточно большого числа независимых, стационарных и ординарных потоков (сравнимых между собой по интенсивности) получается поток, близкий к простейшему.

Для простейшего потока с интенсивностью  $\lambda$  интервал времени  $\Delta t$  между соседними событиями имеет *показательное распределение* с плотностью

$$f(\Delta t) = \lambda e^{-\lambda \Delta t} \quad (4.21)$$

Величина  $\lambda$  в (4.21) называется *параметром* показательного закона.

Если ввести так называемый *коэффициент вариации* для случайного интервала времени  $T$ , равный отношению среднеквадратического отклонения к математическому ожиданию ( $v_T = \sigma_T / m_T$ ), то можно сделать вывод о том, что для простейшего потока событий коэффициент вариации интервалов между событиями равен единице, так как для показательного распределения  $\sigma_T = m_T = 1/\lambda$ .

Очевидно, что для регулярного потока событий ( $\sigma_T = 0$ ), коэффициент вариации  $v_T$  равен нулю. Для большинства потоков событий, встречающихся на практике, коэффициент вариации интервалов между событиями заключен между нулем и единицей и может служить некоторой мерой *степени регулярности потока*.

Простейший поток является наименее регулярным потоком.

Величина числа событий  $N$  за время  $\Delta t$  для простейшего потока подчиняется распределению Пуассона

$$P_{N=k} = \frac{(\lambda \Delta t)^k}{k!} \cdot e^{-\lambda \Delta t} \quad (4.22)$$

В частности, вероятность непоявления на интервале времени  $\Delta t$  ни одного события составляет  $P_{N=0} = e^{-\lambda \Delta t}$ ; более одного события –  $P_{N>0} = 1 - e^{-\lambda \Delta t}$ .

В случае достаточно малого промежутка времени ( $\Delta t \ll 1$ ) величина экспоненты в (4.22)  $e^{-\lambda \Delta t} \approx 1$ . Тогда вероятность появления на таком малом интервале времени ровно одного события равна

$$P_{N=1} = p_{\Delta t} = \lambda \Delta t, \quad (4.23)$$

где  $0 < \Delta t \ll 1$ .

Величину  $p_{\Delta t}$ , входящую в (4.23), называют *элементом вероятности*. Заметим, что элемент вероятности для простейшего потока не зависит от того, сколько событий и когда появилось ранее.

Поток событий называется *рекуррентным (потоком Пальма)*, если он стационарен, ординарен, а интервалы времени между событиями  $T_1, T_2, T_3, \dots$  представляют собой независимые случайные величины с одинаковым *произвольным* распределением. Очевидно, что простейший поток представляет собой частный случай рекуррентного потока, когда интервалы между событиями имеют показательное распределение.

Рассмотрим моделирование марковского случайного процесса. Для этого составим и решим так называемые *уравнения Колмогорова* – особого вида дифференциальные уравнения, в которых неизвестными (искомыми) функциями являются вероятности состояний. Покажем на конкретном примере, как эти уравнения составляются.

Назовем вероятностью  $i$ -ого состояния вероятность  $p_i(t)$  того, что система в момент времени  $t$  находится в состоянии  $S_i$ . Так как число состояний конечно и система должна находиться хотя бы в

одном из них, то в любой момент времени справедливо следующее, так называемое *нормировочное условие*:

$$\sum_{k=1}^N p_k(t) = 1 \quad (4.24)$$

где  $N$  – число состояний системы.

Пусть система  $S$  имеет четыре состояния:  $S_1, S_2, S_3, S_4$ , размеченный *граф состояний* [18] которой показан на рис. 5.8. Переход системы из состояния  $i$  в состояние  $j$  представляет собой простейший поток интенсивностью  $\lambda_{ij}$ . Рассмотрим одну из вероятностей состояния, например  $p_1(t)$  – вероятность того, что в момент  $t$  система будет в состоянии  $S_1$ . Придадим  $t$  малое приращение  $\Delta t$  и найдем  $p_1(t+\Delta t)$  – вероятность того, что в момент  $(t+\Delta t)$  система будет в состоянии  $S_1$ . Это может произойти, очевидно, в двух случаях: либо в момент  $t$  система уже была в состоянии  $S_1$ , а за время  $\Delta t$  не вышла из него; либо в момент  $t$  система была в состоянии  $S_2$ , а за время  $\Delta t$  перешла из него в состояние  $S_1$ .

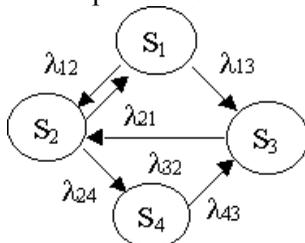


Рисунок 4.8 – Граф состояний

Найдем вероятность развития событий по первому варианту. Система находится в состоянии  $S_1$ . В этом случае возможен переход в состояние  $S_2$  или  $S_3$ . Это несовместные события, т.е. они не могут произойти одновременно. Вероятность появления одного из двух несовместных событий равна сумме вероятностей:  $(\lambda_{12}+\lambda_{13})\Delta t$ . Тогда вероятность того, что система останется в состоянии  $S_1$ , равна  $1-(\lambda_{12}+\lambda_{13})\Delta t$ . Так как вероятность оказаться в состоянии  $S_1$  к моменту времени  $t$  у системы равна  $p_1(t)$  и последующие переходы не зависят от того, каким образом система

пришла к этому состоянию, то вероятность первого варианта равна произведению вероятностей  $p_1(t)[1 - (\lambda_{12} + \lambda_{13})\Delta t]$ .

Рассуждая аналогично, получим вероятность второго варианта  $p_2(t)\lambda_{21}\Delta t$ . Так как первый и второй вариант несовместны (реализуется один из них), то вероятность для системы находиться к моменту времени  $(t+\Delta t)$  в состоянии  $S_1$  равна:

$$p_1(t + \Delta t) = p_1(t)[1 - (\lambda_{12} + \lambda_{13})\Delta t] + p_2(t)\lambda_{21}\Delta t \text{ откуда}$$

$$[p_1(t + \Delta t) - p_1(t)] / \Delta t = \lambda_{21}p_2(t) - (\lambda_{12} + \lambda_{13})p_1(t).$$

Устремив  $\Delta t$  к нулю и переходя к пределу, получим дифференциальное уравнение для  $p_1(t)$ :

$$\frac{dp_1}{dt} = \lambda_{21}p_2 - (\lambda_{12} + \lambda_{13})p_1.$$

Аналогичным образом можно получить остальные три уравнения. В итоге получим систему четырех обыкновенных дифференциальных уравнений Колмогорова с четырьмя неизвестными функциями времени  $p_1, p_2, p_3, p_4$ :

$$\begin{cases} \frac{dp_1}{dt} = \lambda_{21}p_2 - (\lambda_{12} + \lambda_{13})p_1, \\ \frac{dp_2}{dt} = \lambda_{12}p_1 + \lambda_{32}p_3 - (\lambda_{24} + \lambda_{21})p_2, \\ \frac{dp_3}{dt} = \lambda_{31}p_1 + \lambda_{43}p_4 - \lambda_{32}p_2, \\ \frac{dp_4}{dt} = \lambda_{24}p_2 - \lambda_{43}p_4 \end{cases} \quad (4.25)$$

Заметим, что любое из них можно отбросить, заменив соотношением (4.24).

Чтобы решить уравнение Колмогорова и найти вероятности состояний, прежде всего необходимо задать начальные условия. Если мы точно знаем, что в начальный момент система находится в состоянии  $S_1$ , то для этого момента  $p_1(0)=1$ , а все остальные начальные вероятности равны нулю. Теперь полученную систему линейных дифференциальных уравнений можно решить аналитически или численно. В результате решения получим все вероятности состояний как функции времени. Типичное распределе-

ние вероятностей показано на рисунке 4.9. Характерной особенностью является то, что, начиная с некоторого момента времени  $t_c$ , вероятности состояний практически перестают зависеть от времени, принимая некоторое постоянное стационарное значение.

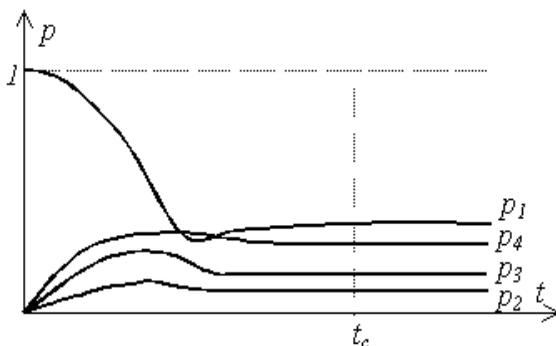


Рисунок 4.9 – Вероятности состояний

В теории случайных процессов доказывается, что если число состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое, то в этом случае существуют так называемые *финальные вероятности* (пределы вероятности состояний при  $t \rightarrow \infty$ , не зависящие от начального состояния системы).

Финальные вероятности будем обозначать теми же буквами, что и сами вероятности состояния ( $p_1, p_2, p_3, p_4$ ), но теперь это уже не функции времени, а постоянные числа. Финальную вероятность состояния  $S_i$  можно рассматривать как среднее относительное время пребывания системы в этом состоянии. Например, если система  $S$  имеет три состояния  $S_1, S_2, S_3$  и их финальные вероятности равны 0,1, 0,2 и 0,7, то это значит, что в предельном стационарном режиме система в среднем одну десятую часть времени проводит в состоянии  $S_1$ , две десятых – в состоянии  $S_2$  и семь десятых – в состоянии  $S_3$ .

Чтобы найти финальные вероятности, воспользуемся тем, что теперь  $p_1, p_2, p_3, p_4$  – постоянны, то есть их производные по времени равны нулю. Теперь из уравнений Колмогорова получим однородную систему линейных алгебраических уравнений. Од-

нако в связи с тем, что в уравнениях нет свободных членов, неизвестные определяются с точностью до произвольного множителя. Поэтому необходимо воспользоваться нормировочным условием (4.24). При этом одно (любое, как правило, наиболее сложное по виду) из уравнений можно отбросить (оно вытекает как следствие из остальных).

## **ГЛАВА 5. ЛИНЕЙНЫЕ И НЕЛИНЕЙНЫЕ МОДЕЛИ**

### ***Введение***

Главное – это из множества проблем выбрать наиболее простые, решение которых позволит выработать допускающие обобщения концепции.

Д. Гильберт

В этой главе авторы преследовали цель на примерах показать читателю, что могут и что не могут описать линейные модели явлений и процессов, происходящих в нашем, в общем-то, нелинейном мире. Для того, чтобы распознать признаки линейности или нелинейности изучаемого явления, ученый должен владеть культурой (набором "штампов") нелинейной науки, с некоторыми понятиями и идеями которой мы попытаемся ознакомить читателя. Эти соображения определили некоторую специфику построения главы, где процедуре построения моделей уделяется меньше внимания, чем характерным признакам, сопровождающим линейные и нелинейные явления. Свои "рассказы" мы постарались выдержать в прикладном духе, чтобы первоначальное назначение математики, "созданной для достижения победы над противником", не отходило далеко на задний план. Авторы смогли написать только о малой части области науки, называемой "нелинейной наукой", более полную картину читатель сможет найти в популярных книгах и многих других цитированных там работах.

## 5.1 О законе Гука и границах линейности

Приступая к созданию физической теории, необходимо отбросить все предшествующие физические модели, а также основанную на них физическую "интуицию".

П. Дирак

С одной из первых линейных моделей – законом Гука

$$F = -kx \quad (5.1)$$

мы знакомимся, только начиная изучать физику. В (5.1)  $F$  – растягивающая сила, приложенная к концам образца,  $x$  – относительное перемещение его концов,  $k$  – некоторая постоянная для данного образца (материала). Этот закон в виде анаграммы "ceiinossttuv" был помещен Робертом Гуком на свободном месте в конце его работы о гелиоскопе в 1676 году. Два года спустя в трактате "О восстановительной способности или об упругости" анаграмма им была расшифрована: "ut tensio sic vis" – "какова сила, таково растяжение". Гук отмечал, что он впервые открыл этот закон еще в 1660 году, но от опубликования его удержало стремление защитить свое изобретение спиральных часовых пружин. Для их конструирования Гук использовал факт независимости частоты колебаний от их амплитуды, вытекающий из линейности связи (6.1).

Упругость в современном понимании означает существование однозначной монотонно возрастающей функции, связывающей напряжение  $\sigma = F/S$  ( $S$  – площадь поперечного сечения образца) и деформацию  $\varepsilon = x/l$  ( $l$  – длина образца):

$$\sigma = f(\varepsilon), f(0) = 0. \quad (5.2)$$

Функция  $f$  в общем случае – нелинейная. Нелинейными упругими свойствами обладают, например, высокоэластичные резины. Резиновый шнур можно растянуть в десять раз ( $\varepsilon = 0,9$ ), а затем отпустить, после чего он восстановит свою длину. При таких больших степенях деформации не следует ожидать линейности связи  $\sigma(\varepsilon)$ . Гук изучал длинные металлические проволоки, которые подвергал малым деформациям ( $\varepsilon \approx 0,001$ ), и обнаружить не-

линейность реакции не удалось. Это означает, что при таких  $\varepsilon$  первый член разложения функции (5.2) (полагая ее аналитической) в степенной ряд

$$\sigma = \frac{\partial f}{\partial \varepsilon} \varepsilon + \frac{1}{2!} \frac{\partial^2 f}{\partial \varepsilon^2} \varepsilon^2 + \dots \quad (5.3)$$

значительно превосходит все остальные. Соотношение (5.1) в этом случае можно переписать в виде

$$\sigma = E\varepsilon, \quad (5.4)$$

где  $E$  – модуль упругости материала при его одноосном растяжении-сжатии.

Открытие Гуком упругого закона вызвало весьма большой интерес научного сообщества; опыты Гука демонстрировались самому королю Англии Карлу II. Поспешно обобщая свои результаты, Гук заявлял: *"Линейную зависимость можно наблюдать не только в этих телах, но и любых других упругих материалах, будь то металл, дерево, камни, спекшаяся глина, волос, шелк, кость, сухожилия, стекло и тому подобное"*. Тем не менее, не все ученые того времени разделяли оптимизм в универсальность линейного закона. В 1687 году Яков Бернулли провел опыты со скрипичной струной, изготовленной из кетгута (cat gut – кошачья кишка), показавшие нелинейность реакции. Готфрид Вильгельм Лейбниц, прочтя трактат Гука, в письме Христиану Гюйгенсу в 1691 г. высказал осторожное недоверие линейному закону и попросил его прислать эксперименты, *"которые Вы проводили по этому предмету"*. Гюйгенс признался, что он согласен с результатами Гука, но только если пружины будут растянуты незначительно. И только в 1849 году Британская королевская комиссия, назначенная для исследования применения железа для железнодорожных сооружений, "отменила" закон (6.4) в пользу параболической зависимости  $\sigma = A\varepsilon - B\varepsilon^2$ .

Заклечение комиссии основывалось на почти двадцатилетнем детальном изучении данных экспериментальных исследований на образцах из железа.

Тем не менее, факт, что при достаточно малых деформациях была обнаружена линейная зависимость между ними и напряжениями в металлах, привел к господству идей *линейной* теории уп-

ругости в течение почти трех столетий. Исключительное значение результатов этой теории подчеркивает Джеймс Фредерик Белл: *"Если бы в 17 веке для твердых тел наблюдались исключительно нелинейные зависимости между напряжениями и деформациями, то большинство достижений в физике и технике, имевших место за прошедшие 200 лет, задержалось бы на несколько столетий"*. Механики-естествоиспытатели получили возможность продвигаться вперед, развивая линейную теорию упругости, а вслед за ней и математический аппарат, пригодный для построения моделей многих других явлений и процессов.

Из этой истории мы можем видеть, что линейная модель упругости занимает только определенную нишу в теории упругости. Этот факт в естествознании – достаточно общий. Как правило, любая линейная теория справедлива в ограниченных пространственных и временных рамках и при малой интенсивности воздействий на изучаемую систему. Например, архитектор без ущерба для своего дела может не принимать во внимание кривизну земной поверхности, поскольку размеры рассматриваемого им участка малы по сравнению с радиусом Земли ( $\approx 6400$  км). Точно так же конструктор авиационной или космической техники в своей работе вполне сможет обойтись ньютоновской классической механикой, не прибегая к теории относительности, поскольку последняя "вступает в силу" при скоростях, соизмеримых со скоростью света ( $c \approx 300\,000$  км/с). При меньших скоростях  $v$  нелинейностью в основных законах можно пренебречь с точностью до малых порядка  $(v/c)^2$ .

Если же мы изучаем явление, не являющееся линейным в интересующих нас рамках, то взяв в качестве первого приближения линейную модель, мы можем не получить даже качественного его описания. Нелинейность в таких явлениях принципиальна. Об этом свидетельствует целый ряд фактов, некоторые из них приведены в настоящей главе.

## **5.2 Поля, сплошные среды и уравнения математической физики. Линейные уравнения и принцип суперпозиции**

Часть, в некотором смысле, больше целого... Общее утверждение должно доказываться легче, чем любой его частный случай.

Дж. Сильвестр

Ранее мы говорили о линейной *функции*. Более богатые модели дают нам уравнения математической физики, которые также могут быть линейными или нелинейными. Можно сказать, что математическая физика – это наука, изучающая математические модели физических явлений. Уравнения математической физики – это чаще всего дифференциальные уравнения в частных производных, описывающие процессы в пространстве и времени. Таким образом, искомые функции должны иметь не менее двух аргументов; дифференциальные уравнения для функции одного аргумента относятся к классу обыкновенных дифференциальных уравнений. В общем случае мы рассматриваем процесс в области (возможно, неограниченной) трехмерного евклидова пространства; то есть, на такой области определено физическое поле, которое изменяется со временем.

Одно из первых уравнений такого типа появилось благодаря Пьеру Симону Лапласу. Лаплас с недоверием относился к закону Всемирного тяготения Ньютона: как планеты, находясь на колоссальных дистанциях, могут действовать друг на друга? Лапласу казалась реальнее концепция контактной передачи усилия (короткодействие) от одной планеты к другой, посредством эфира. Введя в рассмотрение эту субстанцию, *поле*, заполняющую трехмерное пространство, он получил для нее уравнение, которое сейчас называют уравнением Лапласа. Концепция эфира не подтвердилась, но оказалось, что это уравнение описывает многие стационарные процессы в сплошных средах: равновесие линейно-упругого тела, потенциальное течение вязкой жидкости, стационарные распределения электрического и температурного полей в теле и т.д.

Скажем несколько слов о модели *сплошной среды*. Обычно рассматриваемые нами поля определены в области геометрического пространства. Поле при этом оказывается как бы внешним,

удобным для наблюдения и непрерывно заполняющим область математическим объектом. Механики имеют дело с полем материала, вложенного в это пространство, причем на этом материале обычно рассматриваются и другие поля (например, температурное). Физики работают, например, с электромагнитным или гравитационным полями, для которых "механический носитель" не требуется. Необходимо различать внешнее пространство и поле, поскольку последнее изменяется во времени (движется, течет, деформируется) относительно внешнего пространства, с которым (будем считать) жестко связана система отсчета наблюдателя. Оказывается, что реальные материальные среды вовсе не непрерывно заполняют пространство: при рассмотрении их вблизи они оказываются разрывными, состоящими из расположенных в пространстве зерен, гранул, молекул, атомов. Следовательно, и измеряемые на них физические поля с точки зрения внешнего геометрического континуума также не могут считаться непрерывными. Какую бы среду и на каких бы масштабах мы не рассматривали, объем, в котором сосредоточено собственно само вещество, оказывается много меньше объема тела. Но, несмотря на то, что все тела "состоят из пустоты", в практически малых объемах всегда оказывается огромное число частиц. Поэтому мы можем "размазать материю" – то есть полагать материальную среду и поле на ней непрерывными на масштабах, много больших по сравнению с характерными масштабами строения среды и осцилляций поля на ней. Применяя аппарат обычного дифференциального исчисления, мы должны осознавать, что работаем с уже осредненными полями (конечно, осредненное поле далеко не всегда полностью характеризует нам исходное поле). В механике модель сплошной среды применяют к газам, жидкостям и твердым телам. Авторы знакомы с менее обычным сплошнородным телом – газом, состоящим из галактик Вселенной, оказывающимся довольно плотным; газ, состоящий из звезд, оказывается на несколько порядков разреженнее.

Понятие линейности пришло в область дифференциальных уравнений, по-видимому, из геометрии. В линейном пространстве  $X$ , как известно, определены и замкнуты композиции  $\alpha x + \beta y$  для любых элементов  $x, y$  из  $X$  и любых действительных чисел

$\alpha, \beta$ . Рассмотрим пару  $n$ -мерных линейных пространств  $X, Y$  и отображение  $L$ , переводящее элементы  $X$  в  $Y$ . Оператор  $L$  называется линейным, если из соотношений

$$L(x_1) = y_1, L(x_2) = y_2$$

для любых  $\alpha, \beta$  следует

$$L(\alpha x_1 + \beta x_2) = \alpha y_1 + \beta y_2.$$

Но можно рассматривать линейные пространства *функций*, являющиеся бесконечномерными линейными пространствами, а оператор  $L$  – дифференциальным. Тогда на дифференциальное уравнение

$$L(x) = f \tag{5.5}$$

можно смотреть как на оператор, отображающий одно функциональное пространство в другое или в себя, если рассматривать пространство бесконечно гладких (аналитических) функций. В качестве  $L$  мы будем рассматривать линейный дифференциальный оператор в частных производных, а в качестве функций  $x = x(r), f = f(r)$  – функции аргумента  $r$ , четырехмерного пространства-времени. Нетрудно убедиться, что множество решений однородного дифференциального уравнения  $L(x) = 0$  образуют линейное пространство (возможно, состоящее из одной тождественно нулевой функции), являющееся подпространством  $X$  и называемое ядром оператора  $L$ . Обозначим  $x_0$  произвольный элемент ядра. Очевидно, если  $x_f$  – какое-либо решение неоднородного уравнения (5.5), то общее решение (5.5) запишется в виде

$$x = x_0 + x_f.$$

Элемент  $x_0$  конкретизируется из граничных и начальных условий задачи (иначе наша задача может иметь "слишком много" решений, если вообще их имеет).

Если мы разобьем функцию в правой части (5.5)  $f = f_1 + f_2$ , то в силу линейности нашего оператора (в предположении одинаковых граничных условий) решение уравнения запишется в виде  $x = x_0 + x_{f_1} + x_{f_2}$ , где  $x_{f_1}$  и  $x_{f_2}$  – решения уравнений соответственно  $L(x) = f_1, L(x) = f_2$ . Можно ожидать, что, представляя функ-

цию в правой части разложением в ряд Фурье, получая частное решение для каждой его гармоники и составляя аналогичный ряд из частных решений, мы получим частное решение исходного уравнения. Этот путь, использующий *принцип суперпозиции*, действительно приводит к результату.

На этом пути можно пойти еще дальше –научиться решать уравнение, у которого в правой части стоит дельта-функция Дирака:

$$L(x) = \delta(r - r'), \quad (5.6)$$

где  $r'$  –некоторая фиксированная, а  $r$  –по прежнему произвольная точки пространства-времени. Физики "изобрели" и пользовались дельта-функцией для решения уравнений в частных производных давно. Саму дельта-функцию нет нужды определять, поскольку она используется только на промежуточных этапах решения данных уравнений, достаточно знать, что для любой "допустимой" функции  $f(r)$ :

$$\int_{-\infty}^{+\infty} f(r) \delta(r - r') dr = f(r')$$

(о смысле "допустимости" говорить не будем, поскольку здесь нам важно показать идею метода). Это делает возможным представить правую часть (6.5) как бесконечную интегральную сумму  $f(r) = \int f(r') \delta(r - r') dr'$ . Тогда, если  $x(r) = G(r - r')$  - решение (6.6), то вследствие принципа суперпозиции решением (5.5) будет

$$x(r) = \int f(r') G(r - r') dr'.$$

$G(r)$  называют функцией Грина оператора  $L$  (при определенных граничных условиях).

Как мы видим, для решения *линейных* уравнений математической физики разработан аппарат, использующий принцип суперпозиции. Продемонстрируем его, решив одномерное волновое уравнение

$$\frac{\partial^2 v}{\partial t^2} - a^2 \frac{\partial^2 v}{\partial x^2} = 0 \quad (5.7)$$

Это уравнение описывает распространение линейных волн  $v(x,t)$  различной физической природы. Например, (6.7) можно вывести как уравнение малых возмущений совершенного идеально-го баротропного газа (или свободных продольных или поперечных колебаний струны).

### 5.3 Вывод волнового уравнения из законов механики

Нужно сковывать себя ограничениями –и тогда можно свободно выдумывать.

У. Эко

Попутно с выводом волнового уравнения мы покажем, как строятся сплошносредные модели механики и физики. Экспериментально открыт факт, что физические процессы удовлетворяют законам сохранения (массы, количества движения, энергии и др.). Для сплошного деформируемого тела закон сохранения массы в локальной форме принимает вид:

$$\frac{d\rho}{dt} + \rho \nabla \cdot v = 0, \quad (5.8)$$

где  $\rho = \rho(r, t)$  –плотность;  $r = r_i k_i$  –радиус-вектор произвольной геометрической точки;  $k_i$ ,  $i=1,2,3$ ; –декартов ортонормированный репер;  $r_i$ ,  $i=1,2,3$  –координаты радиус-вектора;

$\nabla \equiv k_i \frac{\partial}{\partial r_i}$  –набла-вектор (оператор Гамильтона);  $v=v(r, t)$  –

векторное поле скоростей бесконечно малых материальных частиц (материальных точек). (Очевидно, каждая материальная точка имеет только три степени свободы). Поскольку рассматриваемое тело движется относительно внешней системы координат, для каждой его *материальной точки*  $R$  имеет место закон движения  $r = r(R, t)$ . В отличие от теоретической механики в сплошном теле материальных точек оказывается настолько много, что для их идентификации необходим континуум точек внешней системы

координат:  $r(R,0) = R$ , таким образом,  $R$  есть радиус-вектор геометрической точки, в которой находилась материальная точка в начальный момент времени. Скорость  $v=v(r,t)$  материальной точки, находящейся в момент времени  $t$  в месте  $r$ , определяется как  $v = \frac{dr}{dt}$ . Производя дифференцирование произвольного поля  $a$ ,

можно получить:

$$\left. \frac{da}{dt} \right|_R = \left. \frac{\partial a}{\partial t} \right|_r + v \cdot \nabla a, \quad (5.9)$$

Рассмотрим, например, сплошную среду, *свободно* движущуюся вдоль прямой  $Ox$  со скоростью  $v(x,t)$ . Закон движения любой (материальной) точки этой среды  $x=\varphi(t)=x_0+ut$  ( $u$  – ее скорость) удовлетворяет II закону Ньютона:  $\frac{\partial^2 \varphi}{\partial t^2} = 0$ . По определению скорости материальной точки  $u = \frac{\partial \varphi}{\partial t} = v(\varphi(t), t)$ ; дифференцируя по  $t$ , получаем уравнение, которому удовлетворяет поле  $v(x,t)$ :  $\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = 0$ . Заметим, что за появление нелинейности  $v v_x$

(здесь  $v_x$  означает частную производную скорости по координате  $x$ ) отвечает переход от материальных координат к пространственным (внешним) координатам. Пространственные координаты  $r$  называются эйлеровыми, материальные же координаты  $R$ , позволяющие следить за судьбой отдельных материальных точек, не меняющиеся для каждой материальной частицы в течение всего исследуемого процесса, называются лагранжевыми.

Балансовое уравнение для плотности массы (5.8) с учетом (6.9) запишется в виде:

$$\frac{\partial \rho}{\partial t} + v \cdot \nabla \rho + \rho \nabla \cdot v = 0, \quad (5.10)$$

Другой закон сохранения, необходимый сейчас нам, – количества движения – в локальной форме принимает вид

$$\rho \frac{dv}{dt} = \nabla \sigma + pf, \quad (5.11)$$

где  $f$  – вектор массовых сил, а  $\sigma$  – тензор напряжений. В правой части (5.11) фигурируют силы, действующие на материальную точку, занимающую положение  $r$ . Тензор напряжений отвечает за *контактное* силовое воздействие на точку со стороны окружающего материала, и его следует воспринимать как линейный оператор, ставящий в соответствие проходящей через материальную точку плоскости с нормалью  $n$  вектор поверхностной силы  $t$ :  $t = n \cdot \sigma$ . Левая часть (5.11) соответствует скорости изменения количества движения материальной точки; равенство нулю этой величины в каждой точке сплошного тела описывает его статику. Применяя (5.9) к последнему уравнению, получим

$$\frac{\partial v}{\partial t} + v \cdot \nabla v = \frac{1}{\rho} \nabla \cdot \sigma + f, \quad (5.12)$$

Одного рассмотрения статики или динамики сплошного деформируемого тела обычно оказывается недостаточным для определения физических полей в этом теле (количество неизвестных полевых функций оказывается больше количества уравнений (5.12)). Для получения дополнительных уравнений рассматривают геометрическую или кинематическую сторону вопроса, в результате чего появляются дополнительные уравнения. Полученные динамические и кинематические соотношения остается связать замыкающими уравнениями, которые являются наиболее неформализуемыми и зависящими от особенностей рассматриваемой сплошной среды. Эти замыкающие уравнения называются *определяющими соотношениями*. Можно сказать, что определяющие соотношения – "центральные" уравнения в любой модели.

Мы будем считать, что среда представляет собой идеальный совершенный газ. Первое условие ("идеальность") означает  $\sigma = -pI$ ,  $p$  – гидростатическое давление,  $I$  – тензорная единица. Второе требует, чтобы газ подчинялся уравнению Менделеева – Клапейрона  $p = R\rho T$ , где  $T$  – абсолютная температура,  $R$  – универсальная газовая постоянная. Последнее уравнение и есть определяющее соотношение. Еще сузим круг рассматриваемых процессов: будем полагать, что последнее принимает форму функции

$$p = p(\rho) \quad p = p(\rho), \quad (5.13)$$

Такие процессы, называемые баротропными, существуют: например, для рассматриваемой среды уравнения изотермического и адиабатического процессов записываются в форме (5.13). Уравнения динамики (5.12) (при  $f \equiv 0$ ), кинематики (5.10) и состояния (6.13) в одномерном случае запишутся в виде

$$\begin{aligned} \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + \frac{l}{\rho} \frac{\partial p}{\partial x} &= 0, \\ \frac{\partial \rho}{\partial t} + v \frac{\partial \rho}{\partial x} + \rho \frac{\partial v}{\partial x} &= 0, \\ p &= p(\rho) \end{aligned} \quad (5.14)$$

(три уравнения, три неизвестные функции  $v$ ,  $p$ ,  $\rho$  независимых переменных  $x$  и  $t$ ). Система (5.14) содержит одно алгебраическое и два нелинейных дифференциальных уравнения. Эти дифференциальные уравнения с коэффициентами, зависящими от искомым функций, однако, линейны относительно старших производных (такие уравнения называются *квазилинейными*). Несмотря на обнадеживающее название, линейными уравнения (5.14) не являются и обладают богатой динамикой. Легко проверить, что тройка функций

$$v \equiv 0, \rho \equiv \rho_0, p \equiv p_0 = p(\rho_0) \quad (5.15)$$

является решением системы (5.14). Данное решение соответствует состоянию покоя газа, причем значение давления (либо плотности) необходимо определить из начальных условий. Получим дифференциальные уравнения эволюции малых возмущений состояния покоя системы (5.14). Путь, которым мы пойдем, является достаточно общим для линеаризации нелинейных дифференциальных уравнений вблизи какого-либо их решения.

Определим возмущенное решение (5.15), оставляя для возмущений такие же обозначения, как и у возмущенных полей,

$$v = v, \rho = \rho_0 + \rho, p = p_0 + p \quad (5.16)$$

Возмущения  $v$ ,  $p$ ,  $\rho$  будем считать малыми. Из (5.14)-(5.16) вытекает

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{l}{\rho_0} \frac{\partial p}{\partial \rho} \frac{\partial \rho}{\partial x} &= 0, \\ \frac{\partial \rho}{\partial t} + \rho_0 \frac{\partial v}{\partial x} &= 0, \end{aligned} \tag{5.17}$$

с точностью до членов второго порядка. Функцию (5.13) разложим в ряд вблизи  $\rho_0$ :

$p(\rho_0 + \rho) = p(\rho_0) + p'(\rho_0)\rho + O(\rho^2)$ . Эксперимент показывает, что  $p'(\rho_0) > 0$ , поэтому обозначим  $p'(\rho_0) = a^2$ . Почленным дифференцированием уравнений и удалением подобных систему можно привести к виду двух уравнений вида (6.7) для функций  $v$  и  $\rho$  тех же независимых переменных  $(x, t)$ :

$$\begin{aligned} \frac{\partial^2 v}{\partial t^2} &= a^2 \frac{\partial^2 v}{\partial x^2}, \\ \frac{\partial^2 \rho}{\partial t^2} &= a^2 \frac{\partial^2 \rho}{\partial x^2}. \end{aligned}$$

Окончательно наша модель приняла вид системы дифференциальных уравнений в частных производных с *постоянными* коэффициентами, то есть "чисто линейных" уравнений. Это – обычное дело, если возмущается однородное и стационарное решение. Информация о возмущаемом решении содержится в коэффициенте  $a$ .

## 5.4 Решение волнового уравнения методом Фурье

Я готов биться об заклад, что принцип суперпозиции будет стоять в веках.

Р.Фейнман

Приступим к решению уравнения (5.7)

$$\frac{\partial^2 v}{\partial t^2} - a^2 \frac{\partial^2 v}{\partial x^2} = 0,$$

удовлетворяющего граничным условиям

$$v(0,t) = v(l,t) = 0 \quad (5.18)$$

и начальным условиям

$$v(x,0) = \phi(x), \quad \frac{\partial v}{\partial t}(x,0) = \psi(x) \quad (5.19)$$

Эта задача соответствует динамическому поведению газа в тонком канале длиной  $l$  (или колебаниям струны с закрепленными концами). Задавая функции (6.19) начального возмущения состояния газа, мы не должны забывать, что (6.7) справедливо для малых возмущений. Найдем нетривиальное классическое решение краевой задачи (6.7), (6.18), (6.19). Доказано, что это решение существует и единственно. Будем, следуя Эйлеру, искать его в виде, обычном для линейных уравнений математической физики с постоянными коэффициентами:

$$v \sim \exp(kx + ct), \quad (5.20)$$

где  $k$  и  $c$  могут быть комплексными. Решения типа (6.20) называют нормальными модами или нормальными колебаниями. Характеристическое (дисперсионное) уравнение принимает вид

$$c^2 - a^2 k^2 = 0,$$

откуда

$$c = \sqrt{ak} \quad (5.21)$$

Число решений характеристического уравнения оказалось равным порядку производной по  $t$ . То есть, вместо (5.20) мы должны записать:

$$v = \exp(ct) [A \exp(cx/a) + B \exp(-cx/a)].$$

Применяя к последнему равенству граничные условия, получаем, что при действительном  $c$  имеется только тривиальное решение  $v \equiv 0$ . При  $c = i\gamma$  условие существования нетривиальной пары  $A, B$  примет вид  $\sin(\gamma l/a)$ , откуда

$\gamma = \pi n a / l$ ,  $n = 0, \pm 1, \pm 2, \dots$ . С учетом решения характеристического уравнения функция (5.20) приобретает вид

$$v_n = [F_n \sin(\pi n a t / l) + G_n \cos(\pi n a t / l)] \sin(\pi n x / l) \quad (5.22)$$

Эти функции суть частные решения нашей краевой задачи. В силу принципа суперпозиции функция

$$v = \sum_{n=0}^{\infty} [F_n \sin(\pi n a t / l) + G_n \cos(\pi n a t / l)] \sin(\pi n x / l) \quad (5.23)$$

также будет являться ее решением. Подчиним коэффициенты линейной комбинации (6.23) начальным условиям (5.19):

$$\phi(x) = \sum_{n=0}^{\infty} G_n \sin(\pi n x / l),$$

$$\psi(x) = \sum_{n=0}^{\infty} F_n \frac{\pi n a}{l} \sin(\pi n x / l).$$

Но это есть разложения начальных функций  $\phi(x)$  и  $\psi(x)$  в ряд Фурье. Коэффициенты этого ряда  $G_n$  и  $H_n = F_n \frac{\pi n a}{l}$  находятся как

$$G_n = \frac{2}{l} \int_0^l \phi(\xi) \sin(\pi n \xi / l) d\xi, H_n = \frac{2}{l} \int_0^l \psi(\xi) \sin(\pi n \xi / l) d\xi$$

Как известно, коэффициенты ряда Фурье убывают, и тем быстрее, чем более гладкой является функция. Окончательно решение задачи примет вид

$$v = \sum_{n=0}^{\infty} \left[ \frac{2}{\pi n} \int_0^l \psi(\xi) \sin(\pi n \xi / l) d\xi \sin(\pi n a t / l) + \frac{2}{l} \int_0^l \phi(\xi) \sin(\pi n \xi / l) d\xi \cos(\pi n a t / l) \right] \sin(\pi n x / l)$$

Данный метод, называемый методом Фурье, работает для таких линейных дифференциальных уравнений с постоянными коэффициентами, когда существует дисперсионное соотношение (т.е. для фиксированного  $k$  ограничено  $c$ ), а система частных решений образует полный базис, то есть способна представить начальное условие задачи. Пара  $k, c(k)$  представляет одну базисную функцию (моду), а суммирование по всем модам (по  $k$ ) определяет решение дифференциального уравнения.

Обратим внимание на то, что частные решения (5.22) краевой задачи представимы в виде

$$v = T(t) X(x), \quad (5.24)$$

то есть начальный профиль  $T(0)X(0)$  со временем только растягивается или сжимается, оставаясь подобным самому себе (автомодельным). Подставляя представление (5.24) в (5.7), получим

$$T''X - a^2TX'' = 0,$$

откуда после деления на  $a^2TX$

$$\frac{l}{a^2} \frac{T''}{T} = \frac{X''}{X} = \lambda \quad (5.25)$$

где  $\lambda$ , очевидно, не зависит ни от  $t$ , ни от  $x$ . Одно из уравнений системы (5.25)

$$X'' - \lambda X = 0 \quad (5.26)$$

вместе с условиями

$$X(0) = X(l) = 0, \quad (5.27)$$

вытекающими из (5.18), представляет собой частный случай известной в математической физике задачи Штурма-Лиувилля – задачи нахождения нетривиального (т.е. ненулевого) решения обыкновенного дифференциального уравнения (5.26), удовлетворяющего краевым условиям (5.27). Те значения  $\lambda$ , при которых эта задача имеет нетривиальные решения, называются собственными числами (спектром) задачи, а соответствующие им решения – собственными функциями задачи. Читатель, вероятно, заметил, что задача Штурма-Лиувилля напоминает задачу нахождения собственных векторов  $t$  симметричного тензора второго ранга  $X$  над конечномерным пространством:

$$(X - \lambda I)t = 0, \quad t \neq 0.$$

Формальная замена линейного *симметричного* оператора  $X$  другим линейным *самосопряженным* оператором – координатной частью линейного дифференциального оператора с краевыми условиями, и приводит к задаче Штурма-Лиувилля (напомним, что функции также являются векторами, т.е. элементами линейного пространства). Известно, что эта задача приводит к счетному множеству собственных чисел, все они вещественные, среди них имеется наименьший, а точка сгущения расположена в бесконечности, т.е.  $\lim_{n \rightarrow \infty} \lambda_n = +\infty$ . Каждому собственному значению соответствует единственная нормированная собственная функ-

ция, а полный набор собственных функций образует ортонормированный базис пространства непрерывных функций на отрезке (точнее говоря, речь идет об ортонормированном базисе в гильбертовом пространстве, всюду плотном в  $C[a,b]$ ). Решения задачи (5.26)-(5.27) имеют вид

$$X_n(x) \sim \sin(\sqrt{\lambda_n}x), \quad \lambda_n = \left(\frac{\pi n}{l}\right)^2, \quad n = 0, 1, 2, \dots \quad (5.28)$$

Собственные функции  $\sin(\sqrt{\lambda_n}x)$ ,  $n = 0, 1, 2, \dots$  еще называются модами или гармониками. Решая второе уравнение системы (6.25) с учетом краевых условий, получим дискретный спектр и представление:

$$T_n(t) = F_n \sin(\sqrt{\lambda_n}at) + G_n \cos(\sqrt{\lambda_n}at) \quad (5.29)$$

Обратим внимание на то, что элемент спектра определяет не только форму гармоники (5.28), но и частоту колебаний  $\sqrt{\lambda_n}a$ , называемую собственной частотой этой гармоники. Из (5.24) следует, что если мы выберем начальное возмущение (5.19) соответствующим (5.28), то в системе будут происходить колебания газа с  $n$ -ой собственной частотой.

Рассмотренная нами задача также описывает поперечные колебания линейно-упругой струны, шарнирно закрепленной по концам. Заметим, что амплитуда колебаний любой гармоники в рассмотренной задаче не зависит от частоты ее колебаний. Этот факт является общим для линейных волновых и колебательных моделей (и упоминался нами в начале главы в связи с открытием Гука).

Задача на нахождение собственных значений и собственных функций обобщается на граничные задачи, поставленные для трехмерных тел. Удивительно, что тело, некоторые свойства которого описываются линейным дифференциальным уравнением в частных производных, имеет набор собственных частот и функций ("паспортных данных"), зависящих только от оператора, формы тела и наложенных на его границах условий. Важно, что собственные функции задачи образуют базис некоторого функционального векторного пространства, по которому можно раз-

ложить решение (вопрос – всегда ли? – оставим открытым). Например, для решения задачи со сложным (но линейным) оператором и некоторыми граничными условиями иногда разумнее решить сначала задачу с более простым оператором и теми же условиями, а полученное разложение далее использовать для основной задачи. Отметим, что если мы решаем линейные уравнения для бесконечных областей, спектр оказывается непрерывным и суперпозиция нормальных мод в представлении начальных условий сводится к интегральной сумме по  $k$ .

## 5.5 О характеристиках уравнений математической физики. Решение волнового уравнения методом Даламбера

Долог путь поучений, короток  
и успешен путь примеров

Сенека Младший

Здесь нам понадобится понятие характеристик линейного дифференциального оператора. Начнем с простого примера, использующего понятия эйлеровых и лагранжевых координат. Пусть каждая частица  $\xi$  "одномерной" сплошной среды имеет некоторый локальный параметр  $\vartheta$ , не изменяющийся со временем:

$$\frac{d\vartheta}{dt} = 0 \quad (5.30)$$

Вспоминая формулу (5.9),

$$\frac{\partial \vartheta}{\partial t} + v(x, t) \frac{\partial \vartheta}{\partial x} = 0 \quad (5.31)$$

Поле скоростей  $v(x, t)$  будем считать известным, а  $\vartheta(x, t)$  – разыскиваемым. Дифференциальное уравнение (5.31), линейное относительно старших производных с коэффициентами, зависящими только от координаты и времени  $x, t$ , принадлежит к классу почти линейных уравнений в частных производных I порядка (по терминологии). Как любое дифференциальное уравнение, оно

имеет бесконечное множество решений (в некотором функциональном пространстве) и чтобы выделить из них определенное, следует дополнить это уравнение начальным условием

$$\vartheta(x, 0) = \vartheta_0(x). \quad (5.32)$$

Такая задача называется начальной задачей или задачей Коши.

Однако решать задачу (5.31)-(5.32) легче всего, принимая во внимание ее происхождение (физическое содержание). Из (5.30) вытекает

$$\vartheta = \vartheta(\xi),$$

откуда с учетом (5.32)

$$\vartheta = \vartheta_0(\xi). \quad (5.33)$$

Мы получили (5.33) – решение с точки зрения Лагранжа. Для нахождения поля  $\vartheta(x, t)$  нам придется решить еще одно дифференциальное уравнение

$$\frac{dx}{dt} = v(x, t) \quad (5.34)$$

с начальным условием

$$x(0) = \xi. \quad (5.35)$$

Решение этой задачи Коши  $x = x(\xi, t)$  – закон движения частицы с именем  $\xi$  – в силу свойств континуума мы можем обратить  $\xi = \xi(x, t)$  и окончательно наше решение с точки зрения Эйлера примет вид

$$\vartheta = \vartheta_0(\xi(x, t)).$$

Интегральные линии уравнения (5.34) на плоскости  $x, t$ , то есть траектории движения частиц при заданном поле скоростей  $v(x, t)$ , называются характеристиками уравнения с частными производными (5.31); это не что иное, как координатные линии  $\xi = const$ . Очевидно, поле характеристик плотно заполняет плоскость  $x, t$ . Результат (5.33) можно понимать так: решения (5.31) должны быть постоянными вдоль каждой из его характеристик.

Таким образом, вдоль каждой своей характеристики уравнение в частных производных (5.31) превращается в обыкновенное дифференциальное уравнение.

Посмотрим, сможет ли это дать что-либо для понимания системы (5.17), которую перепишем в виде

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{a^2}{\rho_0} \frac{\partial \rho}{\partial x} &= 0, \\ \frac{\partial \rho}{\partial t} + \rho_0 \frac{\partial v}{\partial x} &= 0. \end{aligned} \tag{5.36}$$

Пусть нам известно, что эта система имеет гладкое решение в некоторой области  $G$  плоскости  $x, t$ . Выберем в этой области точку  $x_0, t_0$  и проведем через эту точку гладкую кривую  $\gamma$ . Предположим, что нам почему-либо известны значения полей  $v, \rho$  вдоль этой кривой. Поставим задачу нахождения значений этих полей вблизи кривой. Конечно, эта задача есть задача Коши. Принятые требования означают, что нам известны производные полей  $v, \rho$  вдоль кривой  $\gamma$  в точке  $x_0, t_0$ :

$$\begin{aligned} \frac{dt}{d\gamma} \frac{\partial v}{\partial t} + \frac{dx}{d\gamma} \frac{\partial v}{\partial x} &= \frac{dv}{d\gamma}, \\ \frac{dt}{d\gamma} \frac{\partial \rho}{\partial t} + \frac{dx}{d\gamma} \frac{\partial \rho}{\partial x} &= \frac{d\rho}{d\gamma}. \end{aligned} \tag{5.37}$$

Условие разрешимости системы (5.36)-(5.37) относительно  $\frac{\partial v}{\partial t}, \frac{\partial v}{\partial x}, \frac{\partial \rho}{\partial t}, \frac{\partial \rho}{\partial x}$  записывается

$$\begin{vmatrix} 1 & 0 & 0 & \frac{a^2}{\rho_0} \\ 0 & \rho_0 & 1 & 0 \\ \frac{dt}{d\gamma} & \frac{dx}{d\gamma} & 0 & 0 \\ 0 & 0 & \frac{dt}{d\gamma} & \frac{dx}{d\gamma} \end{vmatrix} = a^2 \left( \frac{dt}{d\gamma} \right)^2 - \left( \frac{dx}{d\gamma} \right)^2 \neq 0. \tag{5.38}$$

Это условие в силу принятой нами гипотезы существования решения вблизи кривой заведомо выполняется. Таким образом, из (5.36)-(5.37) мы находим первые производные, затем точно

таким же способом последовательно можно найти и смешанные производные высших порядков, что и позволяет построить искомое решение. Но если  $\gamma$  имеет уравнение  $x \pm at = const$ , то наше требование о существовании решения может не иметь смысла. Из линейной алгебры известно, что в этом случае решение может иметь место, когда ранг матрицы в (5.38) совпадает с рангом расширенной матрицы

$$\begin{bmatrix} 1 & 0 & 0 & a^2/\rho_0 & 0 \\ 0 & 0 & 1 & \rho_0 & 0 \\ \frac{dt}{d\gamma} & \frac{dx}{d\gamma} & 0 & 0 & \frac{dv}{d\gamma} \\ 0 & 0 & \frac{dt}{d\gamma} & \frac{dx}{d\gamma} & \frac{d\rho}{d\gamma} \end{bmatrix}$$

Например, должен равняться нулю определитель матрицы:

$$\begin{vmatrix} 0 & 0 & 0 & a^2/\rho_0 \\ 0 & \rho_0 & 1 & 0 \\ \frac{dv}{d\gamma} & \frac{dx}{d\gamma} & 0 & 0 \\ \frac{d\rho}{d\gamma} & 0 & \frac{dt}{d\gamma} & \frac{dx}{d\gamma} \end{vmatrix} = \frac{a^2}{\rho_0} \left( -\rho_0 \frac{dt}{d\gamma} \frac{dv}{d\gamma} - \frac{d\rho}{d\gamma} \frac{dx}{d\gamma} \right) = 0.$$

Вдоль характеристик  $x \pm at = const$  это требование равносильно выполнению соотношений  $d\left(v\mu \frac{a}{\rho_0} \rho\right) = 0$ . В этом примере

мы пришли к понятию характеристик как к кривым, на которых могут не иметь смысла начальные условия задачи Коши. Характеристиками оказались два семейства прямых  $x \pm at = const$ , образующих координатную сетку в плоскости  $x, t$ . Эта система координат обладает особыми свойствами: на каждой ее координатной кривой сохраняется постоянной величина соответствующая

$v\mu \frac{a}{\rho_0} \rho$ . Запишем систему (5.36) в новой системе координат

$$y = x + at, z = x - at :$$

$$\frac{\partial v}{\partial y} + \frac{a}{\rho_0} \frac{\partial \rho}{\partial y} = 0, -\frac{\partial v}{\partial z} + \frac{a}{\rho_0} \frac{\partial \rho}{\partial z} = 0$$

или после преобразований:

$$\frac{\partial^2 v}{\partial y \partial z} = 0, \frac{\partial^2 \rho}{\partial y \partial z} = 0.$$

Любое из уравнений системы легко решается:  $(v_y)_z$  означает, что функция  $v_y$  постоянна вдоль прямых  $y = \text{const}$  на плоскости  $y, z$ . Итак,  $v_y = \theta(z)$ , а значит

$$v = \int \theta(z) dy = f(z) + g(y) = f(x + at) + g(x - at) \quad (5.39)$$

Мы получили общее решение системы (5.7) с точностью до двух произвольных функций, аргументами которых служат пространственно-временные комплексы  $x \pm at$ . Добавляя к системе начальные условия, для удобства принятые в следующем виде:

$$v(x, 0) = \varphi(x), \frac{\partial v}{\partial t}(x, 0) \equiv 0,$$

а также полагая среду бесконечной  $-\infty < x < +\infty$ , легко найти  $f$  и  $g$  и решение задачи

$$v(x, t) = \frac{1}{2} (\varphi(x - at) + \varphi(x + at)) \quad (5.40)$$

Данное решение называется решением типа бегущей волны. Точнее, в решении (5.40) присутствуют две волны одинакового профиля, разбегающиеся в разные стороны от "начальной волны"  $\varphi(x)$ . Существование такого решения является следствием трансляционной инвариантности задачи по временной и пространственной координатам. Это означает существование двух сдвигов  $\Delta x$ ,  $a\Delta t$ , переводящих решение само в себя при фиксированном  $a$ . Мы снова столкнулись с автомодельным решением: профиль каждой из пары волн (5.39) не меняется, меняется только их расположение в пространстве.

Обратим внимание на то, что величина  $a = \sqrt{p'(\rho_0)}$  есть скорость распространения малых возмущений в газе, что видно из (6.40). Учитывая, что константа  $a$  формально присутствует в волновом уравнении, описывающем различные физические волно-

вые процессы, ее нужно трактовать как скорость распространения волн. Здесь нельзя не вспомнить об одном глубоком выводе, сделанном Джеймсом Клерком Максвеллом из открытых им уравнений электромагнитной динамики.

## 5.6 Уравнения Максвелла

Не бог ли эти знаки начертал?  
Таинственен их скрытый дар!

Гете

В отличие от механического, электромагнитное поле не требует для себя какого-либо носителя типа твердого, жидкого или газообразного тела, оно само является материальным полем и может распространять возмущения в вакууме и с конечной скоростью. Эти идеи в середине прошлого века в свете господствующей амперовской электродинамики еще казались абсурдными. Практические успехи "старой" электродинамики впечатляли: к тому времени уже действовал телеграфный кабель через Атлантический океан. Однако она вследствие своей внутренней противоречивости напоминала Максвеллу вавилонскую башню. Он не мог принять мгновенного действия на расстоянии (дальнодействия), заложенного, например, в законе Кулона. Возможно, его воодушевляло то, что в 1849 Арман Ипполит Физо нашел чудовищное (но конечное) значение для скорости света – 313 с лишним тысяч километров в секунду. Максвелл, еще со студенчества воодушевленный работами Майкла Фарадея, впервые введшего абстрактные силовые линии в пространстве, придерживался точки зрения, что главным действующим лицом должно являться электромагнитное поле в пространстве. Эту смелую философию он подкреплял аргументами: *"Кому свойства воздуха незнакомы, тому передача силы посредством этой невидимой среды будет казаться столь же непонятной, как и всякий другой пример действия на расстоянии..."*. Отказавшись от некоторых законов, сохранив и обобщив другие, наделив каждую математическую операцию физическим смыслом и в то же время дав каждой физической величине четкую математическую характеристику, Мак-

свелл пришел к своим знаменитым уравнениям электромагнитного поля:

$$\begin{aligned}
 \nabla \cdot D &= \rho, \\
 \nabla \times E &= -\dot{B}, \\
 \nabla \cdot B &= 0, \\
 \nabla \times H &= j + \dot{D},
 \end{aligned}
 \tag{5.41}$$

$E$  – вектор напряженности электрического поля,  $D$  – вектор электрической индукции,  $B$  – вектор магнитной индукции,  $H$  – вектор напряженности магнитного поля,  $\rho$  – плотность заряда,  $j$  – ток смещения.

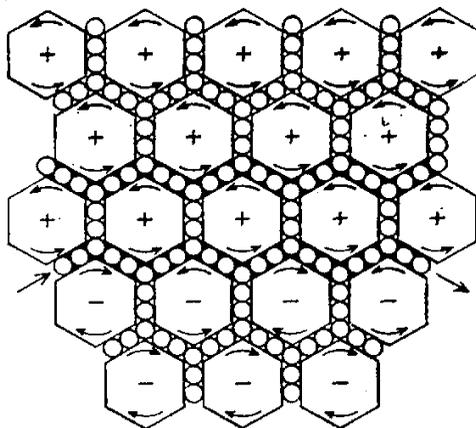


Рисунок 5.1 – Механическая модель электромагнитного поля

Для уяснения смысла некоторых понятий, Максвелл пользовался механическими моделями типа изображенной на рисунке 5.1. В "Трактате об электричестве и магнетизме", главном труде Максвелла, такие схемы он уже не использовал (с точки зрения поля они бессмысленны), но именно они позволили ему многое понять. Нужно заметить, что в век Максвелла ("век пара") доминировали идеи и методы механики, так что уже после выхода "Трактата" многие пытались выяснить механическое содержание его уравнений. Людвиг Больцман думал, что Максвелл имеет механические модели, которыми он обосновывал свои уравнения,

но прячет. Уже после смерти Максвелла Больцман, который им восхищался, излагая на лекциях максвелловскую теорию, предвзял изложение эпиграфом из "Фауста": *"Я должен пот тяжелый лить, чтоб научить тому, чего не понимаю сам"*. Известно, что Максвелл пользовался представлениями Гельмгольца о вихревых движениях идеальной жидкости, но заслуживает изумления смелая аналогия "неосязаемого" континуума с "осязаемым". До сих пор термин "механическое" механики часто заменяют на "материальное", хотя электромагнитные поля материальны ничуть не меньше.

Система уравнений (5.41) – незамкнута, в ней не хватает определяющих соотношений. В простейшем (и имеющем смысл) случае мы можем записать

$$\begin{aligned} D &= E, \\ B &= H, \\ j &\equiv 0. \end{aligned} \tag{5.42}$$

Применяя к (5.41)<sub>4</sub> операцию  $\nabla$  и используя (5.42) и (5.41)<sub>2</sub>, получаем

$$\nabla \times \nabla \times H = -\ddot{H}.$$

Далее, пользуясь известным тождеством  $\nabla \times \nabla \times a = \Delta a - \nabla(\nabla \cdot a)$  и (5.41)<sub>3</sub> вместе с (5.42), окончательно получаем

$$\Delta H = \ddot{H} \tag{5.43}$$

( $\Delta \equiv \nabla \cdot \nabla$  - оператор Лапласа). Точно такое же уравнение получается и для  $E$ . Но (5.43) – это волновое уравнение в трехмерном случае. Это привело Максвелла к предсказанию существования электромагнитных волн. Если попытаться вычислить скорость распространения этих волн, то получится, что она равна отношению электростатической и электромагнитной единиц заряда. Но это величина известная! Ее независимо измерили Кольрауш, Вебер и Столетов и равна она примерно 300000 километров в секунду, то есть скорости света (ввиду такой высокой скорости распространения электромагнитные воздействия и принимали за мгновенные). Максвелл смог придать этому случайному совпадению, "любопытному курьезу" по его словам, глубокий

физический смысл: свет есть не что иное, как электромагнитные волны. Последующая всесторонняя экспериментальная проверка уравнений Максвелла блестяще их подтвердила. Генрих Герц писал об этих уравнениях: *"Трудно избавиться от чувства, что эти математические формулы живут независимой жизнью и обладают своим собственным интеллектом, что они мудрее, чем мы сами, мудрее даже, чем их первооткрыватели, и что мы извлекаем из них больше, чем было заложено в них первоначально"* [49]. Молодой голландский физик Гендрик Антуан Лоренц впоследствии пытался применить теорию Максвелла к движущимся телам, и из этого труда появились "преобразования Лоренца" – важнейшая предпосылка создания теории относительности. Трактат Максвелла об электричестве и магнетизме открыл эпоху новой физики.

## 5.7 О классификации квазилинейных систем

Нет ничего практичнее хорошей теории.

Л.Больцман

Понятие характеристик позволяет классифицировать уравнения математической физики. Рассмотрим систему уравнений в частных производных первого порядка, символически представимую в виде

$$\frac{\partial u}{\partial t} + C \frac{\partial u}{\partial x} = f \quad (5.44)$$

Здесь  $u$  и  $f$  есть  $n$ -мерные векторы неизвестных и известных функций соответственно,  $C$  – матрица коэффициентов, зависящих от  $x, t$ . Эта система – почти линейная. Систему (5.44) сводят к одному уравнению  $n$ -го порядка и  $n-1$  связям (обратный переход мы уже делали с системой (5.17)). Поставим уже известную нам задачу отыскания условий разрешимости задачи Коши вблизи кривой  $g$  в плоскости  $x, t$ . Для нахождения первых производных вблизи этой кривой систему (5.44) следует дополнить уравнениями

$$\frac{dt}{d\gamma} \frac{\partial u}{\partial t} + \frac{dx}{d\gamma} \frac{\partial u}{\partial x} = \frac{du}{d\gamma} \quad (5.45)$$

Разрешимость этой задачи связана с определителем

$$\left\| \begin{array}{cc} I & C \\ \frac{dt}{d\gamma} I & \frac{dx}{d\gamma} I \end{array} \right\| = \left( \frac{dt}{d\gamma} \right)^2 \left\| \begin{array}{cc} I & C \\ I & \frac{dx}{dt} I \end{array} \right\| = (-1)^n \left( \frac{dt}{d\gamma} \right)^2 \left\| C - \frac{dx}{dt} I \right\|.$$

Поэтому характеристиками являются кривые, удовлетворяющие уравнениям

$$\frac{dx}{dt} = k_i(x, t),$$

в которых поля коэффициентов локального наклона характеристик  $k_i(x, t)$ , очевидно, суть собственные числа матрицы  $C$ . Спектр произвольной матрицы с действительными коэффициентами в общем случае есть подмножество комплексных чисел. Если спектр действителен и прост в любой точке рассматриваемой области, то система (5.44) называется гиперболической. Если спектр комплексный (конечно, если  $n$  четно), то система – эллиптическая. В других случаях строения спектра система называется параболической. Таким образом, только гиперболические системы имеют инвариантную систему координат, состоящую из характеристик. Слово "инвариантный" означает связанность с самим уравнением, а не с его конкретной записью в той или иной системе координат. Для таких систем имеет смысл задача Коши в открытой области, согласованной с расположением поля характеристик, и не имеют смысла постановки в ограниченных областях. Гиперболические системы описывают распространение возмущений с конечной скоростью. Примеры таких "чисто линейных" систем нам уже встречались ранее: (5.17), (5.43). Для эллиптических систем характеристики отсутствуют и имеют смысл только задачи на замкнутой области (пусть и бесконечной), а условия на границе называются краевыми. Возмущения граничных условий распространяются этими системами мгновенно ("диффузионным механизмом"). Пример эллиптической системы – упоминавшееся уже уравнение Лапласа  $\Delta v = 0$ .

Для этого уравнения второго порядка различают задачу Дирихле (на границе задано значение функции), задачу Неймана (на границе задана производная по нормали к границе) и смешанную задачу. Параболические системы, имеющие поле характеристик,

не образующее систему координат, допускают постановки в открытой области, согласованной с расположением характеристик. То есть, краевые ("начально-граничные") задачи для параболических систем могут иметь смысл, а могут и не иметь его. Например, уравнение диффузии (теплопроводности)  $\Delta v = \dot{v}$  с условиями, заданными в области  $t \rightarrow 0$  описывает распределение субстанции  $v$  по пространству: начальные особенности со временем сглаживаются. Но если поставить задачу на области  $t \rightarrow 0$ , из первоначально гладкого профиля в течение обратного хода времени неизбежно будут возникать особенности, то есть решения такой задачи будут неустойчивыми. Задачи, в которых нужно установить "прошлое" по настоящему, относятся к некорректным по Адамару задачам. Корректность по Адамару означает существование, единственность и устойчивость решения (в данном случае – по отношению к возмущению начальных условий). Причиной таких свойств параболического уравнения диффузии кроется в его инвариантности по отношению к замене  $t \rightarrow -t$ .

Классификация систем с использованием характеристик обобщается на квазилинейные системы (т.е. системы дифференциальных уравнений, линейные относительно старших производных, но содержащие в коэффициентах искомые функции и их производные низших порядков) и на большее двух число независимых переменных (тогда характеристики являются уже поверхностями). Исследуем, например, систему (5.14), являющуюся квазилинейной, считая начальные "возмущения" конечными. Введением функции давления  $P$ :  $dP = dp / a\rho_0$  эта система сводится к следующей:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + a \frac{\partial P}{\partial x} = 0, \quad a \frac{\partial v}{\partial x} + \frac{\partial P}{\partial t} + v \frac{\partial P}{\partial x} = 0.$$

По образцу, данному выше для квазилинейных систем, (добавляя к независимым переменным  $v$ ) или элементарными сложением и вычитанием уравнений получается

$$\begin{aligned} \frac{\partial}{\partial t}(P+v) + (v+a) \frac{\partial}{\partial x}(P+v) &= 0, \quad \frac{\partial}{\partial t}(P-v) + \\ + (v-a) \frac{\partial}{\partial x}(P-v) &= 0. \end{aligned}$$

Очевидно, что характеристиками системы будут кривые  $\frac{dx}{dt} = v \pm a$ , на которых постоянны соответствующие величины  $P \pm n$ . Здесь в отличие от системы, описывающей малые возмущения,  $a$  представляет собой поле местной скорости звука, а  $\pm a$  – относительные скорости распространения волн по газу. Вдоль первого семейства характеристик распространяется волна разрежения, а вдоль второго – волна сжатия.

## 5.8 Связь непрерывного и дискретного на примерах уравнения колебаний струны и уравнения Шредингера

Лишь первый исчез, как сейчас же в ином положении Новый родится за ним, а нам кажется, – двинулся первый.

Лукреций

Вопросы о движении, переходе постепенных количественных изменений в качественные, появление у целого свойств, которыми не обладает ни одна из его частей, являются одними из ключевых вопросов современного фундаментального естествознания. Эти вопросы имеют глубокие философские корни. Развитие естественных наук заставляет еще и еще раз возвращаться к ним. Ученых XIX века поразило наличие волновых свойств у света, который они представляли как поток дискретных частиц. К глубокому пересмотру фундаментальных понятий привело в XX веке создание квантовой механики. Оказалось, что дискретные и непрерывные свойства материи нельзя противопоставлять друг другу, что они неразрывно связаны между собой. Выяснилось и другое важное обстоятельство. Анализ многих явлений требует сочетания дискретного и непрерывного подходов. И вопрос о соотношении тех и других свойств при построении теории оказывается далеко не простым. От его успешного решения часто зависит, насколько глубоко нам удастся разобраться в изучаемом объекте.

Рассмотрим задачу, при решении которой был развит ряд ключевых идей. Проведем анализ поведения упругой струны, по которой ударили в начальный момент времени. Рассмотрим сначала систему, состоящую из точечного груза массой  $m$ , к которому прикреплены две одинаковые упругие горизонтальные нити длиной  $l_0/2$ , натянутые силой  $F_0$  (сила тяжести отсутствует). При отклонении груза от положения равновесия появляется возвращающая сила, пропорциональная отклонению

$$F = -2F_0 \sin \alpha \approx -4F_0 u / l_0,$$

тогда из второго закона Ньютона получаем:

$$\frac{d^2 u}{dt^2} + \omega^2 u = 0, \quad \omega^2 = 4F_0 / (ml_0). \quad (5.46)$$

Уравнение (5.46) описывает колебания с круговой частотой  $\omega$

$$u = A \sin \omega t + B \cos \omega t,$$

где константы  $A$  и  $B$  определяются начальными положением и скоростью груза. Мы решили задачу о колебании струны, вся масса которой сосредоточена в центре. Для однородной струны с массой  $m$  и длиной  $l_0$  это слишком грубое приближение; разумнее заменить ее набором из  $N$  шариков с массой  $\mu = m/N$ , расположенных на расстоянии  $h=l_0/N$  и соединенных нитями, натянутыми силой  $F_0$  (так рассуждали, в частности, при выводе уравнений колебаний струны Иоганн и Даниил Бернулли). Если  $u_k$  – отклонение  $k$ -го шарика от положения равновесия, то при условии, что разница в отклонениях соседних шариков мала, можно получить

$$\begin{aligned} \mu a_k &= \frac{mh}{l_0} \frac{\partial^2 u_k}{\partial t^2} = F \approx F_0 \frac{u_{k+1} - 2u_k + u_{k-1}}{h} \Rightarrow \frac{\partial^2 u_k}{\partial t^2} = \\ &= \frac{F_0 l_0}{m} \frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} \end{aligned}$$

В пределе при  $h \rightarrow 0$  получаем уже известное нам волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0, \quad c^2 = F_0 l_0 / m.$$

Этот вывод уравнения впервые был сделан Даламбером, который не только записал это уравнение, но и нашел его общее решение (5.39) в виде суперпозиции двух волн

$$u = f(x + ct) + g(x - ct).$$

Однако использовать данное решение для струны конечных размеров непросто. Действительно, после удара по струне вправо и влево идут волны. Они доходят до концов струны, отражаются, идут в обратную сторону, и устанавливается некий режим, описывать который с помощью полученной формулы неудобно. Возможен другой путь, предложенный Фурье. Так как струна совершает колебательные движения, решение задачи ищут в виде  $u = (A \sin \omega t + B \cos \omega t) z(x)$ , что дает в итоге

$$u = \sum_{n=0}^{\infty} (A_n \sin \omega_n t + B_n \cos \omega_n t) \sin \frac{\pi n x}{l_0}.$$

Решение нашлось в виде суперпозиции стоячих волн, для которых на длине струны укладывается целое число  $n$  полуволн. Собственные значения  $l_n$  определяют, с какой частотой может колебаться струна в конфигурации  $n$ -й стоячей волны, форму которой описывает собственная функция  $z_n = \sin(\pi n x / l_0)$ .

Полученные решения по виду сильно отличаются друг от друга. Равноправность их не очевидна. Этот вопрос стал причиной дискуссии в середине XVIII века (спор о струне) между Эйлером, Даламбером и Лагранжем. Дискуссия позволила убедиться не только в эквивалентности двух решений (ведь исходная задача имеет решение и оно единственно!), но и лучше разобраться в уравнениях. Решение, полученное Фурье, дает возможность выяснить соотношение непрерывного и дискретного в этой задаче. Если мы имеем автомодельное решение волнового уравнения в виде стоячей волны

$$u = c_n(t) \sin \frac{\pi n x}{l_0},$$

то окажется, что для разных  $c_n(t)$  получаются уравнения

$$\frac{d^2 C_n(t)}{dt^2} + \omega_n^2 C_n(t) = 0, \quad n = 0, 1, 2, K.$$

Но это опять уравнения колебаний. Значит, струна оказывается эквивалентной бесконечному множеству независимых колеблющихся грузов. Интересно и другое: в непрерывной задаче, описывающей колебания струны, есть дискретный набор собственных частот. Непрерывное и дискретное вновь оказываются тесно связанными.

Глубокая связь дискретного и непрерывного отмечена и в физике микромира, где в одних случаях материю удобно рассматривать как электромагнитную волну, а в других – как поток частиц (квантов). Микрочастицы в некоторых опытах ведут себя как волны, например, испытывая дифракцию и интерференцию. В то же время оказалось, что свет (волна) квантуется, регистрируются дискретные порции света – фотоны. Оказалось, что и здесь довольно гармоничный дуализм может быть описан при помощи *линейного* уравнения математической физики – уравнения Шредингера. Выражение простейшей (плоской) волны, описывающей колебания в пространстве с частотой  $\omega$  и волновым вектором  $k$  записывается как  $\exp(ikx - i\omega t)$ .

Сопоставим этой волне частицу с энергией  $E = \eta\omega$  и импульсом  $p = \eta k$ , где  $\eta$  – постоянная Планка. Выразив  $\omega$  и  $k$  из последних соотношений и подставив их в формулу для волны, получим волновую функцию для частицы с энергией  $E$  и импульсом  $p$ :

$$\psi(x, t) = \exp\left(i\frac{p}{\eta}x - i\frac{E}{\eta}t\right).$$

Эта комплексная функция определяет плотность вероятности нахождения частицы во времени и пространстве. Нетрудно убедиться в справедливости тождеств

$$-\frac{\eta}{t} \frac{\partial \psi}{\partial t} = E\psi \quad \text{и} \quad -\eta^2 \frac{\partial^2 \psi}{\partial x^2} = p^2 \psi.$$

Но это означает, что энергия  $E$  является собственным значением оператора  $-\frac{\hbar}{i}\hat{\psi}$ , а квадрат импульса  $p^2$  – собственным значением оператора  $-\eta^{2\Delta}\psi$ , при этом  $\psi$  в обоих случаях выступает

в качестве собственной функции. Если частица массы  $m$  движется в потенциальном поле  $V(x)$ , то в силу закона сохранения энергии

$$E = \frac{p^2}{2m} + V.$$

Чтобы получить волновой аналог этого соотношения, мы должны заменить  $E$ ,  $p^2$  и  $V$  соответствующими дифференциальными операторами

$$-\frac{\hbar}{i}\psi' = -\frac{\hbar^2}{2m}\Delta\psi + V\psi,$$

откуда и следует фундаментальное уравнение волновой и квантовой механики

$$i\hbar\psi' = -\frac{\hbar^2}{2m}\Delta\psi + V\psi, \quad (5.47)$$

открытое австрийским физиком Эрвином Шредингером в 1926 году. Оно описывает движение частицы в заданном потенциальном поле. Квадрат амплитуды волновой функции

$$\int_G \psi(x,t)\psi^*(x,t) dx \equiv P(G,t)$$

определяет вероятность  $P(G,t)$  нахождения частицы в момент времени  $t$  в области  $G$ . Для оператора (5.47) мы можем поставить задачу на собственные значения

$$-\frac{\hbar^2}{2m}\Delta\psi + V\psi = E\psi, \quad \int_{-\infty}^{\infty} \psi\psi^* dx = l,$$

в результате чего получим спектр  $E_n$  значений энергии частицы, которые она может принимать, двигаясь в заданном потенциальном поле. Соответствующие им собственные функции  $\psi_n$  показывают, с какой вероятностью можно обнаружить частицу в разных точках пространства. Решив уравнение Шредингера для кулоновского потенциала  $V = -e^2/r$  ( $e$  – заряд электрона,  $r$  – расстояние до протона), можно найти энергетические уровни атома водорода, с огромной точностью согласующиеся с экспериментом.

До появления уравнения Шредингера существовала матричная квантовая механика Гейзенберга, использующая в качестве

аппарата пространство бесконечномерных векторов с ограниченным скалярным квадратом. Теория же Шредингера использует векторное пространство функций с ограниченным скалярным квадратом (вводимым через интеграл). С точки зрения функционального анализа эти пространства есть эквивалентные представления гильбертова пространства. Одна из самых замечательных идей Давида Гильберта заключается в том, чтобы рассматривать пространства функций как евклидовы. Шредингер первым увидел в квантованности состояний аналогию с проблемой собственных значений линейного дифференциального оператора. Однако рассказывают, что еще до открытия Шредингером своего уравнения, к Гильберту в Геттинген приезжали физики и задавали ему вопрос о смысле матриц в теории Гейзенберга, на который Гильберт ответил, что обычно такие таблицы появляются при решении некоторых дифференциальных уравнений. Физики решили, что Гильберт просто не понял вопроса, а через год Шредингер открыл свое знаменитое уравнение.

## 5.9 Анализ подобия и размерности

Чрезмерная глубина лишь путает и затуманивает мысль

Э.По

Ученые часто сталкиваются с задачей построения математической модели явления настолько нового, необычного и сложно, что для него еще не найден удовлетворительный механизм и не установлены балансовые и определяющие уравнения. В этих случаях главную роль играют экспериментальные методы исследования, дающие возможность установить простейшие опытные факты. Постановка же эксперимента требует предварительного теоретического анализа. Возможность такого предварительного анализа дает теория размерности и подобия. Несмотря на свою простоту и эффективность, этот анализ, конечно, требует от исследователя определенного опыта и проникновения в сущность изучаемого явления. Самый замечательный итог применения анализа размерности и подобия заключается в том, что он оказы-

вается достаточным для установления особого класса автомодельных решений математической формулировки задачи, отражающих глубокие физические закономерности изучаемого явления. Сформулируем суть метода и приведем несколько примеров его применения.

Пусть имеется какая-то зависимость, определяющая величину  $a$  в функции  $n$  параметров  $a_1, \dots, a_n$ :

$$a = f(a_1, \dots, a_n). \quad (5.48)$$

Если эта зависимость имеет физическое содержание, то соотношение (5.48) должно отражать тот бесспорный факт, что хотя числа  $a, a_1, \dots, a_n$  выражают значения соответствующих величин в определенной системе единиц измерения, физическая закономерность, которую это соотношение представляет, не зависит от произвола при выборе единиц измерения. Анализ размерностей основан на физически содержательном утверждении, сформулированным Э. Бакингамом и известным как П-теорема: пусть существует физическая закономерность, выраженная в виде зависимости некоторой размерной величины от размерных определяющих параметров. Эта зависимость может быть представлена в виде зависимости некоторой безразмерной величины от безразмерных комбинаций, определяющих параметров. Количество этих безразмерных комбинаций меньше общего числа определяющих параметров на число размерных определяющих параметров с независимыми размерностями.

Следует отметить, что П-теорема интуитивно вполне очевидна и ее неявное использование началось задолго до того, как она была явно сформулирована и формально доказана. В этой связи следует прежде всего назвать имена Галилея, Ньютона, Фурье, Максвелла, Рейнольдса, Релея. Практический выигрыш при обработке экспериментальных данных, даваемый этой теоремой, очевиден. Пусть для выяснения зависимости некоторой величины  $a$  от некоторого определяющего параметра  $a_i$  надо измерить эту величину при десяти значениях данного аргумента. Тогда для экспериментального определения величины  $a$  как функции  $n$  определяющих параметров  $a_1, \dots, a_n$  нужно произвести  $10^n$  экспериментов. Согласно П-теореме, если все величины  $a, a_1, \dots, a_n$  выра-

жаются через  $k$  независимых размерностей, дело сводится к определению функции  $n-k$  безразмерных аргументов  $\Pi_1, \dots, \Pi_{n-k}$ , для нахождения которой достаточно  $10^{n-k}$  опытов, т.е. в  $10^k$  раз меньше! Трудоемкость определения искомой функции сокращается на столько порядков, сколько среди определяющих параметров величин с независимыми размерностями.

Применение П-теоремы включает следующую последовательность действий. Разобьем величины  $a_1, \dots, a_n$  на две группы. В первую группу  $a_1, \dots, a_k$  включаются определенные величины с независимыми размерностями (например, длина, скорость, плотность и т.д.). Во вторую группу  $a, a_{k+1}, \dots, a_n$  входят величины, размерности которых можно выразить через размерности величин первой группы. То есть

$$a \square a_1^p, \dots, a_k^q, a_{k+1} \square a_1^{p_{k+1}}, \dots, a_k^{q_{k+1}}, \dots, a_n \square a_1^{p_n}, \dots, a_k^{q_n}.$$

( $p, q, p_i, q_i$  здесь – показатели степени). При этом величины

$$\Pi = \frac{a}{a_1^p, \dots, a_k^q}, \Pi_1 = \frac{a_{k+1}}{a_1^{p_{k+1}}, \dots, a_k^{q_{k+1}}}, \dots, \Pi_{n-k} = \frac{a_n}{a_1^{p_n}, \dots, a_k^{q_n}}$$

оказываются безразмерными, т.к. их значения будут одними и теми же при любом выборе системы единиц измерения. Независимость закономерности, имеющей физический смысл, от выбора единиц измерения означает, что соответствующее ей соотношение (6.48) можно представить в виде

$$\Pi = \Phi(\Pi_1, \dots, \Pi_{n-k})$$

Уменьшение числа аргументов упрощает исследование, иногда – существенно. Рассмотрим показательный пример.

В 1909-1911 годах специалисты в области физической химии Э.Бозе, Д.Рауэрт и М.Бозе опубликовали серию экспериментов. Ими изучалось время  $\tau$  заполнения сосуда данного объема  $Q$  и перепад давления  $P$  на концах трубки при стационарном протекании через трубку различных жидкостей: воды, хлороформа, бромформа, ртути. Результаты опытов были представлены в виде серии зависимостей перепада давления от времени заполнения для разных жидкостей (рисунок 5.2).

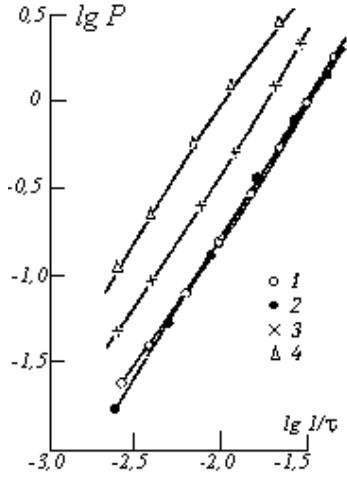


Рисунок 5.2 – Результаты опытов Бозе-Рауэрта в исходном виде для различных жидкостей (1 – вода, 2 – хлороформ, 3 – бромформ, 4 – ртуть)

Эти работы были замечены Теодором фон Карманом, который исследовал их с точки зрения анализа размерностей. Рассуждения Кармана можно представить следующим образом: перепад давления  $P$  на концах трубки должен зависеть от времени заполнения сосуда  $\tau$ , объема сосуда  $Q$  и свойств жидкости (коэффициента вязкости  $\mu$  и плотности  $\rho$ ):

$$P = f(\tau, Q, \mu, \rho)$$

(в данном случае  $n=4$ ). Размерности параметров в классе MLT (масса, длина, время) выражаются следующими соотношениями:

$$[P] = \frac{M}{LT^2}, [\tau] = T, [Q] = L^3, [\mu] = \frac{M}{LT}, [\rho] = \frac{M}{L^3}.$$

Легко увидеть, что первые три размерности определяющих параметров  $\square, Q, \square$  имеют независимые размерности, размерность же четвертого параметра  $\square$  выражается через размерности первых трех  $[\rho] = [\mu][\tau][Q]^{-2/3}$ .

Таким образом,  $k=3, n-k=1$  и анализ размерности дает

$$\Pi = \Phi(\Pi_1), \quad \Pi = \frac{P}{\mu\tau^{-1}}, \quad \Pi_1 = \frac{\rho}{\mu\tau Q^{-2/3}}.$$

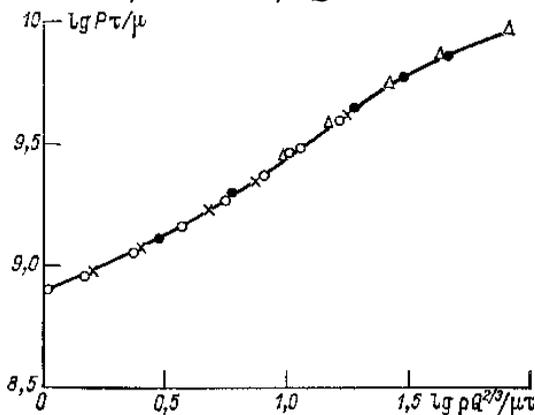


Рисунок 5.3 – Результаты опытов Бозе-Рауэрта в интерпретации Кармана

Следовательно, в координатах  $\Pi, \Pi_1$  все опытные точки должны лечь на единую кривую. Этот вывод блестяще подтвердился (рис.6.3). Ясно, что заранее проведенный анализ размерностей мог бы сократить объем экспериментальной работы во много раз.

Интересен также следующий пример. При атомном взрыве в области, настолько малой, что ее можно считать точкой, быстро (практически мгновенно) выделяется значительная энергия  $E$ . От центра взрыва распространяется мощная ударная волна, давление за которой вначале составляет сотни тысяч атмосфер. Это давление много больше, чем начальное давление воздуха, влиянием которого на первой стадии взрыва можно пренебречь. Таким образом, радиус фронта ударной волны  $r$  через промежуток времени  $\tau$  после взрыва зависит от  $E$ ,  $\tau$  и начальной плотности воздуха  $\rho$ :

$$r = f(E, \tau, \rho).$$

Таким образом,  $n=3$ , а размерности определяющих параметров в классе MLT есть:

$$\Pi = \Phi(\Pi_1), \quad \Pi = \frac{P}{\mu\tau^{-1}}, \quad \Pi_1 = \frac{\rho}{\mu\tau Q^{-2/3}}.$$

Легко видеть, что  $k$  равно трем,  $n-k=0$ , так что функция  $\Phi$  в выражении  $\Pi = \Phi(\Pi_{n-k}, \dots, \Pi_n)$  не зависит ни от одного из аргументов, т.е.  $\Phi = const$ . Далее нетрудно доказать, что

$$\Pi = r(E\tau^2 / \rho)^{-1/5}$$

откуда

$$r = const(E\tau^2 / \rho)^{1/5}.$$

Эта формула показывает, что если измерить тем или иным способом радиус ударной волны в разные моменты времени, то в логарифмических координатах  $(5/2)\lg r, \lg \tau$  экспериментальные точки должны лечь на прямую автомодельного решения, имеющую наклон, равный единице.

Это подтвердил в 1950 году английский механик и физик сэра Джеффри Ингрэм Тэйлор, обработавший кинофильм о распространении огненного шара, снятый во время американских испытаний ядерного взрыва в Нью-Мексико в 1945 году (рис.6.4). Как показывает более детальный анализ, значение  $const \approx 1$ . Зная это, по экспериментальной зависимости радиуса фронта от времени можно определить энергию взрыва. Публикация Тэйлором этой величины, оказавшейся равной примерно  $10^{17}$  Дж, вызвала в свое время смущение в американских официальных кругах.

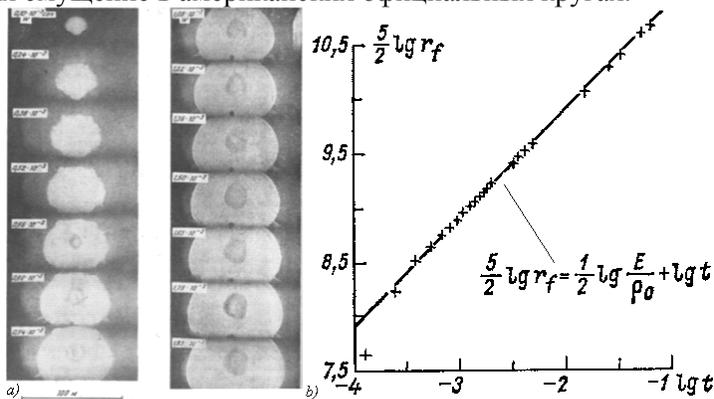


Рисунок 5.4 – Распространение ударной волны при ядерном взрыве: а) последовательные фотографии огненного шара; б) опытные данные из кинофильма ложатся на прямую автомодельного решения

Приведем еще один скорее забавный пример применения анализа размерностей: "докажем" с его помощью теорему Пифагора. Площадь прямоугольного треугольника  $S$  определяется величиной его гипотенузы  $c$  и для определенности меньшим из его острых углов  $\varphi$ , причем из анализа размерности следует  $S = c^2 f(\varphi)$ ,  $f$  – функция, значения которой безразмерны. Высота, перпендикулярная гипотенузе, разбивает основной треугольник на два подобных ему прямоугольных треугольника. Очевидно, что их площади  $S_1 = a^2 f(\varphi)$ ,  $S_2 = b^2 f(\varphi)$ , где  $f(\varphi)$  – та же, что и в случае основного треугольника. Но поскольку  $S_1 + S_2 = S$ , то  $a^2 f(\varphi) + b^2 f(\varphi) = c^2 f(\varphi)$ , откуда следует, что  $a^2 + b^2 = c^2$ .

## 5.10 Автомодельность

Персонажи обязаны подчиняться законам мира, в котором живут. То есть писатель – пленник собственных предпосылок.

У. Эко

В различных областях математической физики (задачах гидродинамики, переноса тепла, диффузии, химической кинетики) появляются своеобразные частные решения, которые при изменении времени преобразуются одно из другого по правилу подобия. Для функции пространственной и временной переменных такое решение имеет вид:

$$f(x, t) = t\alpha g(t\beta x) \quad (5.49)$$

и инвариантно относительно преобразования  $t \rightarrow bt$ ,  $x \rightarrow x/b^\beta$ ,  $f \rightarrow b^\alpha f$ , образующего мультипликативную группу. В геометрии такое

преобразование называется аффинным. Само решение (5.49) называется автомодельным. В ранее рассмотренном автомодельном решении (5.40) волнового уравнения подобие сводится к параллельному переносу начального профиля и не связано с изменением масштаба. В таком решении аргумент  $x \pm at = const$ , устойчивый относительно преобразований трансляции  $t \rightarrow t+b$ ,  $x \rightarrow x \pm at$ , образующих аддитивную группу, может быть приведен к виду (5.49) с помощью некоторой замены. Показатель степени  $\beta = \pm a$  в (5.49) называется показателем автомодельности. Наиболее важное свойство автомодельного решения состоит в том, что зависимость от аргументов входит через единственный комплекс. Поэтому подставляя (5.49) в уравнение в частных производных, мы получим обыкновенное дифференциальное уравнение. Его интегрирование существенно проще получения решения начально-краевой задачи. В сложных нелинейных задачах получение таких решений зачастую остается единственно возможным средством "пробиться" через аналитические трудности и обрести понимание качественных особенностей явления. При обработке опытных данных автомодельность приводит к тому, что, казалось бы, беспорядочное в обычных координатах "облако" опытных точек ложится на единую кривую или поверхность, построенную в автомодельных координатах. Автомодельность привлекает внимание и как глубокий физический факт, свидетельствующий о наличии определенного типа стабилизации исследуемых процессов, имеющей место для достаточно широкого круга начальных и/или граничных условий.

Получим автомодельное решение линейного параболического уравнения диффузии

$$\frac{\partial \rho}{\partial t} = \kappa \frac{\partial^2 \rho}{\partial x^2} \quad (5.50)$$

где  $\rho$  – плотность массы,  $\kappa$  – коэффициент диффузии. Используем для построения автомодельного решения анализ подобия и размерности. Решение не может зависеть от естественного произвола в выборе основных единиц измерения и, как следует из анализа размерности, должно представляться в виде функции

единственно возможного безразмерного комплекса  $x^2/kt$ . Поэтому будем искать решение в виде

$$\rho = t^\alpha g\left(\frac{x^2}{kt}\right) \quad (5.51)$$

т.е. нам остается определить показатель  $\alpha$  и функцию  $g$ . Показатель степени находится из закона сохранения масс  $\int \rho dx = m$ .

Подставляя решение (5.51) в уравнение (5.50), приходим к соотношению

$$\int_0^\infty t^\alpha g\left(\frac{x^2}{kt}\right) dx = \int_0^\infty t^\alpha g(s) \sqrt{kt} \frac{ds}{2\sqrt{s}} = \sqrt{kt}^{\alpha+1/2} \int_0^\infty g(s) \frac{ds}{2\sqrt{s}} = const$$

( $s = x^2/kt$ ). Обыкновенные линейные дифференциальные уравнения 2-го порядка с переменными коэффициентами редко интегрируются в квадратурах, но оператор в нашем случае оказывается представимым в виде произведения двух дифференциальных операторов первого порядка:

$(2sD+1)(4D+1)g=0$ , где  $D$  означает операцию дифференцирования. Интегрируя уравнение  $(2sD+1)f=0$  по частям, получаем

$$f = \frac{B}{\sqrt{s}}.$$

Возвращаясь к функции  $g$ , будем иметь для нее уравнение

$$(4D+1)g = \frac{B}{\sqrt{s}} \Leftrightarrow 4g' + g = \frac{B}{\sqrt{s}},$$

решение которого

$$g(s) = C \exp(-s/4) + \frac{B}{4} \exp(-s/4) \int_0^s \frac{1}{\sqrt{z}} \exp(z/4) dz$$

состоит из двух линейно независимых функций

$$g_1(s) = \exp(-s/4) \int_0^s \frac{1}{\sqrt{z}} \exp(z/4) dz, \quad g_2(s) = \exp(-s/4).$$

Можно показать, что первое из решений приводит к отрицательным значениям для плотности массы на  $x \in [-\infty, 0]$ , что не имеет физического смысла, а второе решение вполне пригодно:

$$\rho(x, t) = t^{-1/2} \exp\left(-\frac{x^2}{4\kappa t}\right). \quad (5.52)$$

Графики полученного автомодельного решения в некоторые последовательные моменты времени показаны на рисунке 5.5. При  $t \rightarrow 0$  и  $x = 0$  имеем  $\rho \rightarrow 0$ , что видно из рисунка 5.5. В пределе получаем "функцию", равную нулю всюду, кроме точки  $x = 0$ , и бесконечности в этой точке, причем  $\int_{-\infty}^{+\infty} \rho(x, +0) dx = m$ , то есть уже встречавшуюся нам "дельта-функцию":

$$\rho(x, +0) \square m\delta(x).$$

Так как масса  $m$  вещества сохраняется, решение (5.52) описывает эволюцию сгустка частиц, первоначально находящегося в начале координат. В пределе при  $t \rightarrow \infty$  мы получим равномерно размазанное по оси  $x$  поле  $\rho \rightarrow 0$  (тенденция эта видна из рисунка 5.5), однако суммарная масса вещества остается равной  $m$ .

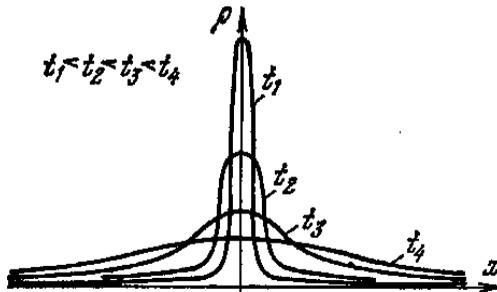


Рисунок 5.5 – Автомодельное решение уравнения диффузии

В разобранный задаче показатель автомодельности  $\beta$  в (5.49) нашелся из соображений размерности. Решения, для которых этот показатель определяется на основании законов сохранения либо размерности, Я.Б. Зельдович отнес к автомодельным решениям первого рода. Если для нахождения этого показателя требуется решение специальной задачи на собственные значения, то

соответствующие решения именуют автомодельными второго рода.

Удивительно то, что в *нелинейных* системах на решения такого рода можно выйти из некоторого класса начальных условий. Такие системы со временем забывают "случайные черты" начальных условий и выходят на автомодельное решение (или одно из автомодельных решений, если их несколько). Об этом далее мы поговорим подробнее.

## 5.11 Самоорганизация и структуры в нелинейных средах

Поистине восхищения заслуживает изобретательность, с которой природа столь тонко и разнообразно маскирует простоту под сложность!

П. Эткинс

Линейное волновое уравнение (5.7) прекрасно описывает распространение звука, когда его громкость невелика (при этом скорость звука не зависит от громкости). Если звук очень громкий, например, от взрыва, может возникнуть ударная волна. Ее скорость зависит от разности давлений за волной и перед ней. Для описания такого поведения требуется формулировать нелинейную модель. Однако получение решения и качественное исследование нелинейной модели требует других подходов, поскольку даже малая нелинейная добавка качественно меняет ситуацию: сумма двух решений уже не удовлетворяет уравнению. Принцип суперпозиции "не работает" и "сшить" общее решение из частных уже не удастся.

Исследуем, к каким качественным изменениям приводит появление нелинейности в простейших математических моделях на примере уравнения теплопроводности, по виду совпадающего с уравнением диффузии (5.50)

$$\frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} \quad (5.53)$$

где  $k > 0$  – коэффициент теплопроводности,  $T \rightarrow 0$  – температура. Оно описывает передачу тепла, диффузию частиц, проникновение магнитного поля в плазму и некоторые другие процессы. Представим, что у нас есть компьютер с соответствующим программным обеспечением, который умеет решать различные уравнения. Остается только задать граничные условия и начальные данные и посмотреть, как ведет себя решение. Будем считать для простоты, что отрезок бесконечен,  $-\infty < x < +\infty$ , причем  $T(k \rightarrow \pm\infty) \rightarrow 0$ . Пусть в начальный момент до высокой температуры нагрет маленький участок (рисунка 5.6а). На рисунке 5.6 а показаны профили температуры в разные моменты времени.

Можно видеть, что максимум температуры уменьшается как  $A \propto t^{1/2}$ , а полуширина профиля на уровне  $A/(2L)$  растет как  $L \propto t^{1/2}$ . Это кажется понятным: от более нагретых участков тепло передается менее нагретым, при этом температура первых уменьшается, а вторых растет. Поскольку количество тепла остается неизменным,  $AL \propto const$ .

Теперь рассмотрим среду, в которой происходит процесс горения. При этом в правой части уравнения теплопроводности появляется новый член, описывающий тепловыделение источников. Будем считать, что интенсивность горения пропорциональна температуре. Это приводит к линейному уравнению

$$\frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} + qT, \quad q > 0. \quad (5.54)$$

Решение этого уравнения для тех же начальных данных, представленное на рисунке 5.6 б, имеет вид: амплитуда  $A$  растет, а полуширина меняется по закону  $L \propto t^{1/2}$ . Таким образом, закономерности очень похожи. В чем причина? Оказывается, что уравнение (5.54) можно привести к виду (5.53), если ввести замену переменной  $T^* = \exp(qt)T$ . Значит,  $A \propto \exp(qt)t^{1/2}$ , что совпадает с результатами расчета.

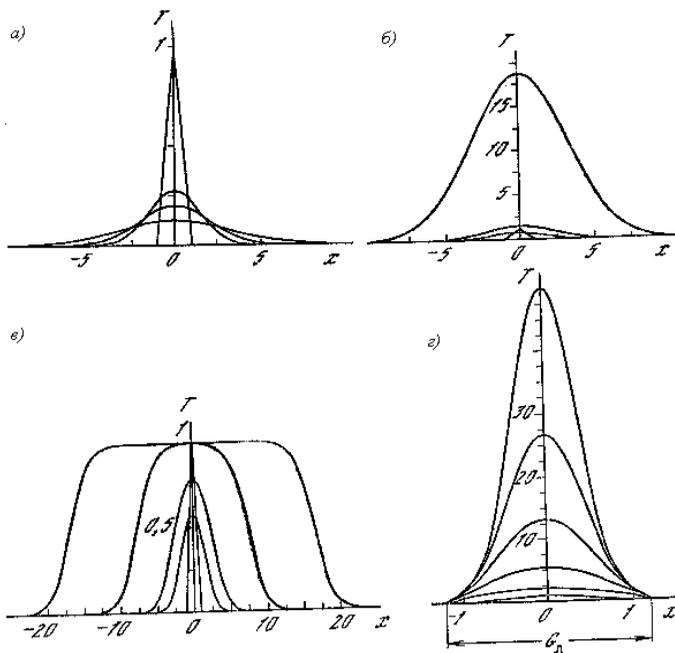


Рисунок 5.6 – Решения уравнения теплопроводности

Посмотрим, что произойдет, если источник нелинеен:

$$\frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} + qT - \alpha T^3; \quad q, \alpha > 0. \quad (5.55)$$

Нелинейный член описывает прекращение горения при больших температурах. Это может быть связано с выгоранием топлива или с большей ролью эндотермических реакций с ростом  $T$ . Уравнение (5.55) возникает не только в теории горения. С его помощью моделируют распространение эпидемий, прохождение импульса по нервному волокну. Посмотрим, как выглядят решения этого уравнения (рисунок 5.6 в). Они разительно отличаются от того, что мы видели раньше: возникает тепловая волна, распространяющаяся с постоянной скоростью, причем амплитуда волны стремится к постоянному значению  $(q/\alpha)^{1/2}$ . Но известно,

что бегущая волна – это автомодельное решение вида  $T_*(\pm \eta t)$ , сохраняющее свою форму. Причем в отличие от линейного уравнения теплопроводности, где существует бесконечный набор собственных функций  $T_1, T_2, \dots$ , здесь функция  $T_*$  только одна и определяется свойствами нелинейной среды (в нелинейных уравнениях собственных функций может оказаться конечное число). Роль функции  $T_*$  в модели тепловых структур велика. Она определяет локализованные конфигурации в пределах, в которых процессы идут согласованно. Именно поэтому ее во многих работах называют собственной функцией нелинейной среды. Однако в отличие от линейных задач она описывает локализованные процессы и никак не связана с краевыми условиями.

Наличие источников и стоков является типичным в так называемых открытых системах, которые в отличие от замкнутых, могут обмениваться с окружающей средой энергией, веществом, информацией. "Забывание" начальных данных, то есть выход с целого класса начальных профилей на одно и то же решение характерно для большого класса открытых нелинейных систем. Причем решение, на которое происходит выход, часто оказывается автомодельным. Такое поведение говорит о возникновении упорядоченности в системе или о самоорганизации. В самом деле, выход на автомодельное решение означает уменьшение числа степеней свободы и выделение нескольких основных (параметров порядка), к которым подстраивается все остальные.

Рассмотрим систему с нелинейным коэффициентом теплопроводности  $k$  и нелинейным источником, обладающим следующим свойством: чем больше отклонение от равновесия, тем быстрее идет процесс

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( k T^\sigma \frac{\partial T}{\partial x} \right) + q T^\beta, \quad k, q > 0, \quad \beta > \sigma + 1 \quad (5.56)$$

Такие модели характерны для физики плазмы, химической кинетики, экологии, разрушения, пластичности. Система (5.56) обладает еще более необычными свойствами. Обратим внимание на два принципиальных отличия от всех других решений, обсуждавшихся выше. Профиль температуры оказывается лока-

лизованным внутри некоторой области  $G\lambda$ , вне которой  $T(x,t)$  равна нулю (рисунок 5.6 г). Из непрерывной среды при этом выделяются ограниченные участки, в пределах которых и происходит горение. Решение существует только в течение ограниченно-го времени  $t_f$ , называемого временем обострения. За это время функция  $T(x,t)$  в конечной области обращается в бесконечность. До недавнего времени математики полагали, что такие решения не представляют особого интереса, что их наличие говорит о несовершенстве модели. Однако развитие физики плазмы, газовой динамики, других областей привело к появлению содержательных задач, в которых ведущими оказываются один или несколько наиболее быстрых процессов. Если посмотреть внимательно на решение, то оказывается, что профили температуры в процессе эволюции остаются подобными себе. Не определяются ли и они автомодельными решениями? Действительно, на развитой стадии (когда выделяется много больше тепла, чем в начальный момент времени) процесс горения описывается формулой

$$T = g(t) f(x/\varphi(t)). \quad (5.57)$$

Здесь  $g(t)$  определяет закон роста амплитуды,  $\varphi(t)$  – полуширины,  $f$  – форму профиля,  $g(t) \rightarrow \infty$  при  $t \rightarrow t_f$ ,  $g, \varphi$  – функции степенного вида

$$g(t) \propto (t-t_f)^{-1/(\beta-1)}, \varphi(t) \propto (t-t_f)^{(\beta-\sigma-1)/(\beta-1)}.$$

Обратим внимание на то, что в пределах области локализации горение происходит согласованно в разных точках пространства. Выход на решение вида (5.57) говорит о спонтанном возникновении упорядоченности, о формировании локализованных структур. Для этой нелинейной среды также характерно "забывание" деталей начальных данных. Это в некотором смысле парадоксально. В среде, где есть только горение и теплопроводность – диссипативный процесс, связанный с рассеянием энергии и обычно уничтожающий всякую упорядоченность, – возникают структуры, сохраняющие свою форму. Чтобы подчеркнуть необычность этого явления, их называют *диссипативными структурами*.

Познакомимся подробнее с тепловыми структурами. Ясно, что у них есть две важные характеристики. Это время их жизни  $t_f$  и область локализации  $G\lambda$ . Если в среде независимо развиваются две локализованные структуры с разными временами обострения  $t_{f2} > t_{f1}$ , то практически горение происходит на временах  $\approx t_{f1}$ , то есть только в первой, и замирает во второй. Реально живут в одном мире только структуры с одинаковыми временами обострения. А если попробовать объединить несколько структур так, что их области локализации пересекутся? Структуры начинают взаимодействовать, возникает волна горения сложной формы, сходящаяся к центру. Однако эта форма с ростом температуры меняется, и в конце концов остается одна простая быстро горящая структура. Встает вопрос – можно ли построить сложные структуры, сохраняющие в процессе эволюции свою форму? Вопрос о наборе форм, заключенном в различных средах, является очень глубоким. И для древних философов, и для многих физиков-теоретиков, занимающихся элементарными частицами, характерно представление о сложном наборе структур, которые потенциально заключены в среде. Некоторые из них встречаются часто, другие могут возникать только в определенных условиях. Однако и те, и другие характеризуют глубокие внутренние свойства системы.

Посмотрим на собственные функции нелинейной среды с другой точки зрения. Начальные условия можно рассматривать как воздействия на нашу систему. Пусть задача состоит в том, чтобы создать в среде сложную структуру желаемой организации. Оказывается, чтобы решить эту задачу, достаточно задать начальный профиль температуры малой амплитуды, согласованный с собственными функциями нелинейной среды. Для того, чтобы воздействие на нелинейную систему было эффективным, оно должно быть согласовано с ее внутренними свойствами. Такое воздействие, даже будучи слабым, окажется действеннее, чем в тысячи раз более сильное, но не согласованное со свойствами системы.

## 5.12 О нелинейных волнах в сплошных средах

*В поисках истины логика нащупывает свой путь по отклонениям от обычного и заурядного*

*Э. По*

Поскольку большинство ранее рассмотренных примеров составили волны, то нельзя не рассказать о нелинейных волнах на мелкой воде, к которым относят цунами. Цунами – необычайно высокие и мощные волны, обрушивающиеся на побережья океанов. Чаще всего цунами возникают в результате сильных подводных землетрясений, сопровождающихся быстрым образованием на дне океана сбросов, обвалов и оползней. Быстрое смещение блоков горных пород, словно гигантский поршень, выталкивает огромные массы воды в виде горба на поверхность океана. Навероятно страшная волна высотой 30-40 метров образовалась при взрыве вулкана Кракатау в 1883 году и в течение нескольких минут смыла с индонезийских островов 35 тысяч человек. Волны эти необычайно устойчивы и, удаляясь от своего (практически точечного) источника на многие тысячи километров, способны пересечь Тихий океан. Например, Чилийское цунами 1960 года нанесло значительный урон Японским и Курильским островам. Эти волны – очень пологие и длинные: длина волны составляет 100-300 километров, но даже сильнейшие цунами имеют в открытом океане высоту не более чем 2 метра. Суда, которые оказались в открытом океане на пути цунами, обычно не ощущают "никакого цунами". Скорость этих волн – колоссальная и составляет в зависимости от глубины океана 400-850 километров в час. Цунами уменьшают свою высоту, проходя над большими глубинами, и увеличивают ее на мелководье. Скорость, напротив, увеличивается на больших глубинах и уменьшается на малых. Подходя к берегу, длина и скорость волны уменьшаются, а ее высота увеличивается (кинетическая энергия преобразуется в потенциальную), волна принимает резко асимметричную форму и опрокидывает свой гребень вперед, тараном обрушиваясь на берег.

Вероятно, читатель сразу отметил, что некоторые детали указывают на нелинейную природу этих волн (например, существование зависимости скорости от амплитуды). Поэтому строить линейные модели здесь бессмысленно, они не будут адекватными. Вероятно, первым естествоиспытателем, очарованным нелинейной волной был англичанин Джеймс Скотт Рассел (кстати, конструктор гигантского океанского лайнера "Грейт Истерн", с которого укладывался первый трансатлантический телеграфный кабель). В 1834 году, прогуливаясь верхом вдоль узкого судоходного канала, он заметил, как от остановившейся баржи отделилось и побежало вперед большое одиночное возвышение ("водяной холм"), которое продолжало свой путь, нисколько не меняя своей формы и не снижая скорости, пока, как писал Рассел, *"через одну или две мили погони я не потерял его из виду"*. Это наблюдение вдохновило Рассела провести экспериментальное изучение поверхностных волн на воде, в результате чего были воспроизведены уединенные волны и получено выражение для скорости их распространения

$$v = \sqrt{g(h + \eta)},$$

где  $g$  – ускорение свободного падения,  $h$  – глубина спокойной жидкости, а  $\eta$  – высота гребня волны над поверхностью спокойной жидкости. Заметим: скорость распространения волны зависит от ее амплитуды! Этот результат наблюдательного ученого не был понят многими известными его коллегами и полемика продолжалась полстолетия, пока голландцы Дидерик Кортевег и Густав де Фриз не вывели своего уравнения мелкой воды

$$v_t + 6vv_x + v_{xxx} = 0.$$

Это уравнение – квазилинейное, параболического типа,  $v(x,t)$  – скорость материальной точки жидкости.

Далее мы изучим другое нелинейное уравнение распространения нелинейных волн, обладающее всеми особенностями волн цунами, – интегро-дифференциальное уравнение в частных производных

$$v_t + v v_x + \int_{-\infty}^{\infty} K(x-s) v_s(s, t) ds = 0, \quad (5.58)$$

предложенное в 1967 году Дж. Уиземом. Но прежде поговорим о важных "атрибутах" нелинейных волн – дисперсии и нелинейности.

Пусть рассматривается линейное дифференциальное уравнение в частных производных по  $x$  и  $t$  (с постоянными коэффициентами), которое в символической форме имеет вид:

$$P\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right)v = 0 \quad (5.59)$$

где  $P$  – некоторый полином. Подставим решение в нормальной форме

$$v = A \exp(ikx - i\omega t), \quad (5.60)$$

где  $k$  – волновое число (вещественное),  $\omega$  – частота (комплексное число) в уравнение (5.59):

$$G(\omega, k) A \exp(ikx - i\omega t) = 0,$$

где  $G$  – полином. Экспонента не может обращаться в нуль при всех  $x, t$ , поэтому нетривиальное решение (5.59) ( $A \neq 0$ ) существует, если  $k$  и  $\omega$  связаны дисперсионным соотношением

$$G(\omega, k) = 0. \quad (5.61)$$

Предположим, что (5.61) позволяет определить  $\omega$  как вещественную функцию волнового числа  $k$ . Для каждой моды (т.е. каждой  $k$ ) точки постоянной фазы перемещаются с фазовой скоростью

$$c = \omega / k. \quad (5.62)$$

Если отсутствует зависимость фазовой скорости от волнового числа  $k$ , то есть

$$\frac{d^2 \omega}{dk^2} = 0 \quad (5.63)$$

то говорят, что волна, описываемая уравнением (5.59), перемещается без *дисперсии*. В противном случае имеет место дисперсия

волн, здесь различные волны уже имеют различные фазовые скорости, и поведение решения зависит от того, как волны интерферируют друг с другом. В этом случае вводят групповую скорость

$$c_g = \frac{d\omega}{dk},$$

значение которой состоит в том, что по истечении достаточно большого времени каждое волновое число доминирует в решении в области  $x \approx c_g t + O(t)$ .

Дисперсионное уравнение для волнового уравнения (5.7) было нами уже получено (5.21):  $\omega = \pm ak$ , откуда  $c = a$ , следовательно, волны распространяются без дисперсии. Для уравнения диффузии (5.53) мы имеем  $\omega = \kappa k^2$ , откуда  $c = c_g / 2 = \kappa k$ , то есть дисперсия волн наблюдается (иное было бы странно для уравнения диффузии). Мы уже знаем, что решения (5.59) могут быть представлены в виде суперпозиции (суммы или интеграла) по  $k$  волн вида (5.60). Поэтому смысл дисперсии волн – в "размазывании" с течением времени волнового возмущения по пространству. Недиспергирующие среды переносят это возмущение без изменения его формы.

Сказанное относилось к линейным уравнениям и к зависимости  $\omega(k)$  (не путать с зависимостью скорости волны от ее амплитуды  $c(A)$ , существующей только для нелинейных волн). В пространственном случае (т.е. когда независимые переменные дифференциального уравнения суть  $t, r$ ), нормальная форма решения берется в виде плоской волны  $v = A \exp(ikr - i\omega t)$  ( $k$  – волновой вектор). В неоднородных или нелинейных средах, уравнение (6.59) становится почти линейным, квазилинейным или сильно нелинейным, и попытка дать определение дисперсии волн затруднительна, поскольку у упомянутых уравнений, как правило, отсутствуют решения в виде плоских волн.

Для исследования дисперсионных свойств уравнения Уизема рассмотрим вспомогательное линейное уравнение

$$v_t + \int_{-\infty}^{\infty} K(x-s)v_s(s,t)ds = 0. \quad (5.64)$$

Подставляя в него решение вида (6.60), получим связь

$$i\omega \exp(ikx) - \int_{-\infty}^{\infty} K(x-s) ik \exp(iks) ds = 0,$$

откуда делением на  $ik \exp(ikx)$  и заменой переменной  $x-s = y$  под интегралом получаем

$$c(k) = \frac{\omega}{k} = \int_{-\infty}^{\infty} K(y) \exp(-iky) dy.$$

Поскольку правая часть представляет собой преобразование Фурье ядра  $K(y)$ , выполняя обратное преобразование, запишем

$$K(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} c(k) \exp(ikx) dk.$$

Таким образом, задаваясь любой желаемой функцией  $c(k)$  (дисперсионной зависимостью), можно определить соответствующее ядро  $K(x)$ , то есть модельное уравнение описывает дисперсию волн.

Теперь остановимся на вопросе нелинейности. Нелинейность уравнения Уизема (6.58) связана с конвективным членом  $v v_x$ . Простейшее уравнение в частных производных, обладающее этой нелинейностью, имеет вид

$$v_t + v v_x = 0 \tag{5.65}$$

и носит название уравнения Эйлера. Снабдим его начальными условиями

$$v(x, 0) = v_0(x). \tag{5.66}$$

Но (5.65) нам уже встречалось раньше. Комбинация  $v_t + v v_x$  часто присутствует в уравнениях динамики жидкости и газа и, как мы показали, появляется при переходе от материальных (лагранжевых) координат к пространственным (эйлеровым). Без такого перехода вместо исходного квазилинейного уравнения вида  $v_t + v v_x + L(v) = 0$  мы имели бы почти линейное уравнение  $\dot{v} + \tilde{L}(v) = 0$ : избавляясь от  $v v_x$ , получаем неоднородные коэффициенты в остальной части оператора –  $\tilde{L}(v)$ . Однако в этом

случае мы получим решение, определенное в материальных (лагранжевых) переменных, координаты которых в пространстве сами требуют определения из системы дифференциальных уравнений (каких?). При описании развитого турбулентного движения второй подход является намного более сложным. К тому же в механике жидкости и газа часто более естественно задаются граничные условия на пространственной (неподвижной) границе области (каналы, резервуары или просто произвольная область пространства, сквозь которую стационарно течет среда).

Ранее мы находили уравнение характеристик для (5.65):

$$\frac{dx}{dt} = v(x, t).$$

Если бы поле  $v(x, t)$  было постоянным, то общее решение (5.65) имело бы вид

$$v = f(x - vt).$$

Подстановка такого представления в исходное уравнение с учетом начальных условий дает нам неявное решение

$$v = v_0(x - vt), \tag{5.67}$$

обнаруживающее некоторую зависимость скорости от амплитуды (а не позволит ли нам это описать свойства волн цунами и опрокидывание волн?).

Качественный ответ на вопрос можно получить с помощью характеристик. Пусть начальная функция  $v_0(x)$  имеет вид рис. 6.7а. Выберем материальную точку с координатой  $x$  и построим проходящую через нее характеристику  $\Gamma\xi: x = \xi + v_0(\xi)t$  с углом наклона  $1/v_0(\xi)$  (рисунок 5.7 б).

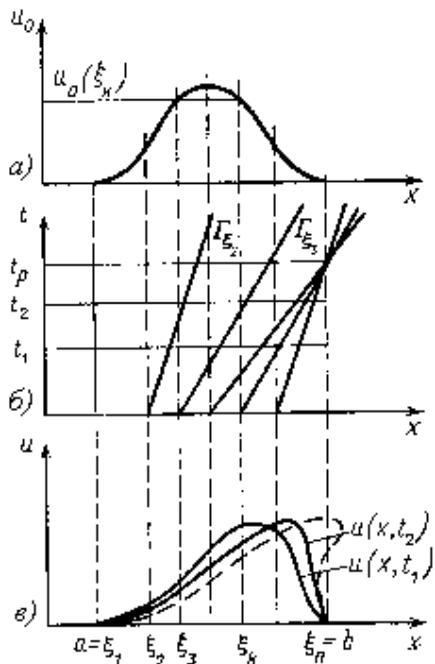


Рисунок 5.7 – Поведение решения уравнения Эйлера

Выберем момент времени  $t$  и определим графически из рисунка 5.7 б эйлерову координату  $x_k$  материальной частицы  $x$ . Теперь точку  $x, v_0(x)$  можно отложить на плоскости  $(x, v)$ . Выполнив данную процедуру со всеми материальными точками  $x$  оси, мы получим профиль  $v(x, t)$  волны на плоскости  $(x, v)$ . Обратимся вновь к рис. 6.7а, б и заметим, что при данном начальном профиле характеристики должны неизбежно пересечься, начиная с некоторого  $t_p$ . Это означает, что при  $t > t_p$  функция  $v(x, t)$  оказывается неоднозначной (рис. 6.7в). Из построенных графиков и (5.67) видно, что чем больше амплитуда начальной скорости  $v_0(x)$  материальной точки в  $x$ , тем с большей скоростью она распространяется. Гребень волны неизбежно опережает ее подошву. То есть, уравнение Эйлера описывает опрокидывание волн! Здесь нужно заметить, что интерпретация этого явления зависит от того, что мы понимаем под смыслом решения уравнения. Если допустить

неоднозначные решения  $v(x,t)$ , то мы столкнемся с опрокидыванием волны, а если считать его однозначным, то уравнение Эйлера опишет разрушение волны, т.е. образование разрыва скорости. Такое решение будет пониматься в ином, неклассическом смысле.

Нетрудно увидеть, что уравнение Уизема содержит в себе оба рассмотренных модельных уравнения. При этом единоразовое взаимодействие механизмов опрокидывания (разрушения) и диссипации волн приводит к весьма богатой динамике уравнения. Теоретические исследования показали, что, во-первых, это уравнение имеет решения типа уединенных волн, распространяющихся с неизменной формой и постоянной скоростью. Скорость таких волн монотонно увеличивается с ростом амплитуды, причем существует предельная амплитуда. При этом оказывается, что гребень волны с предельной амплитудой заостряется ("стоксово заострение"). Уравнение имеет и периодические решения ("рябь"), обладающие теми же особенностями. Во-вторых, существуют режимы, при которых волны разрушаются.

Интересно исследовать процесс набегания волн на берег. Выбирая дисперсионное соотношение (мы имеем такое право) в виде

$$c(k) = \sqrt{gth(kh)/k},$$

где  $h$  – глубина,  $g$  – ускорение свободного падения,  $th$  – гиперболический тангенс, и вычисляя по нему ядро  $K(x)$ , а также предполагая берег пологим (т.е.  $h$  постоянно на каждом этапе эволюции волн), мы имеем следующую картину (рисунок 5.8). Находясь на сравнительно большой глубине, волны являются гладкими периодическими бегущими волнами; нелинейность при этом играет лишь роль фактора, сдерживающего дисперсию. При переходе на меньшую глубину задние гребни волн, имеющие большую скорость распространения, начинают догонять передние, в результате чего происходит уменьшение длины волн. Однако энергия волн, оставаясь неизменной (а это еще одно замечательное свойство этого уравнения), приводит к увеличению амплитуды волн. Процесс происходит до некоторой предельной амплитуды и сопровождается заострением волн (нелинейность здесь проявляет

себя именно в этом). Волны предельной амплитуды, переходя на мелководье, начинают опрокидываться, поскольку при  $h \rightarrow 0$   $K(x) \rightarrow 0$  и нелинейность полностью "побеждает" диссипацию.

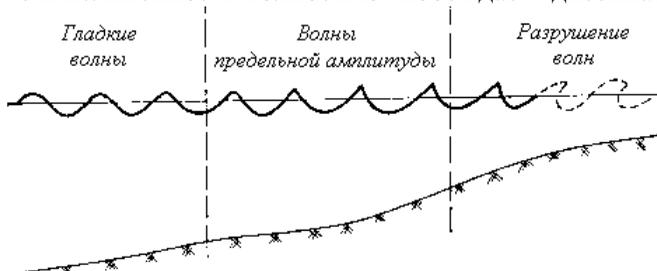


Рисунок 5.8 – Зависимость профиля волнового решения уравнения Уизема от глубины

Необходимо сказать, что здесь нами было предпринято качественное моделирование явления. Мы не выводили уравнения движения волн на поверхности жидкости, а изначально изучали свойства некоторых модельных уравнений, из которых "склеилось" уравнение Уизема. Такой путь, требующий большой физической интуиции, является весьма ценным, но не заменяет собой "честного" пути построения модели: постановки проблемы, принятия гипотез, вывода уравнений и их исследования. Специалист в области математического моделирования должен овладеть обоими подходами.

### 5.13 Иерархические модели турбулентности и многомасштабные функциональные базисы

Когда я предстану перед Творцом, то первое, о чем я спрошу его, будет: "В чем же загадка турбулентности?"

Т. фон Карман

Сложные нелинейные системы часто имеют иерархическую структуру, состоящую в том, что элементы этой системы образуют иерархию уровней, а взаимодействия между ними становятся

существенными только в том случае, если элементы близки друг другу и "по вертикали", и "по горизонтали".

Ярким примером такой системы может служить развитая турбулентность, характерной особенностью которой является многомасштабность. В турбулентном потоке сосуществуют вихри самых различных размеров – от размеров порядка характерного масштаба потока (диаметра трубы, размера обтекаемого тела, размера сосуда и т.д.), называемого интегральным, или макромасштабом турбулентности, до наименьшего, диссипативного, или микромасштаба, на котором становятся существенными вязкие силы, переводящие подводимую к потоку механическую энергию в тепловую. Тот факт, что макромасштабы, на которых подводится энергия, и микромасштабы, на которых энергия диссипирует, могут отличаться в тысячи (или даже миллионы) раз и определяет необходимость формирования *каскада* энергии от больших вихрей к меньшим.

В этом разделе мы расскажем о моделях турбулентности, основанных на идее применения функционального базиса специального типа, наиболее точно соответствующего структуре турбулентных полей. Идея такого базиса впервые была предложена В.Зиминым в конце семидесятых годов и состояла в использовании семейства самоподобных функций прогрессивно убывающего масштаба. Базис был назван иерархическим и на его основе были построены и исследованы многочисленные модели, также названные иерархическими. В конце восьмидесятых годов в научной литературе появилось слово "вейвлет", а к началу девяностых вейвлет-анализ превратился в самостоятельную, хорошо развитую область математической физики. Идеи, лежащие в основе теории вейвлетов, совпадают с идеями иерархического представления турбулентных полей и в терминах этой молодой науки иерархические модели – это модели, построенные с помощью вейвлет-представления описываемых полей.

Поскольку цель этой главы состоит в изложении подходов к моделированию сложных нелинейных процессов, то мы начнем с идей, приведших к иерархическим моделям. В то же время, нельзя не остановиться и на формулировке основных положений вейвлет-анализа, который оказывается чрезвычайно полезным

при анализе временной и пространственной структуры нелинейных систем.

### ***Иерархический базис для турбулентных полей***

Рассматривая численные методы решения уравнений движения жидкости, можно отметить, что чаще всего используются либо сеточные, либо спектральные методы, либо их комбинация. И те, и другие можно отнести к проекционным методам решения уравнений в частных производных, когда для решения используют проекции всех полей на функциональные базисы.

В сеточных методах функции представлены значениями в точках, плотность которых связана со спектральными свойствами рассматриваемых полей (мелкомасштабные вихри не должны проваливаться между точками сетки). Более строго эта связь выражается теоремой Котельникова, согласно которой функция  $f(x)$ , спектр которой ограничен пространственной частотой  $2\pi/h$ , может быть представлена суммой функций отсчетов (синусов), центры которых размещены на сетке с шагом  $h$ . Очевидно, что сеточное представление эффективно при описании локальных структур – мелкомасштабный вихрь описывается небольшим числом точек, находящихся в соответствующей области пространства. В то же время, для описания даже очень простого по структуре крупномасштабного вихря требуется использование всех базисных функций.

Спектральные методы используют разложение по фурье-гармоникам. В этом случае каждая базисная функция описывает, по сути, систему когерентных вихрей, занимающую все пространство. В таком представлении очень просто описать вихрь, занимающий всю область, или периодическую систему вихрей – и в том, и в другом случае достаточно одной базисной функции. Однако, если требуется описать отдельный вихрь, занимающий малую часть рассматриваемой области, то потребуются весь гармонический ряд.

Сеточные методы эффективны при вычислении нелинейных членов, так как позволяют выразить значение в точке через небольшое число соседних точек, но приводят к большим затратам машинного времени при решении уравнения Пуассона, требую-

щего построения итерационного процесса, в который вовлечены все точки области. Спектральные методы, наоборот, делают решение уравнения Пуассона тривиальным, но приводят к очень сложной структуре нелинейных членов.

Проблемы двух функциональных базисов связаны с их локализованностью в физическом и в фурье-пространствах. Сетки строго локализованы в физическом пространстве, но спектр точки (дельта-функции) представляет собой белый шум. Это означает, что функции делокализованы в пространстве Фурье. Обратная ситуация возникает при разложении Фурье. Каждая гармоника представляет строго одну частоту, но соответствующая ей функция занимает все физическое пространство.

В турбулентном потоке сосуществуют вихри самого различного масштаба, но наиболее эффективные взаимодействия происходят между вихрями (структурами), близкими и в физическом, и в фурье-пространстве. Первое очевидно – чтобы вихри взаимодействовали, они должны перекрываться в пространстве. Второе утверждение составляет основу концепции каскадных процессов – взаимодействуют вихри сравнимых размеров (если размеры не сопоставимы, то маленькие вихри просто переносятся большими без обмена энергией). Это заставляет обратиться к поиску специальных функций, более точно соответствующих структуре турбулентного потока.

Еще один недостаток использования рядов Фурье состоит в низкой информативности высоких частот. Хорошо понятен смысл рассмотрения вихрей с характерным размером  $L$ ,  $L/2$ ,  $L/3$ , ..., но отдельное описание масштабов  $L/957$ ,  $L/958$ ,  $L/959$ , ... и так далее мало оправдано. Это соображение наводит на мысль о необходимости использования функций, масштаб которых изменяется прогрессивно, – такое соотношение получается при равномерном разбиении пространства масштабов в логарифмическом представлении. В теории турбулентности важную роль играет и идея масштабного подобия. Это значит, что искомый базис должен быть составлен из подобных функций.

Суммируя сказанное, можно сформулировать требования, которым должен удовлетворять функциональный базис, предназначенный для описания турбулентных потоков:

- функции базиса должны быть локализованы и в физическом, и в фурье-пространствах;
- функции должны быть подобны и описывать иерархию вихрей прогрессивно убывающих масштабов;
- мелкомасштабные вихри должны переноситься в поле вихрей большего масштаба;
- при подстановке в уравнения Навье-Стокса (уравнения движения линейно-вязкой жидкости) функциональный базис должен приводить к слабосвязанной динамической системе.

Попробуем построить базис, удовлетворяющий этим требованиям. Построения будем проводить для двумерного случая, так как это упрощает иллюстрацию результатов и запись функций.

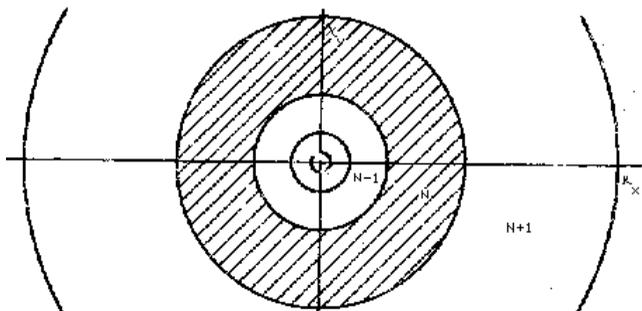


Рисунок 5.9 – Разбиение фурье-плоскости на кольцевые зоны

Итак, имеем двумерное пространство  $r = (x, y)$  и соответствующее ему пространство волновых векторов  $k = (k_x, k_y)$ . Фурье-плоскость разобьем на кольцевые зоны (рис.6.9) таким образом, что для зоны с номером  $N$

$$k_N < |k| < k_{N+1}, \quad k_N = \pi 2^N, \quad N = 0, \pm 1, \pm 2, \quad (5.68)$$

Каждая кольцевая зона включает, таким образом, одну октаву волновых чисел (напомним, что октавой называется интервал, в пределах которого частота изменяется в два раза).

Рассмотрим поле скорости  $v(t, x, y)$  и представим его в виде

$$v(t, x, y) = \sum_N v_N(t, x, y), \quad (5.69)$$

где каждая функция  $v_N$  есть результат фильтрации в фурье-плоскости по соответствующему кольцу (5.68):

$$v_N(t, x, y) = \iint v(t, x', y') g_N(x - x', y - y') dx' dy' \quad (5.70)$$

Здесь  $g_N(r)$  есть функция, фурье-образ которой  $\tilde{g}_N(k)$  локализован в кольце

$$\tilde{g}(k) = \begin{cases} 1, & \text{в кольце } N \\ 0, & \text{вне кольца } N \end{cases} \quad (5.71)$$

В силу определения операции фильтрации (5.70)-(5.71)

$$\int v_N v_M dr \propto \delta_{NM},$$

следовательно, энергия распадается на сумму

$$E = \sum_N E_N = \sum_N \frac{1}{2} \int v_N^2 dr.$$

Таким образом, мы провели первую часть построения – разбили исходное поле по масштабам. На втором этапе нужно провести разбиение полученных полей  $v_N$  на сумму функций, каждая из которых характеризует поле скорости данного масштаба только в определенной области пространства

$$v_N = \sum_n a_{Nn}(t) f_N(r - r_{Nn}),$$

где  $f_N(r)$  есть базисные функции масштаба  $N$ ,  $r_{Nn}$  – радиус-вектор центра вихря (функции). Функции  $f_N(r)$  должны быть подобны, взаимно ортогональны и обеспечивать разреженную матрицу нелинейных взаимодействий  $X_{NnMmLl}$  в уравнении

$$d_t a_{Nn} = \sum X_{NnMmLl} a_{Mm} a_{Ll} + \dots,$$

получающемся при проектировании уравнений Навье-Стокса (уравнений движения линейно-вязкой жидкости) на функциональный базис. Структуру связей, описываемых матрицей  $X_{NnMmLl}$ , иллюстрирует рисунок 5.10. Каждый горизонтальный слой (ярус) соответствует совокупности функций определенного масштаба, а

все вместе описывают иерархическую систему, давшую название моделям этого типа.

Мы не будем останавливаться на свойствах и возможностях иерархических моделей, но поговорим о функциях используемых при их построениях. Увы, удовлетворить всем приведенным требованиям к базисным функциям не удастся. Задача имеет решение в такой постановке только в одномерном случае. Одномерный базис не имеет интереса с точки зрения описания турбулентности, но его построение представляет методический интерес и мы его проведем.

### ***Одномерный иерархический базис***

Рассмотрим функцию  $f(x)$ , для которой существует преобразование Фурье,

$$\tilde{f}(\gamma) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \gamma x} dx.$$

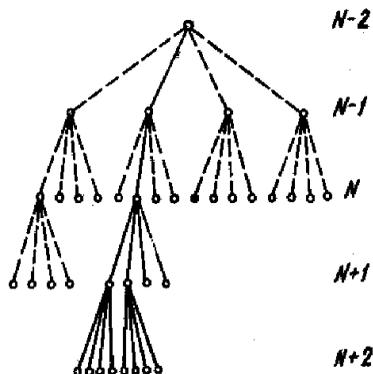


Рисунок 5.10. Структура связей, описываемых матрицей  $X_{NmMl}$

Ось волновых чисел  $\gamma$  (напомним, что  $k = 2\pi\gamma$ ) разбиваем на октавы  $\gamma_N = 2^N$  (рисунок 5.11) и вводим функции

$$\tilde{f}_N(\gamma) = \begin{cases} \tilde{f}(\gamma) & \gamma_N < |\gamma| < \gamma_{N+1} \\ 0 & \text{вне зоны} \end{cases} \quad (5.72)$$

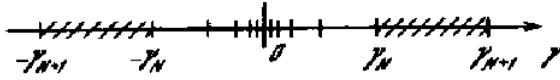


Рисунок 5.11 – Разбиение на октавы оси волновых чисел

Очевидно, что  $\tilde{f}(\gamma) = \sum \tilde{f}_N(\gamma)$ . Полученные функции  $\tilde{f}_N$  обладают замечательным свойством – они допускают периодическое продолжение на всю ось  $\gamma$  с периодом  $2\gamma_N$  (рисунок 5.12)

$$\tilde{F}_N(\gamma) = \begin{cases} \tilde{f}_N(\gamma - 2(m-1)\gamma_N) & (2m-1)\gamma_N < \gamma < 2m\gamma_N \\ \tilde{f}_N(\gamma - 2(m+1)\gamma_N) & 2m\gamma_N < \gamma < (2m+1)\gamma_N \end{cases}$$

Это позволяет разложить функции  $\tilde{F}_N(\gamma)$  в ряд Фурье

$$\tilde{F}_N(\gamma) = \sqrt{h_N} \sum_n A_{Nn} e^{-2\pi i h_N n \gamma} \tag{5.73}$$

где  $h_N = 1/(2\gamma_N)$ . Функции  $h_N e^{-2\pi i h_N n \gamma}$  образуют полный базис в классе функций  $\tilde{F}_N$ , а те же функции, определенные внутри зоны (5.72), – полный базис в классе функций  $\tilde{f}_N$ . Чтобы получить вид базисной функции в физическом пространстве, нужно взять обратное преобразование Фурье. Получается функция вида

$$f_{Nn}(x) = \frac{1}{\sqrt{h_N}} \frac{\sin\left[\frac{\pi}{2h_N}(x - h_N n)\right]}{\frac{\pi}{2}(x - h_N n)} \cos\left[\frac{3\pi}{2h_N}(x - h_N n)\right] \tag{5.74}$$

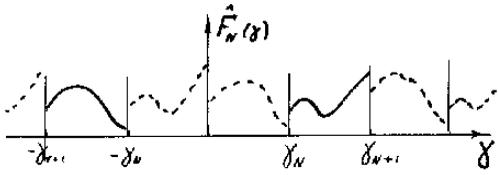


Рисунок 5.12 – Периодическое продолжение функций  $\tilde{f}_N$

Вид функции (5.74) для  $n=0$  показан на рис.6.13. Эти функции известны в математике как функции Литлвуда-Пелли. Функ-

ции медленно убывают в физическом пространстве ( $f_{Nn}(x) \propto x^{-1}$ ), что является результатом обрыва функций в пространстве Фурье. Все базисные функции взаимно ортогональны, то есть

$$\int f_{Nn}(x) f_{Mm}(x) dx = \delta_{NM} \delta_{nm},$$

что следует из ортогональности функций в фурье-пространстве и инвариантности скалярного произведения двух функций относительно преобразования Фурье. Коэффициенты в разложении (6.73) определяются формулой

$$A_{Nn} = \int f(x) f_{Nn}(x) dx.$$

Базисные функции имеют двойную индексацию. Большой индекс отвечает за масштаб, малый – за положение функции в пространстве. Увеличение индекса  $N$  на единицу сжимает функцию вдвое, увеличение индекса  $n$  на единицу сдвигает функцию вдоль оси  $x$  на величину  $h_N$ .

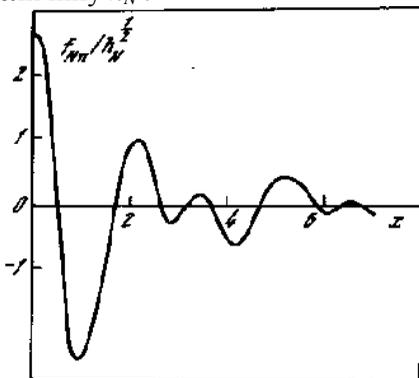


Рисунок 5.13 – Вид функции (5.74) для  $n=0$

### ***Двумерный базис***

Простейший способ получения двумерного базиса состоит в определении двумерной функции как произведения одномерных

$$f_{NnMm}(x, y) = f_{Nn}(x) f_{Mm}(y),$$

однако такие функции не являются изотропными и не удовлетворяют требованию подобия. Последнее обстоятельство не

оставляет надежд на получение простой динамической системы для коэффициентов разложения.

Исходя из локальной изотропии мелкомасштабной турбулентности и стремления получить базис, образованный разномасштабными, но однотипными функциями, можно построить относительно простой, но "не совсем ортогональный" базис. Для этого поле скорости представляется в виде

$$v(t, x, y) = \sum_{Nn} a_{Nn}(t) v_{Nn}(r - r_{Nn}),$$

где  $a_{Nn}$  – зависящая от времени амплитуда,  $v_{Nn}$  – осесимметричная базисная функция, у которой большой индекс отвечает за масштаб, а малый – за положение в пространстве, и  $r_{Nn}$  – радиус-вектор центра функции.

Используем введенное выше разбиение спектральной плоскости на расширяющиеся кольцевые зоны (6.68) и определим базисную функцию так, чтобы ее фурье-образ был равен константе в пределах соответствующего кольца:

$$v_{Nn}(\gamma) = \begin{cases} \frac{i(\gamma \times e)}{\sqrt{3\pi^3} \gamma^2} 2^{-N} \exp(-2\pi i \gamma r_{Nn}) & \gamma_N < |\gamma| < \gamma_{N+1} \\ 0 & \text{вне зоны} \end{cases} \quad (5.75)$$

где  $e$  есть единичный вектор, перпендикулярный рассматриваемой плоскости. Экспоненциальный множитель задает сдвиг центра вихря в физическом пространстве. Числовой коэффициент выбран из условия нормировки.

Чтобы получить вид функции в физическом пространстве, нужно взять обратное преобразование Фурье от (6.75). Соответствующие вычисления дают

$$v_{Nn}(r - r_{Nn}) = \frac{2^{-N} (s \times e)}{\sqrt{3\pi}} \left( \frac{J_0(2s) - J_0(s)}{s} \right) \quad (5.76)$$

где  $s = \pi 2^N |r - r_{Nn}|$ , а  $J_0(s)$  есть функция Бесселя.

Мы оставили без внимания вопрос о количестве базисных функций и об их распределении в пространстве. Плотность функций в физическом пространстве можно оценить исходя из принципа неопределенности. Если области локализации в  $r$  и  $k$

пространства имеют, соответственно, размеры  $\Delta r$  и  $\Delta k$ , то, требуя  $\Delta r \Delta k = 2\pi$ , получаем, что плотность функций заданного масштаба  $r_N$  связана с площадью области локализации функции в пространстве Фурье  $\Delta S_k$  как

$$\rho_N = \frac{\Delta S_k}{4\pi^2} = \frac{3\pi}{4} 2^{2N} \quad (5.77)$$

При вычислении (6.77) учли, что  $\Delta S_k$  есть площадь кольцевой области (6.68). Формула (6.77) отражает тот факт, что число вихрей при переходе от масштаба к масштабу растет в четыре раза (естественно, что в трехмерном случае это отношение будет равно восьми).

Вопрос о распределении функций в пространстве более сложен. Формулируя требования к базису, мы хотели воспроизвести структуру турбулентного потока, в котором мелкие вихри переносятся крупными. Это означает, что радиус-вектор центра функции должен подчиняться уравнению

$$d_t r_{Nn} = \sum_{M < N} \sum_m a_{Mm} v_{Mm} (r_{Nn} - r_{Mm}) \quad (5.78)$$

Подчеркнем, что суммирование в (5.78) ведется только по масштабам, большим данного.

Введенный таким образом базис ортогонален по индексу  $N$ , так как в фурье-пространстве функции различного масштаба занимают неперекрывающиеся области. Неортогональность функций по малому индексу, отвечающему за положение вихрей в пространстве, можно оценить путем вычисления интеграла  $\int v_{Nn} v_{Mm} dr$  для двух вихрей одного масштаба, расположенных друг от друга на расстоянии  $\rho_N^{-1/2}$ , равном среднему расстоянию между вихрями данного масштаба. Такая оценка дает для функций (6.76) значение порядка 0,1.

## 5.14 Вейвлеты

Ошибки в исследованиях обычно проистекают из свойственной человеческому

разуму склонности недооценивать или же преувеличивать значение исследуемого предмета из-за неверного определения его удаленности от нас.

Э.По

В самых разных областях науки возникают задачи, связанные с анализом пространственных полей со сложной, многомасштабной структурой либо временных сигналов с меняющимся со временем спектральным составом. Эти задачи заставляли исследователей делать попытки построения специальных функциональных разложений, близких по своей идеологии описанному выше иерархическому базису. Центральной идеей всех этих подходов было использование базиса, каждая функция которого характеризует как определенную пространственную (временную) частоту, так и место ее локализации в физическом пространстве (во времени).

Слово "вейвлет" (английское слово "wavelet" означает маленькую волну или рябь) было введено А.Гроссманном и Ж.Морле в 1984 году в работе, выполненной в связи с проблемой анализа сейсмических сигналов, в которых требуется выделить и время (положение) всплеска в сигнале, и его спектральный состав (масштаб). В этой статье были сформулированы основные определения и доказаны основополагающие теоремы. Работа вызвала огромный интерес и уже к началу 90-х годов вейвлет-анализ превратился в развитую область математической физики, нашедшей широкое применение в задачах анализа временных сигналов, распознавания образов и синтеза изображений, шифровки и дешифровки информации и многих других.

Как уже отмечалось, вейвлеты используются как при анализе временных сигналов, так и при исследовании структуры пространственных полей. Временные ряды представляют собой одномерный сигнал и все основные идеи проще продемонстрировать на задачах анализа временных последовательностей. По этой причине мы забудем на некоторое время о пространственных полях и переключимся на сигналы вида  $f(t)$ .

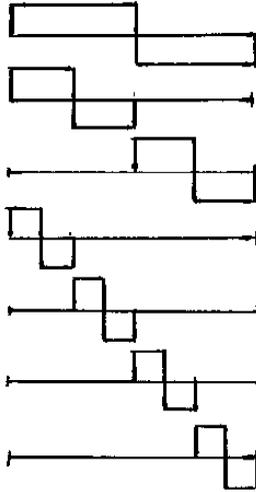


Рисунок 5.14 – Функции Хаара

Первая попытка построить функциональный базис, состоящий из функций, каждая из которых характеризует пульсации определенной продолжительности в определенный момент времени, принадлежит А.Хаару (1909г.). Первые семь функций Хаара, построенные на единичном отрезке, показаны на рисунке 5.14. Каждая функция представляет собой пару следующих друг за другом прямоугольных импульсов с разными знаками и одинаковой длительностью. Среднее значение любой функции равно нулю, а совокупность функций образует полный ортонормированный базис. Каждая функция строго локализована в физическом пространстве (во времени), но характеризуется медленно спадающим спектром частот (как  $1/n$ ).

Следующим шагом стали функции Литлвуда-Пелли (1937г.). Именно это семейство функций получается при построении одномерного иерархического базиса. Функции строятся путем вырезания полосы частот в пространстве Фурье. Это дает строгую локализацию в пространстве частот, но медленное затухание функции в физическом пространстве (во времени): функции описывают осцилляции, амплитуда которых падает как  $1/t$ .

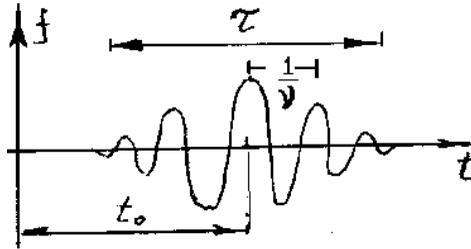


Рисунок 5.15 – Функция Габора

Важным этапом в развитии идеи локального анализа спектральных (частотных) свойств стало преобразование Габора (1946г.), называемое также фурье-преобразованием в окнах. Функции Габора представляют собой гармонический сигнал, модулированный функцией Гаусса. Они хорошо локализованы и во времени и в частотах, но каждая функция Габора характеризуется тремя параметрами: положением центра окна  $t_0$ , шириной окна  $t$  и частотой осцилляций  $\nu$  (рисунок 5.15). При этом функции различного масштаба не являются подобными (имеют различное число осцилляций).

Вейвлеты объединили в себе два важных свойства – подобие и выраженную локализованность в физическом и фурье-пространствах. Сформулируем требования, которым должно удовлетворять семейство функций, чтобы быть вейвлетами.

1. *Допустимость.* Функция  $y(t)$ , которую будем называть *анализирующим вейвлетом* (употребляют также термин *материнский вейвлет*), должна иметь нулевое среднее значение:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (5.79)$$

2. *Подобие.* Все функции семейства получаются из анализирующего вейвлета путем масштабного преобразования и сдвига,

$$\psi_{ab}(t) = \psi\left(\frac{t-b}{a}\right).$$

Таким образом, вейвлеты образуют двухпараметрическое семейство функций, в котором параметр  $a$  отвечает за мас-

штаб (растяжение) функции, а параметр  $b$  – за ее положение (сдвиг).

3. Обратимость. Вейвлет-преобразование должно быть обратимо, то есть должно существовать обратное преобразование, однозначно восстанавливающее исходную функцию по ее вейвлет-представлению.
4. Регулярность. Функция  $y(t)$  должна быть хорошо локализована и в физическом пространстве, и в пространстве Фурье.

Согласно последнему требованию и функции Хаара, и функции Литлвуда-Пелли не попадают под определение вейвлетов. По сути, они являют собой два предельных случая (в одном случае резкие границы в физическом пространстве приводят к бесконечным "хвостам" в пространстве частот и, наоборот, обрыв в пространстве частот дает бесконечные "хвосты" в физическом пространстве).

В отличие от преобразования Фурье, вейвлет-преобразование допускает широкий выбор анализирующей функции. Согласно первому требованию, вейвлет всегда является знакопеременной функцией, включающей обычно небольшое количество осцилляций. Выбор конкретного вида вейвлета зависит от целей проводимого анализа.

Приведем несколько примеров широко используемых вейвлетов. Простым вещественным вейвлетом, широко используемым в задачах, требующих хорошего пространственного разрешения и не требовательных к спектральному разрешению, является вейвлет, получивший название "мексиканская шляпа" (рисунке 5.16, а),

$$\psi(t) = (1-t^2)e^{-t^2/2} \quad (5.80)$$

В задачах, требующих лучшего спектрального разрешения, часто используется вейвлет Морле – комплексная функция вида

$$\psi(t) = e^{-t^2/2} e^{ia_0 t} \quad (5.81)$$

На рис.6.16,б сплошной линией показана его вещественная часть, а пунктирной – мнимая. Сама функция (6.81) совпадает с видом функций, используемых в преобразовании Габора, но семейство вейвлетов отличается от функций Габора тем, что один

раз выбрав частоту  $\omega_0$  для анализирующего вейвлета и задав тем самым число осцилляций, мы в дальнейшем сжимаем или растягиваем функцию как целое, не нарушая подобия отдельных функций семейства.

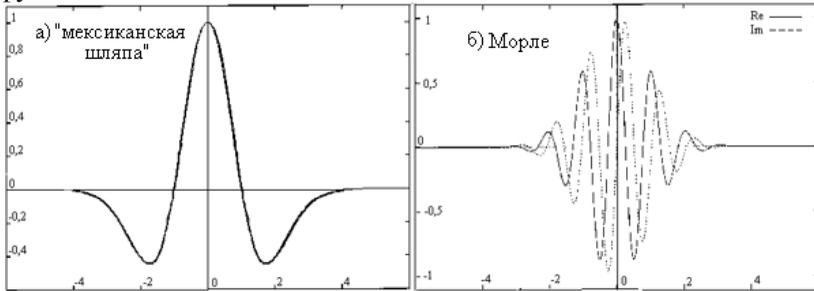


Рисунок 5.16 – Графики вейвлетов а) "мексиканская шляпа"; б) Морле

### ***Непрерывное вейвлет-преобразование***

Непрерывное вейвлет-преобразование одномерной функции  $f(t)$  есть

$$w(a,b) = a^\kappa \int_{-\infty}^{\infty} f(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (5.82)$$

где  $\psi(t)$  – вещественная или комплексная функция, удовлетворяющая требованиям 1-4 (знак  $*$  вводится для комплексно сопряженных величин). Если  $\hat{\psi}(\omega) = \int \psi(t) e^{-i\omega t} dt$  есть фурье-образ анализирующего вейвлета и выполнено условие

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (5.83)$$

то для преобразования (5.82) существует формула обращения

$$f(t) = \frac{1}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} \psi \left( \frac{t-b}{a} \right) w(a,b) \frac{dadb}{a^{3+\kappa}} \quad (5.84)$$

Условие (5.83) эквивалентно условию (5.79), так как интеграл (5.83) расходится при наличии в спектре вейвлета нулевых

частот, что равносильно отличному от нуля среднему значению. В определении (5.82) присутствует параметр  $k$  – показатель степени масштабного множителя. Конкретный выбор этого параметра зависит от целей анализа. Широко используется нормировка  $k=-1$ , при которой равные значения вейвлет-коэффициентов  $w(a,b)$  соответствуют равным амплитудам пульсаций сигнала, независимо от масштаба пульсаций.

Вейвлет-образ  $w(a,b)$  функции  $f(t)$  можно выразить и через ее фурье-образ  $\hat{f}(\omega)$ . Действительно,

$$\hat{f}(\omega) = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty \hat{\psi}(a, \omega) w(a, b) e^{-i\omega b} \frac{da db}{a^{2\kappa}} \quad (5.85)$$

а

$$w(a, b) = \frac{a^{\kappa+1}}{2\pi} \int_{-\infty}^\infty \hat{\psi}^*(a, \omega) \hat{f}(\omega) e^{i\omega b} d\omega \quad (5.86)$$

Пользуясь соотношениями (5.85)-(5.86) и теоремой Парсеваля, несложно получить аналог этой теоремы для вейвлет-преобразования

$$\int_{-\infty}^\infty f_1(t) f_2^*(t) dt = \frac{1}{C_\psi^2} \int_0^\infty \int_{-\infty}^\infty w_1(a, b) w_2^*(a, b) \frac{da db}{a^{3+2\kappa}},$$

из которого, в частности, следует

$$\int_{-\infty}^\infty |f(t)|^2 dt = \frac{1}{4\pi^2} \int_{-\infty}^\infty |\hat{f}(\omega)|^2 d\omega = \frac{1}{C_\psi^2} \int_0^\infty \int_{-\infty}^\infty |w(a, b)|^2 \frac{da db}{a^{3+2\kappa}} \quad (5.87)$$

Напомним, что в фурье-анализе спектральной плотностью энергии является величина  $E(\omega) = |\hat{f}(\omega)|^2$  (называемая также спектром энергии) и введем величину

$$M(a) = \int_{-\infty}^\infty |w(a, b)|^2 db,$$

которая характеризует интенсивность *всех* пульсаций заданного масштаба. Если в определении вейвлет-преобразования положить  $k = -1/2$ , то формулу (5.87) можно переписать в виде

$$E = \int_0^{\infty} E(\omega) d\omega = \int_0^{\infty} M(a) \frac{da}{a^2} \quad (5.88)$$

В этом случае  $M(a)$  описывает распределение энергии пульсаций по масштабам и называется интегральным вейвлет-спектром. Из сказанного следует, что нормировка  $k = -1/2$  должна использоваться, если результаты вейвлет-анализа предполагается сопоставлять с фурье-представлением сигнала. Действительно, если фурье-спектр следует степенному закону  $E(\omega) \propto \omega^a$ , то при этой нормировке интегральный вейвлет-спектр будет иметь тот же степенной закон  $M(a) \propto a^{-a} \propto \omega^{-a}$  (это следует из формулы (5.88) с учетом того, что  $\omega \propto 1/a$ , а  $d\omega \propto -da/a^2$ ).

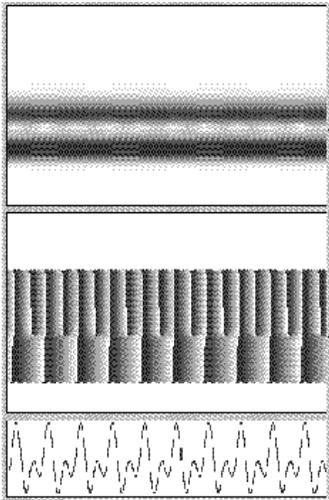


Рисунок 5.17 – Вейвлет-разложение Морле сигнала, изображенного в нижней части рисунка

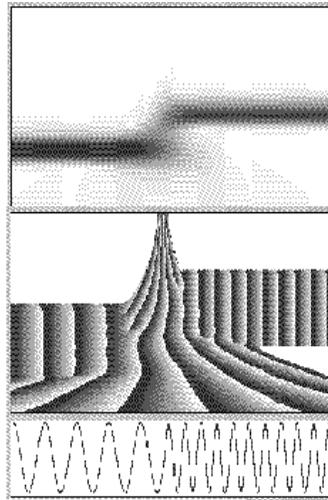


Рисунок 5.18 – Вейвлет-разложение Морле сигнала, изображенного в нижней части рисунка

Вейвлет-преобразование отображает пространство функций одной переменной (время) в пространство функций двух переменных (время и частота, или время и масштаб) и является избыточным. Избыточность непрерывного вейвлет-преобразования выражается в коррелированности вейвлет-коэффициентов, которая тем больше, чем больше рассматриваемый масштаб  $a$ . Иначе говоря, чем больше масштаб, тем меньше независимых точек в вейвлет-разложении. Этот недостаток устраняется в дискретном вейвлет-представлении (пример тому – рассмотренный выше иерархический базис, в котором число функций геометрически уменьшается с ростом пространственного масштаба).

Преимущество вейвлет-преобразования перед преобразованием Фурье состоит в том, что оно позволяет проследить за изменением спектральных свойств сигнала со временем, указать, какие частоты (масштабы) доминируют в сигнале в каждый конкретный момент времени. На рисунках 5.17 и 5.18 показаны два примера вейвлет-разложения простых временных сигналов с помощью вейвлета Морле (5.81). В верхней части каждого рисунка показан модуль вейвлет-разложения на плоскости  $(a, b)$ , а в нижней – фаза. На рисунке 5.17 сигнал представляет собой суперпозицию двух гармоник, а в сигнале на рисунке 5.18 эти же две частоты появляются последовательно друг за другом. Фурье-преобразования этих двух сигналов практически не отличаются друг от друга, так как спектр Фурье теряет всякую информацию о том, когда какая гармоника присутствовала в сигнале. Вейвлет-анализ позволяет восстановить полную эволюцию спектрального состава сигнала во времени. Общее представление о спектрально-временной структуре сигнала можно получить по распределению модуля вейвлет-преобразования. Ширина полосы, получаемой при разложении гармонического сигнала, характеризует спектральное разрешение используемого анализирующего вейвлета. Распределение фазы вейвлет-преобразование менее информативно, особенно для сложных сигналов. В то же самое время, именно фаза дает наиболее точную информацию об особенностях (сингулярностях) в сигнале. Так, на рисунке 5.18 можно видеть, что именно по распределению фазы можно с большой точностью идентифицировать момент смены частоты.

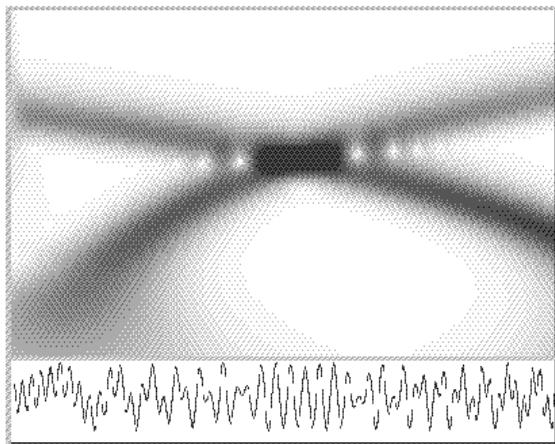


Рисунок 5.19 – Модуль вейвлет-разложения Морле сигнала, график которого приведен в нижней части рисунка

На рисунке 5.19 показан результат вейвлет-разложения сигнала, представляющего собой суперпозицию двух гармонических составляющих с непрерывно меняющимися частотами (снова использован вейвлет Морле). Сам сигнал показан в нижней части рисунка, модуль вейвлет-разложения – в верхней части. Вейвлет-представление позволяет получить точный вид эволюции частоты каждого из двух сигналов.

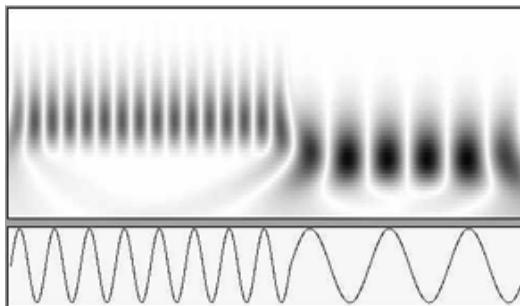


Рисунок 5.20 – Вейвлет-разложение сигнала, показанного в нижней части рисунка, с использованием вейвлета "мексиканская шляпа"

На рисунке 5.20 дан пример использования действительного вейвлета типа (5.80). В качестве сигнала использован тот же временной ряд, что и в примере на рисунке 5.18 (удвоение частоты гармонических колебаний). В этом случае результатом преобразования является действительная величина, модуль которой показан на рисунке. Белые полосы на вейвлет-плоскости, неизбежно появляющиеся при работе с вещественными функциями, соответствуют смене знака вейвлет-коэффициентов и содержат, по сути, информацию, которую в комплексном представлении несет фаза.

В заключение отметим важное свойство вейвлет-представления функций, состоящее в том, что на этапе разложения сигнала по вейвлетам (анализа) и этапе восстановления исходного сигнала по его вейвлет-образу (синтеза) можно использовать различные семейства вейвлетов. Пусть для анализа используется вейвлет  $\psi(t)$ , а для синтеза – вейвлет  $\varphi(t)$ . Тогда прямое преобразование по-прежнему описывается выражением (5.82), а формула восстановления сигнала (5.84) примет вид

$$f(t) \frac{1}{C_{\psi\varphi}} \int_0^\infty \int_{-\infty}^\infty \psi\left(\frac{t-b}{a}\right) w(a,b) \frac{dad b}{a^{3+\kappa}} \quad (5.89)$$

Восстановление (5.89) возможно, если выполнено условие

$$C_{\psi\varphi} = \int_{-\infty}^\infty \frac{|\hat{\psi}(\omega)\hat{\psi}^*(\omega)|}{|\omega|} d\omega < \infty.$$

Это условие "мягче", чем условие (5.83), так как теперь один из двух вейвлетов может и не удовлетворять требованию (5.83) (но при условии, что его "недостатки" компенсирует вейвлет, используемый на втором этапе). Преимущество восстановления по формуле (5.89) состоит в том, что она позволяет использовать на одном из этапов преобразования сингулярную функцию (например,  $\delta$ -функцию), которая сама по себе не попадает под определение вейвлета.

## Список использованной литературы

1. Абловиц М., Сигур Х. Солитоны и метод обратной задачи. — М.: Мир, 1987. 479с.
2. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. — М.: Финансы и статистика, 1983. 471с.
3. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. — М.: Мир, 1987. 358с.
4. Алкок Д. Язык Паскаль в иллюстрациях. — М.: Мир, 1991. 192с.
5. Арнольд В.И. Лекции об уравнениях с частными производными. — М.: Фазис, 1997. 180с.
6. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы: учебное пособие. — М.: Наука, 1987. 600с.
7. Белл Дж.Ф. Экспериментальные основы механики деформируемых тел. В 2-х частях. Часть I. Малые деформации. — М.: Наука-Физматлит, 1984. 600с.
8. Бендат Дж., Пирсол А. Прикладной анализ случайных данных — М.: Мир, 1989. 540с.
9. Биркгоф Г. Математика и психология. М.: Советское радио. 1977. 95с.
10. Блехман И.И., Мышкис А.Д., Пановко Я.Г. Прикладная математика: предмет, логика, особенности подходов. — Киев: Наукова думка, 1976. 270с.
11. Боровков А.А. Теория вероятностей. — М.: Наука, 1986. 432с.
12. Бурсиан Э.В. Задачи по физике для компьютера. — М.: Просвещение, 1991. 256с.
13. Бусленко В.Н. Автоматизация имитационного моделирования сложных систем — М.: Наука, 1977. 240с.
14. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. — М.: Наука, 1988. 480с.
15. Глинский Б.А., Грязнов Б.С. и др. Моделирование как метод научного исследования. — М.: Наука, 1965. 245с.
16. Глушков В.М., Иванов А.С., Яненко К.А. Моделирование развивающихся систем. — М.: Наука, 1983. 276с.

17. Гмурман В.Е. Теория вероятностей и математическая статистика. — М.: Высшая школа, 1977. 479с.
18. Годунов С.К. Уравнения математической физики. — М.: МГУ, 1971. 416с.
19. Горстко А.Б. Познакомьтесь с математическим моделированием. — М.: Знание, 1991. 160с.
20. Грогано П. Программирование на языке Паскаль. — М.: Мир, 1982. 260с.
21. Гулд Х., Тобочник Я. Компьютерное моделирование в физике: В 2-х частях. Часть 2 — М.: Мир, 1990. 400с.
22. Ивченко Г.И., Медведев Ю.И. Математическая статистика. — М.: Высшая школа, 1984. 248с.
23. Калиткин Н.Н. Численные методы. — М.: Наука, 1978. 512с.
24. Клейнен Дж. Статистические методы в имитационном моделировании /Пер. с англ. Ю.П.Адлера и др. Вып. 2 — М.: Статистика, 1978. 335с.
25. Кнут Д. Искусство программирования для ЭВМ, т.2: Получисленные алгоритмы. — М.: Мир, 1977. 370с.
26. Корн Г., Корн Т. Справочник по математике (для научных работников и инженеров). — М.: Наука, 1978. 832с.
27. Круг Г.К., Сосулин Ю.А., Фатуев В.А. Планирование эксперимента в задачах идентификации и экстраполяции. — М.: Наука, 1977. 298с.
28. Крылов В.И., Бобков В.В., Монастырный П.И. Вычислительные методы. Т.2. — М.: Наука, 1977. 400с.
29. Лекции по теории графов /Емеличев В.А. и др. — М.: Наука, 1990. 384с.
30. Лихачев В.А., Малинин В.Г. Структурно-аналитическая теория прочности. — С-Пт.: Наука, 1983. 471с.
31. Майника Э. Алгоритмы оптимизации на сетях и графах. — М.: Мир, 1981. 323с.
32. Максимей И.В. Имитационное моделирование на ЭВМ — М.: Радио и связь, 1988. — 232 с.
33. Мануйлов В.Г. Разработка программного обеспечения на Паскале. — М.: “ПРИОР”, 1996. 240с.

34. Марчук Г.И. Методы вычислительной математики. — М.: Наука, 1980. 536с.
35. Математическое моделирование /Под ред. Дж.Эндрюса и Р.Мак-Лоуна. — М.: Мир, 1979. 250с.
36. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. — М.: Высшая школа, 1989. 367с.
37. Розанов Ю.А. Теория вероятностей, случайные процессы и математическая статистика. — М.: Наука, 1989. 320с.
38. Розенблют А., Винер Н. Роль моделей в науке. В кн. Модели в науке и технике. — Л.: 1984, С.171-175.
39. Сабоннадьер Ж.-К., Кулон Ж.-Л. Метод конечных элементов и САПР. — М.: Мир, 1989. 190с.
40. Самарский А.А. Введение в численные методы. — М.: Наука, 1982. 272с.
41. Самарский А.А., Гулин А.В. Численные методы: Учеб. пособие для вузов. — М.: Наука, 1989. 432с.
42. Татт У. Теория графов. — М.: Мир, 1988. 424с.
43. Тоффоли Т., Марголюс Н. Машины клеточных автоматов. — М.: Мир, 1991. 280с.
44. Уизем Дж. Линейные и нелинейные волны. — М.: Мир, 1977. 622с.
45. Хемминг Р.В. Численные методы для научных работников и инженеров. — М.: Наука, 1968. 400с.
46. Шеннон Р. Имитационное моделирование систем — Искусство и наука. — М.: Мир, 1978. 417с.
47. Шеффе Г. Дисперсионный анализ. — М.: Наука, 1980. 512с.