

**МЕТОДЫ СИСТЕМНОГО АНАЛИЗА  
И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ  
В ЗАДАЧАХ ПРОЕКТИРОВАНИЯ  
ЛЕТАТЕЛЬНЫХ АППАРАТОВ**

**2007**



**САМАРА**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ  
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ  
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЕВА»

МЕТОДЫ СИСТЕМНОГО АНАЛИЗА  
И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ  
В ЗАДАЧАХ ПРОЕКТИРОВАНИЯ  
ЛЕТАТЕЛЬНЫХ АППАРАТОВ

*Утверждено Редакционно-издательским советом университета  
в качестве учебного пособия*

САМАРА  
Издательство СГАУ  
2007

УДК 629.7.001 (075)+519.6(075)  
ББК 68.53+22.161  
М 545



**Инновационная образовательная программа  
"Развитие центра компетенции и подготовка  
специалистов мирового уровня в области аэро-  
космических и геоинформационных технологий"**

Авторы: ***В. В. Салмин, А. С. Кучеров,  
О. Л. Старинова, А. Г. Прохоров***

Рецензенты: д-р техн. наук, проф. Ю. Н. Г о р е л о в  
канд. техн. наук, доц. Е. А. Е с и п о в

М 545 **Методы системного анализа и исследования операций в  
задачах проектирования летательных аппаратов:** учеб. пособие  
/ [В. В. Салмин и др.]. – Самара: Изд-во Самар. гос. аэрокосм.  
ун-та, 2007. – 272 с. : ил.

**ISBN 978-5-7883-0681-0**

Излагаются основы системного анализа и исследования операций. Описана общая методология и основные методы решения проектных задач на основе системного подхода. В качестве примеров приведены решения прикладных задач: синтеза структуры и параметров космической системы наблюдения; системного проектирования низкоорбитальных космических аппаратов, функционирующих в условиях неточного знания силы аэродинамического сопротивления на больших высотах; синтеза парето–оптимальных комбинированных схем выведения на геостационарную орбиту; многокритериальной оптимизации баллистических схем межпланетных перелетов с малой тягой.

Предназначено для студентов, обучающихся по специальностям: «Моделирование и исследование операций в организационно-технических системах», «Космические летательные аппараты и разгонные блоки» и «Ракетостроение». Пособие представляет интерес для инженеров, аспирантов и научных сотрудников, специализирующихся в области системного проектирования и математического моделирования летательных аппаратов.

УДК 629.7.001 (075)+519.6(075)  
ББК 68.53+22.161

**ISBN 978-5-7883-0681-0**

© Салмин В.В., Кучеров А.С.,  
Старинова О.Л., Прохоров А.Г., 2007  
© Самарский государственный  
аэрокосмический университет, 2007

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ	10
1. ОСНОВЫ СИСТЕМНОГО АНАЛИЗА	12
1.1. Системный подход в проектировании	12
1.1.1. Классификация систем	12
1.1.2. Структурные связи, структура и функция системы	15
1.1.3. Цели искусственной системы	18
1.1.4. Структурные уровни	19
1.1.5. Построение модели функциональной структуры	21
1.1.6. Модель «черного ящика»	22
1.1.7. Функционирование и развитие системы	24
1.1.8. Сложная техническая система	26
1.1.9. Жизненный цикл сложной технической системы	29
1.1.10. Проектирование, анализ и синтез сложной технической системы	32
1.1.11. Основные этапы исследования систем путем математического моделирования	35
1.2. Методы анализа и синтеза сложных технических систем	38
1.2.1. Значение и особенности процесса постановки задачи	38
1.2.2. Качество и эффективность функционирования системы	40
1.2.3. Надежность системы	42
1.2.4. Затраты ресурсов на создание и функционирование системы	43
1.2.5. Показатели функционирования системы в условиях неопределенности	44
1.2.6. Проблема многокритериальной оптимизации	46
1.2.7. Построение множества Парето для сужения множества альтернатив	53

1.3. Модели синтеза и анализа сложной технической системы	59
1.3.1. Структура и состав модели	59
1.3.2. Уточнение цели системы в процессе ее исследования	63
1.3.3. Адекватность и работоспособность модели	66
1.3.4. Определяющие параметры и объект моделирования	69
1.3.5. Принцип декомпозиции - субоптимизации	70
1.3.6. Разновидности неопределенностей	73
1.3.6. Функциональные требования и программные продукты	75
1.4. Методы неформализованного анализа	79
1.4.1. Классификационные шкалы	80
1.4.2. Порядковые шкалы	82
1.4.3. Модифицированные порядковые шкалы	83
1.4.4. Интервальные шкалы	84
1.4.5. Периодические или циклические шкалы	85
1.4.6. Абсолютная шкала	85
1.4.7. Шкалы отношений	85
1.4.8. Качественные или не метрические показатели	86
1.4.9. Задача количественной оценки качественных показателей	87
1.4.10. Балльные шкалы	88
1.4.11. Проблема перехода от балльной шкалы к количественной шкале	90
1.4.12. Организация коллективных экспертных оценок	91
1.4.13. Метод мозговой атаки	93
1.4.14. Формирование альтернатив	96
1.4.15. Морфологический анализ и синтез	98
2. МЕТОДЫ И ЗАДАЧИ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ	100

2.1. Основные понятия исследования операций	100
2.1.1. Предмет исследования операций и история становления	100
2.1.2. Основные определения	101
2.1.3. Методология исследования операций	102
2.1.4. Методы моделирования	103
2.1.5. Классификация моделей ИСО	104
2.1.6. Требования к критериям	105
2.1.7. Общая форма критериев. Примеры критериев	106
2.1.8. Типичные классы задач исследования операций	107
2.2. Линейное программирование	109
2.3. Целочисленное программирование	113
2.3.1. Метод отсечения (метод Гомори)	114
2.3.2. Метод ветвей и границ	115
2.4. Динамическое программирование	118
2.4.1. Сущность метода динамического программирования	118
2.4.2. Примеры задач динамического программирования	122
2.5. Метод статистического моделирования	128
2.5.1. Единичный жребий	130
2.5.2. Моделирование работы технического устройства	134
2.6. Использование марковских случайных процессов в исследовании операций	138
2.6.1. Марковские случайные процессы	138
2.6.2. Потоки событий	141
2.6.3. Пуассоновские потоки событий и непрерывные марковские цепи	144
2.6.4. Процесс «гибели и размножения»	146
2.6.5. Циклический процесс	148

2.7. Теория массового обслуживания	150
2.7.1. Классификация систем массового обслуживания	151
2.7.2. Одноканальная система массового обслуживания с отказами	152
2.7.3. Многоканальная система массового обслуживания с отказами	154
2.7.4. Одноканальная система массового обслуживания с очередью	156
2.7.5. Многоканальная система массового обслуживания с очередью	159
2.8. Метод динамики средних	162
2.8.1. Идея метода. Область применения	162
2.8.2. Учет зависимости интенсивностей потоков событий от численности состояний. Принцип квазирегулярности	165
2.9. Задачи принятия решений в условиях риска	167
2.9.1. Одноэтапные процедуры	167
2.9.2. Многоэтапные процедуры	170
2.10. Задачи принятия решений в условиях неопределенности	172
2.11. Основы игровых методов принятия решений	175
2.11.1. Классификация игр	176
2.11.2. Матричные игры двух игроков с нулевой суммой. Решение в чистых стратегиях	178
2.11.3. Матричные игры двух игроков с нулевой суммой. Решение в смешанных стратегиях	180
2.11.4. Игры с платежной матрицей размером $2 \times 2$	181
2.11.5. Матричные игры с платежной матрицей $2 \times n$ и $m \times 2$	183
2.11.6. Свойства смешанных стратегий	186
2.11.7. Решение матричных игр методом линейного программирования	188
2.11.8. Физическая смесь стратегий	189

3. ПРИКЛАДНЫЕ ЗАДАЧИ СИСТЕМНОГО АНАЛИЗА И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ	191
3.1. Проектирование космической системы наблюдения	191
3.1.1. Постановка расширенной задачи	191
3.1.2. Модель функционирования орбитальной группировки	193
3.1.3. Проектирование КСН оптимальной с точки зрения наилучшего показателя периодичности наблюдения	205
3.1.4. Пример решения задачи синтеза орбитальной структуры	207
3.2. Анализ эффективности применения электроракетной двигательной установки для коррекции низких орбит искусственных спутников Земли	209
3.2.1. Коррекция низких орбит космических аппаратов электроракетным двигателем	209
3.2.2. Постановка задачи многокритериального синтеза	212
3.2.3. Процедура совместной оптимизации законов управления и проектных параметров КА	213
3.2.4. Учет неопределенности возмущающих факторов	215
3.2.5. Модель коррекции орбиты спутника Земли без учета ограничений по энергетике	219
3.2.6. Влияние ограничений по энергетике на формирование циклограмм коррекции орбиты	222
3.2.7. Математическая модель коррекции орбиты низкоорбитального спутника Земли с учетом неопределенности плотности верхней атмосферы и ограничений по энергетике	225
3.2.8. Определение гарантирующих затрат рабочего тела на поддержание орбиты	226
3.3. Синтез Парето-оптимальных комбинированных схем выведения полезных нагрузок на геостационарную орбиту	235
3.3.1. Постановка задачи	235
3.3.2. Проектная модель космической транспортной системы	237



3.3.3. Модель управления движением космической транспортной системой	239
3.3.4. Результаты моделирования перелета на ГСО	240
3.3.5. Построение множества Парето	243
3.3.6. Формирование проектного облика МТА с использованием системы твердотельного моделирования Solid Works	245
3.4. Синтез баллистических схем межпланетных перелетов КА с ЭРДУ на основе многокритериальной оптимизации	248
3.4.1. Постановка задачи проектно-баллистической оптимизации межпланетных перелетов	248
3.4.2. Массовая и проектная модель межпланетного КА с малой тягой	251
3.4.3. Модель оценки затрат на осуществление проекта	252
3.4.4. Модели движения, используемые при оптимизации межпланетных перелетов с малой тягой	253
3.4.5. Результаты оптимизации межпланетной миссии к Марсу, осуществляемой КА с солнечным электроракетным двигателем	254
3.4.6. Результаты выбора баллистической схемы межпланетного перелета на основе многокритериальной оптимизации	262
СПИСОК ЛИТЕРАТУРЫ	270

## ПРЕДИСЛОВИЕ

Теория исследования операций как самостоятельная научная дисциплина сложилась в период второй мировой войны в результате обобщения новых подходов к оценке эффективности современных методов и средств ведения боевых действий. Позднее в рамках исследования операций возникла новая дисциплина – «системный анализ». Под системой вообще принято понимать совокупность составных взаимосвязанных частей, реализующих определенную цель. Однако, при анализе систем высокой сложности вообще трудно говорить о критериях и целях, не изучив предварительно основные свойства системы. А они заключаются в том, что взаимосвязанные компоненты системы функционируют согласованно, в результате чего у системы в целом появляются особые свойства, которых нет ни у одного из ее компонентов.

Среди всего многообразия систем особое место занимают сложные технические системы, представляющие собой совокупность объектов различной физической и технической природы, образующих различные иерархические уровни и объединенные единым управлением. В последние десятилетия возникла теория иерархических систем, теория имитационного моделирования, методы неформализованного анализа, которые нашли самое широкое применение в задачах проектирования различных технических систем. Новый импульс к развитию получила и дисциплина «исследование операций», под которой понимается применение математических, количественных методов для обоснования решений во всех областях целенаправленной человеческой деятельности.

В данном учебном пособии изложены основы системного анализа и исследование операций с приложениями к задачам проектирования летательных аппаратов. Вдохновляющим примером для авторов явились великолепные книги профессора А.А. Лебедева. С его согласия часть материала этих книг была включена в это пособие, предназначенное для студентов аэрокосмических специальностей.

Авторы глубоко признательны А.А. Лебедеву и кафедре «Системного анализа и управления» Московского авиационного института за постоянную поддержку и внимание к нашей работе. Авторы благодарят также рецензентов – профессора Ю.Н. Горелова, доцента Б.А. Есипова за их нелегкий труд и полезные замечания.

Авторы

## **ВВЕДЕНИЕ**

В последние годы большое развитие получила дисциплина «прикладная системология», предметом которой является искусство постановки и решения задач проектирования, разработки, создания и эксплуатации сложных технических и организационно-технических систем. Методической основой этой науки является системный анализ. Проектирование сложных систем должно проводиться с учетом функционального назначения системы, ее структуры и свойств подсистем, многообразия внешних и внутренних функциональных связей, эволюции структуры и характеристик системы во времени, управления с элементами самоорганизации, учитывающего априорную неопределенность ряда факторов. Использование традиционных приемов декомпозиции задачи синтеза помогает получить полезные решения с помощью методов исследования операций, методов оптимизации, однако только системный подход во многих случаях позволяет обоснованно решать сложные задачи проектирования и управления.

В этой связи дисциплины «системный анализ», «прикладная системология» и «исследование операций» оказываются тесно взаимосвязанными. Эта идея нашла свое отражение в структуре настоящего учебного пособия. Оно подготовлено на основе курсов лекций, читаемых авторами на протяжении многих лет для студентов специальностей «Космические летательные аппараты и разгонные блоки», «Моделирование и исследование операций в организационно-технических системах».

Основной целью пособия является ознакомление студентов с основными понятиями и методологией системного анализа, обучение методам решения различных задач исследования операций, ознакомление с содержательными примерами решения прикладных задач в области проектирования и математического моделирования летательных аппаратов.

Пособие состоит из трех разделов. В первом разделе излагаются основы системного анализа, системного подхода к проектированию. Описаны методы многокритериального синтеза сложных технических систем, место и роль моделей и моделирования, изложены методы неформализованного анализа.

Второй раздел «Методы и задачи исследования операций» содержит сведения об основах теории исследования операций, методы решения различных классов задач: линейное программирование, динамическое программирование, теорию массового обслуживания. Изложены также процедуры принятия решений в условиях риска, игровые задачи исследования операций.

В третьем разделе даются решения некоторых прикладных задач. К ним относятся: задача синтеза структуры и параметров космической системы глобального мониторинга Земли; задача проектирования системы поддержания низкой орбиты космического аппарата наблюдения; задача синтеза Парето –оптимальных комбинированных схем выведения полезных нагрузок на геостационарную орбиту; задача оптимизации баллистических схем полета к Марсу космического аппарата с электрореактивной двигательной установкой.

Настоящее пособие предназначено для инженеров, поэтому изложение теоретического материала дается в несколько упрощенной форме, без излишней перегруженности текста математическими выкладками. Пособие ориентировано на получение студентами навыков решения широкого класса прикладных задач. Читатели, желающие более глубоко ознакомиться с методами и задачами исследования операций, методологией системного анализа, могут воспользоваться книгами [1 - 11], указанными в списке литературы.

# 1. ОСНОВЫ СИСТЕМНОГО АНАЛИЗА

## 1.1. Системный подход в проектировании

### 1.1.1. Классификация систем

На рис. 1.1 дана двухуровневая классификация систем по их происхождению. Естественные системы возникли в природе без участия человека, искусственные созданы человеком, смешанные объединяют те и другие, а также созданные человеком людские коллективы.

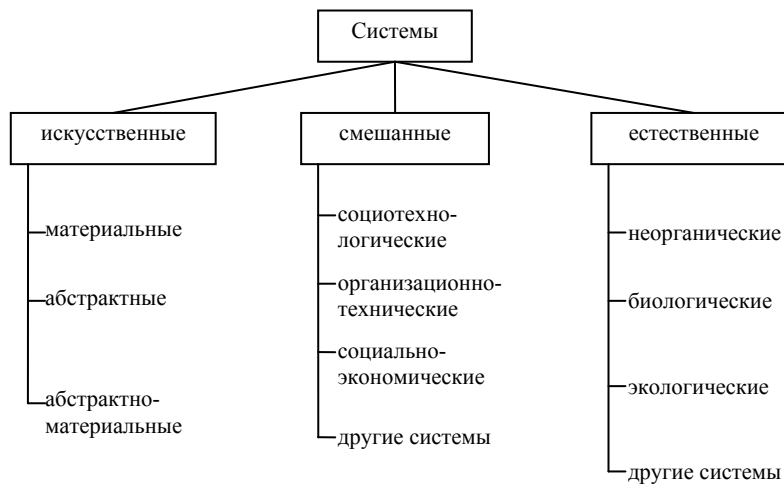


Рис. 1.1. Классификация систем по их происхождению

*Материальные искусственные системы* образованы физическими телами. К ним относятся, например, такие технические изделия, как орудия, механизмы, машины, автоматы, роботы, ЭВМ, ЛА и другие продукты практической деятельности человека.

*Абстрактные* (или идеальные, или символические) *искусственные системы* отражают реальную действительность и являются продуктом как

познавательной, так и практической деятельности человека. Сюда относятся научные знания, произведения литературы и искусства, языки, математические модели, алгоритмы и многое другое.

*Естественные системы* подразделяются на неорганические (неживые), например вулкан, Солнце; биологические (живые), например клетка, живое существо; экологические, в которых растительные и животные организмы взаимодействуют между собой и средой их обитания.

К естественным системам могут быть отнесены и социальные системы — системы совместной деятельности людей: общество в целом и его частные системы, например семья. Они исторически сложились как продукт естественной социальной эволюции.

*Смешанные системы*, как и искусственные, служат человеку средством для достижения его целей. Эти системы различаются по таким признакам: назначение системы; тип, структура и роль человеческого компонента; типы и структуры материальных компонентов, типы взаимодействия этих компонентов.

В результате практической деятельности человек создал разнообразные социотехнологические системы, которые представляют собой совокупность:

а) производственного и обслуживающего персонала, участвующего в функционировании системы и являющегося элементом социальной подсистемы общества;

б) технических средств, преобразующих материально-энергетические и информационные потоки, входящие в систему и выходящие из нее;

в) устройств и сооружений, предназначенных для обеспечения технологического процесса и условий труда персонала;

г) компонентов естественных систем (промысловая акватория для рыболовецкого предприятия, лесные участки для леспромхоза, недра Земли для добычи полезных ископаемых, биологический компонент для биотехнического завода).

Частным случаем является организационно-техническая система, состоящая из людей, рассматриваемых лишь в качестве операторов, и технических средств, предназначенных для реализации назначения системы. Различные автоматизированные системы представляют собой

организационно-технические системы.

САПР – это организационно-техническая система, которая состоит из комплекса средств автоматизации проектирования, взаимодействующего с подразделениями проектной организации, и выполняет автоматизированное проектирование.

Человеческий компонент *организационно-технической системы* является организационной системой, а технический компонент — *технической системой*.

По общему назначению организационно-технические системы разделяют на два подкласса: системы, производящие изделия или технические средства, — производственные системы (завод, производящий самолеты, авиационная промышленность) и системы, выполняющие услуги (самолет с экипажем, аэропорт, космическая система исследования природных ресурсов Земли, система продажи авиабилетов, САПР, системы транспорта и связи).

Подкласс смешанных систем — *социально-экономические системы*, возникающие в результате практической деятельности человека и действия объективных ("естественных") законов общественного развития. К ним, например, относятся государство, экономическая система страны.

Компоненты системы взаимодействуют не только между собой, но и с компонентами внешней среды. Например, система ориентации КА взаимодействует с КА и другими бортовыми системами (двигательной установкой, системой электропитания и др.), поэтому выделение некоторой части взаимосвязанных объектов реального мира в самостоятельный объект-систему должно быть как-то предусмотрено в определении системы. Так, формулировка "система есть отграниченное множество взаимодействующих элементов" частично устраняет указанную неполноту определения системы, но все же остается вопрос о границе "система – не система (внешняя среда)". Ответ на него содержится в первом определении: "система есть совокупность элементов, взаимосвязанных для достижения определенной цели". Другими словами, общая цель функционирования взаимодействующих элементов объединяет их и выделяет из внешней среды в качестве обособленной системы. Эту обособленность от окружающей среды можно назвать внешней целостностью системы, т.е. целостностью по отношению к внешней среде.

Все сказанное представляет собой первый подход к определению понятия системы.

Второй подход связан с представлением о внутренней нецелостности системы. Благодаря общей цели взаимосвязанные компоненты системы функционируют согласованно, в результате чего у системы в целом появляются особые свойства, которых нет ни у одного из ее компонентов, взятых в отдельности. Такие свойства называют системными.

Так, например, компоненты ЛА (самолета) создают силу тяги (двигательная установка), подъемную силу при движении относительно воздуха (крыло), отклоняют органы управления (пилот и система управления) и т.д. При совместном согласованном функционировании всех компонентов ЛА возникает новый эффект — управляемый полет ЛА, который не присущ ни одному компоненту ЛА в отдельности.

Как видно, внутренняя целостность системы проявляется в том, что свойства системы не являются суммой свойств ее компонентов и что при объединении компонентов в целое возникает качественно новое системное свойство, которого не было и не могло быть без этого объединения. Свойство целостности называют в литературе также интегрированным свойством и эмерджентностью (возникновением из ничего, внезапным появлением).

Теперь становится ясным второе определение системы. Система — объект любой природы, обладающий выраженным "системным" свойством (свойствами), т.е. свойством, которого не имеет ни одна из частей системы при любом способе ее членения.

Система, выделенная из окружающей ее среды, является обособленной лишь условно, так как всегда имеет место ее взаимодействие с внешней средой. Поэтому любое исследование системы будет бессмысленным без надлежащего учета ее взаимодействия с внешней средой. Так, например, при изучении полета крылатого ЛА в атмосфере Земли внешней средой для ЛА являются воздушная среда и гравитационное поле Земли. Они соответственно вызывают силы, действующие на ЛА: аэродинамическую силу и силу земного притяжения, без учета которых исследование полета лишено смысла.

Итак, любая система является частью некоторой более обширной системы. Приступая к исследованию или проектированию системы, прежде



всего необходимо выделить ее из окружающих систем, другими словами, определить систему, установив ее состав и границы. Если исследуемая или проектируемая система неправильно выделена из внешней среды, то может оказаться, что при исследовании будут сделаны неправильные выводы, а при проектировании будут приняты нерациональные, а в некоторых случаях и ошибочные решения.

Выделение системы из внешней среды является одним из важнейших элементов системного анализа.

### ***1.1.2. Структурные связи, структура и функция системы***

Среди множества связей в системе имеются связи, необходимые для функционирования системы в направлении достижения цели. Благодаря этим связям совокупность компонента приобретает системные свойства (свойство целостности). Такие связи называют структурными или системообразующими.

На систему действуют внутренние и внешние (со стороны среды) возмущения; они стремятся разрушить систему, препятствуют достижению ее цели. Благодаря структурным связям система сохраняет свои свойства и структуру при воздействии возмущений, изменении внешних и внутренних условий.

Совокупность составляющих систему компонентов и структурных связей будем называть структурой системы.

Компоненты, функционируя, реализуют физические операции, для чего создают потоки преобразуемых или превращаемых веществ, энергии, информации и других факторов. Такие потоки определенным образом объединяют и связывают компоненты и их физические операции. Они называются *потоковыми связями*.

Каждый компонент выполняет определенную физическую операцию. Совокупность компонентов (или физических операций) и потоковых связей между ними образует потоковую структуру.

Существуют два вида построения потоковых структур:

- 1) потоки связывают компоненты;
- 2) потоки связывают физические операции компонентов.

Каждый компонент в составе системы имеет свое назначение.

Свойство компонента исполнять свое назначение называется функцией.

Исполнение компонентом своей функции дает определенный результат вне этого компонента – в другом компоненте или в вышестоящей подсистеме.

Совокупность компонентов и функциональных связей между ними образует *функциональную структуру системы*.

Функция системы проявляется во внешней среде как целевой выход системы, как результат ее функционирования, как реализация цели системы. Носителем функции системы является ее структура.

При исследовании системы иногда целесообразно отвлечься от физической природы компонентов и рассмотреть укрупненные подсистемы, образованные по признаку общности частных задач, решаемых совокупностью технических компонентов, т.е. по функциональному признаку.

Совокупность подсистем, образованных по функциональному признаку, и связей между ними назовем *операционной структурой системы*.

Система определенным образом (целенаправленно) организована. Она объективно состоит из взаимосвязанных частей – компонентов. Каждый компонент является частью другого более обширного компонента, что означает иерархическое построение системы. Если компонент В входит в состав компонента А, то В подчиняется А, т.е. В является компонентом более низкого уровня иерархии, чем А. В результате членения системы на компоненты (декомпозиция системы) исследователь получает модель состава системы, отображающую, из каких компонентов она состоит. Это описательная модель в виде списка, таблицы, схемы.

С позиции исследователя компоненты системы могут быть двух типов – подсистемы и элементы. *Подсистемы* состоят более чем из одного элемента и расчленяются дальше на свои компоненты: подсистемы более низкого уровня и элементы. *Элементы* – это неделимые в данной задаче части системы, хотя они могут иметь весьма сложный состав.

В иерархии компонентов системы каждая подсистема может рассматриваться как система соответствующего более низкого уровня и, наоборот, каждая система может быть представлена как подсистема некоторой системы более высокого уровня – надсистемы. Сама

исследуемая система является частью надсистемы, находящейся во внешней среде по отношению к рассматриваемой системе.

### ***1.1.3. Цели искусственной системы***

Цель искусственной системы – желаемый человеком результат функционирования системы. Так, целью космической системы наблюдения за Землей является периодическое наблюдение природных ресурсов в заданном районе земной поверхности с желаемым количеством и качеством информации для потребителя.

Любые системы (технические, биологические, социальные и т.д.) обладают способностью осуществлять поведение, преследующее достижение определенной цели. Это свойство называют *целенаправленностью*.

Свойство целенаправленности является системным, оно возникает вследствие взаимодействия компонентов, обусловленного структурой системы. Это взаимодействие организуется путем целенаправленного воздействия на функционирование компонентов системы, т.е. путем *управления*.

Для создания системы необходимо описать ее цель. Любое описание цели – это ее модель, словесная или математическая.

Описание назначения состоит из трех компонентов:

- наименование действия, выполняемого системой и приводящего к желаемому результату, т.е. к удовлетворению потребности;
- указание объекта внешней среды, на который направлено это действие;
- указание особых условий и ограничений, при которых выполняется действие.

Иерархической структуре системы соответствует иерархия целей ее компонентов. Благодаря согласованности и взаимной подчиненности своих целей все компоненты работают целенаправленно для достижения общей для них цели системы.

Цель системы достигается путем реализации целей ее компонентов – подсистем.

Цель подсистемы – средство достижения цели системы более

высокого уровня – задача системы более высокого уровня, однако это не простое сложение последних, так как реализация каждой из подцелей приводит к достижению качественно новой цели высшего уровня (свойство целостности системы).

#### ***1.1.4. Структурные уровни***

Любая материальная система является полиструктурной, т.е. имеющей множество компонентов структуры – подструктур.

Каждой подструктуре соответствуют так называемый структурный уровень (не смешивать с иерархическим уровнем) и аспект исследования.

Многоуровневая структура системы определяет функцию системы в целом.

Технический структурный уровень системы, в свою очередь, также является многоуровневым. Декомпозицию, т.е. членение, технической структуры можно провести, воспользовавшись принципом единства структуры и функции.

Функция технического структурного уровня — выполнение назначения системы с помощью технических средств. Она подразделяется на две: 1) непосредственная реализация цели с помощью главного элемента системы; 2) обеспечение неизменности структуры системы, т.е. готовности системы к реализации цели.

Поддержание структуры и реализация цели обеспечиваются путем управления системой. Для управления предназначены специализированные компоненты, которые образуют подсистему, называемую системой управления, и соответствующий структурный уровень управления, который, в свою очередь, подразделяется на подуровни управления поддержанием структуры и реализации цели.

Для этих видов управления необходима информация о состоянии системы и внешней среды, о выполнении компонентами системы своих функций. Информацию воспринимают, передают по каналам связи, преобразуют и отображают соответствующие технические средства. Как видно, в системе имеется информационный структурный уровень и два информационных подуровня, в которых преобразуется информация для

управления поддержанием структуры системы и реализации цели. Если система информационная, например космическая система связи, то также преобразуется специальная информация.

Для функционирования системы необходимы также преобразования вещества и энергии и соответствующие им структурные уровни. Для этого в ее состав должны входить, кроме преобразующих устройств, запасы или источники материальных и энергетических ресурсов или коммуникативные связи, через которые поступают необходимые ресурсы из внешней среды, а также внутренние каналы материального и энергетического снабжения, доставляющие компонентам системы необходимые материалы и энергию. Все эти функции выполняют специализированные компоненты системы.

Для эксплуатации самолета, например, на борту необходимо иметь топливо и систему подачи топлива, систему энергопитания, состоящую из генератора и электросети, а во внешней среде (в аэропорту) – запасные части, запасы топлива, смазочные и другие расходные материалы.

Заметим, что один и тот же компонент системы может принадлежать различным структурным уровням, например БЦВМ (уровни управления, информационный, энергетический).

Итак, сделаем обобщающий вывод.

Система – это ограниченный в окружающей ее внешней среде и взаимодействующий с ней объект, который обладает следующими взаимосвязанными свойствами:

1) имеет цель (назначение), для достижения которой он функционирует;

2) состоит из взаимосвязанных составных частей – компонентов, образующих многоуровневую иерархическую структуру и выполняющих определенные функции, направленные на достижение цели объекта;

3) имеет управление, благодаря которому все компоненты функционируют согласованно и целенаправленно;

4) имеет в своем составе или во внешней среде источники энергии и материалов (вещества) для функционирования;

5) обладает интегративными свойствами, не сводимыми к сумме свойств его компонентов.

### **1.1.5. Построение модели функциональной структуры**

Модель структуры системы описывает компоненты системы, структурные и коммуникативные связи (существенные для целей исследования). Наиболее распространенный вид модели структуры – структурная схема.

*Построение модели функциональной структуры системы удобно вести в форме таблицы.*

Вначале необходимо построить модель состава системы. При этом каждый из компонентов должен иметь минимальное число (но не менее одной) определенных функций.

Для описания коммуникативных связей в модели структуры одновременно с определением состава системы строят вспомогательную модель внешней среды, выделяя компоненты среды, с которыми система находится в существенном функциональном или вынужденном взаимодействии, и связи этих компонентов с компонентами системы.

В первую очередь из среды выделяются объекты, на которые направлено действие системы. Среди всех компонентов системы имеется *главный компонент*, непосредственно взаимодействующий со средой при реализации цели системы.

Одновременно с разделением системы на компоненты и выделением взаимодействующих с ними компонентов среды целесообразно описывать функции компонентов системы.

Структурная схема системы представляет собой граф, вершинами которого являются наименования элементов системы и среды, а ребрами или дугами – функции элементов. Дуги выходят из вершин – элементов, чьи функции они описывают, и заканчиваются в вершинах – элементах, работу которых они обеспечивают, или в вершинах – элементах среды, взаимодействующих с рассматриваемым элементом. Ребра не ориентированы и описывают сложное взаимодействие двух элементов, которое в рассматриваемой задаче не подлежит членению на элементарные функции.

Из каждой вершины – элемента выходит столько дуг и ребер, сколько функций имеет элемент. Вершины, в которых заканчиваются дуги

– ребра, указаны в описании функций. Всем дугам – ребрам на графе присваивают обозначения, совпадающие с обозначениями соответствующих функций элементов.

### 1.1.6. Модель «черного ящика»

В процессе решения новой проблемы исследователь изучает систему постепенно усложняющихся моделей. Вначале важную роль играют модели наглядные, образные в естественно-языковой форме, например в виде описаний, и визуальные в знаковой или графической форме, например в виде схем. В процессе исследования модель системы уточняется, конкретизируется — происходит процесс развития модели до математической формы.

Наипростейшей моделью системы является ее определение. В нем о внутреннем устройстве системы говорится в самой общей форме как о совокупности взаимосвязанных частей, а суть определения сводится к обособлению системы как целостного объекта от взаимодействующей с ней среды.

Следующий шаг в развитии модели системы – построение модели так называемого "черного ящика". Изобразим систему и виде непрозрачного ящика, выделенного из окружающей среды, а связи системы со средой – в виде стрелок (рис. 1.2).

Связи, отображающие воздействия среды на систему, называют входами систем. К ним относятся функционально обусловленные воздействия, необходимые для реализации цели системы, например управляющие воздействия, и возмущающие воздействия.



Рис. 1.2. Модель «черного ящика»

Связи, отображающие воздействия системы на среду, называют выходами системы. Их также можно разделить на две группы. Одни

выходы — это полезные результаты работы системы, предназначенные для воздействия на среду или для использования в среде, т.е. реализация цели системы, другие — вынужденные воздействия, например вредные выбросы промышленных отходов в среду.

Как видно, в модели "черного ящика" указываются только входы и выходы, т. е. связи системы со средой, и отсутствуют сведения о внутреннем содержании системы, а также о границах между системой и средой. Такая модель, несмотря на простоту, отражает два важных свойства системы: обособленность от среды и взаимодействие с ней, и часто оказывается полезной, например, при исследовании естественных и социально-экономических систем, внутреннее устройство которых мало известно.

Для первоначального изучения системы достаточно словесного описания входов и выходов модели, например в виде их списка. При анализе и синтезе системы потребуется уже математическое описание входных и выходных воздействий.

Приведем пример простейшей модели космической системы наблюдения (КСН):

*Входы функциональные:*

- $U$  — информация от объектов наблюдения в виде электромагнитного излучения с определенными спектральными характеристиками;
- центральное гравитационное поле Земли.

*Входы возмущающие:*

- $V$  – воздействия на КА и на количество и качество информации, принимаемой КСН;
- нецентральность гравитационного поля Земли;
- сопротивление атмосферы;
- высота Солнца над местным горизонтом;
- облачность над объектами наблюдения;
- состояние атмосферы (дымка и пр.);
- фон, на котором происходит наблюдение (освещенность и светотеневая структура местности).

*Выход целевой:*



- информация, получаемая потребителем и удовлетворяющая его требованиям по периодичности, оперативности, качеству, производительности и др.

*Выходы вынужденные:*

- засорение космического пространства отработанными КА;
- засорение поверхности Земли отработанными ступенями ракет-носителей.

Построение модели "черного ящика" становится непростой задачей, если предполагается ее использование при построении математической модели для анализа и синтеза конкретной реальной системы. Трудность состоит в том, что надо решить, какие из бесчисленного множества реальных связей следует включить в модель. Критерием оценки должны служить цели исследования, существенность той или иной связи по отношению к этим целям. Всегда возможны ошибки: важные связи могут оказаться неизвестными или отброшенными как несущественные. Особо важное значение имеет определение выходов создаваемой системы, поскольку они описывают цель системы и определяют тем самым ее будущий облик.

### ***1.1.7. Функционирование и развитие системы***

Под динамикой в широком смысле подразумевают любые изменения во времени чего-либо. Системы, в которых происходят какие бы то ни было изменения во времени, называют динамическими. Различают два типа динамики систем: функционирование и развитие.

Под функционированием понимают процессы, которые происходят в системе, стабильно реализующей неизменную цель.

Развитием называют то, что происходит с системой при изменении ее цели или при существенном изменении условий в окружающей среде. В этих случаях существующая структура перестает соответствовать новой цели или новым внешним условиям, и для реализации цели системы приходится изменять состав и структуру системы. В новой структуре остается часть компонентов из старой. Например, в КС для исследования спутника Марса Фобоса использовались многие компоненты КС для

исследования Марса. Система может также функционировать и развиваться одновременно. Например, в функционирующей КСН могут модернизироваться наземные средства обработки специальной информации.

Задачи развития технических объектов различаются новизной и сложностью. В зависимости от степени новизны можно выделить следующие задачи:

- *частичная модернизация* существующей системы — изменение ее параметров (а иногда и структуры) с целью небольшого улучшения одного или нескольких свойств системы;
- *существенная модернизация* или развитие — изменение параметров и структуры системы для значительного улучшения основных свойств системы при решении прежних или новых целевых задач;
- *создание новой системы* для решения новых целевых задач или для получения принципиально новых свойств системы при решении прежних задач.

Отображение процессов, происходящих в системе и окружающей среде, осуществляется с помощью динамических моделей. Модели развития отображают происходящие с течением времени изменения в системе и среде. Модели функционирования описывают поведение системы, последовательность каких-то действий, операций.

При исследовании системы происходит развитие ее динамической модели от описательной до математической формы. Для описания моделей имеет место аналогия между статическими и динамическими моделями. Модели состава системы соответствует состав (перечень) действий, необходимых для перевода начального состояния системы в конечное. Модели структуры соответствует последовательность этих действий, логика функционирования, описываемая логической моделью операции.

Система в процессе функционирования выполняет ряд целенаправленных согласованных и взаимосвязанных действий, совокупность которых называют *операцией*.

Описание действия (элемента операции) формализованно можно представить состоящим из трех компонентов:

$$A \rightarrow E \rightarrow C,$$

где  $A$  и  $C$  – входной и выходной факторы;  $E$  – наименование действия по превращению  $A$  в  $C$ . Результат  $C$  действия  $E$  является событием, т.е. тем, что произошло.

Описание последовательности действий и событий, происходящих в процессе реализации цели системы, представляет собой модель осуществления операции, которую будем называть логической моделью операции, или короче — операционной моделью.

Операционная модель определяет последовательность действий компонентов системы: выходной фактор  $C_i$  действия  $E_i$ , т.е. результат этого действия служит входным фактором  $A_{i+1}$  последующего действия. Входными факторами являются также характеристики компонентов системы и среды, влияющие на результаты действия. Операционная модель может служить основой для построения математической модели операции. Для этого следует действия описать соответствующими математическими моделями компонентов, характеристики системы и среды – параметрами, а события оценивать числовыми характеристиками – показателями.

### ***1.1.8. Сложная техническая система***

В литературе среди всевозможных систем выделяют классы больших и сложных систем. Однако эти понятия окончательно еще не установились. Существует много различных подходов к определению больших и сложных систем. Одни авторы вообще не используют этих понятий, другие применяют их как синонимы, третьи определяют различие между ними по-разному. Многие из этих определений основываются на сложности описания и исследования системы, т.е. обусловлены позицией исследователя. Например, большую систему иногда определяют как систему, математическая модель которой имеет большую размерность. Сложность системы будем оценивать сложностью структуры, которая характеризуется следующими признаками:

- многоуровневой структурой со структурными уровнями различной природы (искусственный, естественный, организационный и т.д.);

- многоуровневой структурой искусственного структурного уровня (уровни реализации цели, сохранения структуры, управления, информационные, энергетические, вещественные);
- большим количеством элементов различной физико-технической природы, в том числе абстрактно-материальных элементов;
- целостной подсистемой на каждом структурном уровне, целями, подчиненными цели сложной системы;
- главной (системообразующей) ролью информационных связей в образовании сложной системы. Так, например, все части ЛА благодаря механическим, физическим и информационным связям образуют единое техническое изделие. Комплексы и системы более высоких уровней иерархии, чем ЛА, приобретают функциональную целостность только благодаря информационным связям между их компонентами.

Из сказанного следует, что сложная техническая система (СТС) представляет собой целостную многоуровневую подсистему сложной смешанной системы на техническом структурном уровне. Примером СТС является космическая система связи, космическая система наблюдения (КСН) (рис. 1.3), космический аппарат (рис. 1.4)



Рис.1.3. Состав космической системы наблюдения

Так, например, космическая система (КС) является надсистемой для КА, а сам КА можно представить состоящим из подсистем двух уровней иерархии и элементов (см. рис. 1.4).

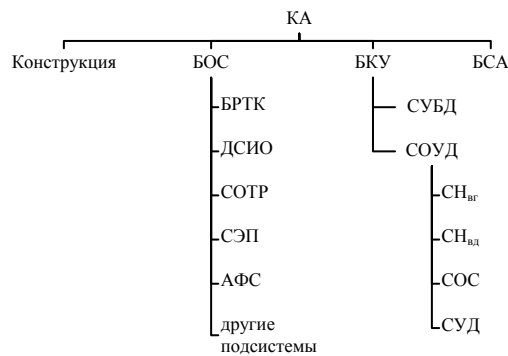


Рис. 1.4. Состав космического аппарата:

БСА – бортовая специальная аппаратура; БКУ – бортовой комплекс управления; СУБС – система управления бортовыми системами; СОУД – система ориентации и управления движением; СНвг – подсистема навигации; СНвд – подсистема наведения; СОС – система ориентации и стабилизации; СУД — подсистема управления движением (система коррекции); БОС – бортовые обеспечивающие системы; БРТК – бортовой радиотехнический комплекс; ДСИО – двигательная система с исполнительными органами; СОТР – система обеспечения теплового режима; СЭП – система электропитания; АФС – антенно-фидерная система

Модель состава ограничена сверху и снизу: на верхней границе (верхнем уровне) находится исследуемая система, отграниченная от среды, на нижней (нижнем уровне) — элементы.

Выделение тех или иных компонентов в системе определяется двумя факторами: с одной стороны, объективно существующими свойствами системы, с другой – субъективным подходом исследователя (его целями, знаниями о системе и исследовательскими возможностями). В результате расчленение реальной целостной системы является относительным, условным: условны как границы модели, так и границы между компонентами. Причины этого состоят в следующем: 1) объективные границы между системой и средой и между компонентами могут быть размыты, нечетки; 2) в зависимости от цели исследования подсистемы и элементы можно определить по-разному: элемент может оказаться подлежащим разделению, а расчленение на подсистемы может происходить по различным признакам.

Сложные технические системы представляют собой для инженеров

принципиально новые объекты исследований, разработки и эксплуатации, потому что компонент может быть оптимальным только в том случае, если он обеспечивает оптимальную СТС в целом. Поэтому при проектировании СТС необходимо определить, какими характеристиками должны обладать оптимальные компоненты.

### ***1.1.9. Жизненный цикл сложной технической системы***

Жизненный цикл СТС может быть представлен в виде следующих основных этапов:

1. Внешнее проектирование СТС.
2. Опытно-конструкторские работы (ОКР):
  - 2.1. Техническое предложение.
  - 2.2. Эскизный проект.
  - 2.3. Рабочий проект.
  - 2.4. Изготовление и экспериментальная отработка опытных образцов элементов.
  - 2.5. Летно-конструкторские испытания опытных образцов
3. Серийное изготовление компонентов и сооружение обеспечивающих средств.
4. Развертывание СТС.
5. Эксплуатация СТС.

*Первый этап.* Его можно условно назвать внешним проектированием. Такое наименование связано с тем, что на этом этапе СТС рассматривается как компонент более крупной системы (надсистемы), учитывается взаимодействие СТС с окружающей средой и, в частности, с другими компонентами надсистемы, оцениваются результаты, которые получит надсистема эксплуатации проектируемой СТС, и затраты ресурсов на ее создание. Этот этап обычно выполняется совместными усилиями заказчика и разработчиков имеет целью разработку общих требований к СТС, которые будут служить исходными данными для ОКР.

Внешнее проектирование СТС начинается с установления необходимости создания новой СТС и формулирования ее назначения и задач. Затем определяются основные ограничения на технические,

организационные и экономические решения, в рамках которых должна в дальнейшем разрабатываться СТС; примерами таких ограничений могут служить тип старта ЛА или вид топлива для его двигательной установки, сроки создания и эксплуатации СТС, допустимые затраты на ее создание. Результатом внешнего проектирования является тактико-техническое задание (ТТЗ) на ОКР.

Результаты внешнего проектирования дают возможность приступить к следующему этапу.

*Второй этап.* Осуществляются опытно-конструкторские работы, целью которых является реализация ТТЗ.

На подэтапе технического предложения работы разворачиваются в условиях, когда на проектируемую СТС в целом уже наложены некоторые требования и ограничения, определенные на этапе внешнего проектирования. В свою очередь, одной из задач технического предложения является формирование требований и ограничений, в рамках которых будет проходить проектирование элементов СТС. Для этого уточняются облик СТС (корректируется ТТЗ) и типы элементов, устанавливаются диапазоны основных характеристик элементов, обеспечивающих выполнение СТС своего назначения. Итогом работ на этом подэтапе является техническое предложение – комплект документации, включающий пояснительную записку в нескольких томах, альбомом чертежей и научно-технические отчеты. В нем содержатся принятые решения по реализации ТТЗ на СТС, технические задания на проектирование элементов, перечень работ, которые необходимо провести при последующем эскизном проектировании. На основании технического предложения принимается решение о проведении эскизного и рабочего проектирования.

На подэтапе *эскизного проекта* проверяются, конкретизируются принципы и положения технического предложения, принимаются подробные технические решения и прорабатываются все части проекта. Итогом этих работ является эскизный проект СТС.

На подэтапе рабочего проекта создается полный комплект опытно-конструкторской, технологической и эксплуатационной документации,

достаточный для изготовления элементов СТС.

На подэтапах изготовления и экспериментальной отработки опытных образцов и летно-конструкторских испытаний по результатам испытаний выявляются возможные ошибки и недоработки проекта, принимаются меры к их устранению. Итогами работ этих подэтапов являются решения о серийном производстве элементов СТС и передача документации на предприятия-изготовители.

*Третий этап.* Содержит техническую подготовку производства и серийное изготовление компонентов и сооружение обеспечивающих средств СТС.

*Четвертый этап.* Это этап развертывания СТС – изготовленные элементы объединяются в систему: между элементами создается необходимое взаимодействие, обеспечивается работоспособность системы в целом. В завершение этапа принимают решение о сдаче СТС в эксплуатацию.

*Пятый этап.* Здесь происходит эксплуатация СТС в соответствии с ее назначением. При этом выполняются также работы, обеспечивающие функционирование СТС, например техническое обслуживание и ремонтные работы. Для военно-технических СТС может иметь место подэтап боевого применения.

Разумеется, что приведенное разделение опытно-конструкторских работ по этапам жизненного цикла является в определенной мере условным, так как это – модель, а в реальной жизни трудно провести безусловные и четкие грани между работами на этапах и подэтапах создания СТС.

Кстати, моделями СТС являются результаты работ на каждом этапе (подэтапе), за исключением, естественно, этапа эксплуатации. Тактико-техническое задание вместе с результатами ОКР: техническим предложением, эскизным проектом, рабочим проектом, опытным образцом составляют иерархический ряд описаний (моделей) создаваемой системы. С каждым последующим этапом ОКР модель системы становится конкретнее, определеннее, детальнее. Рабочий проект с учетом уточнений и поправок, внесенных результатами стендовых и летно-конструкторских испытаний опытных образцов, позволяет уже изготовить систему.



### ***1.1.10. Проектирование, анализ и синтез сложной технической системы***

Проектирование – это комплекс работ с целью получения описаний нового или модернизируемого технического объекта, достаточных для реализации или изготовления объекта в заданных условиях. Объектами проектирования могут быть изделия (например, ЛА, ЭВМ) или процессы (например, процесс управления автоматическим КА). Комплекс проектных работ включает в себя теоретические и экспериментальные исследования, конструирование, разработку технической, технологической и эксплуатационной документации.

Рассмотрим некоторые особенности проектирования, важные для дальнейшего изложения.

*Принятие решений на основе исследований.* Проектирование можно рассматривать как процесс выработки решений по самым разнообразным вопросам (например, из каких элементов должно состоять изделие, какими должны быть параметры каждого элемента и т.д.). Решение принимает лицо, несущее ответственность за это решение, – обычно руководитель организации или руководители подразделений различных уровней. Для принятия решения по сложным вопросам руководителю обычно требуется, кроме имеющейся у него информации, еще и новая информация, которую получают, выполняя необходимые исследования. Поскольку в инженерном деле без количественных данных невозможно принимать проектные решения, исследования представляют собой в основном математическое моделирование и эксперименты. Большую роль в получении количественной информации для принятия решения играет математическое моделирование, так как эксперимент над системой в целом очень сложен или невозможен, как, например, эксперимент над космическим комплексом для исследования кометы Галлея или над системой стратегического ядерного вооружения.

Результатом исследования является информация, необходимая руководителю для принятия решения, часто в форме рекомендаций, так как любое исследование не в состоянии предложить руководителю решение в окончательной формулировке. Это объясняется следующими причинами.

Никакое исследование не в состоянии учесть множество разнообразных факторов, которые требуется принять во внимание для принятия решения. Поэтому необходимы и играют основную роль субъективные суждения руководителя, основанные на его опыте, интуиции, знаниях и творческих способностях, а также на его способности рисковать в условиях неопределенности.

Ситуации выбора всегда противоречивы, а исследование может дать лишь описание противоречий и рекомендацию компромисса между ними. Компромиссное решение выбирает руководитель, принимая во внимание свои (или вышестоящего руководителя) предпочтения, систему ценностей, политику с учетом ответственности за принятое решение.

Исследования при проектировании делятся на процедуры синтеза и анализа.

*Синтез.* Состоит в определении структуры и параметров объекта по его свойствам, сформулированным в техническом задании. Под структурой объекта понимают состав его компонентов и их связи друг с другом. Параметр – это величина, характеризующая некоторое свойство объекта или режим его функционирования. Соответственно различают структурный и параметрический синтез. Примерами структурного синтеза могут служить выбор состава и типов компонентов КА или синтез алгоритма управления КА (его структура определяется составом и последовательностью операций). Параметрический синтез заключается в выборе значений параметров объекта, его компонентов при заданной структуре объекта. Итак, будет называть синтезом СТС исследование, имеющее результатом рекомендации, необходимые для принятия решений по выбору структуры и параметров СТС.

*Анализ.* Состоит в определении свойств объекта по его структуре и значениям параметров объекта и его компонентов. Анализ технических объектов часто осуществляется с помощью расчетов для серии вариантов структуры и параметров объекта – вариантных расчетов. Одна из целей анализа – проверка работоспособности и эффективности объекта, т.е. проверка результатов синтеза. Другая цель – синтез путем обобщения данных, добытых анализом. Так как прямой (непосредственный) синтез СТС является проблемой чрезвычайной сложности, на практике обычно

выбирают предпочтительный вариант структуры и параметров СТС, оценив результаты вариантных расчетов.

*Синтез* СТС, как и ее проектирование, является весьма сложной проблемой выбора рекомендаций по структуре и параметрам СТС. Сложные проблемы возникают в современном мире не только при создании и эксплуатации СТС, но и во всех областях человеческой деятельности – в технике, производстве, народном хозяйстве, экономике, социальной сфере, военном деле и т.д. Необходимость решения этих проблем привела к разработке в разных областях жизни множества приемов и методов, которые в результате теоретического обобщения оформились в научную дисциплину "Системный анализ". Ее можно считать дальнейшим развитием теории управления и исследования операций.

Задачи выбора рекомендаций для принятия решений (соответствующее научное направление называют обычно "Теорией выбора" или "Теорией принятия решений") различаются математическими постановками и методами решения вследствие разнообразных ситуаций, в которых происходит выбор.

Общая для всех ситуаций математическая постановка задачи выбора состоит в следующем:

- 1) задано исходное множество альтернатив, из которых следует выбрать наиболее предпочтительные (сузить множество альтернатив, обычно до одной);
- 2) цели, для достижения которых производится выбор, определены;
- 3) имеется способ или правило сравнения альтернатив между собой и определения наиболее предпочтительной;
- 4) имеется математическая модель для оценки альтернатив.

Это означает, что уже пройдены чрезвычайно важные этапы системного анализа:

- 1) определение цели, ради достижения которой производится выбор;
- 2) порождение множества альтернатив, на котором предстоит осуществлять выбор;
- 3) выбор способа сравнения альтернатив;
- 4) построение модели.

### **1.1.11. Основные этапы исследования систем путем математического моделирования**

Процесс исследования системы можно представить в виде операционной модели. Названия этапов и подэтапов (задач) сформулируем так, чтобы они одновременно обозначали действие, цель этого действия и его результат (табл. 1.1).

*Модель цели* системы представляет собой список функциональных требований к системе, которые должны быть выполнимы посредством выбора предпочтительной альтернативы.

При синтезе системы альтернативами обычно служат возможные пути или средства достижения цели исследуемой системы, например, варианты структуры и значений параметров системы или различные алгоритмы управления функционированием системы. Для возможных вариантов структуры, параметров и алгоритмов управления всегда существуют физические и практические ограничения, налагаемые на эти факторы, которые определяют границы множества альтернатив.

В случае анализа системы аналогом альтернатив служат различные исследуемые факторы – варианты структур, параметров, алгоритмов, влияние которых на свойства системы требуется изучить.

Таблица 1.1. Этапы и подэтапы исследования системы

№ п/п	Этапы	Подэтапы (задачи)
1	Постановка задачи	Составление модели цели системы Формирование множества альтернатив и выявление ограничений этого множества Выбор показателей свойств системы - критериев Построение критерия предпочтения Формулировка постановки задачи
2	Разработка математической модели	Построение модели Разработка алгоритма решения задачи Разработка программы (программного и информационного обеспечения)
3	Вычисления на ЭВМ	Подготовка исходных данных Получение результатов вычислений
4	Выработка рекомендаций	Анализ результатов моделирования Формулирование рекомендаций

Чтобы выбирать альтернативы, необходимо получать количественные оценки степени достижения цели и важнейших свойств

системы, обеспечивающих достижение цели. Для этого выбирают соответствующие числовые характеристики цели и свойств системы: критерии – показатели эффективности, качества, надежности и др. Особую роль играют показатели, оценивающие затраты ресурсов. Для создания и эксплуатации любой системы необходимы определенные затраты ресурсов (денег, оборудования, материалов, энергии, рабочей силы и т.д.). Для каждой из альтернатив требуются свои затраты ресурсов, чаще всего оцениваемые показателем стоимости создания и эксплуатации системы.

Для рекомендации предпочтительной альтернативы выбирают критерии предпочтения – правила, на основе которых по значениям критериев принимают рекомендации по выбору альтернатив.

При анализе системы аналогом критерия предпочтения служат некоторые правила или рекомендации, как следует оценивать результаты исследования.

Значения критериев вычисляются с помощью модели системы, которая представляет собой логико-математическое описание зависимостей между критериями эффективности, качества, надежности и стоимости системы, с одной стороны, и альтернативами достижения цели и характеристиками системы и внешней среды – с другой.

Алгоритм решения задачи на ЭВМ — это точное формальное предписание, однозначно определяющее содержание и последовательность операций, переводящих заданную совокупность исходных данных в искомый результат. Текст описания алгоритма, составленный средствами какого-либо языка программирования, называется программой.

Результаты вычислений служат основой для выработки рекомендаций по выбору предпочтительной альтернативы. Получение рекомендаций является целью исследований системы, на основе рекомендаций и неформальных суждений руководитель, принимает решение.

При выборе (построении или описании) всех элементов исследования (цели, альтернатив, ограничений, показателей, критерия предпочтения, модели, алгоритмов, программы, рекомендаций) исследователь, в свою очередь, принимает решения на основе соответствующих частных исследований и неформальных суждений.

"Модель-алгоритм-программа" представляет собой единство трех

понятий — триаду.

*Постановка задачи.* Является очень важным этапом исследования и включает формулировку цели исследования, описание модели цели системы, выбор множества альтернатив и их ограничений, показателей и критерия предпочтения и подготовку исходных данных для построения модели системы и проведения вычислений. Результатом этапа является формулировка постановки решаемой задачи.

Следует указать две особенности процесса постановки задачи.

1. Процесс постановки задачи является очень сложным и не формализуемым математически. По существу это творческий процесс, обусловленный объективными свойствами исследуемой системы и субъективными данными исследователя. Поэтому различные исследователи создают обычно разные варианты решения одной и той же задачи.

2. Постановка задачи – процесс итеративный. Во-первых, он итеративен сам по себе как итеративный процесс мышления. В процессе выбора постановки задачи у исследователя незаметно возникают колебания в принятии решений, он совершает ошибки, выбирает те или иные подходы, заводящие исследование в тупик, и т.п. Во-вторых, сам процесс исследования является итеративным. Необходимость уточнения постановки задачи может возникнуть на любом этапе исследования в результате появления новой внешней информации, положительных или отрицательных результатов исследования, непредвиденных трудностей или ограничений и т.п.

Процесс постановки задачи полностью завершается вместе с окончанием исследования системы.

*Разработка математической модели.* Это – центральный этап исследования любой системы. От совершенства модели зависит судьба всего исследования. Этап разработки модели можно разделить на три подэтапа:

- 1) построение модели;
- 2) выбор приближенного численного метода решения задачи и разработка алгоритма ее решения;
- 3) разработка программы – создание программного и информационного обеспечения.

Данный этап можно назвать разработкой триады "модель-алгоритм-

программа".

Выработка рекомендаций является последним этапом исследования, на котором анализируются результаты вычислений и принимаются рекомендации по выбору альтернативы решения поставленной задачи.

Часто указанные этапы трудно четко отграничить друг от друга, так как исследование может проходить посредством параллельного выполнения некоторых этапов и неоднократных повторений уже пройденных, т.е. итеративным путем. Например, на каждом из этапов может потребоваться уточнение постановки задачи; при разработке алгоритма оптимизации – упрощение модели, при выработке рекомендаций – корректировка цели системы и т.д. В таких случаях цикл исследований повторяется. Процессы построения модели, разработки алгоритма вычислений и программирования часто идут параллельно, взаимодействуя друг с другом.

## **1.2. Методы анализа и синтеза сложных технических систем**

### ***1.2.1. Значение и особенности процесса постановки задачи***

Первоначально проводится изучение некоторой расширенной задачи. При этом рассматривается или система задач, в которую входит первоначально поставленная задача, или задача более высокого уровня иерархии, или задача того же уровня, но охватывающая более широкий состав рассматриваемых факторов. Параллельно могут рассматриваться системы более высоких уровней и иерархии, чем система, для которой решается задача. Такой подход на ранней стадии постановки задачи способствует развитию творческого подхода к исследованию и помогает рационально сузить задачу до формулировки решаемой задачи.

Процесс постановки задачи можно представить тремя стадиями (или ступенями). Сначала это постановка задачи в первоначальном виде, затем некоторое ее развитие или расширение с целью лучшего понимания задачи и, наконец, некоторое сужение к конкретной решаемой задаче построения и исследования модели.

Основные элементы (действия, операции), из которых может состоять работа исследователя над постановкой задачи:

1. Выбор расширенной задачи – всегда является первым этапом.
2. Описание цели системы – начинается с изучения вышестоящей системы или с изучения и обоснования цели исследуемой системы. При этом возможна корректировка заданной цели. В некоторых случаях может быть выбрана цель, хотя это задача заказчика работы. В любом случае действие заканчивается описанием модели цели системы в виде списка функциональных требований к системе.
3. Описание системы – включает построение моделей состава и структуры системы, которые могут представлять собой описания, таблицы или схемы состава и структуры исследуемой системы, перечень основных функций системы и ее компонентов.
4. Описание внешней среды – указывает условия работы системы во внешней среде, возможные варианты взаимодействия системы со средой, характеристики внешней среды, которые необходимо учесть при моделировании.
5. Описание операций системы во взаимодействии со средой – составляется, например, в виде логической модели операции по реализации цели или поддержанию неизменной структуры системы. Эта модель, как и модель структуры систем, необходима в качестве основы для разработки в последующем математических моделей решаемой задачи.
6. Выбор состава множеств альтернатив – среди них будут выбираться предпочтительные альтернативы.
7. Выбор показателей и критерия предпочтения – является основой принятия рекомендаций по выбору предпочтительной альтернативы.
8. Определение объектов моделирования и синтеза – служит для сужения задачи путем исключения из исследования тех компонентов и параметров системы и внешней среды, которые слабо влияют на решение задачи.
9. Декомпозиция задачи, т.е. разделение на составные части расширенной задачи, – необходима для сужения задачи с обеспечением системного подхода к ее постановке.
10. Подготовка исходных данных.
11. Формулирование постановки решаемой задачи – включает цель решения задачи, альтернативы с ограничениями, критерий предпочтения.



### ***1.2.2. Качество и эффективность функционирования системы***

Среди свойств системы целесообразно выделить две группы: качество и эффективность функционирования системы.

Качество системы – одно из (или совокупность) обязательных (с позиции надсистемы) свойств функционирования системы, обеспечивающих ее работоспособность, т.е. пригодность системы к выполнению своего назначения – достижению цели системы. Качество КСН характеризуется, например, способностью наблюдать заданный район, спектральными характеристиками аппаратуры наблюдения, пространственной разрешающей способностью, обзорностью, периодичностью и т.д.

Эффективность системы – фактический (достигнутый) или будущий (ожидаемый, проектируемый) результат (эффект) ее функционирования на определенном интервале времени, характеризующий степень достижения цели надсистемы.

Например, эффективность КСН может характеризоваться ее производительностью, т.е. объемом достоверной информации, получаемой за определенный промежуток времени, а также оперативностью – временем от момента получения информации на борту КА до момента передачи ее потребителю. В соответствии со сказанным, среди требований, которые предъявляются к системе, целесообразно выделить две группы:

1) требования, определяющие качество системы, невыполнение которых приводит к тому, что система оказывается непригодной для решения целевых задач (для достижения цели);

2) требования, определяющие эффективность системы (степень достижения цели), невыполнение которых ухудшает результаты функционирования, потребные надсистеме.

Различие между требованиями к качеству и эффективности системы состоит в том, что первые позволяют выявить необходимые альтернативы, а вторые – выбрать из них наилучшие. Это различие несколько условно и зависит как от цели системы, так и от позиции исследователя или заказчика системы. Одни и те же требования для разных целей системы могут быть отнесены как к первой, так и ко второй группе. Указанное разделение требований полезно методически для рациональной постановки задачи

синтеза СТС. Составляя список требований к СТС для ее синтеза, исследователь должен убедиться, что предусмотрены все необходимые требования к качеству и эффективности СТС.

Требования к качеству и эффективности системы описываются соответствующими показателями.

Показатели качества  $Q_i$  — количественные характеристики свойств системы, обеспечивающих пригодность системы к выполнению своего назначения, например, периодичность наблюдения и разрешающая способность КСН.

Требования к качеству системы задаются в форме ограничений соответствующих показателей:

$$Q_i = \bar{Q}_i \quad \text{или} \quad Q_i \geq (\leq) \bar{Q}_i.$$

Эти требования обязательны для любых альтернатив системы.

Показатели эффективности  $W_k$  — количественные характеристики свойств системы, служащие для выбора наилучшей альтернативы из всех пригодных. Так, например, показателями эффективности КСН могут быть производительность и оперативность КСН. Показатели эффективности должны иметь экстремальные значения:

$$W_k \rightarrow \max(\min).$$

Эффективность — внешнее свойство СТС: оно проявляется в надсистеме, например в отрасли народного хозяйства.

*Экономическая эффективность.* Если эффективность системы оценивается экономическим показателем, например доходом, то говорят об экономической эффективности. Экономическая эффективность  $\mathcal{E}$  есть экономический результат:  $D$  — доход, приносимый надсистеме функционированием системы и соотнесенный с суммарными затратами ресурсов  $C_\Sigma$  в надсистеме на создание и эксплуатацию системы и на использование системой результатов функционирования системы. Очевидно, что экономическая эффективность является понятием, внешним по отношению к системе, и оценивается двумя показателями: доходом  $D$  и затратами ресурсов  $C_\Sigma$ , причем доход зависит от затрат  $D(C_\Sigma)$ .

Применяются показатели, использующие разность между доходом и

затраченными ресурсами:

1) прибыль  $D - C_{\Sigma}$ ;

2) рентабельность  $\frac{D - C_{\Sigma}}{C_{\Sigma}}$ .

Оценка экономической эффективности при проектировании системы чрезвычайно сложна. Поэтому показатель экономической эффективности целесообразно расчленить на несколько частных показателей, зависящих от структуры и параметров системы.

### **1.2.3. Надежность системы**

Понятие надежности технической системы существенным образом связано с понятием ее качества. Ранее было сказано, что качеством системы называется совокупность свойств, определяющих степень пригодности системы для использования по назначению. С течением времени свойства, составляющие качество системы, видоизменяются в нежелательную сторону. Под надежностью понимается способность системы сохранять качество при определенных условиях эксплуатации.

Понятие надежности является совокупностью трех понятий: безотказности, надежности восстанавливаемой системы, долговечности.

Под безотказностью понимается способность объекта сохранять работоспособность (т.е. не иметь отказов) в течение определенного интервала времени в определенных условиях эксплуатации. Отказ — это частичная или полная утрата или видоизменение таких свойств, которые существенным образом снижают работоспособность или приводят к ее полной потере. Понятие отказа является весьма относительным, так как оно существенным образом зависит от конкретных условий использования системы. Часто под отказом понимают отклонение показателя качества от номинального значения за предельно допустимые границы. В большинстве случаев под надежностью (в узком смысле) понимают вероятность безотказной работы в течение заданного промежутка времени.

Дорогой, сложный и необходимый компонент системы после отказа обычно восстанавливается (заменяется новыми или ремонтируется). При большом сроке службы системы восстановление происходит многократно — имеет место процесс восстановления. Так как каждое восстановление требует некоторого времени, то основной характеристикой надежности

восстанавливаемой системы является готовность системы к эксплуатации в любой момент времени — вероятность того, что в каждый момент времени система находится в исправном состоянии. Как видно, выполнение требований к качеству и надежности системы преследует общую цель — обеспечение пригодности системы к выполнению своего назначения.

Важной составляющей надежности является долговечность системы, под которой понимают способность системы к эксплуатации (с возможным восстановлением) в течение длительного срока службы. В конце срока службы, определяющего долговечность, в системе появляются такие процессы, связанные с износом или старением, устранение которых невозможно или экономически нецелесообразно.

#### ***1.2.4. Затраты ресурсов на создание и функционирование системы***

Надсистема вынуждена "платить" за получение эффекта от функционирования системы затратами ресурсов на создание и эксплуатацию СТС (денег, оборудования, материалов, энергии, рабочей силы и т.д.). Затраты ресурсов на создание и эксплуатацию СТС — внешнее по отношению к системе явление: затраты происходят в надсистеме.

Без учета затрат ресурсов эффективность сама по себе не дает оснований для выбора наилучшей СТС. Можно было бы добиться очень высокой эффективности за счет чрезмерных затрат, но ресурсы всегда ограничены. Если стремиться к наименьшим затратам, то можно получить низкую или даже неудовлетворительную эффективность. Существует много альтернатив достижения цели СТС, для каждой из них требуются различные затраты ресурсов. Следовательно, предпочтительную альтернативу можно выбрать только с помощью критерия, учитывающего как эффективность СТС, так и затраты на ее создание и эксплуатацию.

Для оценки затрат ресурсов необходимы соответствующие показатели. Они могут быть различны. Наиболее общий и распространенный показатель — стоимость создания и эксплуатации СТС. В некоторых случаях может оказаться необходимым учитывать затраты других материальных ресурсов, главным образом дефицитных.

Показатель стоимости складывается из нескольких составляющих. В качестве примера приведем статьи затрат на создание и эксплуатацию орбитального комплекса (ОК) в течение срока службы  $T_\phi$ . В указанную стоимость входят затраты на:

- научно-исследовательские работы (НИР) и опытно-конструкторские работы (ОКР) по разработке как отдельного КА, так и ОК в целом, в том числе затраты на проведение наземных и летных испытаний; эти затраты на НИР и ОКР зависят от массы КА, сложности КА и ОК, стоимости материальной части, предназначенной для проведения испытаний;
- серийное производство КА в количестве  $N_T$ , необходимом для надежного функционирования ОК в течение срока службы  $T_\phi$ ;
- пуск  $N_T$  космических аппаратов за срок  $T_\phi$ , включая затраты на ракеты-носители и обеспечение пусков в наземных комплексах (техническом, стартовом, измерительном);
- эксплуатацию  $N_T$  космических аппаратов за срок службы  $T_\phi$  средствами НКУ и СПОИ (зависят от сложности обслуживания КА).

Как видно из этого примера, затраты на создание и эксплуатацию технического объекта существенно зависят от его надежности и требуемого срока службы  $T_\phi$ .

За повышение всех положительных свойств системы (качества, эффективности, надежности, срока службы) необходимо платить затратами ресурсов. Все перечисленные свойства зависят от стоимости системы.

Обычно в ТТЗ требуется, чтобы стоимость системы была ограничена или минимальна:  $C \leq \bar{C}$  или  $C \rightarrow \min$ .

### ***1.2.5. Показатели функционирования системы в условиях неопределенности***

При выборе показателей функционирования системы (например, качества, надежности, эффективности и стоимости) необходимо обеспечить соответствие показателей той информации о характеристиках системы и внешней среды, которой располагает исследователь. Чем более полным является это соответствие, тем рациональнее можно употребить

имеющиеся ресурсы на обеспечение достаточно высокой эффективности системы. В зависимости от информации о системе и внешней среде, известной исследователю, показатели системы могут формироваться в условиях неопределенности (риска), интервальной неопределенности, противодействия.

Показатели системы формируются в условиях определенности, если вектор  $\nu$  параметров, характеризующих состояние системы и среды, – детерминированная величина. При этом каждой альтернативе соответствует определенное значение вектора показателей  $W(x, \nu)$ . Здесь  $x$  – вектор управляющих переменных (проектных параметров).

Показатели формируются в условиях статистической неопределенности, если вектор  $\nu$  параметров системы и среды – случайная величина с известной плотностью распределения вероятностей  $\rho(\nu)$ . (Неслучайную составляющую вектора  $\nu$  можно рассматривать как частный вид случайной при одном возможном значении с вероятностью единица.)

Рассмотрим некоторые из показателей, ограничившись случаем, когда "полезными", более эффективными, являются меньшие значения показателя  $W$ .

При минимаксном подходе показателем, служащим для выбора альтернативы, является значение показателя  $W^*$ , соответствующее наихудшему сочетанию неопределенных параметров  $\nu^*$  из множества  $V$ :

$$W^* = \max_{\nu \in V} W(x, \nu) = W(x, \nu^*). \quad (1.1)$$

Согласно данному показателю предпочтительное решение  $x^*$  ищется с учетом требования

$$\min_{x \in X} W^* = \min_{x \in X} \left[ \max_{\nu \in V} W(x, \nu) \right]. \quad (1.2)$$

Здесь  $X$  – множество допустимых решений. Можно требовать также выполнения условия

$$W^* = \max_{\nu \in V} W(x, \nu) \leq \bar{W}. \quad (1.3)$$

Условие (1.2) называют минимаксным критерием или критерием

Вальда.

Следует обратить внимание на две особенности "минимаксного решения".

1. Положение "худшей" точки зависит от рассматриваемой альтернативы, и эту точку требуется определять в процессе выбора в соответствии с условиями (1.1).

2. Оптимальное решение, получаемое с помощью минимаксного критерия, с практической точки зрения часто оказывается чрезмерно "осторожным", поскольку жестко связывает полученное решение с положением "худшей" точки  $v^*$  в области  $V$  возможных значений неопределенных параметров, независимо от того, сколь часто это "худшее" состояние может возникать в ходе эксплуатации системы.

Чтобы "смягчить" указанный недостаток минимаксного выбора оптимальной альтернативы в условиях интервальной неопределенности, предложены другие подходы (Лапласа, Сэвиджа, Гурвица). Согласно подходу Лапласа, вводится дополнительное предположение о равномерной плотности распределения  $\rho(v)$  параметров  $v$  в области неопределенности  $V$ . Тогда можно в качестве показателей использовать математическое ожидание  $m_W$  и вероятность  $P(W \leq \bar{W})$  или  $P(W \geq \bar{W})$ .

Область "интервальной неопределенности" может характеризовать не только возможные состояния системы и среды, недостаточно точно известные исследователю, но и неопределенное противодействие "противника" (конкурента), цели которого являются противоположными по отношению к целям исследователя разрабатываемой системы. В этом случае альтернативы сравниваются с помощью методов теории игр.

### ***1.2.6. Проблема многокритериальной оптимизации***

В системном анализе главное внимание уделяется постановкам задач выбора и условиям использования различных методов, так как сами методы решения многих задач относятся к другим дисциплинам (теория оптимизации, исследование операций и др.).

В случае одного критерия выбор сводится к отысканию альтернативы с наибольшим (наименьшим) значением критерия, т.е. к

решению задачи оптимизации. На практике необходимо оценивать альтернативы не по одному, а по двум и более критериям, качественно различающимся между собой.

Многокритериальные задачи оптимизации не имеют математического решения. Поэтому предлагается много разных способов придать многокритериальной задаче частный - однокритериальный вид, обладающий решением. Все эти способы основаны на различных субъективных предположениях. Тогда, если модель, правильно отражающая объективно существующую реальность, построена, единственный критерий установлен, то имеется единственно возможное оптимальное решение.

Рассмотрим классическую двухкритериальную задачу "эффективность – стоимость", которая послужила началом системного анализа. Так как в объективной модели нет связей, устанавливающих наилучшее соотношение между значениями критериев, то возникает вопрос, как учитывать взаимосвязь эффективности и затрат ресурсов для выбора предпочтительной альтернативы. Для упрощения изложения будем в дальнейшем оценивать затраты ресурсов стоимостью  $C$  создания и эксплуатации СТС.

Очевидно, что показатель  $W$  эффективности СТС должен зависеть не только от варианта ее построения, характеризуемого вектором параметров  $x$ , но и от стоимости системы  $C$ :  $W(x, C)$ .

Если полагать, что ресурсы расходуются наиболее рационально, то для каждого значения стоимости  $C$ , можно найти такие параметры системы  $x^*$ , которые дают максимальную эффективность:

$$\max_{x \in X} W(x, C_i) = \bar{W}(x_i^*, C_i). \quad (1.4)$$

Зависимость максимального значения показателя эффективности  $\bar{W}$  от стоимости  $C$  представляет собой множество, которое сузило множество альтернатив до зависимости  $x^*(C)$ .

Зависимость максимального значения показателя эффективности  $\bar{W}$  от стоимости  $C$  является возрастающей (рис. 1.5). Это показывает, что бессмысленно требовать  $\max \bar{W}$  при  $\min C$  и что выбор следует проводить



среди значений  $x^*(C)$ .

Для решения рассматриваемой двухкритериальной задачи возможны два подхода:

1. Эффективность считаем более важным показателем, и тогда затраченные ресурсы должны обеспечивать наивысшую эффективность системы. В этой постановке выбирается вариант системы  $x_1^*$  в соответствии с максимально возможной затратой ресурсов  $C_1 = C_{\max}$ ;

$$C = C_1, \quad \max_{x \in X} W(x, C_1), \quad (1.5)$$

что обеспечивает при заданной стоимости системы  $C_1$  максимальную ее эффективность  $W(x_1^*, C_1)$ .

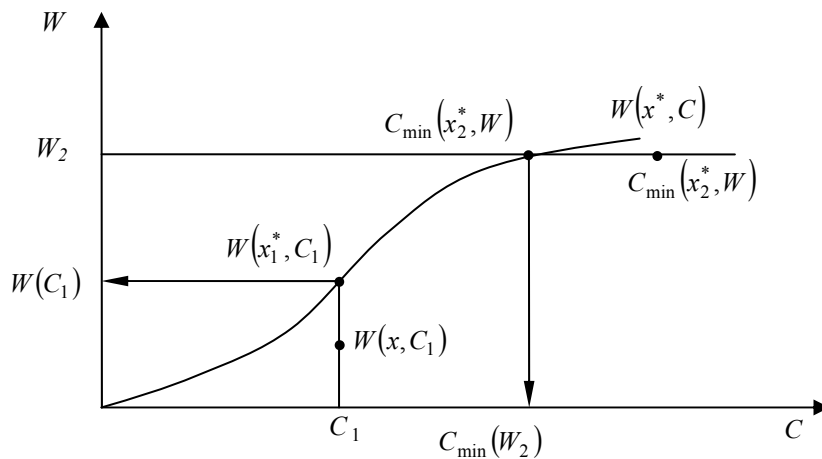


Рис. 1.5. Зависимость оптимального значения показателя эффективности от стоимости системы

2. Во втором подходе эффективность системы должна быть обеспечена при наименьших затратах ресурсов. Тогда, чтобы не допустить неудовлетворительную эффективность, выбирается необходимая эффективность  $W = W_2$  и вариант системы  $x_2^*$  удовлетворяет условию

$$W = W_2, \quad \min_{x \in X} C(x, W_2), \quad (1.6)$$

что обеспечивает при минимальных затратах ресурсов  $C_{\min} = C_2$  заданный уровень эффективности  $W_2 = \bar{W}(x_2^*, C_{\min})$ .

Эти два варианта эквивалентны, если функция  $W_1(x^*, C)$  возрастающая. В модели цели должно быть записано одно из условий: или ограничение по стоимости, или ограничение эффективности. Выбор одного из этих условий в значительной степени зависит от того, что именно – затраты или эффективность – может быть задано с большей степенью точности. Обычно менее точно определяются затраты из-за грубых методик и неточных исходных данных. По этой причине целесообразно задавать стоимость и максимизировать эффективность.

Теперь мы можем видеть, что двухкритериальную задачу одним из двух способов свели к однокритериальной, т.е. к задаче оптимизации. Но для этого потребовалось ввести субъективное ограничение на один из критериев. Такое ограничение может следовать из технических требований, предъявляемых разрабатываемому объекту, либо назначаться руководителем, принимающим решение.

Рассмотренный способ распространяют и на задачу, в которой более чем два критерия. Он весьма распространенный и может применяться, когда критерии различной важности или неравнозначные. Тогда среди них можно выделить главный критерий. Сущность способа вытекает из его названия: выделение главного критерия. Наиболее важный критерий (обозначим его через  $W_1$ ) выбирается с помощью дополнительной информации, и к нему предъявляется условие экстремума:

$$W_1(x, \nu) \rightarrow \max(\min), \quad (1.7)$$

а на остальные критерии накладываются жесткие  $W_k(x, \nu) = \tilde{W}_k$  или нежесткие  $W_k(x, \nu) \leq \tilde{W}_k$  ограничения, причем  $k = 2, \dots, m$ .

Основанием этого способа служит то, что в технических условиях ко многим критериям эффективности системы обычно предъявляются очень жесткие требования в виде ограничений, которые обязательно должны быть выполнены. Это позволяет показатели эффективности перевести в показатели качества при анализе модели цели системы.

*Свертывание критериев в целевую функцию* - это второй способ сведения многокритериальной задачи к однокритериальной, наиболее распространенный на практике, хотя и не самый совершенный. Подход к решению многокритериальной задачи состоит в переходе от нескольких критериев эффективности  $W_k (k = 1, \dots, m)$  к некоторой одной функции критериев  $\Phi(W_k)$ . Этот переход называют *свертыванием*; функцию  $\Phi(W_k)$  будем называть целевой функцией (другие названия: обобщенный показатель, суперпоказатель, критериальная функция, функция предпочтения).

Свертывание основывается на использовании дополнительной информации – объективной или субъективной. Соответственно различают объективное или субъективное свертывание. Объективная информация – это информация об объективных свойствах системы, субъективная – об относительной важности критериев эффективности.

*Объективное свертывание.* В некоторых случаях объективно существует возможность перехода от многих критериев к единственной целевой функции. Такая возможность, например, появляется при переходе к рассмотрению системы более высокого уровня иерархии, в состав которой входит синтезируемая система, если эффективность этой системы более высокого уровня можно оценить с помощью единственного критерия. Конечно, при этом необходимо привлекать некоторую дополнительную информацию, позволяющую рассчитать критерий эффективности системы более высокого уровня при всех альтернативах синтезируемой подсистемы. Например, от двух критериев эффективности системы наведения ЛА (математического ожидания  $m_h$  и дисперсии  $D_h$  промаха  $h$ ) можно перейти к критерию эффективности системы более высокого уровня – самого ЛА, взяв в качестве критерия его эффективности вероятность поражения цели  $P$ . При этом требуется привлечь дополнительную информацию о плотности вероятности промаха  $\rho(m_h, D_h, h)$ , характеризуемой параметрами  $m_h$  и  $D_h$ , а также зависимость условной вероятности поражения цели  $P_\xi(h)$  от промаха  $h$ . Тогда с помощью формулы полной вероятности

$$P = \int_{-\infty}^{+\infty} P_{\xi}(h) \rho(m_h, D_h, h) dh \quad (1.8)$$

от двух критериев  $m_h$  и  $D_h$  переходим к одному критерию  $P$  эффективности ЛА.

К субъективным способам построения целевой функции можно, в частности, отнести: введение обобщенного критерия с помощью весовых коэффициентов, называемое свертыванием; упорядочение показателей по важности и др.

Для введения обобщенного критерия вначале переходят к критериям в безразмерной форме с помощью преобразований вида

$$\tilde{W}_k = \frac{W_k}{W_{k\varepsilon}}, \quad k = 1, \dots, m, \quad (1.9)$$

или

$$\tilde{W}_k = \frac{W_k - W_{k \min}}{W_{k \max} - W_{k \min}}, \quad (1.10)$$

где  $W_{k\varepsilon}$  – эталонное (плановое) значение критерия  $W_k$ ;  $W_{k \max}$  и  $W_{k \min}$  соответственно максимальное и минимальное значения критерия  $W_k$ ,

Затем с помощью выражения вида

$$\Phi(\alpha_k, \tilde{W}_k), \quad k = 1, \dots, m,$$

переходят к обобщенному критерию. Положительные весовые коэффициенты  $\alpha_k$ ,  $k = 1, \dots, m$ , характеризуют относительную важность безразмерных критериев  $W_k$  и устанавливают определенный компромисс между ними. Обычно требуют выполнения условия нормировки:

$$\sum_{k=1}^m \alpha_k = 1.$$

При данном способе свертывания дополнительной субъективной информацией являются: способ приведения критериев к безразмерной форме, весовые коэффициенты  $\alpha_k$ , а также вид оператора  $\Phi$ .

Широко распространенным способом свертки критериев является линейная свертка:

$$\Phi = \sum_{k=1}^m \alpha_k \tilde{W}_k . \quad (1.11)$$

Используются также квадратичная свертка:

$$\Phi = \left[ \sum_{k=1}^m \alpha_k (\tilde{W}_k)^2 \right]^{1/2} , \quad (1.12)$$

мультипликативная:

$$\Phi = \prod_{k=1}^m (\tilde{W}_k)^{\alpha_k} \quad (1.13)$$

и ряд других.

Процесс выбора оператора  $\Phi$  и коэффициентов  $\alpha_k$  является неформальным актом. Он требует проведения тщательного анализа задачи выбора. Коэффициенты важности  $\alpha_k$  могут выбираться экспертами на основе объективной и субъективной информации о предпочтительности некоторых критериев. Предложено много методов расчета весовых коэффициентов, которые также используют дополнительную информацию, как субъективную, так и объективную, об особенностях решаемой задачи.

Объединение нескольких критериев в один имеет следующий недостаток. Даже небольшое изменение коэффициентов важности может привести к тому, что новая оптимальная альтернатива окажется очень сильно отличающейся от старой.

Одним из распространенных способов свертывания служит использование относительных критериев. Примером может служить относительный критерий "стоимость – эффективность", требующий максимальной эффективности на единицу затрат:

$$\max \frac{W(x)}{C(x)} . \quad (1.14)$$

Однако, выбор на основе относительного критерия может привести к принятию неправильных рекомендаций, если не зафиксирована требуемая величина эффективности или требуемая величина затрат.

### 1.2.7. Построение множества Парето для сужения множества альтернатив

Третий способ решения многокритериальной задачи – построение множества Парето или множества «неулучшаемых решений». Этот способ предлагает отказ от поиска единственной наилучшей альтернативы. В результате использования множества Парето происходит сужение множества альтернатив. На этом множестве выбирают наилучшую альтернативу. Примером может служить рассмотренный выше метод решения задачи по критерию "эффективность – стоимость". Кривая на рис. 1.5 является множеством Парето. На этом суженном множестве альтернатив  $x_*$ ,  $C$  с помощью выделения главного критерия  $W_k$  или  $C$  находится оптимальное решение.

Ситуация, когда существует альтернатива, наилучшая для всех критериев, практически не встречается. Различные требования к синтезируемой системе, формулируемые с помощью несколькими критериев, как правило, противоречат друг другу; альтернатива, улучшающая один из критериев, часто обуславливает ухудшение одного или нескольких других критериев.

Среди множества альтернатив можно выделить подмножество альтернатив, обладающих следующим свойством: любая альтернатива этого подмножества лучше какой-либо одной или нескольких других из исходного множества по одному критерию, но хуже их по другим. *Альтернативы, обладающие таким свойством, называют эффективными или Парето-оптимальными.* Выбор с помощью целевой функции оптимальной альтернативы из подмножества эффективных альтернатив позволяет снизить роль субъективных факторов при свертывании критериев.

Приведем примеры эффективных альтернатив. Ограничимся для простоты случаем двух критериев:  $k = 2$ , причем каждый критерий должен стремиться к минимуму. Каждой альтернативе  $x \in X$  можно поставить в соответствие некоторую точку  $\{W_1(x), W_2(x)\}$  на плоскости критериев  $W_1 - W_2$  (рис. 1.6). Таким образом, множество  $X$  отображается в некоторое множество  $Q$ , которое обычно называют множеством

достижимости.

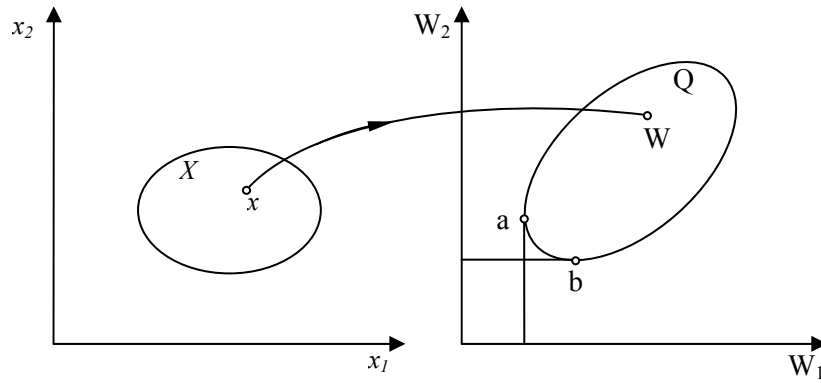


Рис. 1.6. Множество альтернатив и множество достижимости

В случае непрерывного множества  $X$  (рис. 1.7) эффективными являются альтернативы, расположенные на левой нижней границе множества  $X$ . Все остальные альтернативы этого множества "худшие": для каждой из них найдется хотя одна альтернатива  $\tilde{x} \in X$ , которая по всем составляющим векторного критерия будет лучше этой альтернативы.

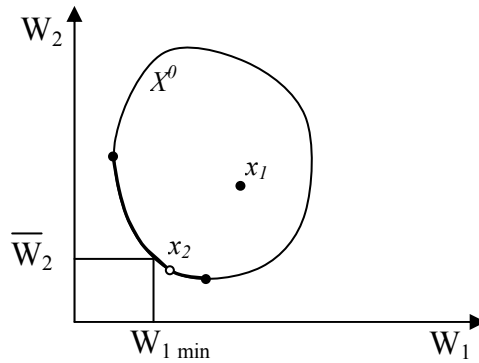


Рис. 1.7. К определению множества Парето-оптимальных альтернатив системы

В частности, на рис. 1.7 альтернатива  $x_1$  хуже, чем  $x_2$ , так как

$$W_k(x_1) > W_k(x_2), \quad k = 1, 2$$

Таким образом, все множество  $X$  можно разбить на два

подмножества:  $X^0$  – худших и  $X^\ominus$  – эффективных альтернатив, причем при синтезе системы альтернатив множество  $X^0$  можно исключить из дальнейшего рассмотрения и решение искать среди альтернатив множества  $X^\ominus$ .

Рассмотрим два метода, которые могут использоваться для построения множества эффективных альтернатив.

1. *Метод рабочих характеристик.* Точки множества  $X^\ominus$  ищутся путем фиксации различных значений всех критериев, кроме одного, и минимизации этого (незафиксированного) критерия. Например, одна точка множества  $X^\ominus$  находится в результате решения следующей задачи оптимизации (рис. 1.7):  $\min W_1$  при  $W_2 = \bar{W}_2$ . Находя решение многократно при различных значениях  $\bar{W}_2$ , получаем зависимость  $W_{1\min}(\bar{W}_2)$ , называемую рабочей характеристикой. Если эта зависимость монотонно убывающая, то множество альтернатив, формирующих рабочую характеристику, совпадает с множеством  $X^\ominus$ .

Случай, когда рабочая характеристика – немонотонно убывающая функция, показан на рис. 1.8. Альтернативы, задающие рабочую характеристику на участке  $x_2x_3$ , неэффективны. Эффективными в данном случае являются альтернативы на участках  $x_1x_2$  и  $x_3x_4$ . В общем случае  $X^\ominus \subset X^P$ , т.е. рабочая характеристика содержит в себе все эффективные альтернативы, но может включать и неэффективные, соответствующие участкам немонотонного убывания рабочей характеристики.

При  $k > 2$  задача построения множества  $X^\ominus$  решается расчетом так называемой рабочей поверхности, сечениями которой являются рабочие характеристики. Точка рабочей поверхности определяется в результате решения задачи оптимизации:

$$\begin{aligned} X^P &= \arg \min_{x \in X} W_k(x, v), \quad k = 1, \dots, m, \\ W_j(x) &= \bar{W}_j, \quad j = 1, \dots, m, \quad j \neq k. \end{aligned} \quad (1.14)$$

Множество  $X^P$  точек  $x^P$ , получаемых в результате решения задачи



(1.14) при всевозможных комбинациях  $\bar{W}_j$ ,  $j=1, \dots, m$ ,  $j \neq k$ , образует рабочую поверхность, совпадающую с эффективной, если все рабочие характеристики, соответствующие этой рабочей поверхности, – монотонно убывающие функции.

2. *Аналитический метод.* Метод применим при  $\kappa=2$ . Рассмотрим задачу минимизации по двум критериям  $W_1(X)$ ,  $W_2(X)$ . Зависимости  $W_1(X)=b_1$ ,  $W_2(X)=b_2$ , где  $b_1$ ,  $b_2$  – константы, описывают линии уровня критериев. Множества точек касания этих линий образуют *компромиссную кривую* [17].

Если вся компромиссная кривая принадлежит множеству допустимых решений  $G$ , то она является множеством Парето. В противном случае в множество Парето входит та часть компромиссной кривой, которая принадлежит множеству  $G$ , а также в него может входить часть границы множества  $G$ , что проверяется непосредственно по определению множества Парето.

Уравнение компромиссной кривой имеет вид

$$\text{grad}W_1(X) = -\lambda \cdot \text{grad}W_2(X), \quad \lambda = \text{const} . \quad (1.15)$$

Рассмотрим *пример*. Пусть в некоторой задаче известны зависимости двух критериев оптимальности от двух подбираемых параметров:

$$W_1 = x^2 + y^2 \rightarrow \min, \quad W_2 = (x-1)^2 + (y-1)^2 \rightarrow \min .$$

Множество допустимых решений задается зависимостями:

- 1)  $G = \{x, y: 0 \leq x \leq 1, 0 \leq y \leq 1\}$ ,
- 2)  $G = \{x, y: 0 \leq x \leq 3/4, 0 \leq y \leq 1\}$ .

Уравнение компромиссной кривой имеет вид

$$\begin{cases} 2x = 2\lambda(1-x) \\ 2y = 2\lambda(1-y) \end{cases} \Rightarrow x = y .$$

Итак, компромиссная кривая представляет собой диагональ квадрата ОА (рис. 1.8).

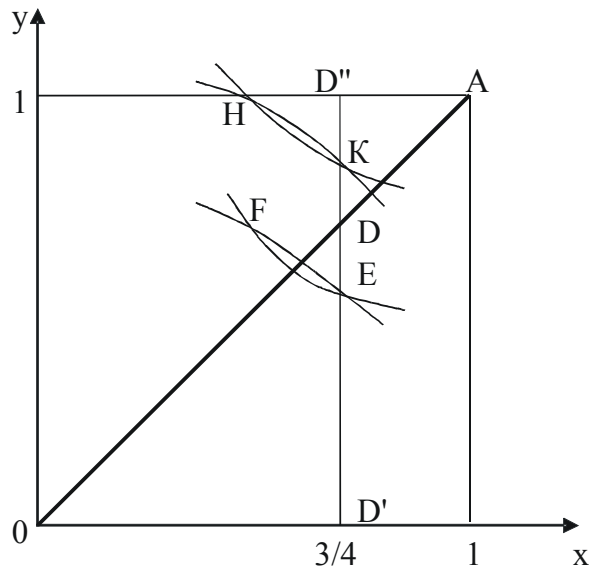


Рис. 1.8. Построение множества Парето аналитическим методом

Для случая 1) множество Парето  $G^* = \{[OA]\}$ , так как прямая  $OA$  целиком лежит в области допустимых решений.

В случае 2) отрезок  $OD \in G$  и входит в множество Парето, но еще нужно проверить границу  $DD''$ . Очевидно, что критерий  $W_1$  есть не что иное, как квадрат расстояния точки с координатами  $x, y$  от точки  $O$ , а критерий  $W_2$  – квадрат расстояния от точки  $A$ .

Проведя линии уровня критериев, видим, что любая точка, принадлежащая сегменту  $FE$ , доминирует точку  $E$ , принадлежащую отрезку  $DD'$ , так как лежит не дальше ее как от точки  $O$ , так и от точки  $A$ . Следовательно, отрезок  $DD'$  не входит в множество Парето.

Напротив, любая точка сегмента  $KH$  доминируется точкой  $K$ , принадлежащей отрезку  $DD''$ , так что отрезок  $DD''$  входит в множество Парето. Таким образом,  $G^* = \{[OD] \cup [DD'']\}$ .

Таким образом, при решении задачи синтеза системы можно сузить множество  $X$  рассматриваемых альтернатив до подмножества  $X^{\ominus}$  эффективных альтернатив, исключив из дальнейшего рассмотрения

определенно худшие альтернативы. Решение о выборе единственной лучшей альтернативы из входящих в множество  $X^{\ominus}$  может быть принято только с использованием дополнительной информации.

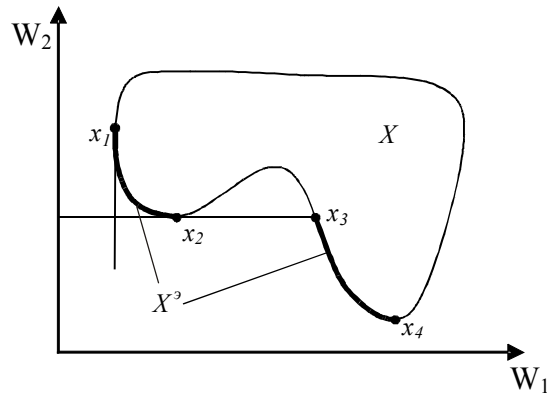


Рис. 1.9. Выпукло-вогнутая "юго-западная" граница области достижимости при оптимизации системы по двум критериям

При построении множества  $X^{\ominus}$  отсеиваются безусловно худшие альтернативы. При этом можно полнее оценить оптимальность решения, найденного в результате свертывания критериев. В ряде случаев оказывается, что все эффективные альтернативы отличаются друг от друга лишь значениями одного или нескольких параметров синтезируемой системы, поэтому для окончательного выбора необходимо меньше дополнительной информации, чем ее потребовалось бы при рассмотрении всего исходного множества альтернатив. Кроме того, если синтезируемая система особенно сложна, то найти единственную строго оптимальную альтернативу часто не удастся, и ограничиваются любым "приемлемым" решением. Желательно, конечно, чтобы это "приемлемое" решение было хотя бы не "худшим", т.е. входило в множество эффективных альтернатив.

Выбор критерия предпочтения тесно связан с выбором метода оптимизации. В достаточно общем виде задача оптимизации сводится к определению такого вектора альтернативы  $x_1^*$  из допустимого множества  $X$ , который обеспечивает минимум целевой функции

$$x_1^* = \arg \min_{x \in X} \Phi(Q(x)), \quad (1.16)$$

если предпочтительны меньшие значения критериев, допустимое множество  $X$  задается совокупностью неравенств (или равенств, или неравенств и равенств одновременно) вида

$$g(x) \leq 0,$$

где  $g(x)$  – вектор-функция, называемая функцией ограничений.

Неравенства  $g(x) \leq 0$  по своему происхождению разделяются на две группы:

а) ограничения на критерии качества, надежности, эффективности и стоимости  $Q_j(x)$ :

$$g_j(x) = Q_j(x) - \bar{Q}_j(x) \leq 0; \quad (1.17)$$

б) непосредственные ограничения альтернатив  $x$ :

$$g_k(x) \leq 0. \quad (1.18)$$

### **1.3. Модели синтеза и анализа сложной технической системы**

#### ***1.3.1. Структура и состав модели***

Как и всякая система, модель для синтеза (анализа) имеет определенную структуру. Эта структура многоуровневая. Каждому критерию (показателю) соответствует свой структурный уровень – блок расчета критерия (показателя). На рис. 1.10 показана схема построения модели для синтеза (анализа) системы по критерию "эффективность – стоимость". Эта модель имеет два структурных уровня – уровень эффективности и уровень стоимости. На каждом структурном уровне можно видеть блоки моделей или отдельные параметры, характеризующие объект синтеза, его среду и операции в аспекте расчета соответствующего критерия (показателя).

Структура каждого уровня модели – иерархическая. На верхнем уровне иерархии находится так называемая "корневая" модель критерия (показателя), параметры которой непосредственно влияют на значение критерия (показателя). Эти параметры, в свою очередь, определяются иерархическими моделями системы, среды и операции.

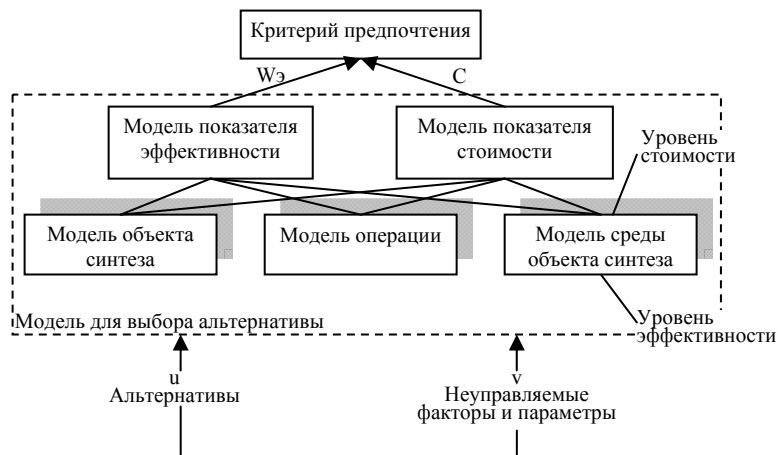


Рис. 1.10. Состав модели синтеза системы по критерию "эффективность-стоимость"

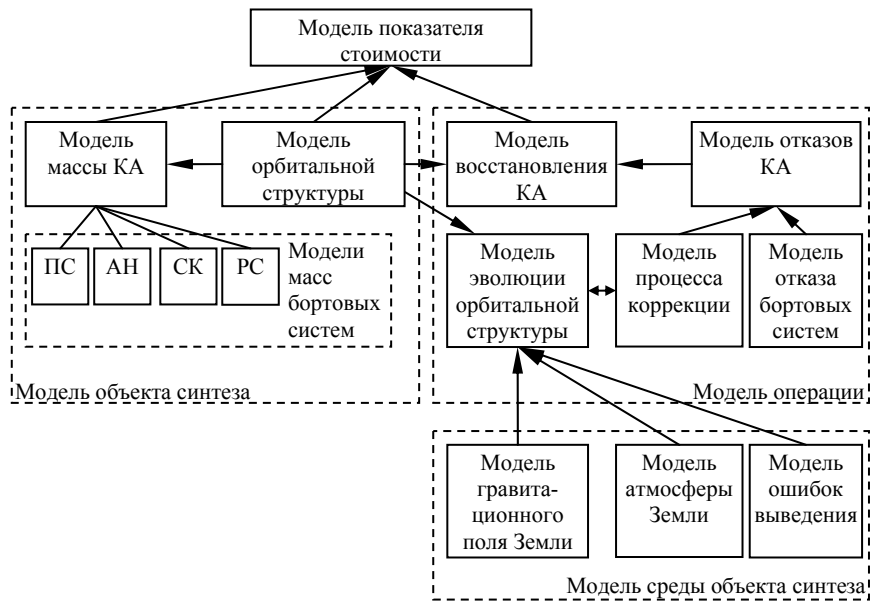


Рис. 1.11. Структура модели стоимости орбитального комплекса (ПС – прочие системы, РС – резервные системы)

В качестве примера приведем схему модели стоимости орбитального

комплекса (рис. 1.11). "Корневой" моделью является  $C_{OK} = C_{KA} N_{KA}$ . Для определения параметров этой модели служат модели системы, операции и внешней среды: стоимость запуска КА вычисляется с помощью модели его массы и удельных характеристик стоимости компонентов КА, а для вычисления количества КА в орбитальном комплексе  $N_{KA}$  за весь срок службы используются модели сети КА, операции и внешней среды.

Таким образом, модель для синтеза (анализа) системы имеет многоуровневую иерархическую структуру. Все структурные уровни модели объединяются в целостную модель посредством критерия предпочтения (сравнения), находящегося на верхнем уровне иерархии и сформированного из критериев (показателей).

Построение модели начинается уже в процессе постановки задачи, когда исследователь получает представление о показателях, характеризующих свойства системы, альтернативах и важнейших определяющих факторах. Разработка модели состоит в установлении системы уравнений и соотношений, описывающих количественные, качественные и логические связи между альтернативами, показателями, формирующими критерий предпочтения, и определяющими факторами и параметрами системы, операции и внешней среды. При разработке алгоритма и программы и проверке модели может происходить корректировка модели.

В простейших случаях в качестве модели могут быть использованы различные математические описания типичных операций и систем с известными характерными свойствами. Подобная ситуация, когда удается воспользоваться готовой математической моделью, на практике встречается крайне редко. В действительности построение модели оказывается достаточно сложной процедурой. При этом используются в качестве компонентов модели уже известные математические модели совместно с вновь разработанными компонентами.

*Целевой характер модели.* Модель, которую создает исследователь,

зависит от моделируемого реального объекта и от того аспекта действительности, который интересует исследователя. Какие именно упрощения примет исследователь, зависит от цели исследования, т.е. от тех вопросов, на которые он хочет получить ответ с помощью модели, и от тех решений, которые предстоит принять руководителю на основе результатов исследований модели. Не существует какой-либо одной универсальной модели для исследуемой системы, которая ответила бы на все вопросы разработки данной системы. Наоборот, исследователь разрабатывает модель применительно к тем или иным вопросам.

Итак, модель, которую создает исследователь, зависит от моделируемого реального объекта и от цели исследования. Для описания одной и той же системы может быть использовано много различных моделей. То, что существенно для одной модели, неважно для другой. Однако все модели обладают следующими общими чертами: они являются целевыми упрощенными отображениями реальности, некоторых аспектов действительности, имитируют реальные системы, учитывая особенности, интересующие исследователя, и пренебрегая другими. Последнее обстоятельство не обесценивает модель, если она используется в соответствии со своим назначением, а только ограничивает область, в которой модель является полезной. Насколько хорошо полученные с помощью модели результаты согласуются с действительностью, зависит от способностей и знаний исследователя, которые проявляются при построении модели – при выборе исходных предпосылок, определяющих параметров и учитываемых взаимосвязей.

Заметим, что модель оказывается всегда полезной для понимания проблемы и использования суждения человека. Она может даже не применяться для расчетов, но при этом быть исключительно полезной для логического анализа. В некоторых случаях анализ системы может быть проведен и без построения модели. Суть анализа состоит тогда в том, чтобы сформулировать возможные альтернативы и сравнить их по показателям свойств системы. Необходимость применения тех или иных

моделей, математических методов и ЭВМ зависит от характера задачи и объема анализируемой информации.

### ***1.3.2. Уточнение цели системы в процессе ее исследования***

Определение цели системы — сложный и практически неформализуемый полностью процесс, для которого требуются интуиция и творческий подход исследователя, выполнение качественного и количественного анализа возможных целей, корректировка, уточнение или изменение целей в процессе исследования системы. Этот итеративный процесс заканчивается одновременно с синтезом системы, и суть его можно охарактеризовать термином "синтез цели системы".

Приведем следующую схему формирования цели системы:

- 1) определение цели системы при ее внешнем проектировании и разработке тактико-технического задания на ОКР по созданию системы;
- 2) описание цели системы по исходным данным из тактико-технического задания в форме списка функциональных требований к системе (формирование модели цели);
- 3) уточнение цели системы в процессе исследования системы путем моделирования.

На рис. 1.12 приведен алгоритм синтеза системы посредством вариантных расчетов. Синтез начинается с назначения в первом приближении структуры и параметров системы. Выбор первого приближения зависит от тех представлений о системе и ее внешней среде, которые имеются у исследователя в начале работы.

Синтез проводится путем последовательных приближений на основе оценки достигаемых значений критериев эффективности, качества, надежности и стоимости по результатам моделирования системы со структурой и параметрами, постепенно приближающимися сначала к удовлетворительным, а затем к предпочтительным (или оптимальным) структурам и параметрам.

Если выясняется, что поставленная перед системой цель не достигается, а ослабление цели недопустимо, то применяют новые средства ее достижения. При этом существенно меняются структура и параметры



системы. Одним из таких средств является расширение состава синтезируемой системы путем включения в нее части компонентов системы вышестоящего уровня.

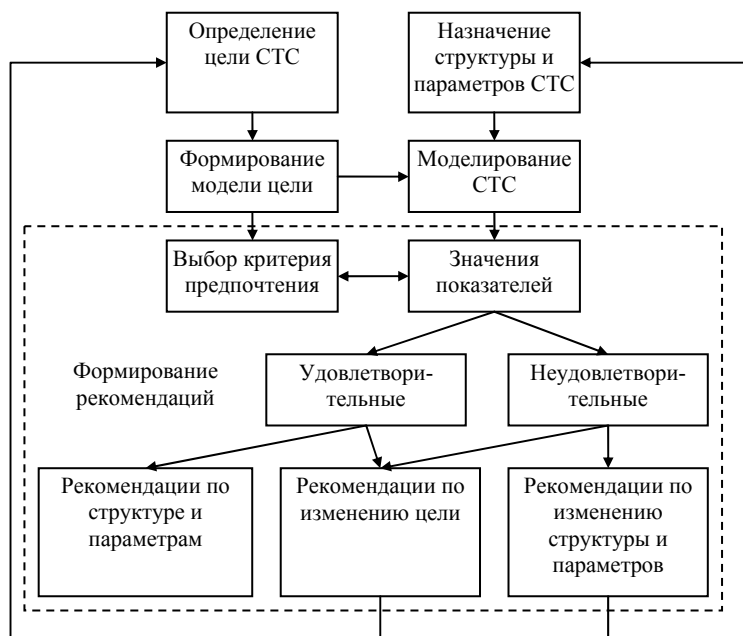


Рис. 1.12. Алгоритм уточнения цели СТС посредством вариантных расчетов

Для примера рассмотрим космическую систему наблюдения, состоящую из ОК, пунктов приема специальной информации (ППИ), центра обработки специальной информации и наземных технических средств. Если при синтезе ОК окажется, что не может быть достигнута требуемая оперативность наблюдения, то в состав синтезируемой системы может быть включена подсистема ППИ. Это означает, что необходимая оперативность будет достигаться путем совместной оптимизации параметров ОК и характеристик ППИ.

Решение задачи синтеза системы будет особенно плодотворным, если исследователь активно ищет новые пути и средства для того, чтобы усилить количественные характеристики цели или чтобы в случае многоцелевой системы сделать совместными некоторые цели, не

являющиеся таковыми при первоначально выбранных средствах. Содействуют ли друг другу цели или мешают, может зависеть от средств, используемых для их достижения. Поэтому часто можно сделать цели совместными, применяя новые средства. При пассивном подходе к синтезу системы приходится жертвовать одной важной целью ради другой, тогда как при некоторой изобретательности можно добиться полезного компромисса между целями системы.

Ранее, когда говорилось о системном подходе, отмечалось требование, чтобы при разработке компонентов системы их цели были подчинены цели системы, другими словами, служили средствами достижения цели системы. При определении цели системы рекомендуется обратное — оценивать возможность достижения цели системы с позиций располагаемых средств (промежуточных целей). В процессе синтеза постоянно происходит выбор и изменение как средств, так и целей системы. Выбранная цель оценивается исходя из необходимых для ее достижения средств. Иначе говоря, выбор цели обусловлен возможностями ее достижения. Эта особенность задачи выбора цели системы заставляет проводить совместный выбор целей системы и ее компонентов в процессе синтеза системы итеративным путем.

Изменение цели в процессе разработки системы и соответственно средств ее достижения вызывается не только трудностями достижения цели, но также изменениями факторов внешней среды и выяснением неопределенностей в ней.

Изменение цели разрабатываемой системы в результате анализа вышестоящих систем, а также изменений во внешней среде или знаний о ней существенно изменяет задачи разработки системы. Особенно сильно на цели создания систем вооружения влияют изменения целей и характеристик систем противника.

Опыт показывает, что цели разрабатываемой системы могут изменяться многократно частично из-за трудностей, возникающих при попытках удовлетворить первоначальному заданию, и частично вследствие того, что изменились цели вышестоящих систем, средством достижения которых должна была служить разрабатываемая система.

Изменение цели может произойти и за время существования системы. Например, научно-технические достижения могут открыть перед

системой новые возможности в результате ее модернизации или сделать цель неосуществимой для данной системы в новых условиях ее работы.

Вышеизложенное свидетельствует о том, что выбор цели системы является слишком сложной задачей, не позволяющей основывать ее решение на интуиции как на единственном средстве. Эта задача решается в процессе анализа и синтеза системы, при этом интуиция может подсказать новые более эффективные средства достижения цели, однако являются ли эти средства действительно необходимыми, на этот вопрос можно ответить только после синтеза системы.

Таким образом, целеобразование – чрезвычайно сложный и ответственный процесс. Сложность определения целей обусловлена трудностями получения необходимых данных о внешней среде, в которой предстоит действовать системе, и о динамике развития внешней среды, а также необходимостью согласования цели системы с целями компонентов в процессе итеративного синтеза системы. Весьма ответственным решением является распределение целей между компонентами систем различных уровней. От целей зависит выбор показателей эффективности и построение моделей, необходимых для проведения математического моделирования и синтеза системы. Цели системы и ее компонентов, в конечном счете, определяют облик, структуру и основные характеристики системы.

### ***1.3.3. Адекватность и работоспособность модели***

Модели – это специальные системы, и при построении модели исследователю следует рассматривать ее как систему. Это означает, в первую очередь, что модель имеет цель, соответствующую цели исследования. Одну и ту же реальную систему можно представлять различными моделями в зависимости от цели исследования. Назначение модели содержится в постановке задачи, где указывается цель исследования, множество альтернатив и критерий выбора предпочтительной альтернативы.

Как объект разработки модель должна удовлетворять определенным конкретным требованиям, зависящим от свойств реальной системы, постановки задачи (цели исследования) и предполагаемого способа

решения задачи.

Рассмотрим основные требования, которым должны удовлетворять все модели вне зависимости от решаемой задачи.

Главное требование — служить для исследователя средством решения задачи. Для этого необходимо, с одной стороны, обеспечить представительность (достоверность) модели — ее способность отражать свойства реальности достоверно, т.е. истинно, достаточно полно и точно, а с другой — обеспечить умеренную сложность модели для ее работоспособности. Эти требования противоречивы — повышение представительности достигается за счет усложнения модели. Требование представительности вынуждает исследователя строить как можно более полную, подробную и точную модель: учитывать на каждом иерархическом уровне системы больше факторов, увеличивать число рассматриваемых уровней иерархии, развивать модель вширь и вглубь. Требование умеренной сложности заставляет сокращать число учитываемых факторов, как по ширине, так и по глубине иерархии системы. Бессмысленно требовать максимальную представительность при минимальной сложности. Компромисс между представительностью и сложностью достигается адекватной моделью, т.е. моделью, вполне соответствующей цели моделирования и реальному объекту.

Такое понимание адекватности не полностью совпадает с понятием представительности (истинности, полноты и точности). Адекватность означает, что требования представительности выполняются не вообще, а лишь в той мере, которая необходима для достижения цели исследования. Задача исследователя — правильно выбрать необходимый уровень представительности. Эта задача решается с помощью принципа существенности: в модели должны учитываться только факторы, существенные по отношению к цели моделирования (вспомним целевой характер модели). Чтобы решить, какие факторы реального объекта включить в модель, а какие — нет, считая их несущественными, исследователь должен искусно сочетать интеллектуальную, творческую работу с использованием научных методов, эвристических приемов и рекомендаций.

В ряде случаев удастся ввести некоторую меру представительности модели, т.е. указать способ сравнения моделей по степени успешности

достижения цели моделирования. Если мера представительности выражена количественно, то можно решать задачу улучшения модели, нахождения наиболее представительной модели, исследовать зависимость представительности от полноты и точности модели и т.д.

Так, например, применительно к задаче синтеза системы принцип существенности можно конкретизировать следующим образом: в модели должны учитываться те факторы и параметры, которые существенно влияют на значения критериев.

Решить вопрос о том, что является в данной модели существенным, а что – нет, можно следующим образом:

- по данным научно-технической литературы;
- неформальным путем на основе качественного анализа реальности;
- посредством эвристических методов;
- с помощью суждений исследователя;
- с помощью вычислений при проверке модели.

Кроме выбора уровня представительности с помощью принципа существенности и, следовательно, необходимого уровня сложности, требуется определить возможный уровень сложности модели. Эта задача решается оценкой работоспособности модели и согласованием модели со средой, в которой ей предстоит работать.

Работоспособность модели означает, что с ее помощью можно получить необходимое решение достаточно легко, т.е. в установленный срок при располагаемых затратах ресурсов. Работоспособность оценивается на этапе проверки модели.

Непротиворечивость логике полученных результатов исследований при изменении важнейших параметров является одним из свидетельств правильности модели. Чтобы проверить это, необходимо проанализировать характер реакции модели на изменения соответствующих входных параметров, в частности, на значения параметров, близких к экстремальным. Совпадения результатов для некоторых частных и экстремальных случаев с известными, обычными результатами также говорят о непротиворечивости модели.

Чувствительность модели к альтернативам означает, что при небольшом изменении альтернатив происходит такое изменение

показателей свойств системы, которое можно обнаружить в условиях погрешности вычислений. Анализ чувствительности модели требует проведения вычислений. Правильная модель должна обладать необходимой чувствительностью.

Реалистичность, т.е. правильность, модели подтверждается ее соответствием тем частным случаям, для которых уже имеются фактические данные или апробированные на практике модели.

Одна из ошибок моделирования состоит в применении модели без проверки ее на достоверность, что нередко встречается в инженерной практике.

#### ***1.3.4. Определяющие параметры и объект моделирования***

Определяющими будем называть такие факторы, которые существенно влияют на результаты решения задачи. Они могут быть качественными и количественными. Примером качественных факторов могут служить структурные связи. Определяющие факторы, описываемые количественно, будем называть определяющими (или представительными) параметрами.

Приведем классификацию определяющих параметров:

1) по характеру математического описания:

- детерминированные величины;
- параметры, характеризующие случайные величины (математическое ожидание, дисперсия, вероятность события и др.);
- границы (максимальные и минимальные значения) неопределенных величин;

2) по отношению к исследованию:

- управляемые (выбираемые);
- заданные, не изменяемые при исследовании;

3) по отношению к системе:

- параметры системы;
- параметры внешней среды.

Упрощение задачи исследования состоит в том, чтобы из множества компонентов системы и ее среды в исследование включить те компоненты, от характеристик которых существенно зависят результаты решения задачи

(например, выполнение требований к системе при ее синтезе).

Совокупность компонентов системы и ее внешней среды, характеристики которых существенно влияют на решение задачи, назовем объектом моделирования. Как видно, в объект моделирования входят те компоненты системы и среды, структуры и параметры которых признаны определяющими.

### ***1.3.5. Принцип декомпозиции - субоптимизации***

Многие задачи исследования и проектирования систем очень сложны. На практике разделяют некоторую общую задачу на ряд частных задач и каждую из них решают как самостоятельную. Операцию разделения целого на части называют декомпозицией. Так как все частные задачи взаимосвязаны, то смежные задачи считаются решенными. Путем итераций, решая одну частную задачу за другой, приходят к решению общей задачи. При этом стремятся обеспечить системность решений частных задач — их согласованность и подчиненность достижению общей цели исследования.

Для решения любой сложной задачи целесообразно воспользоваться методом системного анализа, называемым декомпозицией - субоптимизацией.

В каждой частной задаче выполняется своя частная оптимизация - субоптимизация. Оптимальные решения частных задач, полученные независимо, могут оказаться не согласующимися или даже противоречащими друг другу. Чтобы этого не случилось, необходим системный подход, который в данном случае выражается в подчинении критериев предпочтения в частных задачах критериям в вышестоящих по иерархии задач. В результате будет получена иерархическая система задач и критериев для них, на верхнем уровне которой будет некоторая общая задача и критерий предпочтения для ее решения.

При субоптимизации связи между задачами, т.е. входные величины решаемой частной задачи – выходные из смежной, зафиксированы. Так как на решения частных задач влияют не только критерии, но и связи между задачами, т.е. исходные данные, поступающие из других задач, то путем итераций надо учесть влияние зафиксированных вначале связей – влияние

изменения исходных данных.

Теперь можно сформулировать принцип решения сложной задачи на основе декомпозиции - субоптимизации:

- декомпозиция общей задачи — построение иерархической системы частных задач и критериев предпочтения;
- оптимизация решений частных задач;
- усовершенствование частных решений с учетом взаимосвязей между частными задачами.

В результате совокупность оптимальных частных решений, полученных на основе изложенного системного подхода, будет представлять собой целостное решение общей задачи.

Цель декомпозиции задачи вообще – упростить решение сложной проблемы, обеспечив целостность (системность) совокупности частных решений. Рассматривая отношение проектанта или исследователя к декомпозиции, можно указать более конкретные цели.

Цель декомпозиции – расчлнить общую задачу проектирования СТС, чтобы организовать процесс проектирования, распределив частные задачи между подразделениями ОКБ. При этом декомпозиция общей задачи проектирования основывается на опыте и традициях проектирования в данном ОКБ и на свойствах объекта проектирования.

Так, например, задача проектирования КС для исследования Солнца на этапе технического предложения может разделяться на следующие задачи:

- определение основных характеристик космического комплекса;
- выбор состава и размещение научной аппаратуры;
- проектирование бортового радиотехнического комплекса;
- проектирование бортового комплекса управления;
- разработка программного обеспечения БЦВМ;
- проектирование системы электропитания;
- разработка программы обеспечения надежности;
- анализ эффективности КС и другие.

При решении частной задачи некоторой проблемы исследователям также приходится обращаться к декомпозиции. При этом возможны два варианта декомпозиции.



1. Исследователю может быть поручена достаточно сложная частная задача. Тогда его целью будет упрощение решения своей задачи путем ее декомпозиции - субоптимизации.

2. Исследователю поручена частная задача, не требующая для ее упрощения дальнейшей декомпозиции. Однако, чтобы реализовать системный подход к решению частной задачи и выбрать критерий предпочтения, исследователь должен рассмотреть некоторую надзадачу и выполнить ее декомпозицию так, чтобы порученная частная задача была подчинена надзадаче. Тогда решение порученной исследователю задачи будет частью решения надзадачи.

Основанием декомпозиции обычно является некоторая модель, которая называется «моделью – основанием». Объект декомпозиции сопоставляется с моделью-основанием. Каждому компоненту объекта декомпозиции, например, общей задаче проектирования системы, соответствует компонент модели-основания.

Можно указать следующие виды моделей-оснований.

1. Модель состава или структуры системы. Каждому компоненту системы ставится в соответствие задача его исследования. Структурным связям соответствуют связи между частными задачами.

2. Модель операционной структуры системы, когда декомпозиция системы осуществляется по функциональному признаку, т.е. по общности частных задач, решаемых совокупностью элементов.

3. Логическая модель операции.

4. Организационная структура проектирования системы, т.е. разделения труда между организациями и их подразделениями.

5. Математическая модель системы. В некоторых случаях сложную математическую модель удается разделить на компоненты, которым соответствуют частные задачи исследования. Классическим примером является разделение задачи исследования возмущенного движения ЛА на задачи исследования продольного и бокового движений на основании свойств уравнений возмущенного движения, обусловленных симметрией ЛА относительно продольной плоскости.

6. Модель состава объекта моделирования, полученная из моделей состава системы и среды путем исключения компонентов, не влияющих на решение общей задачи.

### ***1.3.6. Разновидности неопределенностей***

О неточных и неполных знаниях принято говорить как о неопределенности знаний. Одна из основных трудностей при построении моделей состоит в неопределенности многих характеристик системы и внешней среды. Так, например, для создателей КА, спускаемых на Луну, Венеру и Марс, в свое время были неизвестны характеристики грунта Луны, Марса и Венеры и характеристики атмосферы этих планет; при создании систем вооружения обычно недостаточно известны характеристики систем противника.

Следует различать неопределенности в настоящем и будущем. Когда исследуется система, призванная функционировать в будущем, что обычно имеет место при проектировании системы, количество неопределенных факторов существенно возрастает из-за развития во времени внешней среды и технических средств, используемых при создании системы.

В зависимости от полноты сведений о рассматриваемом факторе различают пять ситуаций его описания:

- 1) определенность, когда фактор задается числом (параметром);
- 2) риск – фактор описывается законом распределения плотности вероятности или функцией распределения;
- 3) неопределенность интервальная – фактор описывается диапазоном его возможных значений;
- 4) неопределенность диапазона фактора – известно только, что фактор существует;
- 5) неосведомленность – не известно, имеет ли место какой-либо фактор в изучаемой действительности.

Основной проблемой при подготовке исходных данных для исследования системы является устранение неопределенностей. Можно указать следующие пути устранения неопределенностей параметров системы и среды, не считая такого тривиального способа, как поиск, сбор и обработка дополнительной информации:

1. Проведение специальных исследований, теоретических или экспериментальных, для изучения неопределенных факторов и явлений. Такие исследования проводятся по заказам проектировщиков специалистами в соответствующих областях науки. Для исследований, как

правило, создаются экспериментальные стенды и установки. Однако могут быть такие системы, для которых невозможно проведение экспериментов. Поэтому нужны также теоретические методы устранения неопределенностей исходных данных и в особенности неопределенностей в будущем.

2. Синтез правдоподобных значений неизвестных характеристик внешней среды по некоторому малому объему информации, которым располагает исследователь. Здесь имеет место аналогия с известными методами оценивания с помощью фильтрации параметров динамической системы по данным измерений.

3. Использование для устранения неопределенностей в будущем методов научно-технического прогнозирования.

4. Применение методов экспертных оценок. Экспертные оценки используются для получения не только оценок неопределенных параметров, но и недостающих данных.

Не всегда удается полностью устранить неопределенности, и тогда приходится учитывать оставшиеся неопределенности при выборе критерия предпочтения, построении модели, разработке метода оптимизации, выборе проектных решений при разработке компонентов сложных систем.

На практике чаще всего при учете неопределенностей используют вычисления для различных вариантов возможных значений параметров системы и среды.

Для ситуации интервальной неопределенности разработаны минимаксные методы оптимизации и предложены для принятия рекомендаций критерии Вальда, Лапласа, Сэвиджа и Гурвица. Эти методы базируются на том, что причиной неопределенностей является "природа", которая не стремится причинить вред человеку, использующему систему для достижения своей цели.

Встречаются ситуации, когда источником неопределенностей выступает разумный противник, интересы которого противоречат интересам лица, разрабатывающего систему. В этом случае может использоваться теория игр, которая предлагает модели и методы оптимизации для большого числа разнообразных конфликтных ситуаций.

Если модели, предлагаемые теорией, не являются адекватными рассматриваемой задаче, исследователем могут быть разработаны

специальные модели, учитывающие неопределенности внешней среды и будущей системы. В частности, для построения модели, предусматривающей неопределенность, разрабатываются сценарии возможного, предполагаемого, ожидаемого функционирования системы и вводятся гипотезы, допущения, учитывающие неопределенность.

Одним из перспективных направлений учета неопределенностей является реализация адаптивных проектных решений. Выбор таких решений предусматривает возможность изменения (адаптации) параметров компонентов системы в зависимости от конкретных ситуаций во внешней среде.

### ***1.3.7. Функциональные требования и программные продукты***

Программное обеспечение – это совокупность программ, необходимых для анализа и синтеза системы с помощью ЭВМ и зафиксированных на носителях информации. Программное обеспечение является системой, т.е. объектом, обладающим системными свойствами. Разрабатывая программное обеспечение, к нему следует относиться как к системе, обладающей назначением, составом компонентов с различными функциями, структурой.

Под термином "программа" будем понимать общее название для различного вида программ, их фрагментов, подпрограмм, процедур, подпрограмм-функций, программных модулей. Назначением программы является выдача предписаний ЭВМ на выполнение последовательности действий для решения задачи моделирования с особыми условиями в виде функциональных требований к программе: экономичность, правильность, надежность.

*Экономичность программы* оценивается затратами вычислительных ресурсов — машинного (процессорного) времени и оперативной памяти. Эти свойства зависят от размерности задачи, которая определяется, в первую очередь, свойствами математического обеспечения (совокупности математических моделей, алгоритмов). Неудачная программная реализация может существенно увеличить машинное время и оперативную память. Недостаточная экономичность программы обычно оказывается основным фактором, ограничивающим возможности исчерпывающего анализа и

синтеза систем.

*Правильность программы* — правильность реализации в программе выбранного математического обеспечения.

*Надежность программы* — свойство выполнять заданные функции в заданных условиях. Функции и условия формулируются в терминах той предметной области, к которой относятся исследуемые системы. Основным показателем надежности — вероятности получения правильного результата при использовании программы в условиях, сформулированных потребителем.

Программа может быть правильной, но мало надежной. Несоответствие границ применения математического обеспечения требованиям пользователей снижает надежность, но не влияет на правильность. Например, если в программе безошибочно реализован метод Ньютона, то программа будет правильной, но, возможно, ненадежной, так как условия сходимости метода Ньютона будут выполняться не для любого объекта при произвольном задании начального приближения.

Условия использования программы характеризуются ограничениями на ее применение. Эти ограничения могут относиться к типам и составам исследуемых систем, диапазонам числовых значений параметров системы и среды, перечню выполняемых операций и процедур. Надежность зависит от знания границ применения разработанной программы. Если плохо известны границы применения программы будут уже требуемых границ, то программа окажется в некоторой степени ненадежной. Для достижения высокой надежности надо при отладке и эксплуатации программы тщательнее и полнее выявлять и оговаривать ограничения.

Полностью завершенная, отлаженная программа, выполняющая заданные вычислительные функции, удовлетворяющая функциональным требованиям, может быть преобразована в более полезный продукт, называемый программным продуктом.

*Программный продукт* – это программа, обладающая следующими свойствами:

- тщательно отлаженная с установленными четкими границами области ввода и высокой надежностью;
- универсально используемая в различных ситуациях с расширенными

границами области ввода для многих наборов данных;

- доступная любому пользователю (чтобы программу можно было пропускать, отлаживать, улучшать, расширять);
- снабженная тщательной документацией с тем, чтобы любой пользователь мог применять программу и совершенствовать ее.

Для создания программного продукта необходимо настолько расширить границы области ввода, насколько это позволяет основной алгоритм. Для этого следует программу тщательно отладить, учитывая все влияющие на нее факторы: следует подготовить, пропустить и зафиксировать значительный массив отладочных тестов, изучив область ввода и установив ее границы. После затрат труда не менее чем на разработку исходной программы, последняя перейдет в новое качество — программный продукт.

*Программный комплекс* (ПК) представляет собой совокупность взаимодействующих программ, согласованных по функциям и по форматам, так что их объединение представляет собой единое средство для решения больших задач. Пример состава и функциональной структуры программного комплекса приведен на рис. 1.13.

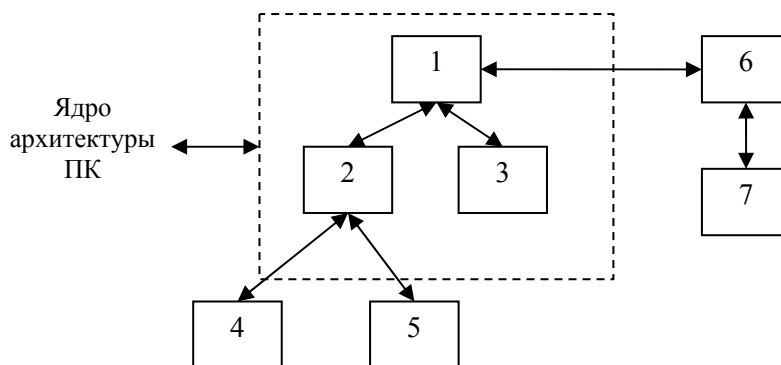


Рис. 1.13. Пример программного комплекса для моделирования движения РН: 1 - управляющий модуль; 2 - расчетный модуль; 3 - графический модуль; 4 - блок моделирования бортовых систем; 5 - блок моделирования воздействий внешней среды; 6 - модуль доступа к базе данных; 7 - база данных

Завершенная программа может стать частью ПК, если она доработана так, что:

- каждый вход и выход соответствуют (по синтаксису и семантике) точно определенным правилам;
- программа использует только отведенные ей ресурсы — объем памяти, устройства ввода-вывода, машинное время;
- программа отлажена во всех возможных сочетаниях.

Эта отладка должна быть очень большой по объему, что требует много времени, ибо очень тонкие ошибки появляются из-за непредвиденных взаимодействий отлаживаемых частей.

*Требования к программному продукту* — это по существу эксплуатационные требования.

*Удобство* использования программы — определяется ее надежностью, наличием проблемно-ориентированных языков и средств диагностики ошибок пользователя.

*Универсальность* программы — чем шире ограничения на применение программы, тем выше ее универсальность и ниже надежность, поскольку границы применения программы недостаточно определены. За повышение степени универсальности приходится платить снижением надежности. Эксплуатация нескольких узкоспециализированных, но надежных программ вместо применения одной универсальной программы повышает надежность.

*Открытость* программы — характеризуется возможностями внесения в нее изменений в процессе эксплуатации. Понятие открытости близко к понятию адаптивности, под которым подразумевается возможность модификации программы для поддержания ее работоспособности и эффективности в изменяющихся условиях применения.

*Сопровождаемость* программы — характеризуется удобством поддержания программы в работоспособном состоянии, взаимодействием с пользователями при внедрении, настройке, адаптации к условиям применения, устранением обнаруженных при эксплуатации ошибок, введением новых функций, модернизацией программы и обеспечивается эксплуатационной документацией.

*Мобильность* или *переносимость* программы — свойство программы, обеспечиваемое системами программирования, которое

заключается в возможности выполнения программ без изменений различными ЭВМ с различной архитектурой. Программы, записанные на машинно-ориентированных языках, не переносимы. Использование языков высокого уровня создает предпосылки для создания переносимых программ.

#### **1.4. Методы неформализованного анализа**

При исследовании сложных систем возникают проблемы, не разрешимые с помощью формальных математических методов. Такие проблемы особенно характерны для слабоструктурированных систем. Тогда используют ряд эвристических методов, неформализованных математически, или пытаются строить модели. При этом необходимы измерительные шкалы для проведения экспериментов с качественными факторами.

Результаты эксперимента регистрируются с помощью измерений — операции, в ходе которой находится отношение измеряемой величины к другой однородной величине, принимаемой за единицу измерения. Это классическое понимание измерений.

Измерения производятся с помощью измерительных шкал. Результаты экспериментов, описания технических объектов, исходные данные для моделирования, результаты моделирования даются в терминах той или иной измерительной шкалы.

Исследователь всегда стремится построить достоверную модель. Первый признак возможности построения достоверной математической модели — возможность надежного измерения всех параметров системы. Такая ситуация обычно имеет место в инженерном деле.

Основная проблема, с которой сталкивается исследователь, пытающийся для слабоструктурированной системы построить достоверную модель, — проблема измерений. Нельзя построить достоверную модель при ненадежном способе измерения.

В слабоструктурированных системах может быть большое число параметров, для которых не существует точных способов количественных измерений, нет эталонов для единиц измерений, например, нет общепринятого количественного измерителя. Такое же положение существовало когда-то и в физике. Например, до изобретения градусника



не было шкалы измерения температуры, поэтому измерения были качественные. Затем были изобретены эталоны для единиц измерений и шкалы.

Современное понимание эксперимента расширилось. Под экспериментом теперь понимается не только опыт с количественными измерениями его результатов, но и наблюдения, результаты которых не допускают числовой меры, но фиксируются в "качественных" шкалах. Эти результаты можно учитывать в моделях, получая качественные, но вполне научные выводы. Современное понимание измерений стало шире классического:

1) измерения могут носить качественный характер, а не только количественный;

2) измерение обычно сопровождается неизбежными погрешностями (погрешности измерений являются не только чем-то чуждым, посторонним для измерений, но и неотъемлемым, естественным и неизбежным свойством самого процесса измерения);

3) интересующая нас величина часто не наблюдаема и поддается лишь косвенным измерениям.

Измерение — операция, которая наблюдаемому состоянию объекта или процесса ставит в соответствие один из элементов подходящей измерительной шкалы.

Измерительные шкалы — множество обозначений (чисел, номеров, символов), используемых для регистрации состояния наблюдаемого объекта.

#### ***1.4.1. Классификационные шкалы***

Пусть эксперимент таков, что мы не можем выполнить никакие количественные измерения, а только самые простые качественные оценки. Простейшая качественная оценка — это когда можно сказать, различимы или нет любые два объекта по какому-либо признаку.

Неотличимым друг от друга объектам по состоянию или признаку присвоим одинаковые обозначения, наименования классов. Это означает, что отношения между состояниями объекта, а также между их обозначениями удовлетворяют аксиомам тождества:

- 1) либо  $A=B$ , либо  $A \neq B$ ;
- 2) если  $A=B$ , то  $B=A$ ;
- 3) если  $A=B$  и  $B=C$ , то  $A=C$ .

(1.19)

Символ "=" обозначает отношение эквивалентности (когда  $A$  и  $B$  — числа, он означает равенство).

Итак, каждому классу состояний — признаков поставим в соответствие обозначение (символ). Указанное множество символов образует шкалу. Для символов, т. е. для обозначения классов, могут быть использованы слова естественного языка.

Примеры шкал: типы автомобилей — легковая, грузовая, автобус, цвета объекта — красный, желтый, зеленый и др.

Такие шкалы называются шкалами наименований или классификационными шкалами.

Измерение в шкале наименований будет состоять в том, чтобы определить принадлежность объекта к тому или иному классу и записать это с помощью символа, обозначающего данный класс.

Обозначения классов — это только символы, даже если для этого использованы номера, которые лишь внешне выглядят как числа, но на самом деле числами не являются, например номера спортсменов, указывающих принадлежность к команде.

С номерами нельзя обращаться как с числами, за исключением определения их совпадения (равенства), т.е. принадлежности к одному классу, или несовпадения, т.е. принадлежности к разным классам.

Поэтому при измерениях в шкале наименований непосредственно с результатами измерений можно выполнять только операцию их совпадения или несовпадения (это первичная обработка), но с результатами этой операции можно выполнять более сложные преобразования (это вторичная обработка):

- считать количество наблюдений  $k$ -го класса  $n_k$ ;
- вычислять относительные частоты классов  $p_k = \frac{n_k}{\sum n_k}$ ;
- сравнивать частоты между собой, находя, например, моду-символ наиболее часто встречающегося класса;

- выполнять различные статистические процедуры над относительными частотами.

Шкалу наименований естественнее всего использовать для дискретных по природе явлений (например, для различных объектов). Необходимость в классификационной шкале возникает и в тех случаях, когда состояния объекта образуют непрерывное множество (например, возраст людей). Тогда задача сводится к предыдущей, если все множество состояний разбивается на конечное число подмножеств, искусственно образуя тем самым классы состояний. Принадлежность состояния к какому-либо классу можно теперь регистрировать в шкале наименований.

Измерения в шкале наименований широко применяются на практике. Такая шкала используется в статистике, например промышленной, при переписи населения, в разных видах учета (товаров, специалистов, приборов), в социологических исследованиях, при проведении выборов и т. д.

#### **1.4.2. Порядковые шкалы**

В тех случаях, когда наблюдаемый (измеряемый) признак состояния имеет природу, позволяющую определять не только принадлежность объекта к одному из классов, но и дающую возможность в каком-то отношении сравнивать разные классы, для измерений можно выбрать более сильную шкалу, чем классификационная. Если не воспользоваться этим, то мы откажемся от части полезной информации.

Однако усиление измерительной шкалы зависит от того, какие отношения между классами существуют в действительности. Это и явилось причиной появления измерительных шкал разной силы.

Следующей по силе за классификационной шкалой идет порядковая или ранговая шкала. Этот класс шкал появляется, если, кроме аксиом тождества 1—3, классы удовлетворяют следующим аксиомам упорядоченности:

$$4) \text{ если } A > B, \text{ то } B < A; \quad (1.20)$$

$$5) \text{ если } A > B \text{ и } B > C, \text{ то } A > C.$$

Обозначив такие классы символами и установив между этими символами те же отношения порядка, что и в аксиомах 4 — 5, получим

порядковую шкалу. Примеры порядковых шкал: нумерация очередности, воинские звания, места в соревнованиях.

Порядковые шкалы применяются для определенного набора сравниваемых объектов, у которых нет общепринятого стандарта.

Отношение порядка ничего не говорит о дистанции между сравниваемыми классами. Поэтому порядковые экспериментальные данные, даже если они изображены цифрами, нельзя рассматривать как числа.

Итак, при измерениях в порядковых шкалах первичная обработка данных должны основываться только на допустимых для этих шкал операциях — вычислении ранга.

Во вторичной обработке с рангами можно работать так же, как с квалификационными классами и, кроме того, определять медиану; находить квантили любого уровня.

Если состояния, которые допускают только порядковые сравнения, измеряются в числовых шкалах через величины, связанные с ними косвенно, то эти измерения все равно остаются измерениями в порядковой шкале. Например, спортивные соревнования, при которых измеряется время, затрачиваемое на преодоление дистанции. Здесь время, хотя и измеряется в числовой шкале, но как мера распределения спортсменов по местам принадлежит порядковой шкале.

### ***1.4.3. Модифицированные порядковые шкалы***

Желание уменьшить относительность порядковых шкал и придать им некоторую объективность вызвало создание модифицированных порядковых шкал с некоторым (чаще всего кажущимся) усилением. Другая причина попыток усиления шкал – многие измеряемые в порядковых (дискретных) шкалах величины имеют непрерывный характер: сила ветра, сила землетрясения, твердость вещества и др. Приведем некоторые примеры таких шкал.

Шкала твердости: из двух минералов тверже тот, который оставляет на другом царапины или вмятины при достаточно сильном соприкосновении (отношение "А тверже В" — типичное отношение порядка). Эталоны: 1 – тальк ,..., 5 – апатит ,..., 10 – алмаз . Оценки

твердости не носят числового характера.

Шкала силы ветра: баллы присваиваются по характеру волнения моря. Эталоны: 0 — штиль,..., 6 - сильный ветер,..., 12 – ураган. Кроме штиля, градации силы ветра имеют условный, качественный характер.

Шкала землетрясений: по последствиям землетрясений (12 - балльная шкала по Рихтеру).

#### ***1.4.4. Интервальные шкалы***

Если упорядочение объектов можно выполнить настолько точно, что известны интервалы (расстояния) между двумя из них, то измерение окажется заметно сильнее, чем в шкале порядка. Естественно выразить все расстояния в единицах, хотя и произвольных, но одинаковых по всей длине шкалы. Это означает, что объективно равные интервалы измеряются одинаковыми по длине отрезками шкалы.

Для измерения интервалов могут быть использованы произвольные единицы их протяженности и произвольные начала отсчета. Построенные таким образом шкалы называются интервальными.

В интервальной шкале измеряются, например, температура – по Цельсию; время – интервал от какого-либо начала, например от начала века, высота местности — интервал от уровня моря, возраст человека — интервал от дня рождения.

Интервальные шкалы не имеют абсолютного нуля, но допускают свободу выбора начала отсчета и единицы длины интервала. В этой шкале только интервалы имеют смысл чисел, над которыми можно выполнять операции сложения, вычитания, умножения и деления на число. Отсчеты в этой шкале не являются числами.

В интервальной шкале новой допустимой операцией над наблюдениями является определение интервала между ними, что является первичной обработкой.

Над интервалами можно выполнять любые арифметические операции, а вместе с ними — использовать подходящие способы статистической и иной обработки данных. Это есть вторичная обработка.

Среднее значение измерений является относительной величиной из-за произвольного начала, но средний интервал уже имеет смысл, а

среднеквадратическое отклонение — объективный смысл, так как разность  $x - x_{cp}$  является интервалом.

#### **1.4.5. Периодические или циклические шкалы**

В этих шкалах также измеряются интервалы. В такой шкале нуль повторяется через период. Приведем примеры таких шкал: измерение направления из одной точки (шкала компаса), измерение времени суток (циферблат), измерение фазы колебаний (в градусах). В пределах одного периода обработка данных здесь производится так же, как и в интервальной шкале.

#### **1.4.6. Абсолютная шкала**

Такая шкала имеет абсолютный нуль и абсолютную единицу. Этими качествами обладает числовая ось, которую естественно назвать абсолютной шкалой, так как она единственна и уникальна. Важная особенность данной шкалы по сравнению с другими шкалами — это безразмерность и абсолютность ее единицы.

Числовая ось используется как измерительная шкала в явной форме только при счете количества предметов, а как вспомогательное средство присутствует во всех остальных шкалах.

#### **1.4.7. Шкалы отношений**

В физике, астрономии, технике и в других точных науках широко используются шкалы отношений. В этих областях научной и практической деятельности большей частью имеют дело с величинами, которые удовлетворяют аксиомам тождества, упорядоченности и следующим аксиомам аддитивности:

- 6) если  $A=P$  и  $B > 0$ , то  $A+B > P$ ;
  - 7)  $A+B=B+A$ ;
  - 8) если  $A=P$  и  $B=Q$ , то  $A+B=P+Q$ ;
  - 9)  $(A+B)+C=A+(B+C)$ .
- (1.21)

Этим аксиомам удовлетворяют также и числа. Следовательно, измерения физических величин, удовлетворяющих этим аксиомам,

являются числами, и с данными измерений можно выполнять любые арифметические действия (такие как вычитание, умножение и деление — частные случаи сложения).

Шкалы для измерения величин, удовлетворяющих аксиомам 1 — 9, называются шкалами отношений. Такое название вытекает из того, что этот класс шкал имеет следующую особенность: отношение двух значений измеряемой величины не зависит от того, в какой из шкал произведены измерения, т.е.

$$\frac{x_1}{x_2} = \frac{y_1}{y_2}.$$

Величины, измеряемые в такой шкале, имеют естественный абсолютный нуль, но при этом существует свобода в выборе единиц.

Приведем примеры величин, природа которых соответствует шкале отношений: длина, вес, электрическое сопротивление, деньги и многие другие.

#### ***1.4.8. Качественные или неметрические показатели***

Среди важных свойств систем, которые необходимо учитывать при решении задачи выбора, встречаются и такие, которые выражаются лишь в качественной форме. Соответствующие этим свойствам характеристики называются качественными или неметрическими показателями.

Рассмотрим, например, воздействие системы на окружающую среду в процессе ее функционирования.

Под экологической совместимостью системы со средой понимается ее функционирование и утилизация без нанесения недопустимого ущерба среде обитания людей и непосредственно самим людям.

При создании, функционировании и утилизации радиосистемы ущерб людям может наноситься самыми разнообразными путями, например:

- генерацией интенсивных электромагнитных излучений или высоких электрических напряжений, наносящих вред здоровью людей, животному или растительному миру;
- какими-либо нарушениями техники безопасности (например, повышенной взрывоопасностью);

- чрезмерным расходом ограниченных природных ресурсов — энергии и веществ (особенно редких, включая золото, серебро, платину и другие редкие металлы и дефицитные вещества);
- загрязнением среды в процессе утилизации системы (т. е. после окончания срока ее службы).

Показатели, которые должны количественно учитывать степень удовлетворения требований экологической совместимости, весьма разнообразны. В первую очередь к ним относятся следующие:

- потребляемая энергия;
- коэффициент полезного действия;
- средняя и пиковая мощность излучения на различных частотах и длительность излучения;
- надежность экранов, препятствующих созданию в окружающей среде сильных электрических и магнитных полей и предотвращающих проникновение в окружающую среду побочных (т. е. вне рабочего диапазона волн) вредных излучений (рентгеновских, гамма-излучения и др.);
- взрывобезопасность;
- массы редких веществ, расходуемых на создание, функционирование и утилизацию данной системы;
- показатели, характеризующие степень засорения окружающей среды в результате ликвидации системы после окончания срока ее службы.

Удовлетворение перечисленным требованиям является одной из важных и сложных проблем, которую должны решать разработчики радиосистемы. Однако на этапе НИР при определении облика системы исследователь вынужден воспользоваться качественным показателем, учитывающим всю вышеуказанную совокупность требований. Таким показателем может быть степень вредного воздействия системы на окружающую среду в процессе функционирования.

#### ***1.4.9. Задача количественной оценки качественных показателей***

Эта задача возникает, когда для принятия решения необходимо учитывать качественные показатели (критерии) наряду с обычными количественными. Задача состоит в том, чтобы качественный показатель



представить (оценить) числом, с которым можно выполнять арифметические действия. Это число, естественно, должно возрастать при улучшении качественного показателя и обращаться в нуль, если показатель имеет очень низкое качество.

Эта задача принципиально не может быть решена формализованными математическими методами. Единственно возможным средством может служить использование человеческого разума в виде экспертных оценок. Чтобы численное представление качественного показателя не было слишком субъективным, целесообразно использовать коллективные экспертные оценки. При правильной организации экспертизы такие оценки могут приемлемо отображать качественные свойства системы.

Человек не способен непосредственно оценивать числом качественное свойство объекта. Он может различать отношения качественных свойств типа "лучше-хуже", т.е. измерять качественный показатель в порядковой шкале. Однако ранг еще не является числом. На помощь приходит способность человека определять в некотором диапазоне степени качественных различий (например, немного лучше, много лучше, значительно лучше).

Так, возникает задача перехода от ранговой шкалы к модифицированным ранговым шкалам, описывающим в баллах изменение степени предпочтительности качественного показателя по мере перемещения по шкале.

#### ***1.4.10. Балльные шкалы***

Возможных балльных шкал очень много. Наиболее предпочтительная из них – девятиуровневая с интервалами между уровнями, равными единице. Есть некоторые соображения в пользу этой шкалы.

1. Она должна иметь конечный диапазон, поскольку использование шкалы в диапазоне от нуля до бесконечности может оказаться бесполезным. Способность человека различать находится в весьма ограниченном диапазоне, и когда имеется значительная несоразмерность (от нуля до бесконечности) между сравниваемыми объектами,

предположения человека тяготеют к тому, чтобы быть произвольными, и обычно оказываются далекими от действительности.

2. Качественные различия обладают "элементом количества", когда эти различия одного порядка и близки друг другу.

3. Способность человека выполнять качественные разграничения выражается тремя или пятью ступенями:

а) три ступени: неприятие, безразличие, принятие, т.е. низкая, средняя (умеренная), высокая оценки;

б) пять ступеней: очень плохая, плохая, посредственная, хорошая, очень хорошая оценки, т.е. равенство, слабое, сильное, очень сильное, абсолютное превосходство.

4. Для более детальной оценки каждую из трех ступеней можно разделить на три: низкую, умеренную, высокую степени. Получается девять оттенков качества. Дополнительно к пяти ступеням можно принять четыре промежуточных ступени между соседними. Получается тоже девять ступеней различий.

Для оценки по 9-балльной шкале сравнивается пара объектов. Получается шкала для парных сравнений. За основу берется показатель с наиболее низким качеством. Получаем 9-уровневую шкалу парных сравнений:

1 - равенство: одинаковая значимость (два показателя вносят одинаковый вклад в качество);

3 - слабое превосходство: некоторое преобладание качества одного показателя перед другим (опыт и суждение отдают легкое предпочтение одному из показателей);

5 - сильное превосходство: существенное или сильное превосходство по качеству (опыт и суждение отдают сильное предпочтение одному из показателей);

7 - значительное превосходство: очень сильное или очевидное преимущество (практически явное превосходство одного из показателей);

9 - абсолютное превосходство (превосходство одного из показателей не вызывает никаких сомнений).

Баллы 2, 4, 6, 8 — промежуточные значения.

Эта шкала является достаточно простой и хорошо подтверждается практикой. Но имеются также и более простые шкалы:

а) 3-балльная шкала:

- 1 - равенство;
- 2 - слабое и сильное превосходство;
- 3 - значительное и абсолютное превосходство;
- б) 5-балльная шкала;
- 1 - равенство;
- 2 - слабое превосходство;
- 3 - сильное превосходство;
- 4 - значительное превосходство;
- 5 - абсолютное превосходство.

Эти шкалы такие же, как и 9-балльная шкала, но без промежуточных значений.

#### ***1.4.11. Проблема перехода от балльной шкалы к количественной шкале***

Если показатели эффективности качественные и их надо свертывать в одну целевую функцию с количественными показателями эффективности, то нужны методы перехода от оценок в 9-балльной или 5-балльной шкале к оценкам в числах, так как балл — это ранг, а не число. Удобна для этих целей количественная шкала с диапазоном от 0 до 1.

Одной из таких шкал, по-видимому, самой простой, является шкала, устанавливающая следующее соответствие между балльными оценками и числовыми интервалами.

В этой шкале имеет место формальный переход от балльной к количественной оценке (табл. 1.2). Для свертывания показателей эффективности количественных и качественных в целевую функцию такой формальный подход может содержать значительный произвол. Это означает, что шкала перехода от баллов к числам должна составлять неотъемлемую часть метода свертывания.

Таблица 1.2. Соответствие качественных и количественных оценок

Качественная оценка	Балл	Количественная оценка
Очень хорошо	5	1,00—0,80
Хорошо	4	0,80—0,63
Удовлетворительно	3	0,63—0,37
Плохо	2	0,37—0,20
Очень плохо	1	0,20—0,00

Следует помнить и учитывать при формулировании рекомендаций, что балльные оценки не вполне надежны, так как каждый эксперт сам определяет свой стандарт качества, сам соотносит баллы и качественные понятия, что является произволом, искажающим первоначальную информацию.

#### ***1.4.12. Организация коллективных экспертных оценок***

В различных областях человеческой деятельности часто прибегают к услугам экспертов, чтобы использовать их интеллект для решения различных слабоструктурированных задач, для ответов на различные вопросы. Сложность проблемы, ограниченные возможности одного человека и субъективность его оценок вынуждают использовать коллективные оценки.

Особенность интеллектуальной деятельности людей состоит в том, что она во многом зависит от внешних и внутренних условий. Поэтому в методиках организации коллективной творческой работы специальное внимание уделяется созданию благоприятных условий и устранению факторов, неблагоприятно влияющих на работу.

На практике используются различные эвристические методы: экспертные оценки, мозговая атака, генерирование альтернатив, морфологический анализ, деловые игры и другие.

Последовательность и содержание этапов экспертизы можно изменять в зависимости от реальных условий и ограничений. Типичный набор этапов:

- формирование группы экспертов;
- создание благоприятных внешних условий для работы экспертов;
- формулирование цели экспертизы, постановка задачи и разработка процедуры опроса;
- проведение опроса;
- обработка информации, полученной от экспертов, и ее анализ;
- формулирование рекомендаций для принятия решения и передача их руководителю.

Методы обработки полученной в результате экспертизы информации (зависят от конкретных задач). К таким методам относятся, в частности:

- формализация информации с помощью измерительных шкал;
- оценка компетентности экспертов;
- оценка фактора одним показателем;
- оценка фактора по нескольким показателям;
- оценка относительной важности фактора (ранжирование);
- оценка согласованности мнений экспертов и многие другие.

*Оценка компетентности эксперта.* Пусть  $i$  - номер эксперта,  $N$  - число экспертов. Коэффициенты компетентности могут быть оценены самими экспертами. Обозначим через  $l$  номер оценивающего эксперта. Чтобы учесть степень компетентности экспертов, принимающих участие в экспертизе, проводится следующая процедура. Каждый эксперт присваивает каждому из своих коллег (в том числе и себе самому) коэффициент компетентности  $k_{il}$  от 1 до 10 или от 0 до 1 (чем больше коэффициент, тем эксперт компетентней). Коэффициент компетентности  $i$ -го эксперта:

$$k_i = \frac{1}{N} \sum_{l=1}^N k_{il} . \quad (1.22)$$

Затем вычисляется  $\alpha_i$  - вес мнения  $i$ -го эксперта – нормированный коэффициент компетентности:

$$\sum_{i=1}^N \alpha_i = 1 .$$

*Оценка фактора (альтернативы, параметра, показателя).* Пусть  $j$  — номер фактора,  $i$  — номер эксперта,  $N$  — число экспертов,  $k_{ij}$  — числовая оценка  $i$ -м экспертом  $j$ -го фактора. Оценки  $k_{ij}$  можно рассматривать как "измерения" искомой "истинной характеристики"  $k_j$ , считая отклонения  $k_{ij} - k_j$  случайными величинами, а  $k_{ij}$  — как выборку из  $N$  элементов. В качестве оценки можно использовать выборочное среднее:

$$k_j = \frac{1}{N} \sum_{i=1}^N k_{ij} . \quad (1.23)$$

С учетом компетентности экспертов:

$$k_j = \frac{1}{N} \sum_{l=1}^N \alpha_l k_{lj} . \quad (1.24)$$

*Ранжирование факторов (альтернатив, параметров).* Пусть  $f_1, \dots, f_m, \dots, f_M$  — список факторов,  $M$  — число факторов. Каждому  $f_m$   $i$ -й эксперт присваивает ранг важности  $k_{mi}$  от 1 до  $M$  ( $M$  — наибольший ранг для наименее важного фактора, 1 - ранг наиболее важного фактора) Матрица мнений  $N \times M$  :

$$\begin{array}{c} f_1 \quad f_2 \quad \dots \quad f_m \quad \dots \quad f_M \\ \begin{array}{c} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \dots \\ \mathcal{E}_i \\ \dots \\ \mathcal{E}_N \end{array} \left\| \begin{array}{cccccc} k_{11} & k_{21} & \dots & k_{m1} & \dots & k_{M1} \\ k_{12} & k_{22} & \dots & k_{m2} & \dots & k_{M2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ k_{1i} & k_{2i} & \dots & k_{mi} & \dots & k_{Mi} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ k_{1N} & k_{2N} & \dots & k_{mN} & \dots & k_{MN} \end{array} \right\| \end{array} . \quad (1.25)$$

Средние ранги:  $\bar{k}_1, \bar{k}_2, \dots, \bar{k}_m, \dots, \bar{k}_M$  .

Сумма рангов всех факторов должна быть одинакова по всем экспертам:

$$\sum_{m=1}^M k_{mi} = \frac{M \cdot (M + 1)}{2}, \quad i = 1, \dots, N \text{ - номер эксперта.} \quad (1.26)$$

Если эксперт считает  $q$  факторов эквивалентными, то присваивает им средний ранг:

$$k_{m+1} = \dots = k_{m+j} = \dots = k_{m+q} = \frac{1}{q} \cdot \sum_{j=1}^q k_{m+j} .$$

Окончательный результат – средние ранги каждого фактора:

$$\bar{k}_m = \frac{1}{N} \sum_{l=1}^N k_{ml} . \quad (1.27)$$

#### 1.4.13. Метод мозговой атаки

Методы мозговой атаки (МА) или мозгового штурма получили широкое применение в инженерной деятельности. Это эмпирически найденные эффективные способы решения творческих задач, основанные

на психологическом эффекте — использовании возможностей подсознания для генерирования идей решения проблемы. Эффект проявляется при коллективной творческой работе в короткое время — отсюда происходят названия: сеанс, шторм, атака. Например, группа из 5—8 человек за 15—30 мин при индивидуальной независимой работе даст в сумме примерно 10—20 идей решения творческой (изобретательской) задачи. Коллективное высказывание идей увеличивает их число до 50—150. Во время сеанса МА происходит как бы цепная реакция идей, приводящая к интеллектуальному взрыву при контакте с мыслями других людей. Можно добавить, что идеи возникают при любых контактах — с книгами и другими носителями информации.

Для всякого значительного целенаправленного действия, операции необходимо поставить задачу. Постановка задачи перед участниками МА может иметь самую различную форму и содержание. Задачу может сформулировать заказчик, руководитель творческой группы или ее участник. Важно, чтобы перед сеансом МА имелась достаточно исчерпывающая четкая постановка задачи, сформулированная просто, ясно, кратко, желательно в документальном виде.

Постановка задачи может быть дана в виде описания проблемной ситуации, в котором должны содержаться ответы на следующие вопросы:

1. В чем состоит затруднение или проблемная ситуация и какова ее предыстория?
2. Что требуется сделать для устранения проблемы, что в итоге желательно получить, какова цель МА?
3. Что мешает устранению проблемы, получению желаемого или достижению цели?
4. Что дает решение задачи? То есть существует следующая формула постановки задачи: проблема, цель, противодействие, эффект.

Главное содержание постановки задачи должно быть четко сформулировано в ответах на вопросы 1—3. Ответы на вопрос 4 должны стимулировать членов творческой группы на активную деятельность, чтобы предлагаемая задача стала для них главной задачей, которую необходимо неотложно решить.

Участники МА должны предложить решение задачи, способы и средства достижения желаемой цели.

Перечислим основные достоинства метода мозговой атаки:

- высокая эффективность;
- легкость и быстрота усвоения, не требуется специальная подготовка;
- естественность для коллективного творчества (группы, бригады, и пр.);
- универсальность, что означает его эффективность при решении творческих задач (изобретательских, рационализаторских) в различных областях и в самых различных постановках (по форме, деятельности, глубине), на различных этапах решения творческой задачи и этапах ОКР, а также в сочетании с другими эвристическими методами.

Универсальность МА позволяет рассматривать почти любую проблему или любое затруднение в сфере человеческой деятельности. Это могут быть задачи из области организации производства, сферы обслуживания, бизнеса, экономики, социологии, уголовного розыска, военных операций и т.д.

Перечислим методы организации МА:

1. Прямая атака — высказываются идеи по устранению проблемы, недостатков.

2. Обратная атака — выявление недостатков.

3. Двойная прямая или обратная атака проводится в два сеанса с перерывом от 2 часов до 2—3 дней. При проведении второго сеанса появляются наиболее ценные идеи или удачное развитие идей первого сеанса, так как во время перерыва продолжает работать подсознание.

4. Обратная и прямая МА (для совершенствования технических объектов). На первом сеансе выявляются недостатки объекта, а на втором предлагаются идеи по устранению выявленных недостатков.

5. Обратная и прямая МА может быть двойной, т.е. выполняется с перерывом для совершенствования идей.

6. Прямая и обратная МА. В первом сеансе высказываются идеи по совершенствованию объекта, а во втором сеансе — прогнозируются недостатки усовершенствованного объекта.

7. Прямая и обратная МА бывает двойной, т.е. с перерывом.

8. Атака с оценкой идей проводится в три этапа для решения сложных конструкторских задач.



#### ***1.4.14. Формирование альтернатив***

Формирование альтернатив является наиболее трудным и наиболее творческим этапом выбора (синтеза). Если в множество альтернатив не включена действительно наилучшая, то никакие методы выбора ее не "найдут". Чем больше предложено альтернатив, тем выше вероятность наилучшего решения.

Процесс формирования альтернатив, т.е. возможных способов достижения цели, является неформальным. Здесь проявляется творчество, талант, умение обобщать предыдущий опыт. Поэтому все рекомендации, как формировать новые альтернативы, являются результатом коллективного опыта по решению творческих задач и являются, по существу, эмпирическими правилами.

Существуют три случая, когда нет необходимости для формирования альтернатив:

- физически возможен или абсолютно очевиден только один путь решения проблемы;
- поставленным срокам решения задачи удовлетворяет только одна альтернатива, причем она оправдана целями более высокого уровня иерархии;
- по условиям материально-технического обеспечения или по пропускной способности материально-технической или производственной базы возможен только один вариант решения задачи.

При формировании альтернатив надо ориентироваться на модель цели системы. Действительно, формирование множества альтернатив с последующим выбором наилучшей производится для достижения цели системы с наивысшей эффективностью, поэтому модель цели системы определяет, в каком направлении следует формировать альтернативы, чтобы выбрать затем одну или несколько самых предпочтительных.

Критерий предпочтения является вариантом модели цели. Если он уже известен, то при формировании альтернатив можно ориентироваться на критерий предпочтения. Обратное, сформированное множество альтернатив может потребовать уточнения или корректировки критерия

предпочтения.

Для формирования альтернатив исследователь может использовать собственные знания, а также искать альтернативы в книгах, журналах, энциклопедиях, справочниках, каталогах выставок и товаров, описаниях патентов и изобретений, рекламных проспектах, в отчетах по НИР и ОКР. На основе имеющихся альтернатив исследователь может их комбинировать, образовывать промежуточные варианты между предложенными ранее альтернативами, модифицировать имеющиеся альтернативы (частичное отличие), включать альтернативы, противоположные предложенным, в том числе нулевую (ничего не делать, т.е. рассмотреть развитие событий без вмешательства).

Если стремиться получить как можно больше альтернатив, то для некоторых проблем их количество может достигнуть многих десятков. Подробное изучение каждой из них приведет к неприемлемым затратам времени и средств. В таких случаях рекомендуется провести "грубое отсеивание", не сравнивая альтернативы количественно, а лишь проверяя их на присутствие некоторых качеств, нежелательных для любой альтернативы.

В первую очередь следует последовательно "отсечь" заведомо непригодные по основным критериям альтернативы: альтернатива не отвечает требованиям эффективности, не соответствует нормам надежности, не укладывается в рамки выделенных средств, требует большого срока реализации. К отсеиванию может привести также обнаружение отрицательных побочных эффектов, например, альтернатива требует привлечения дополнительных дефицитных ресурсов.

Наоборот, можно не отсеивать, а отбирать альтернативы с "хорошими" признаками, например с такими признаками, как адаптивность при изменении условий внешней среды, высокая надежность, экологическая совместимость.

Для сокращения числа альтернатив можно провести ранжирование альтернатив по степени их важности и отсеять наименее важные из них.

### 1.4.15. Морфологический анализ и синтез

Морфологический метод основан на комбинаторике. В техническом объекте (ТО) выделяют группу основных конструктивных или других признаков. Наиболее распространен функциональный подход: в качестве признака берутся функции компонентов  $\Phi_1, \dots, \Phi_i, \dots$ , где  $i$  — номер функции. Для каждого признака — функции предлагают всевозможные варианты способов реализации — альтернативы  $A_i^k$ , где  $k$  — номер альтернативы реализации  $i$ -й функции. Так получают таблицу перечня признаков и альтернатив — морфологическую таблицу (табл. 1.3).

Таблица 1.3. Морфологическая таблица

$\Phi_1$	$\Phi_2$	$\Phi_3$	$\Phi_4$	$\Phi_5$
$A_1^1$	$A_2^1$	$A_3^1$	$A_4^1$	$A_5^1$
$A_1^2$	$A_2^2$	$A_3^2$	$A_4^2$	$A_5^2$
$A_1^3$		$A_3^3$		$A_5^3$
$A_1^4$		$A_3^4$		
$A_1^5$				

Если из каждого столбца-функции взять по одной альтернативе, получим некоторый облик ТО. Например, для пяти функций возможна комбинация  $A_1^3 A_2^2 A_3^4 A_4^1 A_5^3$ . Число возможных конструкций ТО равно произведению количества альтернатив в каждом столбце. Например:  $5 \cdot 2 \cdot 4 \cdot 2 \cdot 3 = 240$ . Таким образом, можно получить множество решений, в том числе представляющих практический интерес.

Итак, суть метода состоит в построении морфологической таблицы (выборе признаков-функций и заполнении таблицы всевозможными альтернативами) и в выборе из всего множества получаемых комбинаций наилучших решений.

1. *Заготовка столбцов таблицы.* Морфологическую таблицу строят на основе функциональной структурной схемы ТО (структурной схемы конкретного прототипа или обобщения ряда прототипов). Таблица состоит

из нескольких столбцов, число которых равно числу анализируемых функций (числу элементов, если рассматривать по одной функции элемента). Число элементов, на которые разделяют ТО, должно быть не меньше 3 и не больше 10. Заглавиями столбцов являются обозначения или описания функций элементов.

*Столбцы* (функции) лучше располагать в таком порядке, чтобы наиболее конструктивно или функционально связанные элементы находились по возможности в соседних столбцах.

Если выполнение функции элемента не имеет альтернатив или вообще не рассматривается для упрощения анализа альтернативы, то для таких элементов в таблице не предусматривают отдельных столбцов. Если элемент ТО выполняет одновременно несколько функций, то целесообразно составлять столбцы для каждой рассматриваемой функции или заменять полифункциональные элементы на элементы, реализующие отдельные функции или их меньшее число.

2. *Заполнение столбцов таблицы альтернативами.* В таблицу сначала вносят реализации функций прототипа. Затем записывают возможные наиболее интересные и эффективные варианты. При этом могут быть использованы методы генерирования альтернатив и мозговой атаки. В каждом столбце путем объединения двух или более альтернатив выделяют комбинированные более эффективные альтернативы (в значительно большей мере устраняющие какой-либо недостаток или улучшающие какое-либо достоинство, чем отдельные альтернативы). Такие комбинации записывают в столбец в качестве дополнительных альтернатив.

3. Для выбора наилучших решений ТО необходима *постановка задачи*, включающая правило выбора — критерий предпочтения.

Системный подход при анализе слабоструктурированных проблем приводит к полезным результатам. Следует помнить, однако, что организация системного исследования должна начинаться с четкого формулирования конечных целей решения проблемы, причем кроме конечных целей существуют цели промежуточные. Построение иерархии (дерева) целей может оказаться весьма важным этапом системного исследования.

## **2. МЕТОДЫ И ЗАДАЧИ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ**

### **2.1. Основные понятия исследования операций**

#### ***2.1.1. Предмет исследования операций и история становления***

Как самостоятельное научное направление исследование операций (ИСО) возникло из потребностей наилучшей организации боевых действий и прогнозирования их исхода при принятии командных решений в условиях неполноты информации о состоянии и возможных действиях противника. С 1935 по 1938 гг. в Великобритании, в местечке Бодси, проводились исследования для подготовки эффективных мер противодействия возрастающей угрозе со стороны ВВС Германии. Именно тогда руководителем группы А. Раузом был впервые употреблен термин «операционное исследование». В дальнейшем соответствующие отделы были сформированы при командовании различных родов войск Великобритании и США. Командование вооруженных сил гитлеровской коалиции методов исследования операций не использовало.

Дальнейшее развитие ИСО связано с НТР, которая сопровождалась бурным развитием и резким усложнением техники, расширением масштабов человеческой деятельности и возрастанием затрат на нее, с развитием вычислительной техники и математических методов в сфере управления.

В 1948 году несколько ученых, активно участвовавших в исследовании военных операций в Англии, создали Клуб исследования операций и стали издавать журнал *Operational Research Quarterly*. В начале 50-х гг. общества исследования операций стали возникать в разных странах – например, в США –ORSA. В 1959 г. возникла Международная федерация обществ по исследованию операций - IFORS. В 1970 г. в СССР создан технический комитет по исследованию операций. На промышленных предприятиях создаются специальные подразделения.

Термин ИСО (operations research или operational research) имеет ряд синонимов, в том числе «наука об управлении». В целом можно определить ИСО как *научный подход к решению задач организационного управления*. Более развернутое определение: *ИСО представляет собой комплексную научную дисциплину разработки и применения математических методов и моделей принятия оптимальных решений для различных аспектов человеческой деятельности с учетом окружающей среды*.

Иногда ИСО понимают только как применение математических, количественных, методов для обоснования решений в различных областях человеческой деятельности. Между тем прежде всего необходимо исходить из содержательных, методологических аспектов решаемой проблемы.

### **2.1.2. Основные определения**

Под операцией понимается совокупность действий или мероприятий, объединенных единым замыслом и направленных на достижение вполне определенной цели. Основными признаками операции являются целенаправленность и управляемость действий.

Основная задача ИСО – обоснование наилучших решений для лица принимающего решения (ЛПР) в предстоящих условиях проведения операции. Примеры операций:

- оптимальное вложение средств в производство различных видов вооружений в интересах обеспечения безопасности государства;
- разработка и реализация мероприятий, направленных на достижение максимальной рентабельности предприятия;
- организация информационного обеспечения САПР;
- разработка и реализация мероприятий, обеспечивающих повышение пропускной способности системы обслуживания заявок;
- наилучшее распределение  $n$  механизмов по  $m$  работам;
- определение оптимальной траектории полета РН;
- построение оптимального плана подготовки к сдаче экзаменов в период сессии.

Эффективность операции есть степень ее приспособленности к выполнению поставленной задачи. Чтобы судить об эффективности операций и сравнивать их, нужно иметь некоторый критерий операций,

или целевую функцию.

Условия, характеризующие обстановку проведения операции и влияющие на ее исход, называются факторами. Факторы могут быть контролируемыми или управляемыми (оперирующая сторона может распоряжаться ими для достижения цели операции) и неконтролируемыми (они могут препятствовать достижению цели операции). Факторы могут быть определенными (детерминированными), случайными, неопределенными.

Решением задачи ИСО называется определенный выбор управляемых факторов (параметров). Допустимым является решение, принадлежащее заданному множеству. Оптимальное решение доставляет экстремум выбранному критерию эффективности операции.

Оперерующей стороной (ОС) называются участники операции, стремящиеся к достижению поставленной цели. В составе ОС выделяется исследователь операции, а также ЛПР.

Ограничения определяют границы и свойства множеств изменения факторов.

Стратегией называется способ реализации действия или мероприятий оперирующей стороны. В зависимости от условий принятия решения, значения контролируемых (управляемых) факторов могут выбираться либо детерминированным образом, либо случайным.

### ***2.1.3. Методология исследования операций***

Использование правильной методики организации ИСО гарантирует, что исследователь не станет тратить усилия на решение неверно поставленной задачи.

Исследование операции подразделяется на ряд этапов, вытекающих из системного подхода к решению организационно-технических проблем.

1. Выявление проблемы. На этой стадии осуществляется анализ и синтез данных, относящихся к проблеме, формируется предварительная целевая установка.

2. Определение цели (целей) операции, постановка ее задач, определение множества возможных вариантов действий, описание необходимых исходных данных (ИД) и возможных условий проведения

операции. Цель может быть сформулирована только с позиций надсистемы.

3. Составление плана проведения исследования: назначение сроков выполнения отдельных работ и распределение их по исполнителям.

4. Формулировка проблемы. На этой стадии отвечают на ряд вопросов:

- нельзя ли представить всю проблему в виде отдельных, частных подпроблем, которые затем можно исследовать независимо одну от другой (параллельно или последовательно);
- построение модели операции в виде математического описания цели и процессов, позволяющей осуществлять анализ результатов операции и сравнивать конкурирующие варианты действий, модель должна быть адекватна (подобна) реальному объекту;
- разработка критериев оптимальности результата операции, позволяющего строить множество оптимальных решений (мероприятий, стратегий), направленных на его достижение;
- разработка математических методов поиска оптимальных стратегий на множестве возможных;
- экспертиза модели и ее программная реализация на ЭВМ;
- разработка методики исследования операции на модели (вычислительного эксперимента);
- обоснование рекомендаций для ЛПР по выбору окончательного решения (мероприятия, стратегии) из множества оптимальных<sup>1</sup>.

После реализации принятого решения осуществляется оценка полученного результата и корректировка модели (замыкание обратной связи).

#### **2.1.4. Методы моделирования**

1. Математические. Математическая модель – это описание (часто приближенное) какого-либо класса явлений реального мира, выраженное с помощью математических символов. Эти модели отличаются от оригинала физической природой и геометрической формой. Сходство с оригиналом заключается в том, что они описываются одними и теми же математическими уравнениями. Достоинства: универсальность методов

---

<sup>1</sup> В общем случае оптимальных решений может быть более одного.



исследования, возможность исследования любых процессов, включая те, которые нельзя осуществить физически.

Примерами математических моделей могут служить формула, система уравнений, описывающие изучаемый объект, алгоритм решения задачи.

2. Физические. Эти модели сходны с оригиналом по физической природе и геометрической форме. Отличие от оригинала заключается в размерах, скорости процесса и других точно учитываемых факторах. Их достоинства:

- возможность исследования процессов независимо от того, поддаются ли они математическому описанию;
- большая наглядность результатов моделирования;
- возможность непосредственного участия человека в моделировании.

Характерным примером является исследование аэродинамических характеристик летательных аппаратов с помощью «продувок» их моделей в аэродинамических трубах.

3. Смешанные. Наиболее трудно описываемая часть процесса здесь представляется физической моделью. Модели этого типа широко используются при разработке изделий ракетно-космической техники. Статистические данные и результаты математического моделирования сочетаются в них с результатами, полученными в ходе натуральных экспериментов.

### **2.1.5. Классификация моделей ИСО**

Модели, применяемые в ИСО, должны позволять вычислять выбранный критерий, иначе решение задачи ИСО невозможно. В общем виде модель ИСО можно представить в виде:

$$W = F(x, z),$$

где  $W$  – критерий операции,  $F$  – оператор модели,  $x$  – вектор управляемых переменных (факторов),  $z$  – вектор неуправляемых факторов.

Модели операции классифицируются по следующим признакам.

1. По форме оператора  $F$ :

- аналитические, в которых критерий связан с векторами переменных аналитическими зависимостями;

- статистические, в которых значение оператора является случайной величиной. Закон распределения и числовые характеристики критерия находятся в результате статистической обработки ряда реализаций операции.

2. По месту на «иерархической лестнице».

Например, модель космической системы наблюдения стоит на высшей ступени иерархии, на следующей ступени – модели космического комплекса и наземного специального комплекса, далее следует модель орбитального комплекса, ниже ее – модель космического аппарата и т.д.

Взаимосвязь моделей должна осуществляться таким образом, чтобы вектор управляемых переменных модели более высокого уровня входил в состав вектора неуправляемых переменных модели более низкого уровня. Это – требование системного подхода к решению задач ИСО.

3. По характеру учета времени:

- статические, в которых не учитывается зависимость факторов, связанных с решаемой проблемой, от времени;
- динамические, в которых такая зависимость учитывается.

4. По числу рассматриваемых этапов:

- одноэтапные (пример: модель функционирования системы ПВО при отражении одного налета);
- многоэтапные (пример: модель воздушного боя).

5. По методу учета действий человека:

- действия человека учитываются в модели непосредственно;
- в модель заранее вводятся возможные варианты действий человека в возникающих ситуациях.

6. По целям исследования:

- модели анализа существующих операций;
- модели синтеза (проектирования) операций.

К моделям операций предъявляются требования, аналогичные требованиям к моделям сложных технических систем.

### ***2.1.6. Требования к критериям***

1. Представительность - приспособленность к оценке эффективности решения основной (а не второстепенных) задачи ИСО.

2. Чувствительность: малые изменения исследуемых факторов должны вызывать заметное изменение значения критерия.

3. Максимальная возможная простота: введение в критерий второстепенных параметров усложняет решение задачи, не приводя к дополнительным выводам. Желательно использование единого критерия.

4. Правильный учет стохастичности. Так, если дисперсия значений критерия в различных реализациях операции изменяется мало, достаточно использовать математическое ожидание критерия, в противном случае приходится строить доверительные интервалы.

### **2.1.7. Общая форма критериев. Примеры критериев**

Общая форма критерия следует из основного постулата ИСО: оптимальным является решение, обеспечивающее достижение заданного эффекта при минимальных материальных затратах (I постановка) либо такое, при котором достигается максимальный эффект при фиксированных материальных затратах (II постановка).

В зависимости от постановки, один из показателей – эффективность или стоимость – выступает в качестве критерия операции  $W$ , другой – как основное ограничение. В любой задаче ИСО (в явной или скрытой форме) всегда присутствует хотя бы одно ограничение, иначе задача теряет смысл. Помимо основных ограничений, в реальных задачах всегда присутствуют дополнительные ограничения вида  $a_j \leq a_{j\text{доп}}$ ,  $j=1, 2, \dots, m$ , где  $a_{j\text{доп}}$  – предельно допустимое значение параметра  $a_j$ ,  $m$  - число ограничений.

Рассмотрим общие формы критериев на примере планирования военной операции. Для таких операций характерно наличие противодействия противника, характеризуемого вектором параметров  $y$ . Затраты на операцию

$$C=C(x, y, z), \text{ эффективность операции } E=E(x, y, z).$$

I постановка. Оптимальным является решение, обеспечивающее в условиях худшего для нас противодействия противника выполнение боевой задачи при минимальных затратах:

$$W = \min_x \max_y C(x, y, z) \text{ при } E(x, y, z) \geq E_{\text{зад}}$$

и дополнительных ограничениях.

II постановка. Оптимальным является решение, обеспечивающее в

условиях худшего для нас противодействия противника выполнение боевой задачи с максимальной эффективностью при фиксированных затратах:

$$W = \min_y \max_x E(x, y, z) \text{ при } C(x, y, z) \geq C_{\text{зад}}$$

и дополнительных ограничениях.

Некоторые типовые критерии эффективности можно рассмотреть на следующем *примере*. Ракетная группировка наносит удар по комплексу площадных целей. Объединим цели в группы так, чтобы в пределах каждой из них можно было считать цели одинаковыми как по значимости, так и по уязвимости. Тогда можно рекомендовать следующие критерии эффективности группировки:

1. Математическое ожидание числа пораженных объектов с учетом их значимости

$$M = \sum_j Q_j \sum_i^{K_j} P_{ij},$$

где  $Q_j$  значимость  $j$ -ой группы целей,  $K_j$  – количество целей в группе,  $P_{ij}$  – вероятность поражения  $i$ -ой цели из  $j$ -ой группы.

2. Математическое ожидание суммарной пораженной площади

$$M = \sum_j \sum_i M_{ji},$$

где  $M_{ji}$  – математическое ожидание пораженной площади  $i$ -ой цели в  $j$ -ой группе.

Этот критерий целесообразно использовать, когда значимости целей мало отличаются или не поддаются определению.

3. Вероятность выполнения следующей боевой задачи: в каждой  $j$ -ой группе необходимо поразить не менее  $a_j$  целей,

$$P = \prod_j P_{aj},$$

где  $P_{aj}$  – вероятность поражения не менее  $a_j$  целей в  $j$ -ой группе.

### **2.1.8. Типичные классы задач исследования операций**

Несмотря на большое разнообразие задач исследования операций, по *содержательной постановке* можно выделить следующие типичные классы задач.

1) *Задачи распределения ресурсов*, заключающиеся в распределении ограниченного запаса средств (денег, энергии, работников и.д.) между рядом операций таким образом, чтобы достигался максимальный суммарный эффект.

2) *Задачи управления запасами*, в которых определяется такой запас материальных средств, чтобы суммарные затраты, связанные с хранением их избыточного количества, с одной стороны, и дефицитом – с другой, были минимальны.

3) *Задачи ремонта и замены оборудования*, в которых определяются сроки восстановительного ремонта и моменты замены оборудования, минимизирующие сумму затрат на обслуживание и обновление оборудования с учетом его старения в процессе эксплуатации.

4) *Задачи массового обслуживания*, связанные с изучением режимов и параметров работы систем массового обслуживания (к этим системам относятся не только такие «бытовые» системы обслуживания, как магазины или билетные кассы, но и железные дороги, системы ПВО, космические системы мониторинга земной поверхности и т.д.).

5) *Задачи теории расписаний*, заключающиеся в определении последовательности операций, производимых над множеством изделий, при которой оптимизируется некий критерий – например, достигается минимальное время изготовления изделий.

6) *Задачи сетевого планирования и управления* – в них определяются оптимальные сроки начала и окончания взаимосвязанных операций с заданной последовательностью, при которых суммарная продолжительность всей совокупности операций (проекта) или затраты на их осуществление минимальны.

7) *Задачи выбора маршрута*, оптимизирующего значение заданного критерия и удовлетворяющего некоторым ограничениям. В их числе можно назвать задачу выбора кратчайшего пути, задачу коммивояжера, задачу об оптимальном потоке.

Приведенная классификация задач не является, конечно, исчерпывающей. Жизнь ставит перед исследователями все новые задачи, которые служат комбинацией рассмотренных классов задач или вообще не

укладываются в приведенную классификацию.

С другой стороны, для выбора методов решения задач ИСО принципиальное значение имеет их классификация *в зависимости от условий окружающей среды и степени информированности лица, принимающего решения*. В данном аспекте различают задачи принятия решений:

- 1) в условиях *определенности* (детерминированные задачи);
- 2) в условиях *риска* (стохастические задачи);
- 3) в условиях *неопределенности*;
- 4) в условиях *конфликта* (игровые задачи).

## 2.2. Линейное программирование

Задачи линейного программирования (ЗЛП) относятся к наиболее исследованным в ИСО. В задаче линейного программирования целевая функция и все ограничения линейны, а все переменные - действительны. В *стандартной форме* ЗЛП имеет следующий вид:

$$\begin{cases} \sum_{j=1}^n c_j x_j \rightarrow \max; \\ \sum_{j=1}^n a_{ij} x_j = b_i, \quad i = \overline{1, m}; \\ x_j \geq 0, \quad j = \overline{1, n}, \end{cases}$$

где  $c_j$ ,  $a_{ij}$ ,  $b_i$  - постоянные,  $x_j$  – управляемые переменные (элементы решения),  $n$  и  $m$  – соответственно количество переменных и количество ограничений. ЗЛП в стандартной форме принято также называть *основной задачей линейного программирования*.

Сумма  $\sum_{j=1}^n c_j x_j = f$  является целевой функцией, а система

ограничений  $\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, \dots, m$ , определяет множество допустимых решений.

Задачи линейного программирования часто встречаются на практике: при решении проблем распределения ресурсов, планирования производства, организации работы транспорта и т.д.

*Пример 2.1.* Предприятие производит изделия трех видов:  $U_1, U_2, U_3$ . По плану должно быть выпущено не менее  $b_1$  единиц изделия  $U_1$ , не менее  $b_2$  единиц изделия  $U_2$  и не менее  $b_3$  единиц изделия  $U_3$ . Условия спроса делают невыгодным производство изделий в количестве, превышающем соответственно  $\beta_1, \beta_2, \beta_3$  единиц. Для изготовления изделий используется четыре вида сырья:  $s_1, s_2, s_3, s_4$ . Запасы сырья каждого вида ограничены значениями  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ . Известно также, что на изготовление одной единицы изделия  $U_j$  необходимо количество единиц сырья вида  $s_i$ , равное  $a_{ij}$ .

При реализации одно изделие  $U_1$  приносит прибыль  $c_1$ ,  $U_2$  – прибыль  $c_2$ ,  $U_3$  – прибыль  $c_3$ . Требуется так спланировать производство, чтобы план был выполнен или перевыполнен (но при отсутствии «затоваривания»), а прибыль была максимальна.

Запишем задачу в виде ЗЛП. Введем переменные («элементы решения»)  $x_1, x_2, x_3$  – количество единиц продукции  $U_1, U_2, U_3$ , планируемых к производству. Условия выполнения плана и условия отсутствия излишней продукции запишутся соответственно в следующем виде:

$$x_1 \geq b_1, x_2 \geq b_2, x_3 \geq b_3, \quad x_1 \leq \beta_1, x_2 \leq \beta_2, x_3 \leq \beta_3,$$

ограничения на количество сырья – в виде

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + a_{31}x_3 \leq \gamma_1, \\ a_{12}x_1 + a_{22}x_2 + a_{32}x_3 \leq \gamma_2, \\ a_{13}x_1 + a_{23}x_2 + a_{33}x_3 \leq \gamma_3, \\ a_{14}x_1 + a_{24}x_2 + a_{34}x_3 \leq \gamma_4. \end{cases}$$

Прибыль от реализации продукции составит

$$f = c_1x_1 + c_2x_2 + c_3x_3.$$

Итак, мы пришли к следующей ЗЛП: найти элементы решения  $x_1, x_2, x_3$ , удовлетворяющие сформулированным ограничениям и максимизирующие функцию  $f$ .

В общем случае ЗЛП может содержать и ограничения типа неравенств; для приведения такой задачи к стандартной форме записи вводятся дополнительные переменные: если для некоторого  $k$  имеет место

неравенство  $\sum_{j=1}^n a_{kj}x_j \leq b_k$ , то введение положительной переменной  $x_{n+1}$

позволяет записать:  $\sum_{j=1}^n a_{kj}x_j + x_{n+1} = b_k$ .

Если же  $\sum_{j=1}^n a_{kj}x_j \leq b_k$ , то  $\sum_{j=1}^n a_{kj}x_j - x_{n+1} = b_k$ .

Задача минимизации  $f \rightarrow \min$  приводится к стандартной форме путем замены  $c_j' = -c_j$ .

*Пример 2.2.* Привести к стандартной форме ЗЛП вида

$$f = x_1 + 2x_2 \rightarrow \min, \begin{cases} 2x_1 - x_2 \geq 2, \\ x_1 + 3x_2 \leq 3, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Введем переменные  $x_3, x_4$  и заменим в записи целевой функции коэффициенты при переменных на противоположные; тогда задача приобретет вид

$$\begin{cases} -x_1 + 2x_2 + 0x_3 + 0x_4 \rightarrow \max, \\ 2x_1 - x_2 - x_3 = 2, \\ x_1 + 3x_2 + x_4 = 3, \\ x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0. \end{cases}$$

*Допустимым* решением ЗЛП называется всякая совокупность неотрицательных переменных, удовлетворяющая системе ограничений; *оптимальным* является то из решений, которое обращает в максимум целевую функцию. Как известно,  $m$  линейно независимых уравнений, задающих ограничения в ЗЛП, можно разрешить относительно неких  $m$  базисных переменных, выразив их через остальные, свободные переменные, число которых равно  $n-m=k$ .

В теории линейного программирования доказывается, что



оптимальное решение (если оно существует) достигается при такой совокупности значений переменных  $x_1, x_2, \dots, x_n$ , где по крайней мере  $k$  из них обращаются в нуль, а остальные неотрицательны.

В случае, когда  $k = 2$ , задача линейного программирования может быть решена графическим методом.

Пусть в качестве свободных выбраны переменные  $x_{m+1}, x_{m+2}$ .

Выразим через них остальные (базисные) переменные  $x_1, x_2, \dots, x_m$ :

$$\begin{cases} x_1 = \alpha_{1,m+1}x_{m+1} + \alpha_{1,m+2}x_{m+2} + \beta_1 \\ x_2 = \alpha_{2,m+1}x_{m+1} + \alpha_{2,m+2}x_{m+2} + \beta_2 \\ \dots \\ x_m = \alpha_{m,m+1}x_{m+1} + \alpha_{m,m+2}x_{m+2} + \beta_m, \end{cases}$$

где  $\alpha_{1,m+1}, \dots, \beta_{1,m+1}, \dots$  - постоянные, полученные при преобразовании системы уравнений. Каждое из уравнений полученной системы описывает прямую линию на плоскости свободных переменных, причем соответствующая базисная переменная равна нулю на этой линии и положительна по одну сторону от нее. Система уравнений в целом описывает некий многоугольник; если все базисные переменные внутри этого многоугольника положительны, то ЗЛП имеет решение.

По аналогии, каждому фиксированному значению целевой функции  $f_0$  также будет соответствовать некая прямая (линия уровня). Вершина многоугольника, описывающего область допустимых значений переменных, в которой значение целевой функции максимально, соответствует оптимальному решению ЗЛП.

*Пример 2.3.* Решить задачу линейного программирования

$$\begin{cases} f = 6x_1 - 5x_2 \rightarrow \max; \\ 2x_1 + 5x_2 + x_3 = 10, \\ 5x_1 + 2x_2 + x_4 = 10, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{cases}$$

Выберем в качестве свободных переменные  $x_1, x_2$ .

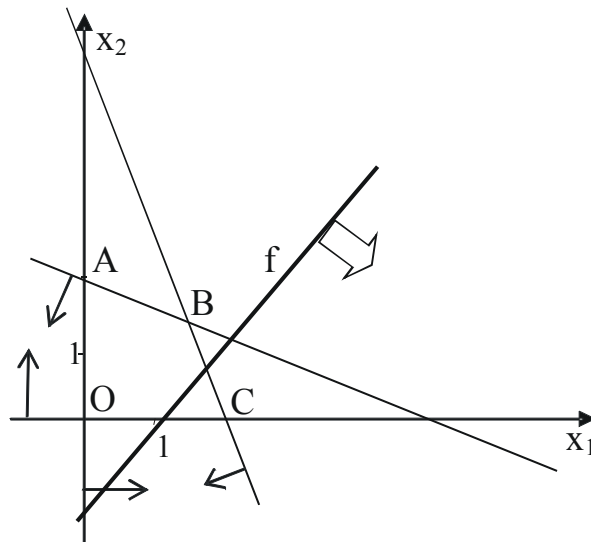


Рис. 2.1. Решение задачи линейного программирования графическим методом

Решение задачи графическим методом показано на рис. 2.1. Областью допустимых значений переменных является четырехугольник  $OABC$  (стрелка у каждой прямой, соответствующей одному из ограничений задачи, показывает полуплоскость, в которой ограничение выполняется).

Оптимальное решение достигается в вершине  $C$  с координатами  $x_1^* = 2; x_2^* = 0$ . Оптимальные значения  $x_3, x_4$  найдем подстановкой полученных значений переменных  $x_1^*, x_2^*$  в систему уравнений ограничений:  $x_3^* = 6; x_4^* = 0$ . Максимальное значение целевой функции равно 12.

Общим методом решения ЗЛП, представленных в стандартной форме, является так называемый *симплекс-метод*, имеющий различные модификации [14].

### 2.3. Целочисленное программирование

По сравнению с обычными задачами линейного программирования, в задачах целочисленного программирования появляется дополнительное ограничение в виде требования целочисленности всех переменных или их части в оптимальном решении. В первом случае задача называется полностью целочисленной, во втором — частично целочисленной. Такие

задачи целочисленного программирования возникают при распределении ресурсов, являющихся штучными изделиями.

На первый взгляд, для получения задачи целочисленного программирования достаточно решить задачу линейного программирования с той же целевой функцией и с теми же ограничениями, но без требования целочисленности переменных (такая задача называется *задачей с ослабленными ограничениями*). Затем, если полученные значения переменных окажутся дробными, округлить их. Однако такой метод не гарантирует получение оптимального решения, что можно видеть на следующем примере. Пусть необходимо решить задачу

$$\begin{cases} x_1 + 1,5x_2 \rightarrow \max; \\ 2x_1 + 4x_2 \leq 17; \\ 10x_1 + 4x_2 \leq 45; \\ x_1, x_2 \in N \cup \{0\}, \end{cases}$$

где  $N$  - обозначение множества натуральных чисел.

Для задачи с ослабленными ограничениями получаем следующее оптимальное решение:  $x_1^* = 3,5$ ;  $x_2^* = 2,5$ ;  $f^* = 7,25$ . Округлим полученные величины соответственно до значений 3 и 2 (округление до значений 4 и 3 недопустимо – при этом, как легко проверить, не выполняются ограничения-неравенства). Подставив округленные значения в выражение для целевой функции, получим  $f=6$ .

Между тем, ограничениям задачи удовлетворяют целые значения  $x_1^{**} = 2$ ;  $x_2^{**} = 3$ , при которых целевая функция достигает большего значения  $f^{**} = 6,5$ .

Для решения задач целочисленного программирования на практике применяются следующие методы.

### **2.3.1. Метод отсечения (метод Гомори)**

Основан на выделении из области  $G$  допустимых решений задачи с ослабленными ограничениями новой области  $G^0$ , для которой выполняются следующие условия:

Все точки с целыми координатами, принадлежащие области  $G$ , принадлежат и области  $G^0$ .

Область  $G^0$  является выпуклым многогранником.  
Все крайние точки области  $G^0$  имеют целочисленные координаты.

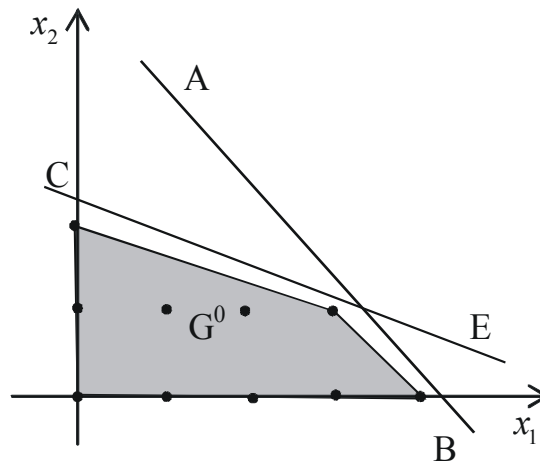


Рис. 2.2. Иллюстрация к методу отсечения

На рис. 2.2 область  $G$  ограничена координатными осями и прямыми  $AB$ ,  $CE$ , область  $G^0$  и принадлежащие ей точки с целочисленными координатами выделены цветом.

Замена множества  $G$  множеством  $G^0$  не приводит к «потере» оптимального решения задачи целочисленного программирования, так как  $G^0$  получено из  $G$  отсечением его подмножества, заведомо не содержащего допустимых целочисленных решений. В этом случае оптимальное решение задачи линейного программирования с ослабленными ограничениями на множестве  $G^0$  будет совпадать с решением соответствующей задачи целочисленного программирования на всем множестве  $G$ .

Выделение множества  $G^0$  осуществляется введением в систему ограничений дополнительных уравнений отсекающих плоскостей [15]. Отметим, что метод отсечений не обеспечивает высокой эффективности вычислений, особенно для задач большой размерности.

### 2.3.2. Метод ветвей и границ

Представляет собой комбинаторный метод, нашедший широкое применение при решении как полностью, так и частично целочисленных

задач. При использовании этого метода вначале решается соответствующая задача с ослабленными ограничениями, затем выбирается одна из переменных  $x_i$ , значение которой  $\beta_i$  не является целым числом. Поскольку интервал  $(\text{int}(\beta_i), \text{int}(\beta_i)+1)$ , где  $\text{int}(\beta_i)$  - целая часть числа  $\beta_i$ , не содержит целых значений  $x_i$ , то любое допустимое целое значение этой переменной в решении задачи целочисленного программирования удовлетворяет одному из неравенств:

$$\text{а) } x_i \leq \text{int}(\beta_i); \quad \text{б) } x_i \geq \text{int}(\beta_i)+1.$$

Таким образом, задача разбивается («разветвляется») на две новые задачи с непересекающимися множествами допустимых решений. Выбрав одну из этих «порожденных» задач, решают соответствующую ей задачу с ослабленными ограничениями – если полученное решение удовлетворяет требованию целочисленности, его фиксируют как наилучшее, и переходят к решению второй порожденной задачи; в противном случае первая задача вновь разбивается на две и т.д. Каждое новое полученное решение сравнивают с зафиксированным ранее и выбирают из них наилучшее. Процесс продолжается до полного исчерпания множества порожденных задач.

*Пример 2.4.* Решить задачу целочисленного программирования

$$\begin{cases} 2x_1 + 3x_2 \rightarrow \max; \\ 5x_1 + 7x_2 \leq 35; \\ 4x_1 + 9x_2 \leq 36; \\ x_1, x_2 \in N \cup \{0\}. \end{cases}$$

В оптимальном решении задачи с ослабленными ограничениями  $x_1^* = 3\frac{12}{17}$ ,  $x_2^* = 2\frac{6}{17}$  обе переменные принимают нецелочисленные значения. Произведем разбиение задачи на две новые по переменной  $x_2$  (рис. 2.3) - этот выбор порождает две новые задачи (задачи 2 и 3) с дополнительными условиями  $x_2 \leq 2$  и  $x_2 \geq 3$ , так как  $\text{int}(x_2^*) = 2$ .

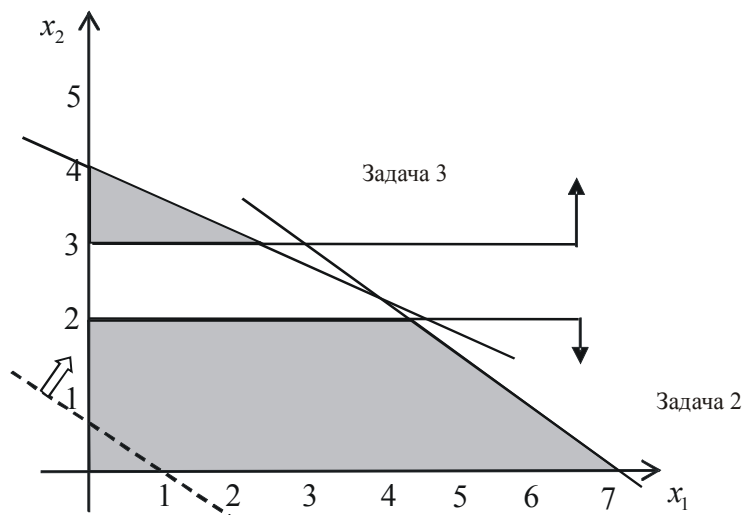


Рис. 2.3. Разбиение задачи 1 на задачи 2 и 3

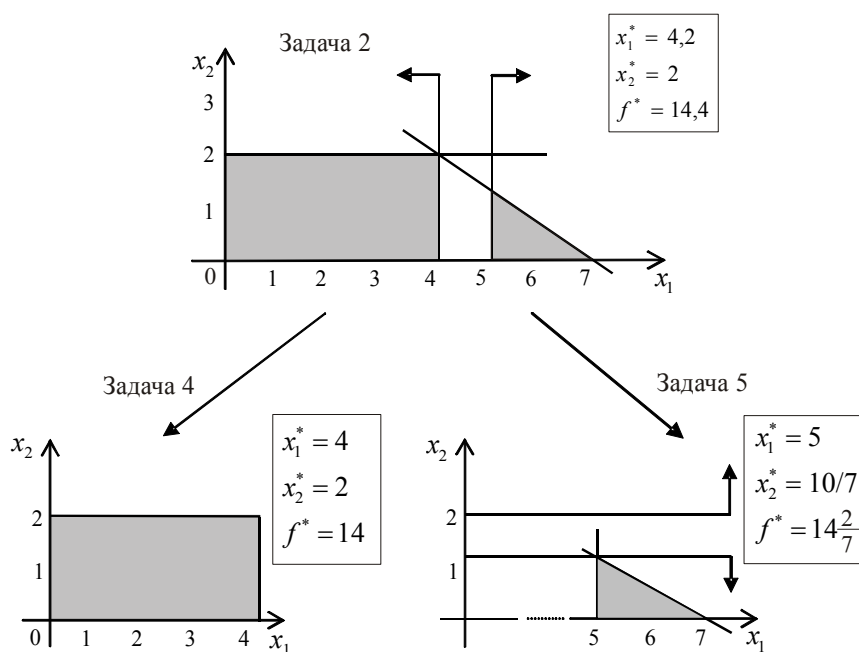


Рис. 2.4. Разбиение задачи 2 на задачи 4 и 5

В результате решения задачи 2, в свою очередь, получим порожденные задачи 4 и 5 – это показано на рис. 2.4, где в рамке рядом с рисунком приведено оптимальное решение.

Решение задачи 4 целочисленное, поэтому зафиксируем его как наилучшее. Производить разделение задачи 5 нет необходимости: действительно, поскольку коэффициенты целевой функции целочисленны, то при решении задач, порожденных задачей 5, не может получиться ее значение, превышающее  $\text{int}\left(14\frac{2}{7}\right) = 14$ .

Поэтому переходим к рассмотрению задачи 3.

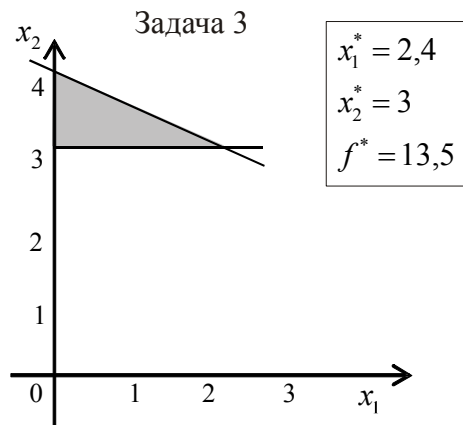


Рис. 2.5. Решение задачи 3

Можно видеть (см. рис. 2.5), что при ее решении получается значение целевой функции, меньшее, чем при решении задачи 4. Поэтому дальнейшее разбиение этой задачи не имеет смысла, если только не требуется найти все без исключения решения задачи целочисленного программирования, а не только оптимальное.

## 2.4. Динамическое программирование

### 2.4.1. Сущность метода динамического программирования

Динамическое программирование представляет собой особый метод решения задач оптимизации, применяемый для многоэтапных

(многошаговых) операций. Типичными задачами, решаемыми с помощью этого метода, являются задачи оптимального многоэтапного распределения ресурсов, задачи динамического управления запасами, задача поиска минимального пути.

В вычислительном аспекте метод динамического программирования представляет собой целенаправленный перебор вариантов решения, при этом число вариантов, подлежащих рассмотрению, получается намного меньше, чем при «простом», ненаправленном переборе. На заре развития вычислительной техники метод динамического программирования представлял собой единственный приемлемый метод точного решения задач многоэтапного принятия решения.

Пусть имеется операция, которая может быть разбита на  $n$  этапов. На каждом этапе могут быть выбраны параметры, от которых зависит как эффективность данного этапа, так и эффективность операции в целом. Управляемый параметр на  $j$ -ом этапе называется *частным, или локальным, управлением*  $x_j$ , а управление в целом характеризуется вектором  $\bar{x} = (x_1, x_2, \dots, x_n)^T$ . При этом в общем случае  $x_j$  – не обязательно число, а может, в свою очередь, являться вектором или функцией.

Для решения подобных задач используется метод динамического программирования, разработанный Ричардом Беллманом в 50-х гг. двадцатого века.

Этот метод основывается на следующем принципе оптимальности:

*Каково бы ни было начальное состояние системы и начальное управление, последующие управления должны быть оптимальны по отношению к состоянию, получающемуся в результате предыдущего решения, то есть по отношению к текущему состоянию системы.*

Рассмотрим сущность метода динамического программирования на примере следующей задачи распределения ресурсов. Пусть необходимо распределить по  $n$  этапам операции ресурсы с общим запасом  $b$ :

$$\sum_{j=1}^n x_j = b, \quad (2.1)$$

где  $x_j$  – количество ресурсов, выделенное для  $j$ -го этапа, при этом эффективность  $j$ -го этапа характеризуется функцией  $W_j(x_j)$ , а



эффективность операции в целом - функцией  $W = \sum_{j=1}^n W_j$ . Очевидно, переменная  $x_j$  здесь является локальным управлением.

Введем функцию

$$F_n(b) = \max_{\{x_j\}} W(x_1, x_2, \dots, x_n), \quad (2.2)$$

выражающую оптимальный доход от распределения данного запаса ресурсов.

Зафиксируем значение  $x_n$ ; тогда оставшееся количество ресурсов  $b - x_n$  целесообразно распределить таким образом, чтобы максимизировать доход по остальным  $n-1$  переменным. По определению функции  $F$ , оптимальный доход от распределения количества ресурсов  $b - x_n$  есть  $F_{n-1}(b - x_n)$ . Поэтому задание значения  $x_n$  приводит к общему доходу

$$W_n(x_n) + F_{n-1}(b - x_n). \quad (2.3)$$

Оптимальным будет такой выбор  $x_n$ , который максимизирует значение (2.3), т.е.

$$F_n(b) = \max_{0 \leq x_n \leq b} [W_n(x_n) + F_{n-1}(b - x_n)].$$

По аналогии можно записать

$$F_{n-1}(b - x_n) = \max_{0 \leq x_{n-1} \leq b - x_n} [W_{n-1}(x_{n-1}) + F_{n-2}(b - x_n - x_{n-1})].$$

В общем виде:

$$F_j(\xi) = \max_{0 \leq x_j \leq \xi} [W_j(x_j) + F_{j-1}(\xi - x_j)], \quad (2.4)$$

где  $j = 1, \dots, n$ ,  $\xi$  - количество ресурсов, распределяемых по этапам.

Выражение (2.4) называется *основным рекуррентным соотношением* метода динамического программирования, а величина  $\xi$  - *параметром состояния*.

Используя соотношение (2.4) последовательно для  $j = n, n-1, \dots$ , придем на последнем шаге к выражению

$$F_1(\xi) = \max_{0 \leq x_1 \leq \xi} W_1(x_1).$$

Чтобы решить задачу динамического программирования с использованием полученного рекуррентного соотношения, процесс вычислений необходимо вести в обратном порядке, начиная с шага номер 1. Варьируя значение параметра состояния  $\xi$  в пределах от 0 до  $b$ , найдем для каждого значения этого параметра значение функции  $F_1(\xi)$  и соответствующее ему *условное оптимальное локальное управление*  $\hat{x}_1(\xi)$ . На практике это можно реализовать путем табулирования функций  $F_1(\xi)$  и  $\hat{x}_1(\xi)$ .

Затем найдем зависимость

$$F_2(\xi) = \max_{0 \leq x_2 \leq \xi} [W_2(x_2) + F_1(\xi - x_2)],$$

используя при вычислении  $F_2(\xi - x_2)$  результаты первого шага, а также *условное оптимальное локальное управление*  $\hat{x}_2(\xi)$ , соответствующее функции  $F_2(\xi)$ .

Аналогичным образом процесс продолжим вплоть до получения зависимостей  $F_{n-1}(\xi)$  и  $\hat{x}_{n-1}(\xi)$ . В построении зависимостей  $F_n(\xi)$  и  $\hat{x}_n(\xi)$  необходимости нет: на последнем шаге параметр  $\xi$  фиксирован и равен общему запасу ресурсов  $b$ , поэтому сразу же получаем значение  $F_n(\xi)$  и *безусловное оптимальное локальное управление*  $x_n^*(b)$ .

Далее, используя уже полученную зависимость  $\hat{x}_{n-1}(\xi)$ , найдем  $x_{n-1}^* = \hat{x}_{n-1}(b - x_n^*)$ , затем  $x_{n-2}^* = \hat{x}_{n-2}(b - x_n^* - x_{n-1}^*)$ , ...,  $x_1^* = \hat{x}_1(b - x_n^* - x_{n-1}^* - \dots - x_2^*)$ .

В итоге будет получено оптимальное управление  $\bar{x}^* = (x_1^*, x_2^*, \dots, x_n^*)^T$  и  $F_n(b) = \max_{\{x_j\}} W(x_1, x_2, \dots, x_n)$ , то есть искомое решение.

Во многих задачах динамического программирования порядок их решения несуществен. Однако существуют задачи, где решение должно осуществляться в строго определенной последовательности. Для них возможны две схемы:

1) решение в прямом направлении, когда первый шаг решения

задачи соответствует первому по времени принимаемому решению в данной операции;

2) решение в обратном направлении, когда первый шаг решения задачи соответствует последнему по времени принимаемому решению в данной операции.

Первая схема применяется в случае, когда известно конечное состояние системы, для нее рекуррентное соотношение имеет вид (2.4). Вторая схема используется в случае, когда известно начальное состояние системы; для нее рекуррентное соотношение принимает вид

$$F_j(\xi) = \max_{0 \leq x_j \leq \xi} [W_j(x_j) + F_{j+1}(\xi - x_j)]. \quad (2.5)$$

#### 2.4.2. Примеры задач динамического программирования

Рассмотрим применение метода динамического программирования на следующих конкретных примерах.

*Пример 2.5.* Задача об использовании рабочей силы.

Требуется составить план набора работников на  $n$  месяцев, при этом требуемое число работников на каждый месяц составляет соответственно  $m_1, m_2, \dots, m_n$ . Число работников в начале работ -  $m_0$ . Затраты работодателя, связанные с наймом работников и их увольнением, определяются функцией

$$f_j(x_j - x_{j-1}) = \begin{cases} a \cdot (x_j - x_{j-1}), & \text{если } x_j \geq x_{j-1} \\ b \cdot (-x_j + x_{j-1}), & \text{если } x_j < x_{j-1} \end{cases}, f_j(0) = 0;$$

затраты, связанные с отклонением фактической численности работников от требуемой, определяются функцией

$$g_j(m_j - x_j) = \begin{cases} c \cdot (x_j - m_j), & \text{если } x_j \geq m_j \\ d \cdot (-x_j + m_j), & \text{если } x_j < m_j \end{cases}, g_j(0) = 0,$$

где  $x_j$  - количество работников в  $j$ -ом месяце,  $a, b, c, d$  - заданные числа.

Решим задачу при следующих исходных данных:  $n=4, m_1=2, m_2=5, m_3=3, m_4=1, a=10, b=7, c=8, d=11$ ; число рабочих в начале работ  $m_0=2$ .

Используем метод динамического программирования. Будем решать

задачу в обратном направлении, используя информацию о начальном состоянии системы ( $x_0 = m_0 = 2$ ).

Введем обозначения:

$$F_n(\xi) = \min_{x_n} [f_n(x_n - \xi) + g_n(x_n - m_n)];$$

$$\Omega_j(\xi, x_j) = f_j(x_j - \xi) + g_j(x_j - m_j) + F_{j+1}(x_j).$$

Основное рекуррентное соотношение, согласно (2.5), будет в данном случае иметь вид

$$F_j(\xi) = \min_{x_j} \Omega_j(\xi, x_j), \quad j = 1, \dots, n.$$

При этом  $F_{n+1} = 0$ .

*Шаг 1.* Для отыскания  $F_4(\xi)$  рассчитаем значения вспомогательной функции  $\Omega_4(x_4, \xi)$ . Так как функции  $f_j, g_j$  выпуклы по переменным  $x_j$ , то и функция  $\Omega_4(x_4, \xi)$  также выпукла по  $x_4$  для всех  $\xi$ . Поэтому для нахождения  $F_4(\xi)$  достаточно определить первый локальный минимум  $\Omega_4(x_4, \xi)$  по  $x_4$ , который будет также глобальным; после достижения этого первого минимума дальнейшее вычисление функции  $\Omega_4(x_4, \xi)$  не требуется. Расчет значений  $\Omega_4(x_4, \xi)$  приведен в табл. 2.1.

Таблица 2.1. Результаты первого шага решения

1. $\xi=0, x_4=0, \Omega_4=0+11(1-0)=11$ ; $\xi=0, x_4=1, \Omega_4=10(1-0)+0=10$ ; $\xi=0, x_4=2, \Omega_4=10(2-0)+8(2-1)=28$ .	2. $\xi=1, x_4=0, \Omega_4=7(1-0)+11(1-0)=18$ ; $\xi=1, x_4=1, \Omega_4=0+0=0$ .
3. $\xi=2, x_4=0, \Omega_4=7(2-0)+11(1-0)=25$ ; $\xi=2, x_4=1, \Omega_4=7(2-1)+0=7$ ; $\xi=2, x_4=2, \Omega_4=0+8(2-1)=8$ .	4. $\xi=3, x_4=0, \Omega_4=7(3-0)+11(1-0)=32$ ; $\xi=3, x_4=1, \Omega_4=7(3-1)+0=14$ ; $\xi=3, x_4=2, \Omega_4=7(3-2)+8(2-1)=15$ .
5. $\xi=4, x_4=0, \Omega_4=7(4-0)+11(1-0)=39$ ; $\xi=4, x_4=1, \Omega_4=7(4-1)+0=21$ ; $\xi=4, x_4=2, \Omega_4=7(4-2)+8(2-1)=22$	6. $\xi=5, x_4=0, \Omega_4=7(5-0)+11(1-0)=46$ ; $\xi=5, x_4=1, \Omega_4=7(5-1)+0=28$ ; $\xi=5, x_4=2, \Omega_4=7(5-2)+8=29$ .

Полученные результаты приведены в табл. 2.2 и 2.3.

Таблица 2.2. Сводные результаты первого шага решения

$\xi$	$x_4$	$\Omega_4$	$\xi$	$x_4$	$\Omega_4$
0	0	11	3	0	32
	1	10		1	14
	2	28		2	15
1	0	18	4	0	39
	1	0		1	21
				2	22
2	0	25	5	0	46
	1	7		1	28
	2	8		2	29

Таблица 2.3. Оптимальные значения  $x_4^*$

$\xi$	$F_4$	$x_4^*$
0	10	1
1	0	1
2	7	1
3	15	1
4	21	1
5	28	1

*Шаг 2.* Значения  $\Omega_3(x_3, \xi)$ ,  $F_3(x_3)$  и  $x_3^*$  вычисляются по аналогии. Результаты приведены в табл. 2.4 и 2.5.

Таблица 2.4. Результаты второго шага вычислений

$\xi$	$x_3$	$\Omega_3$	$\xi$	$x_3$	$\Omega_3$	$\xi$	$x_3$	$\Omega_3$	
0	0	40	2	2	18	4	4	29	
	1	32		3	24				
	2	38							
1	0	50	3	0	64	5	2	39	
	1	22		1	36				
	2	2		28	2				25
					3				14
2	0	57	4	2	32				
	1	29		3	21				

Таблица 2.5. Оптимальные значения  $x_3^*$ ,  $x_4^*$

$\xi$	$F_3$	$x_3^*$	$x_4^*$
0	32	1	1
1	22	1	1
2	18	2	1
3	14	3	1
4	21	3	1
5	28	3	1

*Шаг 3.* Результаты вычислений  $\Omega_2(x_2, \xi)$ ,  $F_2(x_2)$  и  $x_2^*$  приведены в табл. 2.6.

Таблица 2.6 Результаты второго шага решения и оптимальные значения  $x_2^*$ ,  $x_3^*$ ,  $x_4^*$

$\xi$	$x_3$	$\Omega_3$	$\xi$	$x_3$	$\Omega_3$
0	0	87	3	1	52
	1	76		2	58
	2	71		3	36
	3	66		4	42
1	2	61	4	2	65
	3	56		3	43
	4	62		4	32
2	2	51	5	4	39
	3	46		5	28

$\xi$	$F_2$	$x_2^*$	$x_3^*$	$x_4^*$
0	66	3	3	1
1	56	3	3	1
2	46	3	3	1
3	36	3	3	1
4	32	4	3	1
5	28	5	3	1

*Шаг 4.* Используя начальное условие и приравнивая  $\xi = x_0 = 2$ , получаем значения, приведенные в табл. 2.7.

Таблица 2.7

Далее из табл. 2.6 по значению  $\xi = x_1^* = 2$  находим минимальные значения остальных переменных:  $x_2^* = 3, x_3^* = 3, x_4^* = 1$ . При этом минимальные суммарные затраты составят 46.

$x_0$	$x_1^*$	$F_1$
2	1	74
	2	46
	3	54

*Пример 2.6.* Задача о замене оборудования.

Имеется некое оборудование, которое должно эксплуатироваться в течение  $n$  лет. Даны:  $c$  – начальная стоимость оборудования;  $z(i)$  – стоимость содержания в течение одного года оборудования, которое к началу рассматриваемого года эксплуатировалось в течение  $i$  лет;  $l(i)$  – ликвидная стоимость оборудования, возраст которого  $i$  лет.

В начале каждого года может быть применено одно из двух локальных управлений:  $x_1$  – продолжить эксплуатацию старого оборудования,  $x_2$  – продать старое оборудование по его ликвидной стоимости и приобрести новое. Требуется найти стратегию замены оборудования, минимизирующую суммарные затраты за  $n$  лет.

Состояние системы «оборудование» будем характеризовать переменной  $S_{jk}$ , где  $j$  – возраст оборудования, которое эксплуатируется в данный момент,  $k$  – количество лет с момента первого приобретения оборудования. Тогда под действием управления  $x_1$  система перейдет из состояния  $S_{jk}$  в состояние  $S_{j+1,k+1}$ , а под действием управления  $x_2$  – в состояние  $S_{1,k+1}$ .

Задачу удобно описать в терминах теории графов, поставив в соответствие состояниям системы вершины ориентированного графа, а переходам ее из одного состояния в другое – дуги. В качестве «длины» каждой дуги будем рассматривать суммарную стоимость соответствующего перехода.

Обозначим через  $F_{jk}$  длину минимального пути от вершины  $S_{00}$  до вершины  $S_{jk}$ ; тогда рекуррентное соотношение метода динамического программирования будет иметь вид:  $F_{jk} = \min_{S_{\alpha\beta} \in G} (F_{\alpha\beta} + c_{\alpha\beta,jk})$ , где  $G$  – множество вершин графа, в которых начинаются дуги, оканчивающиеся в вершине  $S_{jk}$ ,  $c_{\alpha\beta,jk}$  – длина дуги между вершиной  $S_{\alpha\beta}$  и вершиной  $S_{jk}$ .

Введем фиктивное состояние  $S_f$ , в которое переходит система в конце эксплуатации в результате продажи имеющегося оборудования. Очевидно, что критерием оптимальности рассматриваемой операции

является  $F_f$ .

Решим рассмотренную задачу при следующих исходных данных:  
 $n=4$ ,  $c=8000$ ,  $l(1)=4000$ ,  $l(2)=3000$ ,  $l(3)=1000$ ,  $z(0)=1400$ ,  $z(1)=1500$ ,  
 $z(2)=1700$ .

Граф, соответствующий задаче, представлен на рис. 2.6.

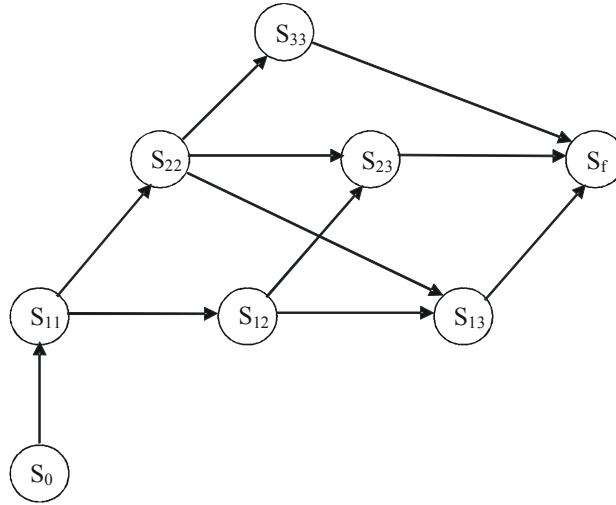


Рис. 2.6. Граф состояний системы «Оборудование»

Вычислим «длины» дуг:

$$c_{00,11} = c + z(0) = 9400; \quad c_{11,12} = c_{12,13} = c - l(1) + z(0) = 5400; \quad c_{11,22} = z(1) = 1500;$$

$$c_{12,23} = z(1) = 1500; \quad c_{13,f} = -l(1) = -4000; \quad c_{22,13} = c - l(2) + z(0) = 6400;$$

$$c_{23,f} = -l(2) = -3000; \quad c_{22,33} = z(2) = 1700; \quad c_{33,f} = -l(3) = -1000.$$

Далее рассчитаем значения  $F_{jk}$ :

$$F_{11} = c_{00,11} = 9400; \quad F_{12} = F_{11} + c_{11,12} = 14800; \quad F_{22} = F_{11} + c_{11,22} = 10900;$$

$$F_{13} = \min\{F_{22} + c_{22,13}; F_{12} + c_{12,13}\} = 17300; \quad F_{23} = F_{12} + c_{12,23} = 16300;$$

$$F_{33} = F_{22} + z(2) = 12600; \quad F_f = \min\{F_{13} + c_{13,f}; F_{23} + c_{23,f}; F_{33} + c_{33,f}\} = \\ = \min\{13300; 13300; 11600\} = 11600.$$

Таким образом, минимум целевой функции достигается при



движении в вершину  $S_f$  из вершины  $S_{33}$  и кратчайший путь состоит из последовательности вершин  $S_{00} \rightarrow S_{11} \rightarrow S_{22} \rightarrow S_{33} \rightarrow S_f$ , которой соответствует набор управлений  $\{x_1, x_1, x_1\}$ , то есть оптимальная стратегия заключается в продолжении эксплуатации старого оборудования в течение всех трех лет.

## 2.5. Метод статистического моделирования

Когда исход операции зависит от случайных факторов, не всегда удается построить аналитическую модель этой операции. В этом случае прибегают к методу статистического моделирования (используются также названия «метод статистических испытаний», «метод имитационного моделирования», «метод Монте-Карло»). Идея метода заключается в следующем. Осуществляется единичная реализация случайного явления (розыгрыш) с помощью процедуры, дающей случайный результат. Повторяя такие розыгрыши многократно, получают материал для последующей обработки методами математической статистики.

Этим методом может быть решена любая вероятностная задача. Однако его применение целесообразно только в тех случаях, когда реализация розыгрышей проще, чем аналитическое решение задачи.

*Пример 2.7.* По мишени произведено 4 выстрела с вероятностью попадания  $p=0,5$ . Мишень поражена, если в нее попали не менее двух раз.

Сымитировать событие поражения мишени весьма просто: берут 4 «правильных» монеты (таких, которые с равной вероятностью падают вверх «орлом» и «решкой» и никогда не падают на ребро) и подбрасывают. Выпадение «орла» означает попадание в мишень. Проведя серию таких опытов, подсчитывают количество тех из них, в которых выпало не менее двух «орлов», и делят это число на общее число опытов. Полученный результат даст частоту поражения мишени, приближенно равную искомой вероятности.

Наряду с тем, в данном случае вычислить вероятность поражения мишени проще аналитически. Действительно, случайное число попаданий

в мишень подчиняется биномиальному закону распределения. Пусть вероятность попадания при одном выстреле равна  $p$  - тогда вероятность поражения мишени

$$P_{пор} = 1 - P_0 - P_1 = 1 - \left[ C_4^0 \cdot (1-p)^4 \cdot p^0 + C_4^1 \cdot (1-p)^3 \cdot p \right] =$$

$$= 1 - \left[ \frac{4!}{4!0!} \cdot (1-0,5)^4 + \frac{4!}{4!1!} \cdot (1-0,5)^3 \cdot 0,5 \right] = 0,688,$$

где  $P_0, P_1, C_4^0, C_4^1$  - вероятности отсутствия попаданий, ровно одного попадания и соответствующие биномиальные коэффициенты.

*Пример 2.8.* На площадную цель сбрасывают  $n$  бомб, радиус поражения каждой бомбы равен  $R_n$ . Цель считается пораженной, если отношение пораженной площади  $S_{пор}$  к общей площади цели  $S_ц$  не менее заданного значения  $\alpha$ . Точки попадания распределяются по заданному закону. Требуется определить вероятность поражения цели.

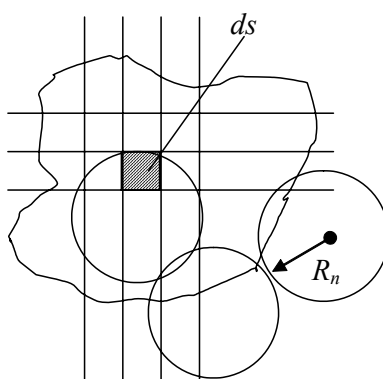


Рис. 2.7. Имитация бомбометания

Эта задача, при сложной конфигурации цели, не поддается аналитическому решению, но может быть довольно легко решена методом статистического моделирования. Площадь цели покрывается координатной сеткой, и с использованием заданного закона распределения вычисляются координаты точек попадания при одном бомбометании (рис. 2.7). Затем проводится сканирование цели: выбирается элементарный участок  $ds$  и проверяется условие  $ds \in \Omega_i$ , где  $\Omega_i$  - зона поражения  $i$ -ой бомбы. Если

условие выполняется, то значение счетчика числа пораженных элементарных площадок увеличивается на единицу:  $k := k + 1$ .

Проверяя все площадки, подсчитываем  $S_{нопj} = \sum_{i=1}^k ds_i = k \cdot ds_i$  - площадь, пораженную при  $j$ -ом эксперименте-«бомбометании». В каждом эксперименте проверяется выполнение условия  $\frac{S_{нопj}}{S_{ц}} \geq \alpha$ . Если оно выполняется, то цель поражена. В этом случае увеличивается на единицу значение другого счетчика:  $m = m + 1$ . Вероятность поражения цели:  $P_{ноп} \approx \frac{m}{N}$ , где  $N$  – общее число экспериментов.

### 2.5.1. Единичный жребий

Любой процесс статистического моделирования складывается из единичных реализаций случайного явления, которые часто называются единичным жребием. Единичный жребий сводится к получению значения случайной величины  $R$ , распределенной с равномерной плотностью на отрезке  $[0; 1]$ .

Рассмотрим ряд задач, решаемых с помощью единичного жребия.

1. Необходимо установить, произошло ли случайное событие  $A$  с вероятностью  $P_A$ .

Разыграем случайное число  $R$ . Тогда вероятность того, что полученное значение окажется меньше, чем  $P_A$ , равна

$$P(R < P_A) = \int_0^{P_A} f(r) dr,$$

$$\text{где } f(r) = \begin{cases} 1, & R \in [0; 1] \\ 0, & R \notin [0; 1] \end{cases} \Rightarrow P(R < P_A) = \int_0^{P_A} 1 \cdot dr = P_A.$$

Таким образом, если полученное значение  $R$  меньше, чем  $P_A$ , то событие  $A$  считается наступившим (рис. 2.8).

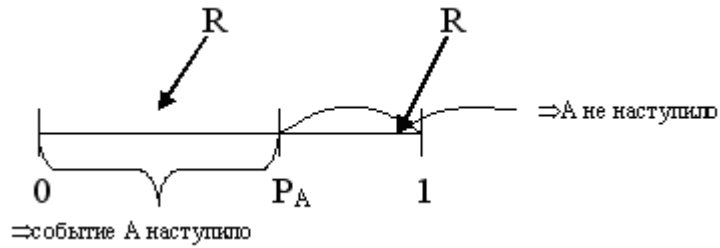


Рис. 2.8. Розыгрыш события A

2. Необходимо установить, какое из несовместных событий:  $A_1, A_2, \dots, A_n$  - произошло в опыте. Пусть  $P_i$  - вероятность события  $A_i$ .

Для решения задачи разыгрывается случайное число  $R$  и определяется тот из интервалов  $[0; P_1], [P_1; P_1 + P_2], \dots, [P_1 + P_2 + \dots + P_n; 1]$ , которому принадлежит  $R$  (рис. 2.9).

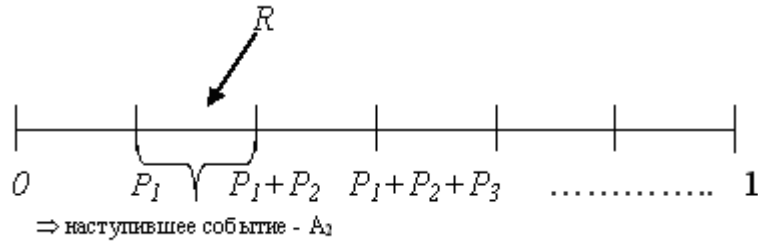


Рис. 2.9. Розыгрыш группы событий

Действительно:

$$P\left(\sum_{j=1}^{i-1} P_j < R < P_i + \sum_{j=1}^{i-1} P_j\right) = \int_{\sum_{j=1}^{i-1} P_j}^{P_i + \sum_{j=1}^{i-1} P_j} f(r) dr = \int_{\sum_{j=1}^{i-1} P_j}^{P_i + \sum_{j=1}^{i-1} P_j} 1 \cdot dr = \sum_{j=1}^{i-1} P_j + P_i - \sum_{j=1}^{i-1} P_j = P_i.$$

3. Необходимо установить: какое значение приняла случайная величина  $X$  в результате опыта.

Если  $X$  дискретна, то есть принимает значения  $X_1, X_2, \dots, X_n$  с вероятностями  $P_1, P_2, \dots, P_n$ , то схема розыгрыша ее значения сводится к случаю 2.

Покажем, как можно разыграть значение  $X$ , если это – непрерывная

величина с заданным законом распределения  $F(x)$  (рис. 2.10).

Возьмем случайную величину  $X$  и найдем функцию ее распределения, т.е. вероятность  $P(X < x)$ . Разыграем на отрезке  $[0;1]$  оси ординат случайную величину  $R$ . Видно (см. рис. 2.10), что для того, чтобы выполнялось неравенство  $X < x$ , величина  $R$  должна принять значение, меньшее  $F(x)$ :  $P(X < x) = P(R < F(x))$ .

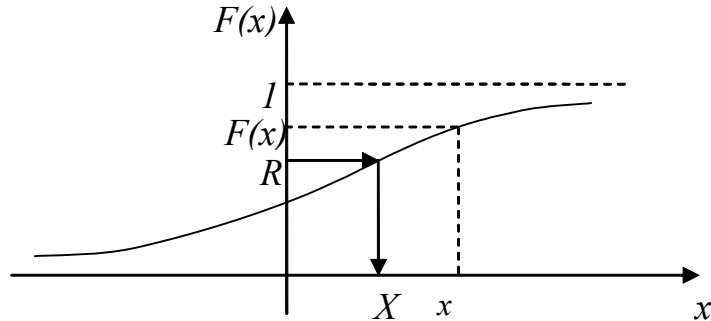


Рис. 2.10. Розыгрыш значения непрерывной случайной величины

Но поскольку случайная величина  $R$  имеет равномерную плотность распределения, то

$$P(X < x) = \int_0^{F(x)} f(r) dr = \int_0^{F(x)} 1 \cdot dr = F(x).$$

Отсюда следует: если взять случайное число  $R$ , равномерно распределенное на отрезке  $[0;1]$ , и найти то значение  $X$ , при котором  $F(X) = R$ , то полученная случайная величина  $X$  будет распределена по закону  $F(x)$ . При этом искомое значение  $X$  можно найти как обратную функцию от величины  $R$ :  $X = F^{-1}(R)$ .

*Пример 2.9.* Пусть случайная величина распределена по показательному закону  $f(x) = \lambda e^{-\lambda x}$ , Тогда  $F(x) = \int_0^x f(x) dx = 1 - e^{-\lambda x}$ , и значение такой случайной величины находится из следующих

соотношений:

$$e^{-\lambda x} = 1 - R \Rightarrow x = -\frac{1}{\lambda} \ln(1 - R).$$

4. Необходимо определить, какую совокупность значений примет система случайных величин  $(X_1, X_2, \dots, X_N)$ .

Если  $X_i$  независимы, то

$$f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n),$$

и розыгрыш совокупности значений системы сводится к последовательному розыгрышу каждой случайной величины независимо от остальных.

Если  $X_i$  зависимы, то

$$f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2/x_1) \cdot \dots \cdot f(x_n/x_1, \dots, x_{n-1}),$$

и схема розыгрыша будет следующей: сначала разыгрывается значение величины  $X_1$ , затем значение величины  $X_2$  при полученном фиксированном значении  $X_1$  и так далее.

*Пример 2.10.* Разыграем значение системы случайных величин  $(X_1, X_2)$  со следующими законами распределения (рис. 2.11):

$$f(x_1) = \begin{cases} 2x_1, & \text{если } x_1 \in [0; 1] \\ 0, & \text{если } x_1 \notin [0; 1]; \end{cases}$$

$$f(x_2/x_1) = \begin{cases} \frac{1}{2}, & \text{если } x_2 \in [x_1 - 1; x_1 + 1] \\ 0, & \text{если } x_2 \notin [x_1 - 1; x_1 + 1]. \end{cases}$$

$$F(x_1) = \int_0^{x_1} 2x dx = x_1^2,$$

$$F(x_2) = \int_{x_1-1}^{x_2} \frac{1}{2} dx = \frac{-x_1 + x_2 + 1}{2}.$$

Розыгрыш осуществляется по следующей схеме: получается случайная величина  $R_1$  и находится  $x_1 = \sqrt{R_1}$ , затем получается величина  $R_2$  и вычисляется  $X_2 = 2R_2 + x_1 - 1 = 2R_2 + \sqrt{R_1} - 1$ .

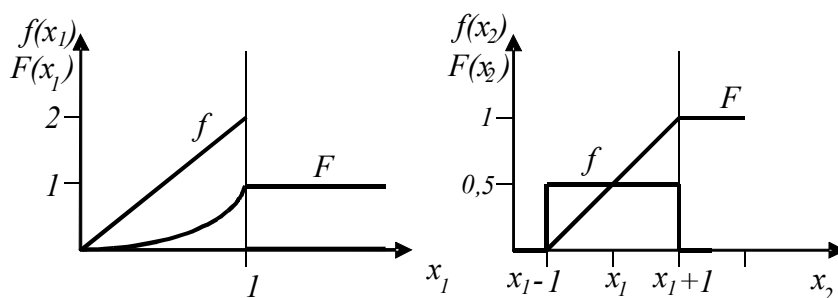


Рис. 2.11. Розыгрыш значения системы случайных величин

Из приведенных примеров можно видеть, что для решения различных задач методом статистического моделирования необходимо получать реализации случайной величины с равномерным законом распределения. Для этого могут использоваться таблицы или генераторы случайных величин, а на ЭВМ - специальные подпрограммы или функции, входящие в состав всех языков программирования высокого уровня. Так, в языке PASCAL для этого используется функция RANDOM.

### 2.5.2. Моделирование работы технического устройства

Техническое устройство состоит из трех узлов:  $V_1$ ,  $V_2$ ,  $V_3$  - и должно работать в течение времени  $\tau$ . Чтобы устройство было работоспособным, необходимо, чтобы работали одновременно узлы  $V_1$  и  $V_2$ .

Узел  $V_3$  обеспечивает работу узлов  $V_1$  и  $V_2$ , при отказе  $V_3$  режим работы этих узлов становится более напряженным. Случайное время работы узлов распределено с плотностями  $f_1(t)$ ,  $f_2(t)$ ,  $f_3(t)$  соответственно.

Узлы  $V_1$  и  $V_2$  в случае отказа заменяются запасными, при этом имеются два запасных узла  $V_1$  и три запасных узла  $V_2$ . Узел  $V_3$  не восстанавливается. При отказе  $V_3$  плотность распределения времени безотказной работы узла  $V_1$  изменяется и становится равной  $f_1(t/t_1)$ , где  $t_1$  - это время, в течение которого узел  $V_1$  уже работал до момента выхода из строя узла  $V_3$ . По аналогии плотность распределения времени

безотказной работы узла  $Y_2$  становится равной  $f_2(t/t_2)$ . Время замены узлов  $Y_1$  и  $Y_2$  распределено соответственно с плотностями  $\varphi_1(t)$  и  $\varphi_2(t)$ .

Требуется найти:

1.  $P_{ii}(t)$  – вероятность исправной работы в текущий момент времени.
2.  $P_{y_i}$  – вероятность того, что отказ устройства произойдет из-за нехватки запасных узлов  $Y_1$ .
3.  $\bar{t}_p$  – среднее время работы устройства.
4.  $\bar{y}_1, \bar{y}_2$  – среднее количество израсходованных узлов  $Y_1$  и  $Y_2$  за все время работы.

Воспользуемся для решения задачи методом статистического моделирования. Введем четыре оси времени: на первой будем отмечать состояние узла  $Y_1$ , на второй – узла  $Y_2$ , на третьей – узла  $Y_3$  и на четвертой – состояние системы в целом (рис. 2.12).

Вначале разыграем случайное время  $T_3$ , которое пройдет до момента отказа узла  $Y_3$ . Используя функцию распределения  $F_3(t) = \int_0^t f_3(t) dt$ ,

найдем:

$$T_3 = F_3^{-1}(R_3), \text{ где } R_3 \in [0;1]. \quad (2.6)$$

Отложим это значение на третьей оси.

Далее разыграем  $T_1$  – случайное время до отказа узла  $Y_1$ :

$$T_1 = F_1^{-1}(R_1), \text{ где } F_1(t) = \int_0^t f_1(t) dt. \quad (2.7)$$

Этот момент времени отложим на первой оси.

Разыграем случайное время, которое уйдет на замену узла  $Y_1$ :

$$\tau_1 = \Phi_1^{-1}(R_1), \text{ где } \Phi_1(t) = \int_0^t f_1(t) dt. \quad (2.8)$$

Если  $T_1 + \tau_1 < T_3$ , то есть в момент начала работы запасного узла  $Y_1$  третий узел еще работает, то снова разыграем время работы запасного узла с помощью соотношений (2.7), а затем – время его замены (соотношения (2.8)).



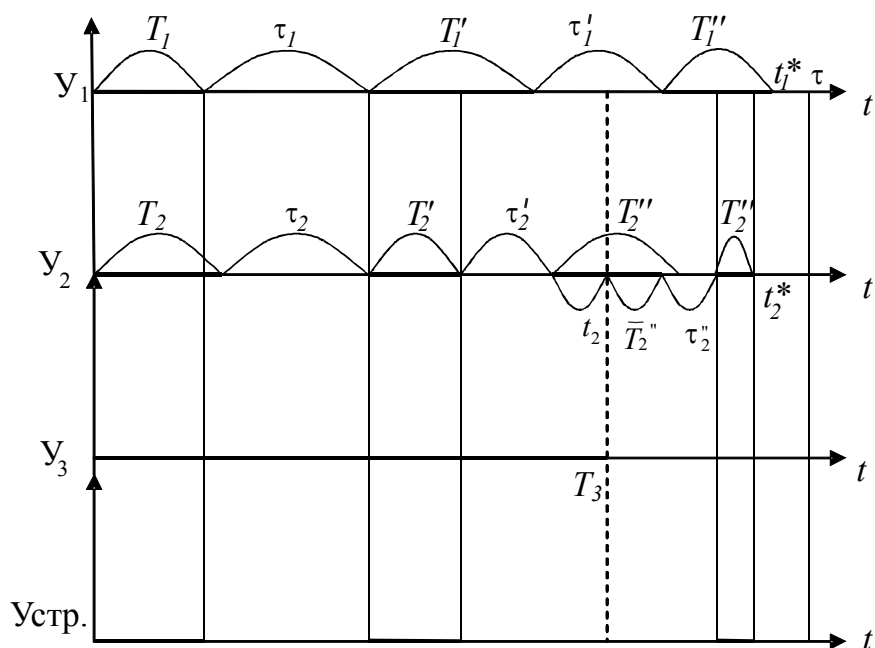


Рис. 2.12. Состояния устройства и его узлов

Пусть  $T_1 + \tau_1 + T_1' + \tau_1' > T_3$  - новый запасной узел начинает работать в условиях отказа узла  $Y_3$ . Тогда для розыгрыша времени безотказной работы запасного узла  $Y_1$  уже нужно использовать не функцию  $F_1(t)$ , а

условную функцию  $F_1(t/0) = \int_0^t f_1(t/0) dt$ , учтя, что до момента отказа узла

$Y_3$  новый узел  $Y_1$  не работал ( $t_1 = 0$ ).

Допустим, что, как показано на рисунке, для момента  $t_1^*$  отказа последнего запасного узла  $Y_1$  выполняется условие  $t_1^* < \tau$  - это значит, что в данный момент все устройство отказало ввиду исчерпания запаса узлов  $Y_1$ .

По аналогии, с использованием соответствующих функций

$F_2, \Phi_2, F_2(t/t_2)$ , разыгрываются состояния узла  $U_2$ . Пусть момент отказа узла  $U_3$  пришелся на период работы второго запасного узла  $U_2$ . Тогда в величину  $T_2$  нужно внести поправку, разыграв время работы узла  $U_2$ , оставшееся после отказа узла  $U_3$ , по условному закону  $F_2(t/t_2)$ .

Полученное при этом время  $\tilde{T}_2$  необходимо прибавить к уже проработанному времени  $t_2$ .

После завершения моделирования определяются интервалы, на которых работает устройство в целом – они выделены на последней оси.

Таким образом, получается одна реализация процесса функционирования технического устройства.

Введем индикаторную величину:

$$X_i(t) = \begin{cases} 1, & \text{если при } i\text{-ой реализации устройство работает в момент } t; \\ 0 & \text{- в противном случае.} \end{cases}$$

Тогда вероятность исправного состояния в момент времени  $t$  определяется по формуле:

$$P_u(t) \cong \frac{\sum_{i=1}^N X_i(t)}{N}, \text{ где } N - \text{общее число реализаций.}$$

На рис. 2.13 показана тенденция изменения величины  $P_u(t)$

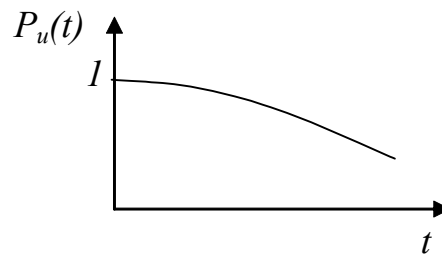


Рис. 2.13. Изменение вероятности исправной работы устройства

2. Вероятность отказа из-за нехватки узлов  $U_1$  можно найти, если в формуле для расчета  $P_u$  величину  $X_i$  заменить величиной.

$$Z_i(t) = \begin{cases} 1, & \text{если при } i\text{-ой реализации } t_1^* < \tau; \\ 0 & \text{- в противном случае.} \end{cases}$$

3. Среднее время работы устройства  $\bar{t}_p \cong \frac{\sum_{i=1}^N \bar{t}_{p_i}}{N}$ , где  $\bar{t}_{p_i}$  - время

работы устройства в  $i$ -ой реализации (равное сумме длин отрезков, выделенных на оси «Устр.»).

4. Среднее количество израсходованных узлов  $\bar{y}_1$  и  $\bar{y}_2$ :

$$\bar{y}_1 \cong \frac{\sum_{i=1}^N y_{1i}}{N}, \quad \bar{y}_2 \cong \frac{\sum_{i=1}^N y_{2i}}{N},$$

где  $y_{1i}$ ,  $y_{2i}$  - соответственно количество запасных узлов  $Y_1$  и  $Y_2$ , использованных в  $i$ -ой реализации.

## 2.6. Использование марковских случайных процессов в исследовании операций

### 2.6.1. Марковские случайные процессы

Если некая система случайным образом переходит из одного состояния в другое, то говорят, что в системе протекает случайный процесс.

Случайный процесс, протекающий в системе, называется *марковским*, если для любого момента времени  $t_0$  вероятностные характеристики этого процесса в будущем зависят только от состояния системы в момент времени  $t_0$  и не зависят от того, когда и как система пришла в это состояние (рис. 2.14).

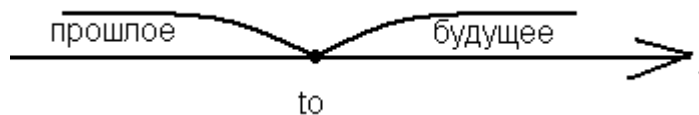


Рис. 2.14. К определению марковского процесса

Случайный процесс называется случайным процессом с дискретным временем, если изменение состояний системы происходит в заранее фиксированные определенные моменты времени, а между этими моментами состояние системы не меняется. Процесс называется процессом с непрерывным временем, если изменения состояний системы возможны в любые случайные моменты времени.

Рассмотрим процесс с дискретным временем. Пусть имеется некая система с возможным числом состояний  $m$ . Обозначим состояние системы в момент времени  $t_k$  как  $S_j^{(k)}$ , если система находится в  $j$ -ом состоянии. В этом случае последовательность событий  $S_j^{(k)}$  образует цепочку, причем события  $S_j^{(k)}$  для любого  $k$  будут несовместны.

Если для любого  $k$ , а также любых  $i, j$  вероятность перехода из состояния  $S_i^{(k)}$  в состояние  $S_j^{(k)}$  не зависит от того, когда и как система пришла в состояние  $S_i^{(k)}$ , то такая цепь событий называется марковской цепью.

Введем следующие обозначения:  $p(k) = (p_1(k), p_2(k), \dots, p_m(k))^T$  - вектор вероятностей состояний системы в момент  $t_k$ , (иначе говоря, на  $k$ -ом шаге);  $P = \|p_{ij}\|$  - матрица переходных вероятностей (размерности  $m \times m$ ). Компонента матрицы  $p_{ij}$  есть условная вероятность того, что система, находящаяся на  $k-1$ -ом шаге в состоянии  $S_i$ , на  $k$ -ом шаге перешла в состояние  $S_j$ . Компоненты матрицы называются переходными вероятностями. Если значения переходных вероятностей не зависят от номера шага, то марковская цепь называется однородной, иначе - неоднородной.

Поскольку в любой момент времени состояния системы образуют полную группу несовместных событий, то

$$\sum_{i=1}^m p_i(k) = 1; \sum_{j=1}^m p_{ij}(k) = 1, \forall k, i \quad (2.9)$$

Будем рассматривать события  $S_1^{(k-1)}, S_2^{(k-1)}, \dots, S_m^{(k-1)}$  как гипотезы. Тогда, по формуле полной вероятности,

$$p_j(k) = \sum_i p_i(k-1) p_{ij}(k). \quad (2.10)$$

*Пример 2.11.* По цели производится два выстрела. Возможные состояния цели:  $S_1$  – невредима;  $S_2$  – повреждена;  $S_3$  – поражена. Процесс изменения состояний системы можно рассматривать как марковский.

Удобно представить его в виде следующего графа состояний (рис. 2.15); рядом со стрелками, обозначающими смену состояний, указаны соответствующие переходные вероятности. В начальный момент времени  $p(0) = (1, 0, 0)^T$  – цель не повреждена.

Согласно (2.9),  $p_{11} = 1 - p_{12} - p_{13} = 0,4$ ;  $p_{22} = 1 - p_{23} = 0,5$ ;  $p_{33} = 1$ .

Соответственно, матрица переходных вероятностей имеет вид:

$$P = \begin{pmatrix} 0,4 & 0,4 & 0,2 \\ 0 & 0,5 & 0,5 \\ 0 & 0 & 1 \end{pmatrix}.$$

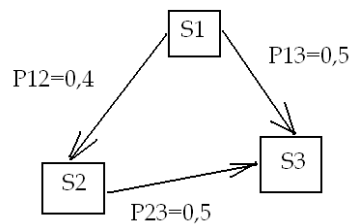


Рис. 2.15. Состояния цели

Согласно (2.10), после первого выстрела

$$p_1(1) = 1 \cdot 0,4 + 0 \cdot 0 + 0 \cdot 0 = 0,4;$$

$$p_2(1) = 1 \cdot 0,4 + 0 \cdot 0,5 + 0 \cdot 0 = 0,4;$$

$$p_3(1) = 1 \cdot 0,2 + 0 \cdot 0,5 + 0 \cdot 1 = 0,2.$$

После второго выстрела

$$p_1(2) = 0,4 \cdot 0,4 + 0 = 0,16;$$

$$p_2(2) = 0,4 \cdot 0,4 + 0,4 \cdot 0,5 + 0 = 0,36;$$

$$p_3(2) = 0,4 \cdot 0,2 + 0,4 \cdot 0,5 + 0,2 \cdot 1 = 0,48.$$

### 2.6.2. Поток событий

*Потоком событий* называется последовательность событий, следующих друг за другом в какие-то случайные моменты времени (рис. 2.16). Потоки используются в системах с дискретными состояниями и непрерывным временем [12].

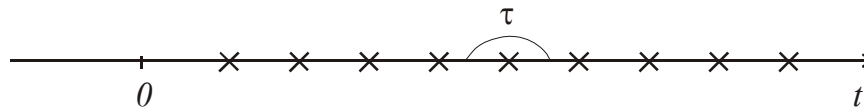


Рис. 2.16. Поток событий

Поток называется *регулярным*, если события следуют друг за другом через строго определенные промежутки времени.

Поток называется *стационарным*, если вероятность попадания того или иного числа событий на участок длиной  $\tau$  (см. рис. 2.16) зависит только от длины участка и не зависит от того, где именно на оси  $0t$  расположен этот участок.

Поток событий называется потоком *без последействия*, если для любых пересекающихся участков времени число событий, попадающих на один из них, не зависит от того, сколько событий попало на другой (другие).

Поток событий называется *ординарным*, если вероятность попадания на элементарный участок  $\Delta t$  двух и более событий пренебрежительно мала по сравнению с вероятностью попадания одного события.

Поток, обладающий свойствами стационарности, ординарности и отсутствия последействия, называется *простейшим или стационарным пуассоновским*.

Поток, обладающий свойством отсутствия последействия и

ординарный, но не стационарный, называется *нестационарным пуассоновским*.

*Интенсивность* событий в потоке  $\lambda$  называется среднее число событий в единицу времени.

В пуассоновском потоке число событий  $m$ , попадающих на интервал времени заданной длины, распределено по закону Пуассона:  $P_m = \frac{a^m}{m!} e^{-a}$ ,

где  $a = \lambda\tau$ ,  $\tau$  - длина интервала.

Выведем формулу для распределения случайной длины  $T$  интервала времени между двумя последовательными событиями в потоке Пуассона. Возьмем на оси  $t$  точку  $t_0$  - момент, в который появилось событие в потоке, и обозначим случайную длительность интервала времени до появления следующего события буквой  $T$ .

По определению, функция распределения величины  $T$   $F_T(t) = P(T < t)$ , то есть это вероятность того, что на интервал длиной  $t$ , примыкающий к точке  $t_0$ , попадет хотя бы одно событие. Выразим эту вероятность через вероятность противоположного события  $P_0(t)$ :

$$F(t) = 1 - P_0(t) = 1 - \frac{a^0}{0!} e^{-a}, \quad a = \lambda t \Rightarrow F(t) = 1 - e^{-\lambda t},$$

тогда плотность распределения величины  $T$   $f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t}$ .

Математическое ожидание длины интервала между двумя последовательными событиями есть  $m_t = \int_0^{\infty} t \lambda e^{-\lambda t} dt$ .

Интегрированием по частям найдем, что  $m_t = \lambda^{-1}$ , то есть это величина, обратная интенсивности потока. По аналогии можно найти, что дисперсия длины интервала  $D_t = \lambda^{-2}$ , а среднее квадратичное отклонение  $\sigma_t = \lambda^{-1}$ .

Выведем формулу для вероятности  $P_1(\Delta t)$  появления события на малом интервале  $\Delta t$ . Эта вероятность  $P_1(\Delta t)$  называется *элементом*

вероятности. Учитывая, что поток ординарный, появление на нем более, чем одного события практически невозможно, поэтому

$$P_1(\Delta t) \approx 1 - P_0(\Delta t) = 1 - e^{-\lambda \Delta t}.$$

Разлагая  $e^{-\lambda \Delta t}$  в ряд по степеням  $\lambda \Delta t$  и пренебрегая величинами высшего порядка малости, получим:

$$P_1(\Delta t) \approx 1 - (1 - \lambda \Delta t) = \lambda \Delta t. \quad (2.11)$$

Представляет также практический интерес поток событий, получающийся из простейшего потока путем его «просеивания»: сохраним в исходном потоке каждое  $k+1$ -ое событие, а остальные выбросим. Полученный поток называется *потоком Эрланга  $k$ -го порядка ( $\mathcal{E}_k$ )*. Очевидно, что простейший поток можно рассматривать как поток Эрланга нулевого порядка.

Можно показать [12], что плотность закона распределения случайного интервала времени между событиями в потоке  $\mathcal{E}_k$  выражается следующей формулой  $f_k(t) = \frac{\lambda(\lambda t)^k}{k!} e^{-\lambda t}$ .

Математическое ожидание и дисперсия интервала времени в потоке  $\mathcal{E}_k$  имеют вид

$$m_t = \frac{k+1}{\lambda}, \quad D_t = \frac{k+1}{\lambda^2}.$$

Выразив интенсивность через  $m_t$ , можно записать:

$$D_t = \frac{m_t^2}{k+1}. \quad (2.12)$$

Зафиксировав  $m_t$  и изменяя  $k$ , можно получить поток с любой степенью последствия. Действительно, при  $k=0$  получаем простейший поток, в котором последствие отсутствует, а при  $k=\infty$  - регулярный поток, в котором моменты появления событий жестко связаны между собой.

Часто бывает удобно заменить реальный поток событий потоком Эрланга, имеющим примерно такие же значения  $m_t$  и  $D_t$ .

*Пример 2.12.* Статистической обработкой данных о значениях



промежутков между событиями в потоке получены оценки для математического ожидания и дисперсии величины  $T$ :  $m_t = 2$  мин,  $D_t = 0,8$  мин<sup>2</sup>. Заменяем этот поток потоком Эрланга с теми же характеристиками. Используя выражение (2.12), найдем  $k + 1 = m_t^2 / D_t \approx 5$ . Таким образом, исходный поток можно заменить потоком Эрланга четвертого порядка с интенсивностью  $\lambda = 1/m_t = 0,5$  (мин<sup>-1</sup>).

### 2.6.3. Пуассоновские потоки событий и непрерывные марковские цепи

В отличие от дискретных марковских цепей, здесь рассматривается система с конечным числом состояний, в которой переход из состояния в состояние может произойти в произвольный момент времени.

Эти переходы можно рассматривать как происходящие под воздействием пуассоновского потока событий: как только появляется очередное событие потока, система переходит в другое состояние. Поскольку пуассоновский поток не имеет последствий, то состояние системы в будущем обусловлено только появлением события в потоке и не зависит от предыстории, следовательно, в системе протекает марковский случайный процесс.

Получим уравнения, описывающие изменение вероятностей состояний системы. Для этого возьмем произвольную непрерывную марковскую цепь, изображенную на рис. 2.17, где рядом со стрелками, обозначающими возможные переходы, указаны их интенсивности.

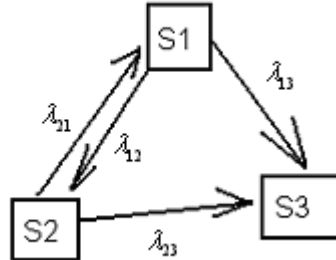


Рис. 2.17. Граф изменения состояний системы

Обозначим вероятности состояний  $S_1, S_2, S_3$  в момент времени  $t$  соответственно  $P_1(t), P_2(t), P_3(t)$  и найдем вероятность состояния  $P_1(t + \Delta t)$ . В момент времени  $t + \Delta t$  система может оказаться в состоянии  $S_1$  в результате одного из двух событий:

- событие  $A$ : в момент  $t$  система находилась в состоянии  $S_1$  и за время  $\Delta t$  не перешла в другое состояние; с учетом формулы для элемента вероятности (2.11), вероятность этого события  $P_A = P_1(t) \cdot (1 - \lambda_{12}\Delta t - \lambda_{13}\Delta t)$ ;
- событие  $B$ : в момент  $t$  система находилась в состоянии  $S_2$  и за время  $\Delta t$  перешла в состояние  $S_1$ ; эта вероятность равна  $P_B = P_2(t) \cdot \lambda_{21} \cdot \Delta t$ .

Поскольку события  $A$  и  $B$  несовместны, то искомая вероятность равна сумме вероятностей этих событий:

$$P_1(t + \Delta t) = P_1(t) \cdot (1 - \lambda_{12}\Delta t - \lambda_{13}\Delta t) + P_2(t) \cdot \lambda_{21} \cdot \Delta t.$$

После преобразований, устремив  $\Delta t$  к нулю, получим:

$$\frac{dP_1(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} = -(\lambda_{12} + \lambda_{13})P_1(t) + \lambda_{21}P_2(t).$$

Рассуждая аналогично, для вероятностей состояний  $S_2$  и  $S_3$  получим еще два дифференциальных уравнений. Для рассматриваемой марковской системы в целом получим систему уравнений

$$\begin{cases} \frac{dP_1(t)}{dt} = -(\lambda_{12} + \lambda_{13})P_1(t) + \lambda_{21}P_2(t), \\ \frac{dP_2(t)}{dt} = -(\lambda_{21} + \lambda_{23})P_2(t) + \lambda_{12}P_1(t), \\ \frac{dP_3(t)}{dt} = \lambda_{13}P_1(t) + \lambda_{23}P_2(t). \end{cases} \quad (2.13)$$

Полученные дифференциальные уравнения для вероятностей состояний непрерывной марковской цепи называются *уравнениями Колмогорова*. Заметим, что с учетом нормировочного условия

$$P_1(t) + P_2(t) + P_3(t) = 1,$$

любое из уравнений системы (2.13) при расчетах можно отбросить.

Представляет интерес вопрос о существовании вероятностей состояний при  $t \rightarrow \infty$ . Если эти пределы существуют, то они называются

предельными (финальными) вероятностями состояний.

Доказано, что если число состояний системы конечно и из каждого из них можно (за конечное число шагов) перейти в любое другое, то предельные вероятности существуют.

При  $t \rightarrow \infty$  в системе устанавливается предельный стационарный режим, в котором вероятности состояний уже не зависят от времени, следовательно,  $\frac{dP_i}{dt} = 0$ ,  $i = 1, \dots, n$ , где  $n$  – общее число состояний системы.

Поэтому, приравняв правые части уравнений Колмогорова нулю, получим систему алгебраических уравнений для вероятностей предельных состояний  $P_i$ . Поскольку эти уравнения однородны (не имеют свободного члена), то они определяют неизвестные с точностью до постоянного множителя.

Однако использование нормировочного условия  $\sum_{i=1}^n P_i = 1$

обеспечивает однозначность решения этой системы. При этом любое из уравнений системы можно отбросить – оно будет выполняться автоматически. Физический смысл предельной вероятности  $P_i$  следующий – это относительное время пребывания системы в  $i$ -ом состоянии после установления стационарного режима.

#### 2.6.4. Процесс «гибели и размножения»

В марковской системе протекает процесс «гибели и размножения», если все состояния этой системы можно расположить в цепочку так, что каждое из них, кроме крайних, связано прямой и обратной связью с двумя соседними, а первое и последнее – с одним соседним (рис. 2.18).

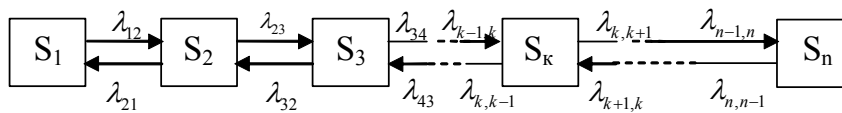


Рис. 2.18. Смена состояний системы в процессе гибели и размножения

Для этой системы вероятности состояний подчиняются уравнениям Колмогорова. Но на практике основной интерес представляют предельные вероятности состояний.

Используя граф состояний системы, можно записать уравнения Колмогорова; приравняв их правые части нулю, получим следующие соотношения:

$$\begin{cases} \lambda_{12}P_1 = \lambda_{21}P_2, \\ \lambda_{21}P_2 + \lambda_{23}P_2 = \lambda_{12}P_1 + \lambda_{32}P_3, \\ \lambda_{32}P_3 + \lambda_{34}P_3 = \lambda_{23}P_2 + \lambda_{43}P_4, \\ \dots \\ \lambda_{k,k-1}P_k + \lambda_{k,k+1}P_k = \lambda_{k-1,k}P_{k-1} + \lambda_{k+1,k}P_{k+1}, \\ \dots \\ \lambda_{n,n-1}P_n = \lambda_{n-1,n}P_{n-1}. \end{cases}$$

Последовательно подставляя первое уравнение во второе, второе в третье и т.д., приходим к следующим выражениям:

$$\begin{cases} P_2 = \frac{\lambda_{12}}{\lambda_{21}} P_1, \\ P_3 = \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} P_1, \\ \dots \\ P_n = \frac{\lambda_{12}\lambda_{23}\dots\lambda_{n-1,n}}{\lambda_{21}\lambda_{32}\dots\lambda_{n,n-1}} P_1. \end{cases} \quad (2.14)$$

Учитывая нормировочное условие, получим:

$$P_1 + \frac{\lambda_{12}}{\lambda_{21}} P_1 + \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} P_1 + \dots + \frac{\lambda_{12}\lambda_{23}\dots\lambda_{n-1,n}}{\lambda_{21}\lambda_{32}\dots\lambda_{n,n-1}} P_1 = 1,$$

$$\text{откуда } P_1 = \left( 1 + \frac{\lambda_{12}}{\lambda_{21}} + \frac{\lambda_{12}\lambda_{23}}{\lambda_{21}\lambda_{32}} + \dots + \frac{\lambda_{12}\lambda_{23}\dots\lambda_{n-1,n}}{\lambda_{21}\lambda_{32}\dots\lambda_{n,n-1}} \right)^{-1}. \quad (2.15)$$

Уравнения (2.14), (2.15) позволяют найти значения всех предельных вероятностей состояний системы.

*Пример 2.13.* Найдем предельные вероятности состояний системы с графом состояний, изображенным на рис. 2.21.

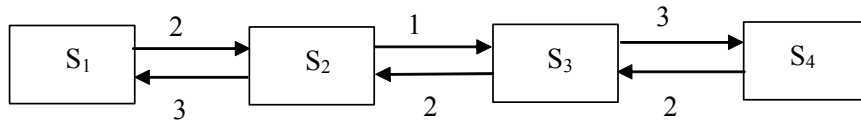


Рис. 2.21. Граф состояний системы из примера 2.13

По формулам (2.14), (2.15) получаем предельные вероятности:

$$P_1 = \left( 1 + \frac{2}{3} + \frac{2 \cdot 1}{3 \cdot 2} + \frac{2 \cdot 1 \cdot 3}{3 \cdot 2 \cdot 2} \right)^{-1} = \frac{2}{5},$$

$$P_2 = \frac{2}{3} \cdot \frac{2}{5} = \frac{4}{15},$$

$$P_3 = \frac{1}{3} \cdot \frac{2}{5} = \frac{2}{15},$$

$$P_4 = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5}.$$

### 2.6.5. Циклический процесс

Марковский случайный процесс называется циклическим, если состояния системы, в которой он протекает, связаны между собой в цикл с односторонними переходами (рис. 2.20).

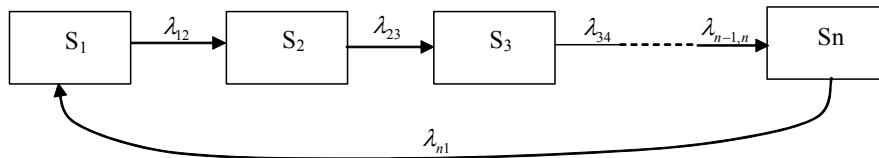


Рис. 2.20. Смена состояний системы в циклическом процессе

Уравнения для предельных вероятностей состояний в этом случае запишутся следующим образом:

$$\left\{ \begin{array}{l} \lambda_{12}P_1 = \lambda_{23}P_2, \\ \lambda_{23}P_2 = \lambda_{34}P_3, \\ \dots\dots\dots \\ \lambda_{n-1,n}P_{n-1} = \lambda_{n,1}P_n, \\ \lambda_{n,1}P_n = \lambda_{12}P_1. \end{array} \right. \Rightarrow \left\{ \begin{array}{l} P_2 = P_1 \frac{\lambda_{12}}{\lambda_{23}}, \\ P_3 = P_2 \frac{\lambda_{23}}{\lambda_{34}} = P_1 \frac{\lambda_{12}\lambda_{23}}{\lambda_{23}\lambda_{34}} = P_1 \frac{\lambda_{12}}{\lambda_{34}}, \\ \dots\dots\dots \\ P_k = P_1 \frac{\lambda_{12}}{\lambda_{k,k+1}}, \\ \dots\dots\dots \\ P_n = P_1 \frac{\lambda_{12}}{\lambda_{n,1}}. \end{array} \right. \quad (2.16)$$

Подставив (2.16) в нормировочное условие, получим выражение для вероятности  $P_1$ :

$$P_1 = \left( 1 + \frac{\lambda_{12}}{\lambda_{23}} + \frac{\lambda_{12}}{\lambda_{34}} \dots + \frac{\lambda_{12}}{\lambda_{n,1}} \right)^{-1}, \quad (2.17)$$

которое совместно с формулами (2.16) позволяет найти предельные вероятности всех состояний.

Формулам (2.16), (2.17) можно придать более простой вид, используя следующее обстоятельство: если в системе протекает марковский случайный процесс, то время, в течение которого система *еще останется* в данном текущем состоянии  $S_i$ , не зависит от того, как долго она *уже в нем находится*. Иначе говоря, на систему в состоянии  $S_i$  действует простейший «поток уходов» с интенсивностью  $\lambda_{i,i+1} = 1/\bar{t}_i$ , где  $\bar{t}_i$  – среднее время пребывания системы в состоянии  $S_i$ . Подставляя в формулы (2.16), (2.17) соответствующие значения интенсивностей, получим:

$$P_i = \frac{\bar{t}_i}{\sum_{i=1}^n \bar{t}_i}. \quad (2.18)$$

*Пример 2.14.* Устройство может находиться в одном из следующих состояний:

- $S_1$  – исправно, работает;
- $S_2$  – неисправно, ведется диагностика;

$S_3$  – неисправность обнаружена и устраняется;

$S_4$  – ремонт закончен, ведется подготовка к пуску устройства.

Все потоки событий – простейшие. Среднее время безотказной работы равно 0,5 суток, среднее время поиска неисправности – 0,5 часа, ремонт занимает в среднем 6 часов, а подготовка к пуску – 1 час.

Очевидно, что в системе протекает циклический процесс. Выражая в сутках все значения среднего времени пребывания устройства в различных состояниях и используя формулу (2.18), найдем:

$$P_1 = \frac{1/2}{1/2 + 1/48 + 1/4 + 1/24} = 0,615; \quad P_2 = 0,026; \quad P_3 = 0,308; \quad P_4 = 0,051.$$

## 2.7. Теория массового обслуживания

Теория массового обслуживания занимается изучением систем массового обслуживания (СМО). Примеры СМО: автоматизированная телефонная станция, автозаправочная станция, библиотека.

СМО состоит из некоторого числа обслуживающих устройств, называемых каналами. СМО предназначена для обслуживания потока заявок, поступающих в систему в случайные моменты времени. Обслуживание каждой заявки занимает случайное время, после чего канал освобождается.

Предмет теории массового обслуживания – это установление зависимостей между характером потока заявок, числом каналов, их производительностью, правилами работы СМО и эффективностью обслуживания.

В качестве показателя эффективности, в зависимости от условий, задачи и целей исследования, могут использоваться:

- среднее число заявок, обслуживаемых в единицу времени;
- средний процент заявок, получивших отказ и покинувших систему необслуженными;
- среднее время ожидания в очереди;
- средний доход, приносимый системой в единицу времени и т.д.

К СМО можно отнести системы, не занимающиеся обслуживанием в бытовом смысле, такие как система противовоздушной обороны (ПВО), сеть ЭВМ, космическая система мониторинга земной поверхности.

Исследование СМО значительно облегчается, если протекающий в ней процесс является марковским [12, 16]. В противном случае характеристики СМО можно оценить только приблизительно. При этом, чем больше в СМО каналов, тем точнее оценки, получаемые приближенно с использованием теории марковских процессов.

### 2.7.1. Классификация систем массового обслуживания

Основные критерии классификации СМО:

- система с отказами (заявка, заставшая все каналы занятыми, уходит не обслуженной) и система с очередью (заявка, застав все каналы занятыми, встает в очередь и ждет, пока освободится один из каналов);
- система с неограниченным временем ожидания в очереди и система с ограниченным временем ожидания;
- разомкнутые и замкнутые СМО;
- СМО без взаимопомощи и с различными типами взаимопомощи между каналами и др.

Если не оговорено иначе, будем считать, что все потоки событий, переводящие систему из состояния в состояние, являются пуассоновскими. Если систему из данного состояния выводит сразу несколько потоков событий, то время пребывания системы в этом состоянии распределено по закону  $f(t) = \lambda e^{-\lambda t}$ , где  $\lambda$  – суммарная интенсивность всех исходящих потоков.

### 2.8.2. Одноканальная система массового обслуживания с отказами

Система может находиться в одном из двух состояний:  $S_0$  – канал свободен,  $S_1$  – канал занят (рис. 2.21).

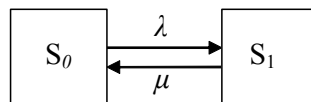


Рис. 2.21. Состояния одноканальной СМО с отказами



Интенсивность потока заявок  $\lambda = \frac{1}{\bar{t}_{cp}}$ , где  $\bar{t}_{cp}$  - средний интервал времени между моментами поступления заявок.

Интенсивность обслуживания  $\mu = \frac{1}{\bar{t}_{об}}$ , где  $\bar{t}_{об}$  - среднее время обслуживания одной заявки.

Поведение системы можно описать уравнениями Колмогорова:

$$\begin{cases} \frac{dP_0}{dt} = -\lambda p_0 + \mu p_1; \\ \frac{dP_1}{dt} = -\mu p_1 + \lambda p_0. \end{cases}$$

Учитывая условие  $p_0 + p_1 = 1$ , первое уравнение можно переписать в виде

$$\frac{dP_0}{dt} = -\lambda p_0 + \mu(1 - p_0), \text{ или } \frac{dP_0}{dt} = -(\lambda + \mu)p_0 + \mu.$$

Разделяя переменные и интегрируя, получим:

$$\frac{dp_0}{-(\lambda + \mu)p_0 + \mu} = dt \Rightarrow \frac{d[(\lambda + \mu)p_0 - \mu]}{(\lambda + \mu)p_0 - \mu} \cdot \frac{1}{(\lambda + \mu)} = -dt \Rightarrow$$

$$\ln[(\lambda + \mu)p_0 - \mu] \Big|_0^{p_0(t)} = -(\lambda + \mu)t \Big|_0^t.$$

С учетом начального условия  $p_0(0) = 1$ ,

$$\ln \frac{(\lambda + \mu)p_0(t) - \mu}{\lambda} = -(\lambda + \mu)t,$$

$$(\lambda + \mu)p_0(t) - \mu = \lambda e^{-(\lambda + \mu)t} \Rightarrow P_0(t) = \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} + \frac{\mu}{\lambda + \mu}.$$

Для вероятности состояния  $S_1$  получим:

$$P_1(t) = 1 - P_0(t) = 1 - \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} - \frac{\mu}{\lambda + \mu} \Rightarrow$$

$$P_1(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t}.$$

Примерный вид полученных зависимостей показан на рис. 2.22.

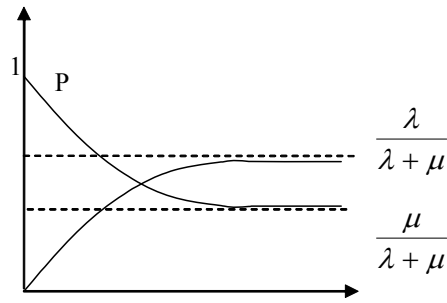


Рис. 2.22. Изменения значений вероятностей состояний

Представляют интерес предельные вероятности состояний:

$$\lim_{t \rightarrow \infty} P_0(t) = \frac{\mu}{\lambda + \mu}; \quad \lim_{t \rightarrow \infty} P_1(t) = \frac{\lambda}{\lambda + \mu}.$$

К основным характеристикам одноканальной СМО с отказами относятся следующие:

1. Относительная пропускная способность  $q$ , равная отношению среднего числа заявок, обслуживаемых системой, к среднему числу заявок, поступивших в систему в единицу времени.

Очевидно,  $q$  равно вероятности того, что канал будет свободен, т.е.  $q = P_0$ . Вероятность отказа  $P_{отк} = P_1 = 1 - q$ .

2. Абсолютная пропускная способность  $A$  – это среднее число заявок, обслуживаемых системой в единицу времени. Поскольку среднее число заявок, поступающих в систему в единицу времени, равно  $\lambda$ , то  $A = q\lambda$ .

*Пример 2.15.* Рассмотрим одну телефонную линию с  $\lambda = 0,8$  мин<sup>-1</sup> и  $\bar{t}_{об} = 15$  мин,  $\mu = \frac{1}{\bar{t}_{об}} = 0,667$  мин<sup>-1</sup>;  $q = p_0 = \frac{\mu}{\lambda + \mu} = 0,455$ .

$A = \lambda q = 0,8 \cdot 0,455 = 0,364$  1/мин. Все потоки простейшие. Вероятность отказа  $P_{отк} = 1 - 0,455 = 0,545$ .

Отметим, что номинальная абсолютная пропускная способность  $A_{ном} = \frac{1}{\bar{t}_{об}} = 0,667$ , т.е. фактическая пропускная способность линии примерно вдвое ниже ее номинальной пропускной способности - это вызвано тем, что заявки поступают в систему не регулярно, а случайным образом.

### 2.7.3. Многоканальная система массового обслуживания с отказами

Пусть в системе имеется  $n$  обслуживающих каналов. Граф состояний системы изображен на рис. 2.23.

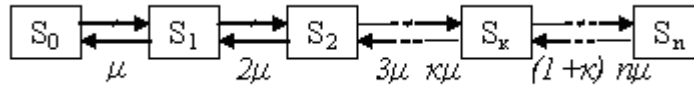


Рис. 2.23. Состояния многоканальной СМО с отказами

Поток обслуживания – пуассоновский, с интенсивностью  $\mu$ , поток заявок – также пуассоновский, с интенсивностью  $\lambda$ .

Состояния системы:  $S_0$  – все каналы свободны,  $S_1$  – занят один канал и т.д.,  $S_n$  – все каналы заняты. Заявка, заставшая все каналы занятыми, уходит не обслуженной. Поток, переводящий систему из  $S_l$  в  $S_0$ , имеет интенсивность  $\mu$ ; из  $S_2-S_1 - 2\mu$ , потому что освободиться может любой из двух занятых каналов, и т.д.

Запишем уравнения Колмогорова [12] и условие нормировки:

$$\left\{ \begin{array}{l} \frac{dP_0}{dt} = -\lambda P_0 + \mu P_1 \\ \frac{dP_1}{dt} = -(\lambda + \mu) P_1 + \lambda P_0 + 2\mu P_2 \\ \dots\dots\dots \\ \frac{dP_k}{dt} = -(\lambda + k\mu) P_k + \lambda P_{k-1} + (k+1)\mu P_{k+1} \\ \dots\dots\dots \\ \frac{dP_n}{dt} = -(\lambda + n\mu) P_n + \lambda P_{n-1} \\ P_1 + \dots + P_n = 1 \end{array} \right.$$

Начальные условия:  $P_0(0) = 1, P_1(0) = \dots = P_n(0) = 0$ .

Аналитическое решение системы сложно, поэтому ограничимся получением выражения для вероятностей предельных состояний. Для этого используем схему процесса гибели и размножения [12], записав в числителе выражения для соответствующей предельной вероятности состояния произведение всех интенсивностей, идущих слева направо до данного состояния, а в знаменателе - произведение всех интенсивностей, идущих от данного состояния налево, и умножим дробь на  $P_0$ :

$$P_k = \frac{\lambda \cdot \lambda \cdot \dots \cdot \lambda}{\mu \cdot 2\mu \cdot \dots \cdot k\mu} P_0 = \frac{\lambda^k}{\mu^k \cdot k!} P_0, \quad k = \overline{1, n}.$$

Учитывая условие нормировки, найдем:

$$P_0 = \left[ 1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \dots + \frac{\lambda^n}{n!\mu^n} \right]^{-1}.$$

Введем понятие приведенной интенсивности потока  $\rho = \lambda/\mu$ .

Очевидно,  $\rho$  представляет собой среднее число заявок, поступающих в систему за время обслуживания одной заявки.

$$\lambda = \frac{1}{\bar{t}_z}; \quad \mu = \frac{1}{\bar{t}_{об}} \Rightarrow \rho = \frac{\bar{t}_{об}}{\bar{t}_z}.$$

Теперь можно записать искомые вероятности в виде:

$$P_k = \frac{P_0 \rho^k}{k!}, \quad k = \overline{1, n}; \quad P_0 = \left[ 1 + \rho + \frac{\rho^2}{2} + \dots + \frac{\rho^n}{n!} \right]^{-1}.$$

Пропускные способности системы:  $q = 1 - P_{отк.} = 1 - P_n$ ;  $A = \lambda(1 - P_n)$ .

Среднее число занятых каналов  $\bar{K}$  найдем как математическое ожидание случайного числа занятых каналов  $K$ :

$$\bar{K} = 0 \cdot P_0 + 1 \cdot P_1 + 2P_2 + \dots + nP_n.$$

Эту величину можно выразить иначе, если учесть, что в единицу времени в среднем системой обслуживается  $A$  заявок, а одним каналом —  $\mu$  заявок, поэтому

$$\bar{K} = \frac{A}{\mu} = \frac{\lambda(1 - P_n)}{\mu} = \rho(1 - P_n).$$

*Пример 2.16.* Рассмотрим трехканальную СМО с интенсивностями

потоков  $\lambda=0,8 \text{ мин}^{-1}$  и  $\mu=0,667 \text{ мин}^{-1}$ . В этом случае  $\rho = \frac{0,8}{0,667} = 1,2$ ,

$$P_0 = \left( 1 + \rho + \frac{\rho^2}{2} + \frac{\rho^3}{6} \right)^{-1} = \left( 1 + 1,2 + \frac{1,44}{2} + 0,288 \right)^{-1} = 0,312.$$

$$P_1 = \rho P_0 = 0,374; P_2 = \frac{\rho^2}{2} P_0 = 0,224; P_3 = \frac{\rho^3}{6} P_0 = 0,090.$$

Получаем  $q = 1 - P_3 = 0,910$ ,  $A = \lambda q = 0,728$ ,  $\bar{k} = \rho A = 1,09$ .

Сравнивая с одноканальной системой, видим, что повышение относительной пропускной способности вдвое достигается ценой того, что из трех каналов в среднем два постоянно простаивают.

#### 2.7.4. Одноканальная система массового обслуживания с очередью

Пусть в очереди имеется  $m$  мест. Заявка, заставшая единственный канал занятым, встает в очередь. Заявка, пришедшая, когда все места и канал заняты, покидает систему не обслуженной. Граф состояний системы показан на рис. 2.24:

- $S_0$  – канал свободен, заявок в очереди нет;
- $S_1$  – канал занят, очереди нет;
- $S_2$  – канал занят, одна заявка в очереди;
- $S_{m+1}$  – канал занят,  $m$  заявок в очереди.

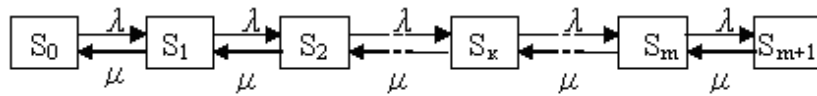


Рис. 2.24. Состояния одноканальной СМО с очередью

Запишем выражения для предельных вероятностей состояний:

$$P_k = \frac{\lambda^k}{\mu^k} P_0 = \rho^k P_0, \text{ где } P_0 = \left[ 1 + \rho + \rho^2 + \dots + \rho^{m+1} \right]^{-1}.$$

Использував формулу суммы геометрической прогрессии, стоящей в скобках, получим:

$$P_0 = \left[ \frac{1 - \rho^{m+2}}{1 - \rho} \right]^{-1} = \frac{1 - \rho}{1 - \rho^{m+2}}.$$

При  $\rho = 1$ ,  $P_0 = \frac{0}{0}$ . Раскроем неопределенность по правилу

Лопиталья:

$$P_0 = \lim_{\rho \rightarrow 1} \frac{(1 - \rho)'}{(1 - \rho^{m+2})'} = \frac{-1}{-m-2} = \frac{1}{m+2}.$$

Отказ заявка получает, когда заняты все  $m$  мест в очереди:

$$P_{отк} = P_{m+1}; \quad q = 1 - P_{m+1} = 1 - \frac{\rho^{m+1} \cdot (1 - \rho)}{(1 - \rho^{m+2})}.$$

Введем в рассмотрение случайное число заявок в очереди  $R$ . Среднее число заявок в очереди

$$\bar{r} = M(R) = 1 \cdot P_2 + 2 \cdot P_3 + \dots + m P_{m+1}.$$

$$\bar{r} = \rho^2 [1 \cdot P_0 + 2 \rho P_0 + \dots + m \rho^{m-1} P_0] = P_0 \rho^2 \sigma,$$

где  $\sigma = 1 + 2\rho + 3\rho^2 + \dots + m\rho^{m-1}$ . Очевидно, что  $\sigma = \frac{\partial \Sigma}{\partial \rho}$ , где

$\Sigma = \rho + \rho^2 + \rho^3 + \dots + \rho^m$  - сумма геометрической прогрессии со знаменателем  $\rho$ ;  $\Sigma = \frac{\rho - \rho^{m+1}}{1 - \rho}$ .

$$\begin{aligned} \sigma &= \frac{[1 - (m+1)\rho^m](1 - \rho) + \rho - \rho^{m+1}}{(1 - \rho)^2} = \frac{1 - (m+1)\rho^m + (m+1)\rho^{m+1} - \rho + \rho - \rho^{m+1}}{(1 - \rho)^2} = \\ &= \frac{1 - (m+1)\rho^m + m\rho^{m+1}}{(1 - \rho)^2} = \frac{1 - \rho^m(m+1 - m\rho)}{(1 - \rho)^2}. \end{aligned}$$

Окончательно,

$$\bar{r} = \rho^2 P_0 \frac{1 - \rho^m(m+1 - m\rho)}{(1 - \rho)^2} = \frac{\rho^2(1 - \rho)}{1 - \rho^{m+2}} \cdot \frac{1 - \rho^m(m+1 - m\rho)}{(1 - \rho)^2}.$$

Введем величину  $\bar{z}$  - среднее число заявок, связанных с системой.

$\bar{z} = \bar{r} + \bar{w}$ , где  $\bar{w}$  - среднее число обслуживаемых заявок. Очевидно  $\bar{w} = 0 \cdot P_0 + 1 \cdot P_{об}$ , где  $P_{об}$  - вероятность обслуживания.

$$P_{об} = 1 - P_0 = P_1 + P_2 + \dots + P_{m+1}.$$

$$\bar{w} = 1 - P_0 = 1 - \frac{1 - \rho}{1 - \rho^{m+2}} = \frac{\rho - \rho^{m+2}}{1 - \rho^{m+2}}.$$

Получим выражение еще одной характеристики - среднего времени ожидания в очереди. С вероятностью  $P_0$  время ожидания равно нулю, с вероятностью  $P_1$  пришедшая заявка застает канал занятым и ждет среднее время  $\bar{t}_{об} = 1/\mu$ . С вероятностью  $P_1$  заявка ждет, пока освободится канал и пока будет обслужена заявка, стоявшая перед ней, то есть время  $\bar{t}_{об} = 2/\mu$ , и т.д. С вероятностью  $P_m$   $\bar{t}_{об} = m/\mu$ . С вероятностью  $P_{m+1}$ , заявка застает все места в очереди занятыми и уходит не обслуженной  $\bar{t}_{об} = 0$ .

Таким образом,

$$\begin{aligned} t_{ож} &= P_1 \cdot \frac{1}{\mu} + P_2 \cdot \frac{2}{\mu} + \dots + P_m \cdot \frac{m}{\mu} = \\ &= \frac{\rho P_0}{\mu} (1 + 2\rho + 3\rho^2 + \dots + m\rho^{m-1}) = \frac{\rho P_0}{\mu} \sigma. \end{aligned}$$

Сравнивая с выражением для  $\bar{r}$ , видим, что  $\bar{t}_{ож} = \frac{\bar{r}}{\rho\mu} = \frac{\bar{r}}{\lambda}$ .

Среднее время пребывания заявки в системе  $\bar{t}_c = \bar{t}_{ож} + \bar{t}_{об}$ . Очевидно, что среднее время обслуживания заявки  $\bar{t}_{об} = q/\mu$ . Поэтому

$$\bar{t}_c = \frac{\bar{r}}{\lambda} + \frac{q}{\mu}.$$

*Пример 2.18.* АЗС имеет одну колонку; площадка при заправке допускает очередь из трех машин. Машины прибывают в среднем один раз в минуту:  $\bar{t}_3 = 1$  мин. Среднее время обслуживания  $\bar{t}_{об} = 1,25$  мин. Определить характеристики системы обслуживания.

Находим:  $\mu = 1/\bar{t}_{об} = 1/1,25 = 0,8$  мин<sup>-1</sup>,  $\lambda = 1$  мин<sup>-1</sup>,  $\rho = \lambda/\mu = 1,25$ .

$$P_0 = \frac{1-\rho}{1-\rho^5} = 0,122; P_1 = \rho P_0 = 0,152; P_2 = \rho^2 P_0 = 0,191;$$

$$P_3 = \rho^3 P_0 = 0,238; P_4 = \rho^4 P_0 = 0,703; A = q\lambda = 0,703 \text{ мин}^{-1}.$$

$$\bar{r} = \frac{(1,25)^2 [1 - 1,25^3 \cdot (3 + 1 - 3 \cdot 1,25)]}{(1 - 1,25^5)(1 - 1,25)} = 1,56;$$

$$\bar{w} = \frac{1,25 - 1,25^5}{1 - 1,25^5} = 0,88; \bar{z} = \bar{w} + \bar{r} = 2,44;$$

$$\bar{t}_{ож} = \frac{\bar{r}}{\lambda} = \bar{r} = 1,56 \text{ мин}; \bar{t}_c = 1,25 + \frac{0,703}{0,8} = 2,44 \text{ мин}.$$

Пусть длина очереди неограниченно возрастает. Если  $\rho \geq 1$ , то предельного установившегося режима не существует; в случае же  $\rho < 1$  установившийся режим будет иметь место. Предельным переходом при  $m \rightarrow \infty$  получаем:  $P_0 = 1 - \rho$ ,  $P_k = \rho^k (1 - \rho)$ . При этом каждая заявка рано или поздно будет обслужена ( $q=1$  и  $A=\lambda$ ).

Формулы для расчета  $\bar{r}$  и  $\bar{t}_{ож}$  приобретают вид:

$$\bar{r} = \lim_{m \rightarrow \infty} \rho P_0 \cdot \frac{(1 - \rho^m)(m + 1 - m\rho)}{(1 - \rho)^2} = \frac{\rho^2}{1 - \rho};$$

$$\bar{t}_{ож} = \frac{\bar{r}}{\lambda} = \frac{\rho \cdot \frac{\rho}{\lambda}}{1 - \rho} = \frac{\rho}{\mu(1 - \rho)}.$$

### 2.7.5. Многоканальная система массового обслуживания с очередью

Пусть число каналов СМО равно  $n$ , число мест в очереди равно  $m$ :

$S_0$  – все каналы свободны;

$S_1$  – занят один канал, очереди нет;

.....

$S_n$  – заняты все каналы, очереди нет;

$S_{n+m}$  – заняты все каналы и все места в очереди.

Тогда граф состояний системы имеет следующий вид (рис. 2.25):



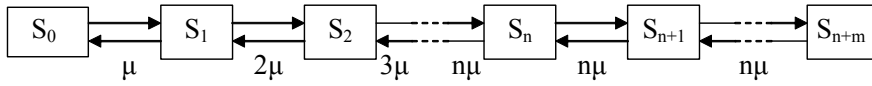


Рис. 2.25. Состояния многоканальной СМО с очередью

Используя схему процесса гибели и размножения, запишем:

$$\begin{aligned}
 P_0 &= \left\{ 1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \dots + \frac{\lambda^n}{n!\mu^n} + \dots + \frac{\lambda^n}{n!\mu^n} \left[ \frac{\lambda}{n\mu} + \frac{\lambda^2}{(n\mu)^2} + \dots + \frac{\lambda^m}{(n\mu)^m} \right] \right\}^{-1} = \\
 &= \left\{ 1 + \rho + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!} \left[ \frac{\rho}{n} + \frac{\rho^2}{n^2} + \dots + \frac{\rho^m}{n^m} \right] \right\}^{-1} = \\
 &= \left\{ 1 + \rho + \dots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!} \cdot \frac{\rho - \left(\frac{\rho}{n}\right)^{m+1}}{1 - \frac{\rho}{n}} \right\}^{-1} \\
 P_\kappa &= \begin{cases} P_0 \frac{\rho^\kappa}{\kappa!}, & \text{если } \kappa \leq n, \\ P_0 \frac{\rho^n}{n!} \left(\frac{\rho}{n}\right)^{\kappa-n}, & \text{если } \kappa > n. \end{cases}
 \end{aligned}$$

$P_{отк} = P_{n+m}$ ,  $q = 1 - P_{n+m}$ ,  $A = \lambda(1 - P_{n+m})$ . Среднее число занятых каналов

$$\kappa = \frac{A}{\mu} = \rho(1 - P_{n+m}).$$

Найдем среднее число заявок, стоящих в очереди:

$$\begin{aligned}
 r &= 1 \cdot P_{n+1} + 2 \cdot P_{n+2} + \dots + m \cdot P_{n+m} = P_0 \frac{\rho^n}{n!} \left[ \frac{\rho}{n} + 2 \left(\frac{\rho}{n}\right)^2 + \dots + m \left(\frac{\rho}{n}\right)^m \right] = \\
 &= \frac{P_0 \rho^{n+1}}{n!n} [1 + 2\chi + 3\chi^2 + \dots + m\chi^{m-1}],
 \end{aligned}$$

где  $\rho/n = \chi$ . Выражение в скобках не что иное, как вычисленная в предыдущем разделе сумма, где вместо  $\rho$  стоит величина  $\chi$ . Окончательно:

$$r = \frac{\rho^{n+1} \cdot P_0}{n!n} \cdot \frac{1 - (m+1)\chi^m + m\chi^{m+1}}{(1-\chi)^2}.$$

Получим формулу для среднего времени ожидания в очереди. Если заявка пришла в систему, находящуюся в состоянии  $S_n$ , она будет ждать, пока освободится один из каналов, в течение времени  $1/n\mu$ , в состоянии  $S_{n+1}$  - ждать в течение времени  $2/n\mu$ , в состоянии  $S_{n+m-1}$  - в течение  $m/n\mu$ , в состоянии  $S_{m+n}$  (заняты все каналы обслуживания и все места в очереди) заявка получает отказ и уходит. Поэтому

$$\bar{t}_{ож} = \frac{1}{n\mu} P_n + \frac{2}{n\mu} P_{n+1} + \dots + \frac{m}{n\mu} P_{n+m-1} = \frac{\rho^n P_0}{n \cdot n! \mu} (1 + 2\chi + \dots + m\chi^{m-1}).$$

Учитывая, что  $\rho\mu = \lambda$  и сравнив с формулой для числа заявок в очереди, получаем:  $\bar{t}_{ож} = \bar{r}/\lambda$ . Время пребывания в системе  $\bar{t}_c = \bar{t}_{ож} + \bar{t}_{об} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu}$ .

При неограниченной очереди ( $m \rightarrow \infty$ ) для  $\chi < 1$  (иначе ряд расходится),

получаются следующие выражения:  $q=1$ ;  $A=\lambda$ ;  $\bar{r} = \frac{\rho^{n+1} \cdot P_0}{n \cdot n! (1-\chi)^2}$ ;  $\bar{k} = A/\mu$ .

*Пример 2.18.* На АЗС имеется две колонки ( $n=2$ ), интенсивность прибытия машин  $\lambda=0,8$  1/мин, среднее время обслуживания  $\bar{t}_{об} = 2$  мин, очередь неограничена ( $m \rightarrow \infty$ ).

Тогда  $\mu=0,5$  1/мин,  $\rho=1,6$ ,  $\chi = \rho/n = 0,8 < 1$ .

Устанавливается предельный стационарный режим с вероятностями

состояний:  $P_0 = \left[ 1 + 1,6 + 1,28 + \frac{1,6^3}{2(2-1,6)} \right]^{-1} = 0,111$ ;  $P_1 = 1,6P_0 = 0,178$ ;

$P_2 = 1,28P_0 = 0,142$ ;  $P_3 = \frac{1,6^3}{2 \cdot 2} P_0 = 0,114$ ;  $\bar{k} = A/\mu = \lambda/\mu = 1,6$ .

Вероятность отсутствия очереди:  $P_{от.оч.} = P_0 + P_1 + P_2 = 0,431$ ,

$\bar{r} = \frac{1,6^3 \cdot 0,111}{2 \cdot 2 \cdot (1-0,8)^2} = 2,84$ ;  $\bar{z} = \bar{r} + \bar{k} = 2,84 + 1,6 = 4,44$  - общее число заявок

в системе,  $\bar{t}_{ож} = \frac{\bar{r}}{\lambda} = 3,55$  мин,  $\bar{t}_{сист} = \bar{t}_{об} + \bar{t}_{ож} = 5,55$  мин.

## 2.8. Метод динамики средних

### 2.8.1. Идея метода. Область применения

Аппарат непрерывных марковских цепей удобно использовать, когда число возможных состояний системы невелико; иначе решение не только дифференциальных уравнений Колмогорова, но даже алгебраических уравнений для предельных вероятностей состояний затруднено, а полученные результаты труднообозримы, для их осмысления приходится пользоваться некими осредненными характеристиками. Поэтому резонно попытаться составить и решить уравнения непосредственно для этих средних характеристик, вместо того, чтобы вычислять их через вероятности состояний. Такими задачами и занимается метод динамики средних (МДС).

Основной применимости МДС является именно то, что препятствует изучению явления более подробными методами: большое число элементов, участвующих в изучаемом процессе.

Как и во всех областях применения методов теории вероятностей, массовость изучаемых явлений позволяет установить в них сравнительно простые закономерности.

Пусть система  $S$  состоит из  $N$  однородных элементов, которые могут случайным образом переходить из состояния в состояние под действием пуассоновских потоков, имеющих в общем случае интенсивности, зависящие от времени. Тогда в системе протекает марковский процесс с  $n$  возможными состояниями. Так как в силу большого значения  $N$  даже перечислить все состояния системы затруднительно, сосредоточим внимание на поведении одного элемента системы.

Пусть  $X_k(t)$  – число элементов системы, пребывающих в состоянии  $E_k$ . Справедливо следующее условие нормировки:

$$X_1(t) + X_2(t) + \dots + X_n(t) = N.$$

Граф переходов системы приведен на рис. 2.26.

Введем случайную величину

$$X_k^{(i)} = \begin{cases} 1, & \text{если } i\text{-ый элемент находится в состоянии } E_k; \\ 0 & \text{- в противном случае.} \end{cases}$$

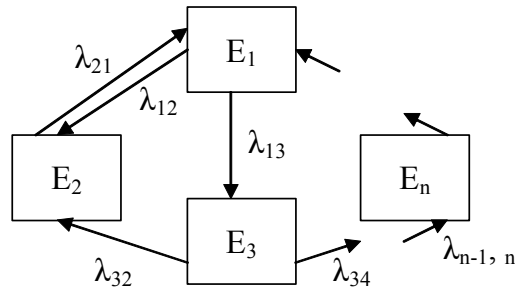


Рис. 2. 26. Граф переходов системы

Тогда, очевидно,  $X_k = \sum_{i=1}^N X_k^{(i)}$ . Если интенсивности переходов известны и неслучайны, то отдельные состояния независимы между собой. Соответственно независимы величины  $X_k^{(1)}, X_k^{(2)}, \dots, X_k^{(n)}$ ,  $k = 1, \dots, n$ . Тогда математическое ожидание числа элементов, находящихся в состоянии  $E_k$ , равно  $m_k(t) = \sum_{i=1}^N M[X_k^{(i)}]$ , а дисперсия  $D_k(t) = \sum_{i=1}^N D[X_k^{(i)}]$ . Математическое ожидание  $M[X_k^{(i)}] = 1 \cdot P_k + 0 \cdot (1 - P_k) = m_k^{(i)} = P_k$ , где  $P_k$  – вероятность пребывания элемента в  $k$ -ом состоянии. Дисперсия  $D[X_k^{(i)}] = P_k(1 - m_k)^2 + (1 - P_k)(0 - m_k)^2 = P_k(1 - P_k)^2 + (1 - P_k)(-P_k)^2 = P_k(1 - P_k)$ .

Так как элементы однородны, то значения  $P_k$  одинаковы для всех элементов. Подставив полученные значения в выражения для  $m_k(t)$  и  $D_k(t)$ , получим:

$$M[X_k(t)] = m_k(t) = N \cdot P_k(t);$$

$$D[X_k(t)] = D_k(t) = m_k(t) \left( 1 - \frac{m_k(t)}{N} \right).$$

Таким образом, численности состояний пропорциональны их вероятностям. Поэтому уравнения Колмогорова можно записывать

непосредственно для величин  $m_1(t), m_2(t), \dots, m_n(t)$ . Если число элементов  $N$  велико, можно считать численности состояний распределенными по нормальному закону, что позволяет оценить диапазон возможных значений этих численностей. Вероятность попадания величины  $m_k$  в интервал  $[\alpha; \beta]$  составляет

$$P(X_k \in [\alpha, \beta]) = \Theta\left(\frac{\beta - m_k}{\sigma_k}\right) - \Theta\left(\frac{m_k - \alpha}{\sigma_k}\right),$$

где  $\sigma_k = \sqrt{D_k}$ ,  $\Theta(\cdot)$  - функция Лапласа.

*Пример 2.19.* Система состоит из 200 однородных элементов, каждый из которых может находиться в двух состояниях:  $E_1$  – исправен,  $E_2$  – неисправен. Граф состояний системы имеет вид (рис. 2.27).

Запишем уравнения Колмогорова:

$$\begin{cases} \frac{dm_1}{dt} = -2m_1 + 3m_2; \\ \frac{dm_2}{dt} = 2m_1 - 3m_2. \end{cases}$$

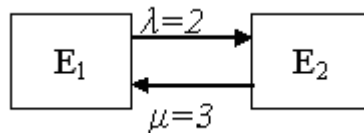


Рис. 2.27. Состояния системы

Выражая из условия  $m_1 + m_2 = 200$  значение  $m_1$  через  $m_2$ , получаем уравнение с разделяющимися переменными, решение которого дает следующий результат:

$$m_2(t) = \left( \frac{\lambda}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} \cdot e^{-(\lambda + \mu)t} \right) N = \left( \frac{3}{5} - \frac{2}{5} e^{-5t} \right) \cdot 200,$$

$$m_1(t) = N - m_2(t).$$

При  $t \rightarrow \infty$ ;  $m_1=120$ ,  $m_2=80$ ,  $D_1=D_2=48$ .

### **2.8.2. Учет зависимости интенсивностей потоков событий от численности состояний. Принцип квазирегулярности**

Выше предполагалось, что интенсивности потоков событий заданы и неслучайны. Но это не всегда так. Например, при ограниченной пропускной способности ремонтной службы интенсивность потока восстановлений (ремонтов) зависит от численности ремонтируемых элементов, которая случайна.

С другой стороны, при повышении числа отказавших элементов нагрузка на работоспособные элементы возрастает, что приводит к повышению интенсивности их отказа. Если интенсивности потоков событий зависят от численностей состояний, то непосредственное применение МДС становится невозможным, так как сами численности событий неизвестны. Эту трудность можно преодолеть, если предположить, что интенсивности зависят не от численностей состояний, а от их математических ожиданий. Это допущение, называемое принципом квазирегулярности, позволяет получить приближенное решение задачи. Ошибки могут стать существенными в том случае, когда  $N$  - число элементов - мало, так как в этом случае действительные численности состояний могут существенно отличаться от своих математических ожиданий. Существенным образом влияет на ошибку вид зависимостей интенсивностей от численностей (чем зависимости ближе к линейным, тем меньше ошибка).

*Пример 2.20.* Система состоит из  $N$  одинаковых элементов, каждый из которых, в свою очередь, состоит из двух узлов (первый – основной, второй – резервный, не работающий в случае, когда основной узел исправен).

На работающий узел действует поток отказов с интенсивностью  $\lambda_1$ , на неработающий –  $\lambda_2$ . Пусть количество узлов, находящихся в ремонте, –  $Y$ , интенсивность потока восстановлений –  $\mu(Y)$ . Функция  $\mu(Y)$  определена только в целочисленных точках (рис. 2.28).

Однако, при использовании принципа квазирегулярности, численности состояний заменяются их математическими ожиданиями, которые могут быть и нецелыми.

Поэтому в данной задаче  $\mu(Y)$  можно аппроксимировать

непрерывной зависимостью.

Граф состояния элемента системы показан на рис. 2.29:

$E_1$  – оба узла исправны;

$E_2$  – основной узел отказал и ремонтируется, резервный узел работает;

$E_3$  – оба узла ремонтируются.

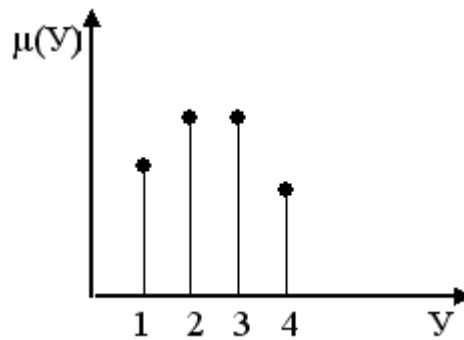


Рис. 2.28. Возможный вид функции  $\mu(Y)$

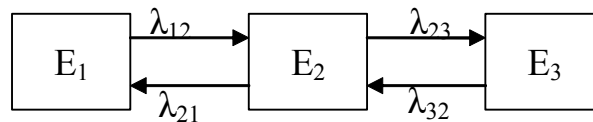


Рис. 2.29. Состояния элемента системы

Определим интенсивности переходов  $\lambda_{12} = \lambda_1 + \lambda_2$ ;  $\lambda_{23} = \lambda_1$ .

Интенсивность потока восстановления равномерно распределяется между всеми ремонтируемыми узлами, так что на один узел приходится интенсивность, равная  $\frac{\mu(Y)}{Y}$ . Поэтому  $\lambda_{21} = \frac{\mu(x_2 + 2x_3)}{(x_2 + 2x_3)}$  (в состоянии  $E_2$

число элементов равно  $x_2$ , а в состоянии  $E_3$  –  $2x_3$ ), а  $\lambda_{32} = \frac{2\mu(x_2 + 2x_3)}{(x_2 + 2x_3)}$ ,

т.к. в состоянии  $E_3$  поток восстановлений действует на число узлов, вдвое большее, чем в состоянии  $E_2$ .

Применяя принцип квазирегулярности, запишем:

$$\lambda_{21} \cong \frac{\mu(m_2 + 2m_3)}{m_2 + 2m_3}; \quad \lambda_{32} \cong \frac{2\mu(m_2 + 2m_3)}{m_2 + 2m_3}.$$

Это позволяет записать уравнения для средних численностей состояний; учитывая условие нормировки ( $m_1 + m_2 + m_3 = N$ ), уравнение для состояния  $E_2$  опустим, как наиболее сложное. Получим замкнутую систему:

$$\begin{cases} \frac{dm_1}{dt} = (\lambda_1 + \lambda_2)m_1 + \frac{\mu(m_2 + 2m_3)}{m_2 + 2m_3}m_2; \\ \frac{dm_3}{dt} = \frac{2\mu(m_2 + 2m_3)}{(m_2 + 2m_3)}m_3 + \lambda_1m_2; \\ m_1 + m_2 + m_3 = N. \end{cases}$$

Эту систему уравнений можно решить численно. Формула для дисперсии, полученная в предыдущем разделе, уже неприменима, для вогнутых функций, подобных  $\mu(V)$ , она дает заниженные значения. Уравнения для дисперсии также можно составить, используя принцип квазирегулярности, однако они громоздки, и здесь не рассматриваются.

## 2.9. Задачи принятия решений в условиях риска

### 2.9.1. Одноэтапные процедуры

В задачах принятия решений в условиях риска наиболее распространенным является критерий *ожидаемого значения*. Пусть доход от операции зависит от случайных факторов и при различных реализациях операции принимает случайные значения  $W$ . Тогда в качестве ожидаемого значения критерия естественно использовать его математическое ожидание  $m(W)$ . Если исследователь располагает выборкой значений  $W_1, W_2, \dots, W_n$ , неизвестное математическое ожидание заменяется его оценкой:

$$m(W) \cong \hat{m}(W) = \frac{1}{n} \sum_{i=1}^n W_i.$$



Когда число реализаций операции невелико, замена математического ожидания ее оценкой может привести к большим ошибкам, поэтому в качестве «гарантированного» значения критерия используется величина

$$\hat{m}(W) - K \cdot \hat{\sigma}(W), \quad (2.19)$$

где  $\hat{\sigma}(W) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n [W_i - \hat{m}(W)]^2}$  - оценка среднеквадратического

отклонения дохода,  $K$  - *уровень несклонности к риску*, назначаемый лицом, принимающим решения.

Если в качестве критерия оптимальности операции фигурирует не доход, а ущерб, то в формуле (2.19) знак «минус» заменяется на «плюс».

*Пример 2.21.* Производитель выпускает продукцию партиями. Партия может содержать 8, 10, 12 или 14% брака с вероятностями 0,4; 0,3; 0,25; 0,05 соответственно. Продукция может поставляться заказчикам А, В или С, в контрактах с которыми оговорены следующие условия:

- процент брака для потребителей А, В, С не должен превышать 8, 12 и 14% соответственно;

- за 1% брака, превышающий допустимое значение, производитель платит штраф 100 у.е.

До отправки потребителю партия изделий не проверяется. При производстве партии более высокого качества, чем требуется заказчиком, производитель несет дополнительные затраты 50 у.е. за 1% снижения брака.

Требуется решить, кому из заказчиков выгоднее всего поставлять продукцию?

В данной задаче «система» - партия изделий - может находиться в одном из четырех состояний: содержать 8, 10, 12 или 14% брака соответственно. Доступные стратегии:  $x_1$  - поставлять продукцию потребителю А,  $x_2$  - поставлять продукцию потребителю В,  $x_3$  - поставлять продукцию потребителю С. Критерием оптимальности операции является математическое ожидание расходов, которое необходимо минимизировать.

Затраты представим в виде матрицы  $\|v_{ij}\|$  размером  $3 \times 4$ , элемент которой  $v_{ij}$  есть затраты, которые несет производитель, применяя свою  $i$ -ую стратегию, когда система находится в  $j$ -ом состоянии.

Если система находится в первом состоянии (8% брака), то применение первой стратегии (поставка потребителю А, которого данное качество продукции удовлетворяет) не приведет к убыткам, поэтому  $v_{11} = 0$ . Поставка этому же заказчику партии, содержащей 10% брака, приведет к штрафу в размере  $(10-8) \cdot 100 = 200$  (у.е.), так что  $v_{12} = 200$ . Напротив, поставка партии с 8% брака заказчику В (стратегия  $x_2$ ) приведет к затратам, связанным с обеспечением «избыточного» качества в размере  $(12-8) \cdot 50 = 200$  (у.е.), и  $v_{21} = 200$ . Аналогично вычисляются остальные элементы матрицы затрат, которая будет иметь следующий вид:

$$\|v_{ij}\| = \begin{pmatrix} 0 & 200 & 400 & 600 \\ 200 & 100 & 0 & 200 \\ 300 & 200 & 100 & 0 \end{pmatrix}.$$

Ожидаемый ущерб от применяемых стратегий составит:

$$m(W(x_1)) = 0 \cdot 0,4 + 200 \cdot 0,3 + 400 \cdot 0,25 + 600 \cdot 0,05 = 190;$$

$$m(W(x_2)) = 200 \cdot 0,4 + 100 \cdot 0,3 + 0 \cdot 0,25 + 200 \cdot 0,05 = 120;$$

$$m(W(x_3)) = 300 \cdot 0,4 + 200 \cdot 0,3 + 100 \cdot 0,25 + 0 \cdot 0,05 = 190.$$

Таким образом, оптимальной является стратегия  $x_2$  - поставлять продукцию заказчику В.

*Пример 2.22.* Фирма может выпускать три вида продукции. Доходы от продаж продукции в 10 различных городах за период, предшествующий плановому, приведены в табл. 2.9. Спрашивается, какую продукцию выгоднее всего производить фирме? Учитывая малый объем выборки, по которой можно оценить средний доход от продажи продукции, следует использовать критерий «ожидаемое значение-дисперсия». Решение задачи при уровнях несклонности к риску  $K=1, 3, 5$  приведено в табл. 2.10.

Таблица 2.9. Доход от продаж

Продукция	Доход, тыс. у.е.									
	А	101	98	103	89	95	97	86	109	114
В	120	117	129	135	125	146	179	116	118	121
С	78	86	73	89	91	85	75	93	84	86

Таблица 2.10. Значения критерия оптимальности

Продукция	$\hat{m}(W)$ , у.е.	$\hat{\sigma}(W)$ , у.е.	$\hat{m}(W) - K \cdot \hat{\sigma}(W)$		
			K=1	K=3	K=5
А	97.10	10.54	86.56	65.49	44.42
В	130.60	19.42	111.18	72.34	33.50
С	84.00	6.68	77.32	63.95	50.58

Из полученных результатов заключаем, что при уровнях несклонности к риску  $K$ , равных 1 и 3, оптимальным является решение о выпуске продукции  $B$ , в то время как при  $K=5$  – решение о выпуске продукции  $C$ .

### 2.9.2. Многоэтапные процедуры

Многоэтапные процедуры принятия решений в условиях риска могут быть представлены графически в виде *дерева решений*. Вершины дерева подразделяются на *решающие*, изображаемые в виде квадрата, и *случайные*, изображаемые в виде круга. Ветви, исходящие из вершин, представляют собой различные варианты развития процесса. Для принятия решения необходима информация о вероятностях реализации вариантов, исходящих из случайных вершин, а также о доходах (расходах), получаемых в результате принятия каждого допустимого решения. В качестве критерия для выбора оптимального решения используется суммарный ожидаемый доход (в задачах минимизации - минимальный ожидаемый убыток). Вычисления выполняются последовательно, начиная с последнего этапа.

*Пример 2.23.* Требуется принять решение о строительстве крупного или малого предприятия. Малое предприятие может быть расширено по истечении двухлетнего периода. Спрос на выпускаемую продукцию с вероятностью 0,75 может быть высоким и с вероятностью 0,25 – низким. Информация о затратах  $C$  и годовых доходах  $D$  приведена на дереве принятия решений (рис.2.30). Необходимо принять решение, обеспечивающее получение максимального общего дохода за 10 лет.

Начнем вычисления со второго этапа принятия решений, которому

соответствует решающая вершина 4. В этой вершине имеется две стратегии:  $X_{41}$  - расширение предприятия - и  $X_{42}$  - отказ от расширения; им соответствуют ожидаемые доходы за 8-летний период, начинающийся после принятия решения

$$E_{41} = (900000 \cdot 0,75 + 200000 \cdot 0,25) \cdot 8 - 4200000 = 1600000;$$

$$E_{42} = (250000 \cdot 0,75 + 200000 \cdot 0,25) \cdot 8 = 1900000.$$

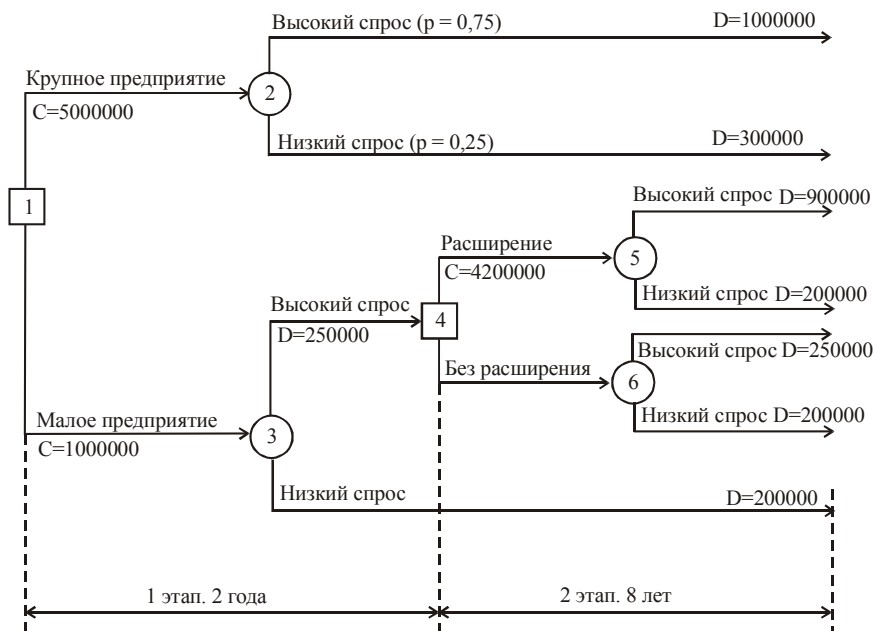


Рис. 2.30. Дерево принятия решений

В качестве оптимальной в вершине 4 выбираем стратегию  $x_{42}$  и для последующего анализа оставляем только ветвь, соответствующую данной стратегии. В вершине 1 имеются стратегии  $x_{11}$  - строительство крупного предприятия - и  $x_{12}$  - строительство малого предприятия.

Им соответствуют ожидаемые доходы за 10-летний период:

$$E_{11} = (1000000 \cdot 0,75 + 300000 \cdot 0,25) \cdot 10 - 5000000 = 3250000;$$

$$E_{12} = 250000 \cdot 0,75 \cdot 2 + E_{42} + 200000 \cdot 0,25 \cdot 10 - 1000000 = 1775000.$$

Таким образом, оптимальным является решение о строительстве крупного предприятия, при этом решающая вершина 4 из рассмотрения исключается.

## 2.10. Задачи принятия решений в условиях неопределенности

Основное различие рассмотренных здесь критериев определяется стратегией поведения лица, принимающего решения. Не существует общих правил для выбора тех или иных критериев принятия решения в условиях неопределенности. Специфическая особенность рассматриваемых в данном разделе задач заключается в отсутствии у лица, принимающего решения, разумного противника. Часто в роли «противника» выступает природа.

Необходимая для принятия решений информация обычно представляется в виде матрицы,  $i$ -ая строка которой соответствует  $x_i$  - стратегии из множества допустимых решений  $G$  ( $x_i \in G$ ), а  $j$ -ый столбец соответствует  $S_j$  - состоянию изучаемой системы. Элемент  $v_{ij}$ , стоящий на пересечении соответствующих строки и столбца, соответствует выигрышу или проигрышу лица, принимающего решения, в данной ситуации.

Рассмотрим наиболее часто применяемые критерии.

1. *Критерий Лапласа* основан на принципе недостаточного обоснования: поскольку вероятности состояний  $P(S_j)$  неизвестны (а иначе мы имели бы задачу принятия решений в условиях риска), то и нет оснований считать эти вероятности различными. Поэтому, когда  $v_{ij}$  - это

выигрыш, то критерием выбора стратегии является величина  $\max_{x_i \in G} \frac{1}{n} \sum_{j=1}^n v_{ij}$ ,

где  $n$  - число состояний системы. Когда  $v_{ij}$  - проигрыш, то критерием

является величина  $\min_{x_i \in G} \frac{1}{n} \sum_{j=1}^n v_{ij}$ .

Пример 2.24. Матрица  $\|v_{ij}\|$  имеет вид  $\|v_{ij}\| = \begin{pmatrix} 5 & 10 & 18 & 25 \\ 8 & 7 & 8 & 23 \\ 21 & 18 & 12 & 21 \\ 30 & 22 & 19 & 15 \end{pmatrix}$ .

Найти оптимальную стратегию для двух случаев: а)  $\|v_{ij}\|$  - матрица выигрышей; б)  $\|v_{ij}\|$  - матрица проигрышей.

Подсчитаем оценки

$$\frac{1}{4} \sum_{j=1}^4 v(x_1, S_j) = \frac{1}{4} (5+10+18+25) = 14,5;$$

$$\frac{1}{4} \sum_{j=1}^4 v(x_2, S_j) = \frac{1}{4} (8+7+8+23) = 11,5;$$

$$\frac{1}{4} \sum_{j=1}^4 v(x_3, S_j) = \frac{1}{4} (21+8+12+21) = 18;$$

$$\frac{1}{4} \sum_{j=1}^4 v(x_4, S_j) = \frac{1}{4} (30+22+19+15) = 21,5.$$

Следовательно, в случае а) оптимальной является стратегия  $x_4$ , а в случае б) - стратегия  $x_2$ .

2. Минимаксный (максиминный) критерий является наиболее «осторожным», или «пессимистичным». Если  $\|v_{ij}\|$  - это матрица потерь (проигрыша), то при выборе стратегии  $x_i$  наибольшие потери на всем множестве состояний составят  $\max_{S_j} v(x_i, S_j)$ . Чтобы минимизировать эту величину, нужно выбрать стратегию, обеспечивающую получение  $\min_{x_i} \max_{S_j} v(x_i, S_j)$ .

Если  $\|v_{ij}\|$  - матрица выигрышей (доходов), то при выборе  $x_i$  минимальный выигрыш на всем множестве состояний составит  $\min_{S_j} v(x_i, S_j)$ . Оптимальной будет стратегия, при которой выигрыш составит  $\max_{x_i} \min_{S_j} v(x_i, S_j)$ .

Для примера 2.24:  $\min_{x_i} \max_{S_j} v(x_i, S_j) = \min_{x_i} (25, 23, 21, 30) = 21$ , и оптимальной для случая минимизации ущерба является стратегия  $x_3$ .

Для случая максимизации дохода  $\max_{x_i} \min_{S_j} v(x_i, S_j) = \max_{x_i} (5, 7, 12, 15) = 15$  оптимальной является стратегия  $x_4$ .

3. Критерий Сэвиджа. Применение минимаксного (максиминного) критерия, в силу его чрезмерного «пессимизма», может порой привести к

нелогичным выводам. Это видно из следующего примера. Пусть решается задача минимизации ущерба, когда матрица потерь имеет вид

$$\|v_{ij}\| = \begin{pmatrix} 11000 & 90 \\ 10000 & 10000 \end{pmatrix}.$$

При использовании критерия 2 мы рискуем проиграть значительную сумму (10000). Чтобы устранить этот недостаток, вводится величина  $r(x_i, S_j)$  равная

$$v_{ij}(x_i, S_j) - \min_{x_i} v(x_i, S_j), \text{ если } \|v_{ij}\| - \text{матрица потерь,}$$

$$\max_{x_i} v(x_i, S_j) - v_{ij}(x_i, S_j), \text{ если } \|v_{ij}\| - \text{матрица дохода.}$$

Матрица  $\|r(x_i, S_j)\|$  называется матрицей сожалений, поскольку ее элементы характеризует сожаления лица, принимающего решения, по поводу того, оно не выбрало для случая, когда система оказалась в состоянии  $S_j$ , оптимальную стратегию.

Для данного примера  $\|r(x_i, S_j)\| = \begin{pmatrix} 1000 & 0 \\ 0 & 9910 \end{pmatrix}$ . Применив для матрицы сожалений минимаксный критерий, найдем, что оптимальной стратегией является  $x_1^*$ . Отметим, что по критерию Лапласа оптимальной также является стратегия  $x_1^*$ .

Построим матрицы сожалений для условий примера 2.24:

$$\text{а) } \|v_{ij}\| - \text{матрица выигрышей; тогда } \|r\| = \begin{pmatrix} 25 & 12 & 1 & 0 \\ 22 & 15 & 11 & 2 \\ 9 & 4 & 7 & 4 \\ 0 & 0 & 0 & 10 \end{pmatrix}.$$

Применяя к ней максиминный критерий, получим, что оптимальной стратегией является  $x_3$ ;

$$\text{б) } \|v_{ij}\| - \text{матрица проигрышей; матрица сожалений имеет вид}$$

$$\|r\| = \begin{vmatrix} 0 & 3 & 10 & 10 \\ 3 & 0 & 0 & 8 \\ 16 & 11 & 4 & 6 \\ 25 & 15 & 11 & 0 \end{vmatrix}.$$

Применяя минимаксный критерий, находим оптимальную стратегию -  $x_2$ .

4. Критерий Гурвица устанавливает баланс между пессимистическим и оптимистическим подходами путем их взвешивания:

а)  $\|v_{ij}\|$  - матрица доходов, тогда в качестве оптимальной выбирается стратегия, обеспечивающая получение дохода

$$\max_{x_i} \left( \alpha \max_{S_j} v_{ij} + (1-\alpha) \min_{S_j} v_{ij} \right),$$

где  $\alpha$  - весовой коэффициент, назначаемый лицом, принимающим решения,  $0 \leq \alpha \leq 1$ ;

б)  $\|v_{ij}\|$  - матрица проигрышей, тогда в качестве оптимальной берется стратегия, при которой проигрыш составляет

$$\min_{x_i} \left( \alpha \min_{S_j} v_{ij} + (1-\alpha) \max_{S_j} v_{ij} \right).$$

Для задачи минимизации в условиях примера 2.24 и при  $\alpha=0,5$  получим значения представленные в табл. 2.11.

Таблица 2.11. Значения выигрышей для примера 2.24

$x_i$	$\min_{S_j} v_{ij}$	$\max_{S_j} v_{ij}$	$\alpha \min_{S_j} v_{ij} + (1-\alpha) \max_{S_j} v_{ij}$
$x_1$	5	25	15
$x_2$	7	23	15
$x_3$	12	21	16,5
$x_4$	15	30	22,5

Оптимальной, очевидно, является стратегия  $x_1$  или  $x_2$ .

## 2.11. Основы игровых методов принятия решений

Специальным случаем теории принятия решений в условиях неопределенности является теория игр. Она представляет собой математическую теорию конфликтных ситуаций, в которых действуют две



или более стороны, преследующие разные (часто противоположные) цели, причем выигрыш каждой стороны зависит от поведения других сторон. Поэтому, в отличие от ранее рассмотренных ситуаций, здесь лицо, принимающее решения, имеет дело с разумным противником (конкурентом).

Конфликтной можно в ряде случаев считать ситуацию, когда необходимо принять решение с одновременным использованием нескольких различных критериев. Для описания конфликтной ситуации используется ее модель, называемая игрой.

*Игра* характеризуется набором правил, определяющих возможное поведение участников (игроков) и ее исход. *Игроками* называются конфликтующие стороны. Развитие игры представляется в виде ряда последующих ходов. *Ход* – это выбор игроком одного из предоставленных правилами действий и его осуществление. Ходы бывают *личные* и *случайные*. Личные ходы выбираются игроком сознательно, для осуществления случайного используется механизм случайного выбора (например, рулетка).

Игра, состоящая только из случайных ходов, называется *азартной*, такими играми теория игр не занимается. Цель теории игр – оптимизация поведения игрока в игре, где (возможно, наряду со случайными) имеются личные ходы. Такие игры называются *стратегическими*.

*Стратегия* - это совокупность правил, определяющих выбор варианта действий при каждом личном ходе игрока в зависимости от сложившейся ситуации, при этом не имеет значения, участвует ли игрок в игре лично. Он может разработать стратегию заранее и передать ее некоему незаинтересованному лицу или автомату. *Оптимальная стратегия* обеспечивает игроку максимальный выигрыш в данной игре. Если игра содержит случайные ходы и многократно повторяется, то оптимальная стратегия обеспечивает получение максимального среднего выигрыша.

### **2.11.1. Классификация игр**

Игры могут классифицироваться по различным критериям.

1. В зависимости от числа игроков, различают:

- а) игры одного игрока;
- б) игры двух игроков;
- в) игры  $n$  игроков.

Наиболее исследованы игры двух игроков. С повышением числа игроков трудность решения игры возрастает.

2. По количеству стратегий игры делятся на конечные и бесконечные.

Если каждый игрок имеет конечное количество стратегий, то игра – конечная; если хотя бы один игрок имеет бесконечное число стратегий (счетное множество или континуум), то игра – бесконечная.

Таким образом, конечность или бесконечность игры связана не с ее продолжительностью, а с количеством стратегий.

3. По характеру взаимоотношений между игроками различают игры:

а) бескоалиционные, в которых игрокам запрещается вступать в соглашения;

б) коалиционные, где такое соглашение возможно;

в) кооперативные, в которых коалиции (союзы) оговариваются заранее.

4. По характеру выигрышей бывают игры:

а) с нулевой суммой;

б) с ненулевой суммой.

В первом случае сумма выигрышей в каждой партии равна нулю, так что общий капитал не изменяется, а перераспределяется в зависимости от получающихся исходов. Частный случай игры а) - это игра двух игроков, в которой первый получает выигрыш за счет проигрыша второго. Такие игры называются антагонистическими.

В игре с ненулевой суммой в выигрыше могут оказаться все игроки. К таким играм, например, относятся лотереи, в которых все игроки за право участия в игре платят определенную сумму некоторому постороннему лицу.

5. По виду функций выигрыша различают:

а) матричные игры – в такой игре двух игроков с нулевой суммой выигрыш первого игрока задается в виде матрицы, выигрыш второго равен проигрышу первого, и наоборот. Теория матричных игр хорошо разработана, в частности, доказано, что каждая матричная игра имеет решение и может быть сведена к задаче линейного программирования;

б) биматричные игры: выигрыш каждого игрока задается отдельной матрицей;

в) игры с непрерывными функциями – в подобной игре выигрыш

каждого игрока задается непрерывной функцией от стратегий всех игроков;

г) игры типа дуэлей, в которых стратегия заключается в выборе момента осуществления хода, причем вероятность получения выигрыша зависит от времени, прошедшего с начала игры до момента хода.

6. По количеству ходов (шагов) игры делятся на одношаговые и многошаговые.

Одношаговая игра заканчивается после того, как каждый игрок делает по одному ходу. Многошаговые, далее, делятся:

- на позиционные (в них каждый игрок может последовательно делать несколько шагов);

- стохастические (в них с определенной вероятностью происходит возврат на предшествующую позицию);

- дифференциальные (в них поведение игроков подчиняется определенным условиям, описываемым дифференциальными уравнениями).

7. В зависимости от состояния информации различаются игры:

а) с полной информацией (игрок на каждом ходу знает все выборы всех участников, сделанные на предыдущих ходах; пример - игра в шахматы);

б) с неполной информацией (пример - покер).

Доказано, что всякая игра с полной информацией имеет решение в чистых стратегиях.

### **2.11.2. Матричные игры двух игроков с нулевой суммой. Решение в чистых стратегиях**

Рассматривается игра, в которой первый игрок («белые») имеет  $m$  стратегий  $(x_1, x_2, \dots, x_m)$ , второй («черные») –  $n$  стратегий  $(y_1, y_2, \dots, y_n)$ . Каждая отдельно взятая стратегия игрока называется чистой стратегией.

При выборе «белыми» чистой стратегии  $x_i$ , а «черными» – чистой стратегии  $y_j$  «белые» получают выигрыш  $a_{ij}$ , а «черные» столько же проигрывают. Таким образом, задана платежная матрица  $A = \|a_{ij}\|$ ,  $i = 1, \dots, m, j = 1, \dots, n$ .

*Пример 2.25.* У «белых» имеется три вида зенитного оружия, у «черных» - три типа самолетов. Значения  $a_{ij}$  – вероятностей поражения самолета  $j$ -го типа оружием  $i$ -го типа – приведены в табл. 2.12.

Таблица 2.12. Вероятности поражения к примеру 2.25

	$Y_1$	$Y_2$	$Y_3$
$x_1$	0,5	0,6	0,8
$x_2$	0,9	0,7	0,8
$x_3$	0,7	0,5	0,6

Обозначим:  $\alpha_i = \min_j a_{ij}$ ,  $\beta_j = \max_i a_{ij}$ . «Белым» выгодно играть так, чтобы получить выигрыш  $\alpha = \max_i \alpha_i = \max_i \min_j a_{ij}$ , а «черным» - так, чтобы уступить противнику выигрыш  $\beta = \min_j \beta_j = \min_j \max_i a_{ij}$ .

Определенная таким образом величина  $\alpha$  есть чистая нижняя цена игры, а  $\beta$  – чистая верхняя цена игры. Получим соотношение между  $\alpha$  и  $\beta$ .

*Теорема.* Пусть  $f(x,y)$  – вещественная функция двух переменных  $x \in A$  и  $y \in B$  и существуют  $\alpha = \max_{x \in A} \min_{y \in B} f(x,y)$ ,  $\beta = \min_{y \in B} \max_{x \in A} f(x,y)$ , тогда  $\alpha \leq \beta$ .

Доказательство. По определению минимума и максимума функции,

$$\min_{y \in B} f(x,y) \leq f(x,y) \leq \max_{x \in A} f(x,y). \quad (2.20)$$

Поскольку в левой части (2.20)  $x$  любое, то

$$\max_{x \in A} \min_{y \in B} f(x,y) \leq \max_{x \in A} f(x,y).$$

Так как в правой части неравенства (2.19)  $y$  любое, то

$$\max_{x \in A} \min_{y \in B} f(x,y) \leq \min_{y \in B} \max_{x \in B} f(x,y),$$

что и требовалось доказать.

В частном случае, когда  $f(x,y)$  задается матрицей  $\|a_{ij}\|$ , справедливо:

$$\alpha = \max_i \min_j a_{ij} \leq \min_j \max_i a_{ij} = \beta.$$

Таким образом, в матричной игре нижняя чистая цена игры не превосходит ее верхней чистой цены.

Существуют игры, в которых  $\alpha = \beta = \nu$ ; они называются играми с седловой точкой, а  $\nu$  - чистой ценой игры. В матричной игре седловой точке соответствует такой элемент, который одновременно является наименьшим в своей строке и наибольшим в своем столбце.

В нашем примере это элемент  $a_{22}$ .

Седловой точке соответствует пара чистых оптимальных стратегий, дающих оптимальное решение игры. В нашем случае это стратегии  $x_2$  и  $y_2$ .

Оптимальное решение обладает следующим свойством: если первый игрок придерживается своей оптимальной чистой стратегии, то для другого игрока не может быть выгодным отклонение от его оптимально чистой стратегии, так как это может только ухудшить его положение.

Таким образом, игры с седловой точкой обладают устойчивостью, и пара оптимальных чистых стратегий образует точку равновесия.

Чистая цена игры  $V$  здесь является тем значением, которое первый игрок никак не может увеличить, а второй – уменьшить, даже если каждый из них знает, какую именно стратегию выберет противник. К играм с седловой точкой относятся, в частности, игры с полной информацией.

### **2.11.3. Матричные игры двух игроков с нулевой суммой. Решение в смешанных стратегиях**

Чаще всего  $\alpha \neq \beta$ , и игра не имеет седловой точки. Возникает вопрос – нельзя ли в этом случае выиграть больше, чем  $\alpha$  (и проиграть меньше, чем  $\beta$ ), если применять не одну, а чередовать несколько стратегий. Очевидно, чередовать стратегии следует случайным образом, чтобы противник не мог угадать заранее, в какой обстановке ему придется действовать.

Если у первого игрока имеются чистые стратегии  $x_1, x_2, \dots, x_m$  и  $p_i$  – вероятность применения стратегии  $x_i$ , то полный набор вероятностей  $S_x = (p_1, p_2, \dots, p_m)$  называется смешанной стратегией первого игрока. По аналогии для второго игрока с чистыми стратегиями  $y_1, \dots, y_n$  смешанная стратегия  $S_y = (q_1, q_2, \dots, q_n)$ , где  $q_j$  – вероятность применения стратегии  $y_j$ .

*Основная теорема теории игр.* Каждая конечная игра имеет по крайней мере одно решение, возможно, в области смешанных стратегий, при этом цена игры  $v$  лежит в диапазоне от  $\alpha$  до  $\beta$  ( $\alpha \leq v \leq \beta$ ).

*Доказательство.* Если «белые» применяют свою смешанную

стратегию, а «черные» – чистую стратегию  $y_j$ , то средний выигрыш «белых» составит  $h_j = \sum_{i=1}^m a_{ij} p_i$ . Если «белые» применяют чистую стратегию

$x_i$ , а «черные» – смешанную, то средний выигрыш составит  $K_i = \sum_{j=1}^n a_{ij} q_j$ .

Если оба игрока применяют свои смешанные стратегии, то средний выигрыш равен

$$E = \sum_{i=1}^m \sum_{j=1}^n a_{ij} p_i q_j .$$

*Пример 2.26.* Задана игра с платежной матрицей  $A = \begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}$ .

Пусть «белые» применяют смешанную стратегию  $S_x = (0,2 \quad 0,8)$ , а «черные» - смешанную стратегию  $S_y = (0,4 \quad 0,6)$ . Тогда средний выигрыш  $E = 5 \cdot 0,2 \cdot 0,4 + 1 \cdot 0,2 \cdot 0,6 + 3 \cdot 0,8 \cdot 0,4 + 4 \cdot 0,8 \cdot 0,6 = 3,4$ .  
 $\alpha = 3, \beta = 4, \alpha < E < \beta$ .

Пусть  $S_x^* = (p_1^*, p_2^*, \dots, p_m^*)$ ,  $S_y^* = (q_1^*, q_2^*, \dots, q_n^*)$  - оптимальные смешанные стратегии, причем некоторые из вероятностей могут быть равны нулю.

Активными называются стратегии игрока, которые входят в оптимальную смешанную стратегию с вероятностями, не равными нулю.

*Теорема об активных стратегиях.* Если один из игроков придерживается своей оптимальной смешанной стратегии, то выигрыш остается неизменным и равным цене игры, независимо от того, что делает второй игрок, если только он не выходит за пределы своих активных стратегий.

#### **2.11.4. Игры с платежной матрицей размером 2 x 2**

В такой игре каждый игрок имеет по две стратегии, и при отсутствии седловой точки обе они являются активными.

Пусть в игре с платежной матрицей  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  первый игрок

придерживается своей оптимальной смешанной стратегии  $S_x^* = (p_1^*, p_2^*)$ .

Тогда, по теореме об активных стратегиях, независимо от действий второго игрока будет выполняться:  $a_{11}p_1^* + a_{21}p_2^* = v$ ,  $a_{12}p_1^* + a_{22}p_2^* = v$ , кроме того  $p_1^* + p_2^* = 1$ . Получаем систему уравнений

$$\begin{cases} (a_{11} - a_{12})p_1^* + (a_{21} - a_{22})p_2^* = 0; \\ p_1^* + p_2^* = 1. \end{cases}$$

Решим ее по методу Крамера:

$$\Delta = (a_{11} - a_{12}) - (a_{21} - a_{22});$$

$$\Delta_{p_1} = \begin{vmatrix} 0 & a_{21} - a_{22} \\ 1 & 1 \end{vmatrix} = -(a_{21} - a_{22}) = a_{22} - a_{21};$$

$$\Delta_{p_2} = \begin{vmatrix} a_{11} - a_{12} & 0 \\ 1 & 1 \end{vmatrix} = a_{11} - a_{12};$$

$$p_1^* = \frac{\Delta_{p_1}}{\Delta} = \frac{a_{22} - a_{21}}{(a_{11} - a_{12}) + (a_{22} - a_{21})};$$

$$p_2^* = \frac{\Delta_{p_2}}{\Delta} = \frac{a_{11} - a_{12}}{(a_{11} - a_{12}) + (a_{22} - a_{21})}; \quad (2.21)$$

$$v = a_{11} \frac{a_{22} - a_{21}}{(a_{11} + a_{12}) - (a_{12} + a_{21})} + a_{21} \frac{a_{11} - a_{12}}{(a_{11} + a_{22}) - (a_{12} + a_{21})} = \frac{a_{11}a_{22} - a_{21}a_{12}}{(a_{11} + a_{22}) - (a_{12} + a_{21})}.$$

Если второй игрок придерживается смешанной стратегии  $S_y^* = (q_1^*, q_2^*)$ , то справедливы уравнения

$$\begin{cases} a_{11}q_1^* + a_{12}q_2^* = v; \\ a_{21}q_1^* + a_{22}q_2^* = v; \\ q_1^* + q_2^* = 1. \end{cases}$$

Решив их, получим:

$$q_1^* = \frac{a_{22} - a_{12}}{(a_{11} + a_{22}) - (a_{12} + a_{21})};$$

$$q_2^* = \frac{a_{11} - a_{21}}{(a_{11} + a_{22}) - (a_{12} + a_{21})}. \quad (2.22)$$

Для примера 2.27 можно получить:  $p_1^* = \frac{4-3}{(5+4)-(1+3)} = \frac{1}{5}$ ;  
 $p_2^* = \frac{4}{5}$ ;  $q_1^* = \frac{3}{5}$ ;  $q_2^* = \frac{2}{5}$ ;  $v = \frac{5 \cdot 4 - 1 \cdot 3}{5} = \frac{17}{5} = 3,4$ .

### 2.11.5. Матричные игры с платежной матрицей $2 \times n$ и $m \times 2$

Рассмотрим игру с платежной матрицей размером  $2 \times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix}.$$

В игре без седловой точки первый игрок имеет только две стратегии, причем обе активные. Обозначим смешанную стратегию белых  $S_x = (p_1, p_2)$ . Если черные применяют свою  $j$ -ую стратегию, то средний выигрыш белых составит  $K_j = a_{1j} \cdot p_1 + a_{2j} \cdot p_2$ , или

$$K_j = a_{1j} \cdot p_1 + a_{2j}(1 - p_1) = a_{2j} + p_1(a_{1j} - a_{2j}).$$

Последнее выражение есть не что иное, как уравнение прямой линии в координатах  $p_1, K$ . Это позволяет решить игру графически. Построим зависимости  $K_j(p_1)$  для случая применения вторым игроком каждой своей стратегии (рис. 2.31).

Минимальный выигрыш «белых» соответствует нижней огибающей семейства прямых (линия  $AOB$ ), а его максимум достигается при значении  $p_1 = p_1^*$ , соответствующем абсциссе точки  $O$ , ордината этой точки дает цену игры  $v$ . Для «черных» активной является пара стратегий, соответствующих прямым, пересекающимся в точке  $O$  – в данном случае стратегий  $y_2, y_n$ . Очевидно, что подобные рассуждения справедливы и для случая, когда платежная матрица имеет размер  $m \times 2$ :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \dots & \dots \\ a_{m1} & a_{m2} \end{pmatrix}.$$



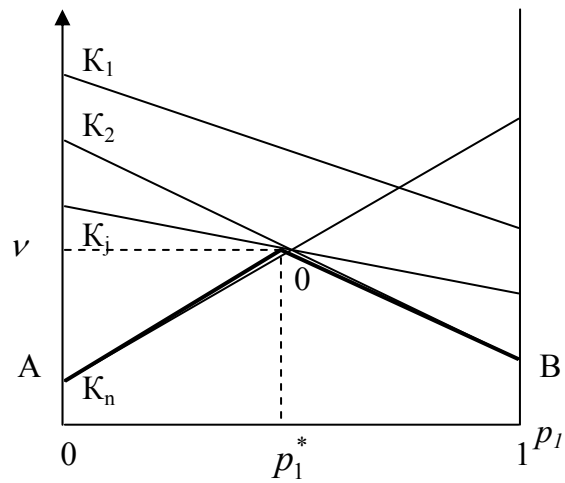


Рис. 2.31. Графическое решение игры  $2 \times n$

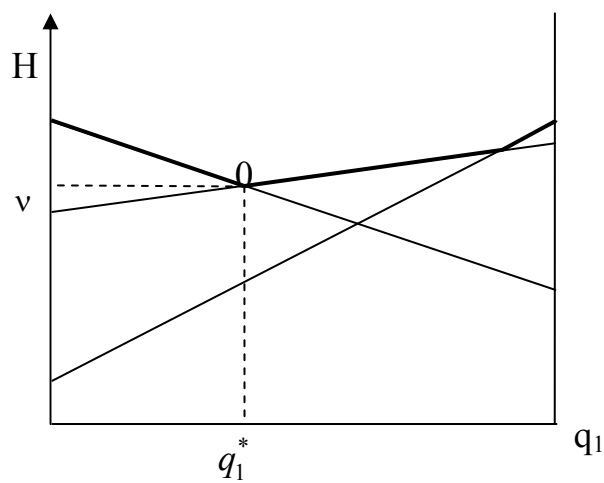


Рис. 2.32. Графическое решение игры  $m \times 2$

Уравнение среднего ожидаемого проигрыша второго игрока в случае применения им смешанной стратегии  $S_y = (q_1, q_2)$ , а белыми — своей  $i$ -ой чистой стратегии есть  $H_i = q_1(a_{i1} - a_{i2}) + a_{i2}$ . Максимальный выигрыш белых будет соответствовать верхней огибающей семейства прямых

$H_i(q_1)$ ,  $i = 1, \dots, m$ , а оптимальная вероятность применения черными своей первой чистой стратегии  $q_1^*$  будет соответствовать нижней точке  $O$  этой огибающей (рис. 2.32).

Графическое решение является приближенным. Чтобы получить точное решение, нужно по графику определить пары активных стратегий каждой стороны, что сводит игру  $2 \times n$  или  $m \times 2$  к игре  $2 \times 2$ , а затем воспользоваться формулами (2.21), (2.22).

*Пример 2.27.* Урожайность сельскохозяйственной культуры зависит от внесенных удобрений и погоды. Рассматривается два вида погоды: 1) сухая; 2) влажная. Возможные варианты внесения удобрений: 1) согласно определенной норме; 2) на 30% выше нормы; 3) на 40% ниже нормы.

Будем считать первым игроком сельскохозяйственное предприятие, а вторым – природу. У первого игрока имеется три стратегии, соответствующие вариантам внесения удобрений, у второго – две стратегии, соответствующие видам погоды. Платежная матрица, выражающая прибыль предприятия в зависимости от стратегий игроков,

имеет вид: 
$$A = \begin{pmatrix} 4 & 2 \\ 2 & 4 \\ 3 & 2,5 \end{pmatrix}.$$

Найдем решение полученной игры  $3 \times 2$  графическим методом. Выигрыши первого игрока при его различных стратегиях выражаются зависимостями:

$$H_1 = (4 - 2)q_1 + 2 = 2q_1 + 2;$$

$$H_2 = (2 - 4)q_1 + 4 = -2q_1 + 4;$$

$$H_3 = (3 - 2,5)q_2 + 2,5.$$

Из рис. 2.33, где полученные зависимости изображены графически, видим, что оптимальное решение дается точкой  $O$ . Соответствующие ему стратегии первого игрока – это  $x_1$  и  $x_2$ . Исключив третью стратегию первого игрока, как неактивную, получим игру  $2 \times 2$ , аналитическое решение которой с помощью формул (2.21), (2.22) позволяет получить:

$$p_1^* = p_2^* = q_1^* = q_2^* = 0,5; \quad v = 3.$$

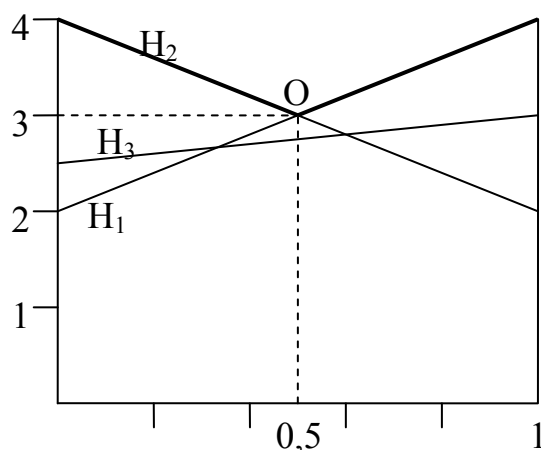


Рис. 2.33. Решение примера 2.27

Таким образом, оптимальная стратегия сельскохозяйственного предприятия есть  $S_x^* = (0,5; 0,5; 0)$ , максимальный ожидаемый доход при этом составит 3 единицы.

### 2.11.6. Свойства смешанных стратегий

Введем обозначение  $GA$  для игры с платежной матрицей  $A$ .

*Теорема об оптимальных стратегиях:* Для того, чтобы в игре  $GA$  с ценой  $v$  смешанная стратегия первого игрока  $S_x^*$  была оптимальной, необходимо и достаточно, чтобы для любой смешанной стратегии второго игрока  $S_y$  выполнялось неравенство  $E(S_x^*, S_y) \geq v$ .

Для того, чтобы смешанная стратегия второго игрока  $S_y^*$  была оптимальной, необходимо и достаточно, чтобы для любой смешанной стратегии первого игрока  $S_x$  выполнялось неравенство  $E(S_x, S_y^*) \leq v$ .

Из данной теоремы следуют свойства 1 и 2 смешанных стратегий.

*Свойство 1.* Для того, чтобы  $S_x^* = (p_1^*, p_2^*, \dots, p_m^*)$  была оптимальной смешанной стратегией первого игрока в игре  $GA$  с ценой  $v$ , необходимо и достаточно выполнение следующих неравенств:  $\sum_{i=1}^m a_{ij} p_i^* \geq v, \quad j = \overline{1, n}$ .

Для того, чтобы смешанная стратегия второго игрока

$S_y^* = (q_1^*, q_2^*, \dots, q_n^*)$  была оптимальной, необходимо и достаточно выполнение следующих неравенств:  $\sum_{j=1}^n a_{ij} q_j^* \leq v, \quad i = \overline{1, m}$ .

*Свойство 2.* Если  $S_x^*$  и  $S_y^*$  - оптимальные смешанные стратегии игроков, и для некоторого  $i \sum_{j=1}^n a_{ij} \cdot q_j^* < v$ , тогда  $p_i^* = 0$ ; если для некоторого  $j \sum_{i=1}^m a_{ij} \cdot p_i^* > v$ , тогда  $q_j^* = 0$ .

Помимо приведенных свойств, также имеют место следующие важные свойства, облегчающие решение ряда матричных игр.

*Свойство 3.* Если платежная матрица игры кососимметрична ( $a_{ij} = -a_{ji}, i = \overline{1, m}, j = \overline{1, m}$ ), то оптимальная смешанная стратегия первого игрока одновременно является оптимальной смешанной стратегией второго игрока, и наоборот.

*Свойство 4.* Если в игре  $ГВ$  каждый элемент платежной матрицы  $v_{ij} = ca_{ij} + d$  ( $c > 0, d$  - любое действительное число), то решение игры  $ГВ$  (смешанные оптимальные стратегии игроков) совпадает с решением игры  $ГА$ , а цена игры  $v_B = v_A \cdot c + d$ .

Если в игре  $ГА$  для чистых стратегий первого игрока  $x_k, x_i$  выполняется неравенство  $a_{ij} \geq a_{kj}, j = \overline{1, n}$ , причем хотя бы для одного значения  $j$  это неравенство строгое, то говорят, что стратегия  $x_i$  доминирует стратегию  $x_k$ . Если для чистых стратегий второго игрока  $y_j, y_1$  выполняется неравенство  $a_{il} \leq a_{ij}, l = \overline{1, m}$ , причем хотя бы для одного значения  $l$  это неравенство строгое, то  $y_j$  доминирует стратегию  $y_1$ .

*Свойство 5.* Пусть в матричной игре  $Г_A$  с платежной матрицей порядка  $m \times n$   $i$ -я стратегия первого игрока доминирует  $k$ -ю стратегию. Пусть  $A'$  - матрица, получаемая из  $A$  путем исключения из нее  $k$ -ой строки и  $\Gamma'_A$  - матричная игра с матрицей  $A'$ . Тогда цена игры  $\Gamma'_A$  совпадает с ценой игры  $Г_A$ , всякая оптимальная смешанная стратегия второго игрока в игре  $\Gamma'_A$  есть также его оптимальная смешанная стратегия в игре  $Г_A$ ; если

$$S_x^* = (p_1^*, \dots, p_i^*, \dots, p_{k-1}^*, p_{k+1}^*, \dots, p_m^*)$$

есть оптимальная смешанная стратегия первого игрока в игре  $\Gamma'_A$ , то его смешанная стратегия  $S_x^* = (p_1^*, \dots, p_i^*, \dots, p_{k-1}^*, 0, p_{k+1}^*, \dots, p_m^*)$  является оптимальной в игре  $\Gamma_A$ .

По аналогии, если из платежной матрицы исключить столбец, соответствующий доминирующей стратегии второго игрока, то цена игры не изменится, и в оптимальном решении доминирующей стратегии будет соответствовать нулевая вероятность ее применения.

*Пример 2.28.* Рассмотрим игру с платежной матрицей

$$A = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 0 & 2 & 3 & 2 \\ 4 & 3 & 1 & 0 \end{pmatrix}. \text{ Стратегия } x_1 \text{ первого игрока доминирует его вторую}$$

стратегию (в случае применения стратегии  $x_2$  игрок выигрывает не больше, чем в случае применения стратегии  $x_1$ ); вычеркивая вторую

строку матрицы, приходим к игре  $A' = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 1 & 0 \end{pmatrix}.$

В этой игре стратегия второго игрока  $y_3$  доминирует его четвертую стратегию (в случае применения стратегии  $y_3$  игрок проигрывает больше, чем в случае применения стратегии  $y_4$ ); исключив третий столбец,

приходим к игре с матрицей  $A'' = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 0 \end{pmatrix}$ , которая может быть решена

графическим методом. В полученном оптимальном решении необходимо положить вероятности  $p_2^*$  и  $q_3^*$ , соответствующие вычеркнутым элементам матрицы, равными нулю.

### **2.11.7. Решение матричных игр методом линейного программирования**

В этом методе предполагается, что цена игры положительна. Данное условие не нарушает общности, так как, согласно свойству 4 смешанных стратегий, всегда можно подобрать число  $c$ , прибавление которого ко всем

элементам платежной матрицы даст матрицу с положительными элементами; при этом цена игры станет положительной, а оптимальные смешанные стратегии игроков не изменятся.

Пусть задана матричная игра с матрицей  $A = \|a_{ij}\|$  порядка  $m \times n$ . Согласно свойству 1 смешанных стратегий, для оптимальной стратегии

$$S_x^* = (p_1^*, p_2^*, \dots, p_m^*) \text{ выполняется: } \sum_{i=1}^m a_{ij} p_i^* \geq v, \text{ причем } \sum_{i=1}^m p_i^* = 1.$$

Поделив оба соотношения на положительное число  $v$  (цену игры), получим:

$$\begin{cases} \sum_{i=1}^m a_{ij} \frac{p_i^*}{v} \geq 1; \\ \sum_{i=1}^m \frac{p_i^*}{v} = \frac{1}{v} = \bar{v}, \end{cases} \text{ или } \begin{cases} \sum_{i=1}^m a_{ij} x_i^* \geq 1; \\ \sum_{i=1}^m x_i^* = \bar{v}, \end{cases} \text{ где } x_i^* = \frac{p_i^*}{v}, \bar{v} = \frac{1}{v}.$$

Поскольку первый игрок стремится найти такие значения  $x_i^*$ , чтобы цена игры была максимальна, а второй игрок - наоборот, то решение игры сводится к задаче линейного программирования

$$\begin{cases} \sum_{i=1}^m a_{ij} x_i^* \geq 1; \\ \sum_{i=1}^m x_i^* \rightarrow \min. \end{cases} \quad \begin{cases} \sum_{j=1}^n a_{ij} y_j^* \leq 1; \\ \sum_{i=1}^n y_i^* \rightarrow \max. \end{cases}$$

### 2.11.8. Физическая смесь стратегий

После нахождения решения игры в смешанных стратегиях, возникает вопрос о практической реализации этих стратегий. Рассмотрим, как это выполняется в случае организации боевых действий. При решении тактических задач применение смешанной стратегии означает применение гибкой тактики, выбираемой с найденной оптимальной вероятностью. Иначе обстоит дело при решении технических задач. Например, при планировании разработки нового вида вооружения случайный выбор образца вооружения вряд ли целесообразен: поскольку решение принимается на длительную перспективу, то противник получает

возможность раскрыть стратегию оперирующей стороны, и, поступая с учетом ее знания, свести ее выигрыш к минимуму.

В таких задачах применяется физическая смесь стратегий, а именно такая их смесь, при которой одновременно применяется несколько стратегий в определенных пропорциях. Например: разрабатывается сразу несколько видов вооружения в оптимальном количественном соотношении.

*Пример 2.29.* Первый игрок планирует производство зенитных управляемых ракет (ЗУР). Имеется 4 образца ЗУР и 5 типов самолетов противника, против которых эти ракеты предполагается использовать. Матрица вероятностей поражения разных самолетов различными типами ракет имеет следующий вид:

$$A = \begin{pmatrix} 0,2 & 0,4 & 0,6 & 0,4 & 0,7 \\ 0,3 & 0,4 & 0,6 & 0,5 & 0,8 \\ 0,4 & 0,5 & 0,6 & 0,5 & 0,8 \\ 0,7 & 0,3 & 0,5 & 0,2 & 0,1 \end{pmatrix}.$$

Требуется обосновать пропорции, в которых нужно выпускать различные типы ЗУР.

Рассматривая матрицу, видим, что третья стратегия первого игрока доминирует его стратегии 1 и 2, поэтому исключаем из матрицы первые две строки. С точки зрения противника, следует исключить его вторую и третью стратегию, как невыгодные по сравнению с четвертой. Получим матрицу вида  $A' = \begin{pmatrix} 0,4 & 0,5 & 0,8 \\ 0,7 & 0,2 & 0,1 \end{pmatrix}$ . Решая полученную игру графически, получаем:

$$S_x^* = (0 \quad 0 \quad 5/6 \quad 1/6), \quad v = 0,45.$$

Таким образом, оптимальный план производства ракет заключается в следующем: ЗУР первого и второго типов производить не следует, а ЗУР третьего и четвертого типов производить в соотношении 5:1.

Применительно к задаче о внесении удобрений (пример 2.27) оптимальная физическая смесь стратегий заключается в том, чтобы на половине посевных площадей удобрения вносить согласно норме, а на другой половине – с 30-процентным превышением нормы.

### **3. ПРИКЛАДНЫЕ ЗАДАЧИ СИСТЕМНОГО АНАЛИЗА И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ**

#### **3.1. Проектирование космической системы наблюдения**

##### ***3.1.1. Постановка расширенной задачи***

Космическая система наблюдения (КСН) используется для изучения земной поверхности в видимой и ближней инфракрасной области спектра. Это – визуальные наблюдения, фото и телевизионная съемка. Методы, регистрирующие невидимую часть электромагнитного спектра излучения Земли, включают в себя ряд специальных съемок: инфракрасную, радиолокационную, спектрометрическую и другие. Можно указать свыше 40 задач, интересующих потребителя [1]. Средством достижения этих целей наблюдения Земли служит КСН.

В качестве основных требований, предъявляемые к КСН потребителями космической информации, можно использовать:

- требования к спектральным характеристикам аппаратуры наблюдения (АН) – обеспечиваются только выбором соответствующей аппаратуры наблюдения;
- требования к району наблюдения – обеспечиваются только выбором орбитальной структуры, то есть количеством КА и параметрами их орбит;
- требования периодичности наблюдения – определяются параметрами орбит;
- требования к пространственному разрешению и обзорности изображения обуславливается совместным выбором характеристик АН и орбитальной структуры;
- требования к оперативности – обеспечиваются выбором АН, орбитальной структуры и других характеристик КСН;
- требования к эффективности – обеспечиваются выполнением перечисленных требований, а также затратами, оцениваемых стоимостью



разработки, изготовлением, эксплуатации КСН и сроком её службы.

*Структура космической системы наблюдения.* Космическая система – совокупность функционально связанных КА и наземных технических средств, предназначенных для решения целевых задач (рис. 3.1).

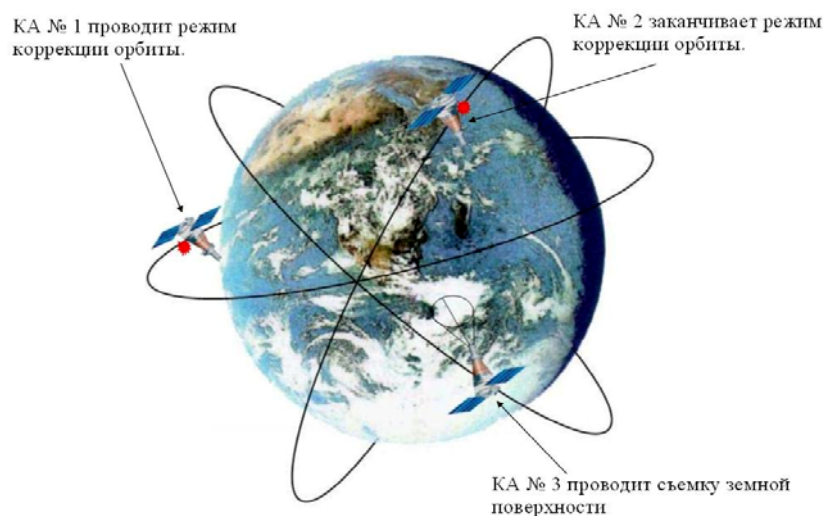


Рис. 3.1. Схема функционирования орбитальной группировки спутников

КА в составе КСН представляют собой подсистему, которая состоит из функционально взаимосвязанных КА, совместно выполняющих целевую задачу. Выведение КА на орбиту для построения системы ее восполнения осуществляется наземным ракетно-техническим комплексом. Для управления космическими аппаратами одного типа служит наземный комплекс управления. Для приема, автоматизированной обработки и анализа поступающей от КА специальной информации служит наземный специальный комплекс (НСК).

*Проектные параметры космической системы наблюдения.* Средствами достижения целей КСН, то есть управляемыми факторами и параметрами, являются: характеристики АН; орбитальная структура; срок эксплуатации КСН и требуемые ресурсы.

*Аппаратура наблюдения,* используемая на КА, служит для получения изображений высокого качества. Её важнейшими характеристиками являются пространственное разрешение фотоаппаратуры

и угол обзора.

*Срок эксплуатации* КСН и требуемые ресурсы. В реальной космической системе происходит эволюция орбитальной структуры под действием возмущений, связанных: с погрешностью выведения КА на орбиту; с нецентральнойностью поля тяготения, аэродинамическим торможением и давлением солнечного света; погрешностью коррекции траекторий.

### **3.1.2. Модель функционирования орбитальной группировки**

При проектировании КСН из условия наилучшего показателя периодичности наблюдений, задачу условно можно разделить на два этапа: задачу построения орбитальной структуры КСН из условия наилучшего показателя периодичности и задачу планирования программы съемки при выбранной орбитальной структуре.

Целью *построения орбитальной структуры* КСН является выбор баллистических параметров входящих в нее космических аппаратов, обеспечивающих наилучшие значения показателей периодичности наблюдений. Объекты наблюдения (ОН) могут быть одиночными, объединенными в группы или равномерно распределенными по поверхности Земли. При решении задачи синтеза орбитальной структуры КСН используются следующие модели: модель орбитального движения КА; модель определения межвиткового сдвига; модель прохода КА над объектом наблюдения; модель облачности и модель освещенности.

Целью *задачи планирования программы съемки* является формирование программы съемки заданных объектов наблюдения при заданных параметрах орбитальной структуры. При этом считается, что перечень объектов наблюдения определен заранее. При решении этой задачи требуется учитывать наличие облачного покрова и освещенность ОН, максимальную скорость программного разворота, объем передаваемой специальной информации, размер буфера целевой аппаратуры.

При решении задачи планирования программы съемки используются следующие модели: модель орбитального движения КА; модель определения межвиткового сдвига; модель прохода КА над объектом наблюдения; модель облачности; модель освещенности; модель оценки

времени пребывания КА в зоне видимости ОН и модель оценки минимального угла программного разворота КА по углу крена.

Так как при решении обеих задач используются одинаковые модели, может быть создан универсальный алгоритм решения задачи построения орбитальной структуры и задачи планирования программы съемки.

*Модель орбитального движения КА.* Моделирование орбитального движения системы КА – основной инструмент получения информации для принятия решений при проектировании спутниковых систем различного назначения. С помощью моделирования оценивается эволюция системы и орбитальная обстановка на длительных интервалах времени, определяются моменты съемки объектов, вычисляются различные показатели эффективности системы. Средства моделирования должны обеспечивать достаточную для принятия решений полноту информации при различной точности вычислений и располагаемых ресурсах вычислительных средств.

Введем вектор оскулирующих элементов орбиты  $X = \{a, e, \omega, \Omega, i, u\}$ , где  $a$  - большая полуось орбиты,  $e$  - эксцентриситет орбиты,  $\omega$  - аргумент перигея,  $\Omega$  - долгота восходящего узла,  $i$  - наклонение плоскости орбиты,  $\vartheta$  - истинная аномалия КА.

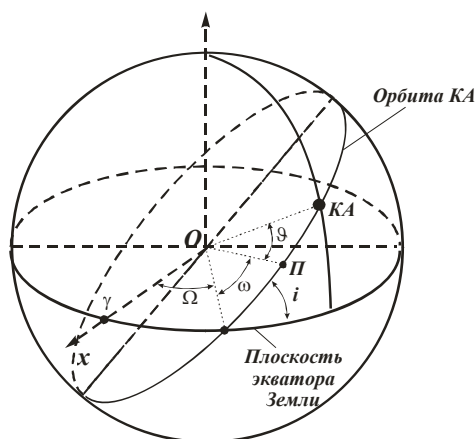


Рис. 3.2. Оскулирующая система координат

Для моделирования орбитального движения используется система дифференциальных уравнений, записанная в оскулирующих элементах [21]:

$$\begin{aligned}
\frac{d\Omega}{dt} &= W \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i}, \\
\frac{di}{dt} &= W \frac{r}{\sqrt{\mu p}} \cos u, \\
\frac{dp}{dt} &= 2 \cdot \sqrt{\frac{p}{\mu}} \cdot Tr, \\
\frac{d\omega}{dt} &= \sqrt{\frac{p}{\mu}} \left( -S \frac{\cos \nu}{e} + \left( 1 + \frac{r}{p} \right) \frac{\sin \nu}{e} T - \frac{r}{p} W \operatorname{ctg} i \sin u \right), \\
\frac{de}{dt} &= \sqrt{\frac{p}{\mu}} \left( S \sin \nu + \frac{e \cos^2 \nu + 2 \cos \nu + e}{1 + e \cos \nu} \right), \\
\frac{d\nu}{dt} &= \sqrt{\frac{p}{\mu}} \left( \frac{\mu}{r^2} + S \frac{\cos \nu}{e} - T \left( 1 + \frac{r}{p} \right) \frac{\sin \nu}{e} \right),
\end{aligned} \tag{3.1}$$

где  $\mu = 398602 \text{ км}^3/\text{с}^2$  - гравитационный параметр Земли,  $S, T, W$  - проекции возмущающих ускорений на оси орбитальной системы координат. В качестве возмущающих ускорений могут рассматриваться гравитационные возмущения от нецентральности поля тяготения Земли, аэродинамические возмущающие ускорения, и др.

*Модель межвиткового сдвига.* При движении КА по орбите непрерывно изменяется положение подспутниковой точки (точки пересечения радиус-вектора, направленного из центра Земли на КА, с поверхностью Земли). Кривую, образованную на поверхности Земли из совокупности подспутниковых точек в процессе полета КА, называют трассой (следом орбиты). Вследствие возмущений и вращения Земли трасса КА в общем случае не является замкнутой кривой [22]. Форма трассы определяется наклоном, эксцентриситетом орбиты и периодом движения КА по орбите  $T$ .

Вследствие вращения Земли и прецессии линии узлов географическая долгота начала и конца витка не совпадают. За период обращения КА Земля повернется на угол  $\Delta\lambda_3 = \omega_3 T$ , где  $\omega_3 = 7.29211 \cdot 10^{-5} \text{ с}^{-1}$  - угловая скорость вращения Земли. Из-за прецессии линии узлов орбиты, конец витка сместится на дополнительный угол  $\Delta\lambda_{II} = \Delta\Omega \cdot T$ . Если учитывать только вторую зональную гармонику в разложении гравитационного поля Земли, то скорость прецессии линии

узлов можно определить по формуле  $\Delta\Omega = -A \frac{\cos i}{p^2}$ .

Угловое расстояние  $\Delta\lambda = \Delta\lambda_3 + \Delta\lambda_{II}$  называют смещением спутника по долготе за виток, или межвитковым интервалом [1].

*Модель прохода КА над объектом наблюдения.* Для правильной оценки периодичности необходимо знать время первого выхода КА на геодезическую широту ОН. Поскольку при съемке нужно учитывать освещенность ОН, время первого выхода на геодезическую широту даст точную привязку начального положения КА к времени и дате. Поскольку построение полосы обзора (традиционная модель) не дает всей полноты картины нахождения КА в зоне видимости, в предлагаемом методе оценки периодичности рассчитываются точки входа и выхода в зону видимости, а также время пребывания КА в зоне. Все расчеты ведутся с учетом поворотов оптической оси телескопа относительно надира и программных разворотов КА по углу крена.

Для упрощения расчетов времени выхода КА на геодезическую широту ОН, а также времен входа и выхода КА в зону видимости ОН пересчитаем геодезическую широту ОН в аргумент широты для текущих параметров орбиты КА. Это избавляет нас от необходимости на каждом витке искать проекцию подспутниковой точки на земной геоид, т.е. строить трассу КА с относительно малым шагом.

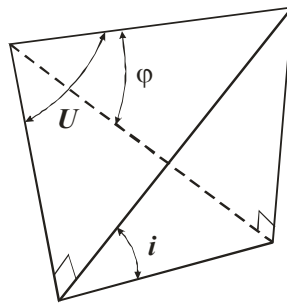


Рис. 3.3. Приведение геодезической широты объекта к аргументу широты

Геодезическая широта подспутниковой точки  $\varphi$  зависит только от аргумента широты ОН  $u$  и наклона орбиты  $\sin u = \frac{\sin \varphi}{\sin i}$  (рис. 3.3).

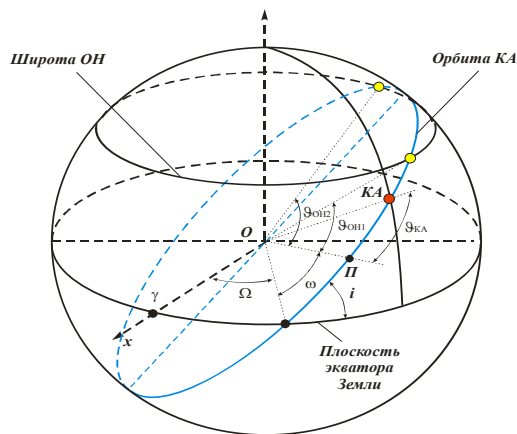


Рис. 3.4. Оценка времени выхода КА на геодезическую широту ОН

Для правильной оценки периодичности необходимо знать время первого выхода КА на геодезическую широту ОН. Поскольку при съемке нужно учитывать освещенность ОН, время первого выхода на геодезическую широту даст точную привязку по времени к начальному положению КА и дате (рис. 3.4). При планировании съемки как одного, так и нескольких ОН, время выхода на широту позволит оценить время программного разворота КА по углу крена и, соответственно, возможность съемки того или иного ОН каждым КА.

Для решения задачи воспользуемся уравнениями связи времени движения КА с элементами орбиты, которые в случае эллиптической орбиты имеют вид

$$t - \tau = \frac{a^{3/2}}{\sqrt{\mu}} [E - e \sin E], \quad (3.2)$$

где  $t$  - время движения КА по орбите,  $\tau$  - время прохождения КА через перигей,  $E$  - угол эксцентрической аномалии:

$$\operatorname{tg} \frac{E}{2} = \sqrt{\frac{1-e}{1+e}} \operatorname{tg} \nu, \quad (3.3)$$

где  $\nu$  - истинная аномалия КА.

Зная начальное положение КА, можно определить время движения КА от перицентра до точки, в которой он находится в начальный момент времени, время движения КА от перицентра до геодезической широты ОН на восходящем витке орбиты и время движения КА от перицентра до геодезической широты ОН на нисходящем витке орбиты.

Время движения КА от начального положения до точки пересечения трассы КА с параллелью, на которой расположен ОН, определится по формулам, для восходящего и нисходящего витков соответственно

$$t_1 = t_{ОН1} - t_{КА0}, \quad t_2 = t_{ОН2} - t_{КА0},$$

где  $t_{ОН1}$  - время движения КА от перигея до широты ОН на восходящем витке орбиты;  $t_{ОН2}$  - время движения КА от перигея до широты ОН на нисходящем витке орбиты;  $t_{КА0}$  - время движения КА от перигея до точки, в которой он находится в начальный момент времени.

Зона обзора КА с учетом максимального угла разворота оптической оси ОЭТК в любом направлении от надира представляет собой сферическую поверхность. Рассмотрим участок орбиты, с которого виден ОН с учетом максимального отклонения ОЭТК от надира (рис. 3.5). Для относительно низких высот полета КА можно пренебречь сферичностью Земли и считать эту поверхность плоской. В первом приближении будем считать участок орбиты, с которого виден ОН, прямой линией. В этом случае зона видимости будет представлять собой круг с радиусом, который нетрудно рассчитать из геометрических соображений  $S = (R - R_3) \operatorname{tg} \gamma$ , где  $S$  - радиус зоны видимости,  $\gamma$  - максимальный угол отклонения оси ОЭТК от надира,  $R$  - радиус орбиты,  $R_3$  - средний радиус Земли.

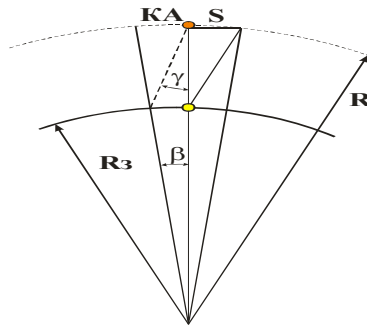


Рис. 3.5. К вопросу об определении зоны видимости

Поскольку построение полосы обзора (традиционная модель) не дает всей полноты картины нахождения КА в зоне видимости, в предлагаемом методе оценки периодичности рассчитываются точки входа и выхода, а также время пребывания КА в зоне (см. рис. 3.5).

Преимущество данного метода заключается в том, что он позволяет определить помимо точек входа и выхода и времени нахождения КА в зоне, кратчайшее угловое расстояние от трассы КА до ОН. С помощью этого углового расстояния можно оценить время программного разворота КА по углу крена, которое является важным параметром при планировании программы съемки.

При построении модели эллиптичностью орбиты можно пренебречь и считать, что радиус орбиты в точке входа в зону, на широте  $OH$  и в точке выхода из зоны одинаковы (рис. 3.6, 3.7).

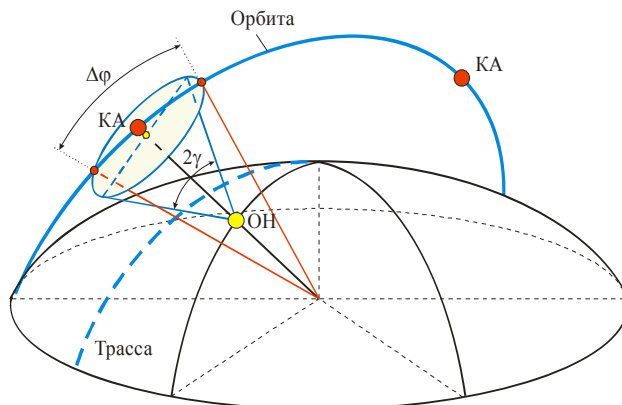


Рис. 3.6. Зона видимости ОН с КА и пересечение её орбитой КА

Сдвиг по долготе относительно ОН, при выходе КА на широту ОН  $a = L_{ОН} - L_i$ , где  $L_{ОН}$  - долгота ОН;  $L_i$  - текущая долгота КА при выходе на широту ОН.

Угловой размер хорды, образованной пересечением линии границы зоны видимости и трассы определится как

$$\Delta\varphi^* = \Delta\varphi_1^* + \Delta\varphi_2^*,$$

$$\cos \Delta\varphi_1^* = \frac{1}{(d^* + d^{**})^2}, \quad \cos \Delta\varphi_2^* = \frac{1}{(d^* - d^{**})^2},$$

$$d^* = \sqrt{S^2 - b^2}, \quad d^{**} = a \cos i, \quad b = a \sin i.$$



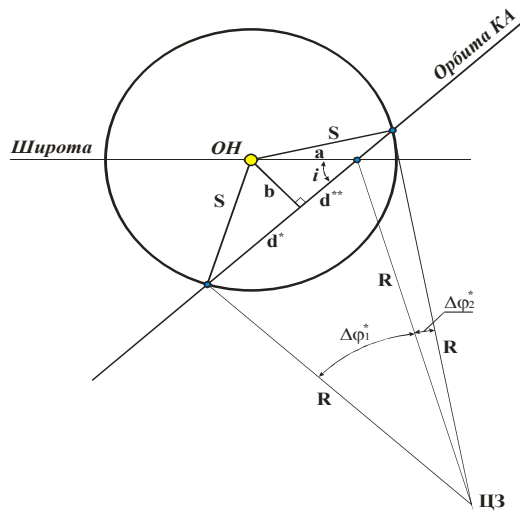


Рис. 3.7. Определение точек входа и выхода в зону видимости объектов наблюдения

Зная угловой размер хорды, можно определить время пребывания КА в зоне видимости. Время движения КА по орбите можно определить из уравнения связи между временем движения КА и элементами орбиты [23].

Время, которое КА находится в зоне видимости ОН, определится как

$$t_{ОН} = t_1 - t_2.$$

где  $t_1$  - время движения КА от перигея орбиты, до точки выхода КА из зоны видимости;  $t_2$  - время движения КА от перигея орбиты, до точки входа КА в зону видимости. Руководствуясь рис. 3.6 и 3.7 и допущениями, сделанными ранее можно найти минимальный угол программного разворота КА по углу крена, необходимый для съемки ОН (рис. 3.8).

Поскольку величина  $b$  использовалась в промежуточных расчетах, будем исходить из того, что нам известно ее значение  $b/H = tg\delta$ , где  $H$  - высота орбиты;  $\delta$  - угол программного разворота КА по углу крена. Если выполняется условие попадания КА в зону видимости  $i$ -го ОН  $|L_{ОНi} - L_i| < S_i$ , то рассчитываются координаты входа, выхода, время пребывания КА в зоне и минимальный угол программного разворота (табл. 3.1). Подобные таблицы составляются для всех ОН.

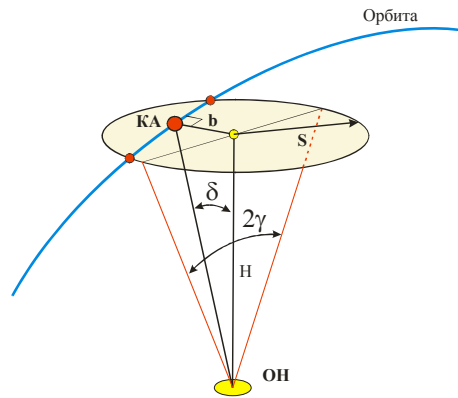


Рис. 3.8. Определение минимального угла программного разворота КА по углу крена

Таблица 3.1. Пример заполнения таблицы учета попаданий.

Объект наблюдения №1				
№пп	№ спутника	Время выхода на ОН	Время наблюдения	Минимальный угол программного разворота
1	1	$t_1$	$T_1$	$\alpha_1$
2	2	$t_2$	$T_2$	$\alpha_2$
...	...	...	...	...
n		$t_n$	$T_n$	$\alpha_n$

Сортировка таблицы по возрастанию элементов столбца «Время выхода на ОН» даст наглядную картину по интервалам между выходами КА на ОН.

Для спутниковых систем, осуществляющих наблюдение Земли из космоса в видимой и инфракрасной зонах спектра, возможность наблюдать определенный участок земной поверхности и качество изображения зависят: от освещенности этого участка солнечным светом и от состояния облачного покрова над объектом наблюдения. Поэтому *модель факторов внешней среды* включает в себя модель освещенности Земной поверхности и модель облачности.

Для осуществления процессов съемки необходима работа специального электрооборудования, требующего определенной электрической мощности. Электроэнергия на борту вырабатывается панелями солнечных батарей (СБ) и она тоже изменяется при различных условиях освещенности.

Освещённость панелей солнечных батарей определяется углом  $\alpha$  - между нормалью к поверхности батареи и направлением на Солнце. С технической точки зрения этот угол  $\alpha$  определяется принятой на рассматриваемом участке полета схемой перекладки панелей СБ. Схема перекладки - это определенное положение панели СБ относительно корпуса КА, характеризующее допустимыми диапазонами углов установки панели по каналам крена и тангажа (рис. 3.9). Каждая схема перекладки соответствует определенному режиму функционирования КА, а также расположению КА относительно Солнца. Различают схемы перекладок для таких режимов функционирования КА, как маневры, сбросы контейнеров, зондирование районов, передача информации потребителю, полет в тени Земли, полет на подходе к зениту Солнца, полет после прохода зенита Солнца.

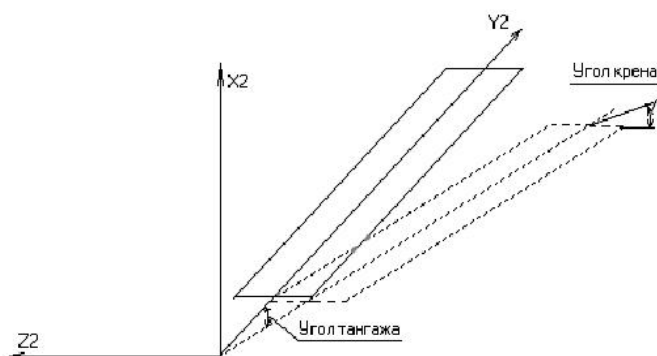


Рис. 3.9. Углы крена и тангажа в связанной с КА системе координат

Таким образом, для заданной схемы перекладки панели СБ необходимо отыскать такие углы установки панели по каналам тангажа и крена, которые обеспечивали бы максимальную средне интегральную мощность снимаемой энергии. Расписание схем перекладок легко определяется по расписанию работы КА. Таким образом, нахождение максимального энергосъема связано с решением двух задач: составление расписания работы КА и определение текущих углов установки панелей СБ, максимизирующих энергосъем [24].

Примем за положительное направление вращения панели СБ по каналу крена вращение плоскости СБ вокруг оси  $Oy_2$  по ходу часовой

стрелки, если смотреть с конца оси  $Oy_2$ , а за положительное направление вращения панели по каналу тангажа вращение плоскости СБ вокруг оси  $OZ_2$  против хода часовой стрелки, если смотреть с конца  $OZ_2$ .

Значение  $\cos \alpha$  определим из соотношения:

$$\cos \alpha = (\bar{S}, \bar{n}) = S_{x2} \cdot n_{x2} + S_{y2} \cdot n_{y2} + S_{z2} \cdot n_{z2}, \quad (3.4)$$

где  $\bar{S}$  - единичный вектор направление на Солнце;  $\bar{n}$  - единичный вектор нормали к панели солнечной батареи;  $S_{x2}, S_{y2}, S_{z2}, n_{x2}, n_{y2}, n_{z2}$  - проекции векторов на оси связанной системы координат.

Найдём проекции вектора  $\bar{S}$  в геоцентрической инерциальной СК:

$$S = \begin{bmatrix} S_x = \cos a_c; \\ S_y = \sin a_c \cdot \cos \varepsilon_c; \\ S_z = \sin a_c \cdot \sin \varepsilon_c. \end{bmatrix},$$

где  $a_c$  - угол между направлением на точку весеннего равноденствия и линией земля Солнце, приближенно равен  $a_c = \omega_z \cdot t$ .

Проекции вектора  $\bar{S}$  в орбитальной СК (рис. 3.10) определяются  $\bar{S}_1 = L \cdot \bar{S}$ , где  $S_1$  - матрица столбец содержащая координаты конца вектора направления на Солнце в орбитальной СК;  $L$  - матрица перехода от геоцентрической инерциальной СК к орбитальной СК.

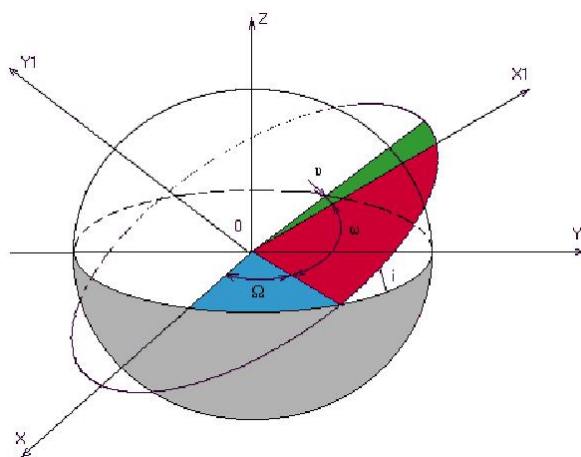


Рис. 3.10. Оси геоцентрической инерциальной  $OXYZ$  и орбитальной  $O_1X_1Y_1Z_1$  систем координат

$$L = \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix};$$

$$L_{11} = \cos \Omega \cdot \cos \omega - \sin \Omega \cdot \sin \omega \cdot \cos i; \quad L_{23} = -\cos \Omega \cdot \sin i;$$

$$L_{12} = -\cos \Omega \cdot \sin \omega - \sin \Omega \cdot \cos \omega \cdot \cos i; \quad L_{31} = \sin \omega \cdot \sin i;$$

$$L_{13} = \sin \Omega \cdot \sin i; \quad L_{32} = \cos \omega \cdot \sin i;$$

$$L_{21} = \sin \Omega \cdot \cos \omega - \cos \Omega \cdot \sin \omega \cdot \cos i; \quad L_{33} = \cos i.$$

$$L_{22} = -\sin \Omega \cdot \sin \omega - \cos \Omega \cdot \cos \omega \cdot \cos i.$$

Перейдём в связанную СК по формуле  $\bar{S}_2 = D \cdot \bar{S}_1$ , где D – матрица перехода от орбитальной к связанной СК (рис. 3.10).

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix}, \quad \begin{array}{lll} d_{11} = \cos \vartheta; & d_{12} = -\sin \vartheta; & d_{13} = 0; \\ d_{21} = \sin \vartheta; & d_{22} = \cos \vartheta; & d_{23} = 0; \\ d_{31} = 0; & d_{32} = 0; & d_{33} = 1. \end{array}$$

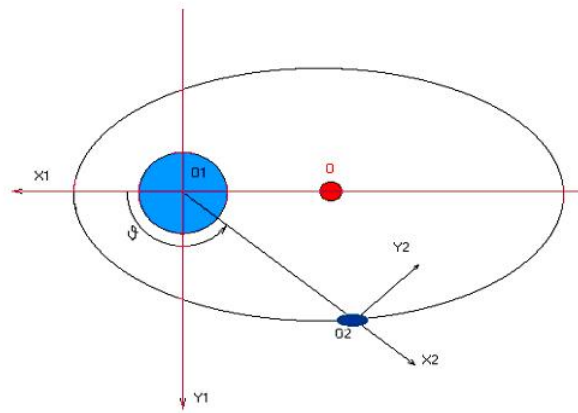


Рис. 3.10. Оси координат  $O_1 X_1 Y_1 Z_1$  и  $O_2 X_2 Y_2 Z_2$

Найдем проекции вектора  $\bar{n}$  на оси связанной СК:

$$n_{x2} = \cos \varphi \cdot \cos \sigma, \quad n_{y2} = \sin \varphi, \quad n_{z2} = -\sin \varphi \cdot \sin \sigma, \quad (3.5)$$

где  $\varphi$  – угол крена;  $\sigma$  – угол тангажа;  $n_{x2}$ ,  $n_{y2}$ ,  $n_{z2}$  – проекции вектора  $\bar{n}$  на оси связанной СК (рис. 3.11).

Максимальный энергосъём с панелей солнечных батарей возможен

при  $\cos \alpha = 1$ , следовательно, при  $\bar{S} = \bar{n}$  получим уравнения, описывающие оптимальные углы тангажа и крена:

$$\begin{cases} S_{x2} = \cos \varphi \cdot \cos \sigma \\ S_{y2} = \sin \varphi \\ S_{z2} = -\sin \varphi \cdot \sin \sigma \end{cases} \Rightarrow \sigma = \arcsin S_{y2}, \quad \varphi = \operatorname{arctg} \frac{-S_{z2}}{S_{x2}}. \quad (3.6)$$

### 3.1.3. Проектирование КСН оптимальной с точки зрения наилучшего показателя периодичности наблюдения

*Расчет показателя периодичности наблюдений.*

Для расчета показателя периодичности для  $i$ -го объекта наблюдения можно использовать максимальное время между двумя последовательными выходами КА на ОН при которых была возможность проведения съемки

$$П = \max (t_i - t_{i-1}) \Big|_{i=2..n} \quad (3.7)$$

где  $n$  - число выходов на ОН,  $t_i$  - время  $i$ -го выхода КА на ОН,  $t_{i-1}$  - время  $i-1$  выход КА на ОН.

При наличии нескольких объектов наблюдения, используют различные показатели периодичности. Одним из вариантов является использование стратегии гарантированного результата. В этом случае показателем периодичности является *максимальный, среди максимальных интервалов* каждого ОН.

$$П = \max_k \max_i (t_i - t_{i-1})_k, \quad k = 1..N, i = 2..n_k, \quad (3.8)$$

где  $n_k$  - число выходов КА на  $k$ -й ОН,  $N$  - число наблюдаемых объектов.

Второй вариант расчета показателя периодичности не обеспечивает гарантированный результат, но дает более осредненную характеристику. На практике этот вариант может оказаться более востребованным и целесообразным и называется *«средний из максимальных»*.

$$П = \frac{1}{N} \sum_{k=1}^N \max_i (t_i - t_{i-1})_k, \quad i = 2..n_k. \quad (3.9)$$

Как было предложено в [23], в качестве показателя периодичности может быть предложена величина *средней периодичности*

$$P = \frac{1}{N} \sum_{k=1}^N \left[ \frac{1}{n_k - 1} \sum_{i=2}^{n_k} (t_i - t_{i-1}) \right]. \quad (3.10)$$

Традиционно для построения модели КСН выбираются круговые солнечносинхронные, изомаршрутные орбиты, и спутники равномерно разносятся в плоскости орбиты со сдвигом по аргументу широты. Это хотя и упрощает задачу моделирования и оценки параметров периодичности и оперативности КСН, но не позволяет учитывать все варианты построения КСН. Использование только круговых орбит осложняется энергоемкостью их поддержания, поскольку из-за нецентральности поля тяготения Земли и действия атмосферы на низких орбитах параметры орбиты КА непрерывно меняются.

Этих недостатков можно избежать, если решать задачу *выбора основных параметров КСН* методами прямого математического программирования.

Оптимизационная задача построения орбитальной структуры КСН сводится к общей задаче нелинейного программирования, которая в данном случае формулируется следующим образом:

$$X^* = \operatorname{arg\,min} f(X), \quad (3.11)$$

при ограничениях на пространственное разрешение

$$g_1(X) = R_n - R \geq 0, \quad (3.12)$$

и параметры рабочей орбиты

$$g_2(X) = H_n - H_{\min} \geq 0,$$

$$g_3(X) = e_n - e_{\min} \geq 0,$$

$$g_4(X) = e_{\max} - e_n \geq 0,$$

где  $X = \{\Omega_n, i_n, p_n, \omega_n, e_n, \nu_n\}$ ,  $n = 1..N$  - вектор параметров орбиты  $n$ -го КА,  $N$  - число КА в КСН,  $f(X)$  - выбранный показатель.

При решении задачи приняты следующие допущения:  $R_n \geq R$ , где  $R_n$  - пространственное разрешение при съемке,  $R$  - допустимое разрешение;  $0,0001 \leq e \leq 0,9999$ ; высота орбиты  $H \geq 250$  км.

Разработанная модель расчета периодичности наблюдений реализована в универсальной целевой функции, обеспечивающей поиск экстремума по трем различным показателям периодичности. Учет ограничений при решении задачи оптимизации осуществляется с помощью штрафных функций.



Рис. 3.12. Блок-схема алгоритма формирования начальных условий баллистической структуры КСН

Следует отметить, что при решении сложных многопараметрических оптимизационных задач нелинейного математического программирования, к которым относится синтез орбитальной структуры КСН, большое значение имеет правильное формирование вектора исходных данных. Для определения начальных параметров баллистической структуры КСН используется следующий алгоритм (рис. 3.12).

### 3.1.4. Пример решения задачи синтеза орбитальной структуры

С помощью разработанного программного комплекса произведен синтез орбитальной структуры КСН, состоящий из модернизированного КА класса «Ресурс», для различных исходных данных:

1. КСН глобального мониторинга земной поверхности.
2. Наблюдение за крупнейшими городами мира.
3. Наблюдение за группой объектов.
4. Наблюдение за точечным объектом.

При моделировании задавались значения параметров орбитальной структуры. Для синтеза параметров КСН использовался специально разработанный алгоритм. Все результаты расчетов сведены в табл. 3.2. В качестве результатов программа также строила трассы КА (рис. 3.13).



Таблица 3.2. Результаты синтеза баллистической структуры КСН

Задача	N	Периодичность			Параметры орбитальной структуры					
		Макс., час	Среднее от макс. час	Среднее значение час	$\Omega$ , град	$i$ , град	$p$ , км	$\omega$ , град	$e$	$\vartheta$ , град
Равномерно распр. поле 91 объектов $\varphi \in [-60;60]$	1	74,11	69,01	30,65	11,04	108,6	6697,75	0,73	0,011	-0,74
	2	30,22	30,18	14,98	12,22	107,88	6697,62	-0,23	0,003	0,53
					89,23	110,44	6691,76	4,87	0,009	180,5
Группа объектов. Положение центра: 60с.ш. 10в.д. Радиус области 1000 км.	1	58,52	53,78	27,59	18,02	98,08	6704,44	-5,89	0,016	5,47
	2	24,69	24,67	13,95	22,59	104,95	6681,12	-4,85	0,001	9,29
					97,19	105,27	6681,39	4,66	0,001	188,0
	3	17,25	17,25	8,85	-1,17	103,27	6691,81	-0,6	0,003	6,41
					98,79	94,99	6711,22	-10,87	0,026	184,5
				57,98	100,09	6703,89	0,4	0,001	87,72	
Точечный объект. Положение: 60с.ш. 10в.д.	1	34,44	34,44	25,26	10,42	109,39	6688,01	-8,75	0,04	12,75
	2	17,57	17,57	9,79	7,39	117,21	6706,44	-1,05	0,02	5,23
					120,5	107,54	6696,24	-31,65	0,062	183,3
	3	14,8	14,8	7,38	6,06	116,09	6707,51	1,11	0,012	5,92
					117,9	108,05	6695,23	-31,48	0,064	183,9
				64,83	109,79	6699,15	-2,41	0,003	-0,38	
208 крупных городов	1	59,14	58,68	30,48	17,85	109,57	6704,13	7,03	0,012	56,88
	2	28,01	24,11	14,31	13,22	107,85	6707,62	-0,01	0,003	0,53
					91,36	108,24	6701,76	5,65	0,001	182,1
Традиционный подход к синтезу орбитальной структуры КСН										
208 крупных городов	2	30,21	25,63	15,37	13,22	107,85	6707,62	-0,01	0,003	0,53
					103,2	107,85	6707,62	-0,01	0,003	0,53
Улучшение периодичности		<b>0,64</b>	<b>1,29</b>	<b>1,06</b>						

Традиционный подход к синтезу орбитальной структуры КСН предусматривает равномерное разведение орбит по долготе восходящего узла. К примеру, если воспользоваться оптимальной орбитальной структурой, синтезированной для одного КА из группировки, потом развести орбиты на 90 градусов по долготе восходящего узла, то для задачи наблюдения за 208 городами группировкой из двух спутников получены результаты: максимальный показатель периодичности 30,21 час.; средний от максимальных 25,63 час.; средний 15,37 час.

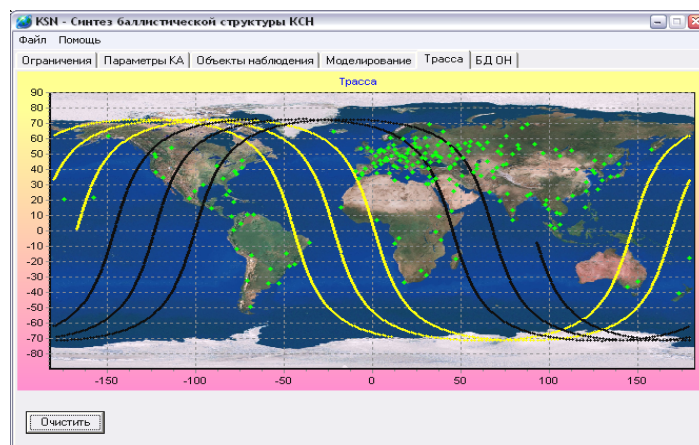


Рис. 3.13 Пример построения трассы

Если для этой задачи сравнить результаты расчетов, которые получены с использованием специального подхода на основе разработанных методов, с результатами традиционного подхода, то получаем: улучшение по показателю максимальной периодичности на 0,64 часа; по показателю среднего от максимальных на 1,29 часа; по показателю средней периодичности на 1,06 часа.

Используя табл. 3.2, можно выбрать параметры орбитальной структуры в зависимости от интересующей потребителя задачи.

## **3.2. Анализ эффективности применения электроракетной двигательной установки для коррекции низких орбит искусственных спутников Земли**

### **3.2.1. Коррекция низких орбит космических аппаратов электроракетным двигателем**

Основной задачей КСН за земной поверхностью является распознавание объектов различных размеров  $h$  с периодичностью  $t_{пер}$  от нескольких часов до нескольких лет.

Желательным вариантом съемки наземных объектов с высоким разрешением является размещение КА на низких орбитах, высотой порядка 300...500 км. Проблемой для всех низких орбит является значительное торможение КА, обусловленное атмосферой Земли, что вызывает

достаточно быстрое снижение высоты орбиты КА. Использование для коррекции орбиты обычной химической двигательной установки (ХДУ) приводит к большие затраты топлива. Альтернативным решением этой проблемы является использование электрических ракетных двигателей (ЭРД) для компенсации аэродинамического сопротивления.

Высокая скорость истечения, обеспечиваемая ЭРД (10000...25000 м/с), позволяет снизить запас топлива для поддержания высоты орбиты и увеличить срок службы КА. Срок существования зависит от количества коррекций, их средней продолжительности, запаса топлива, которые определяются типом ракетного двигателя.

Повышение удельного импульса ЭРД при той же массе топлива позволяет увеличить срок службы КА за счет существенно более низкого расхода топлива. Однако, чем выше удельный импульс, тем больше мощность двигателя и тем более мощная энергоустановка требуется. Следствием этого является возрастание площади солнечных батарей (СБ). В условиях сопротивления атмосферы увеличение площади солнечных батарей приводит к необходимости более частых и продолжительных коррекций.

Указанные противоречия требуют проведения анализа возможности использования ЭРД для поддержания более низких орбит в течение длительного срока службы орбитального комплекса.

Можно выделить область применения ЭРД для поддержания низких орбит. Математически она описывается следующим неравенством

$$\int_0^{T_{сущ}} \frac{P}{M_{КА}} \cdot \alpha \cdot dt \geq \int_0^{T_{сущ}} \sigma_{ср} \cdot \rho \cdot V^2 \cdot dt, \quad (3.13)$$

где  $T_{сущ}$  - время существования КА,  $P$  - тяга ЭРД,  $M_{КА}$  - масса космического аппарата,  $\sigma_{ср} = \frac{C_x \cdot S}{2 \cdot M_{КА}}$  - его средний баллистический коэффициент,  $\rho$  - плотность атмосферы,  $V$  - скорость полета КА,  $\alpha$  - относительное время работы ЭРД на витке ( $\alpha = \frac{T_M}{T_{об}}$ ,  $\alpha \in [0,1]$ , где  $T_M$  - моторное время,  $T_{об}$  - период обращения). Левая часть неравенства

отражает интегральные управляющие воздействия, правая часть – интегральные возмущающие воздействия. Видно, что в неравенстве (3.13) присутствуют проектные параметры, характеризующие облик КА ( $M_{КА}, \sigma_{ср}, P$ ) и параметры, отражающие факторы воздействия внешней среды ( $\rho = \rho(t), V(t)$ ) и параметр, задающий циклограмму коррекций  $\alpha$ .

Область потребных значений корректирующего ускорения от ЭРДУ зависит от баллистического коэффициента и высоты полета (рис. 3.14.). Следует отметить, что плотность атмосферы неоднозначно связывается с высотой полета, а определяется совокупностью факторов, главным из которых является уровень солнечной активности  $F_{10.7}$ .

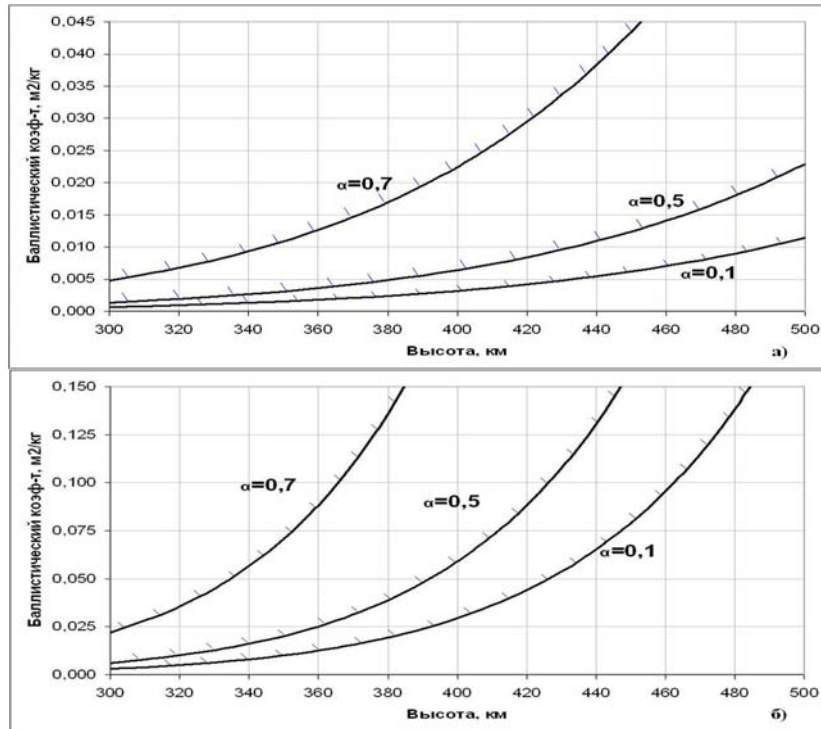


Рис 3.14. Область применения корректирующей ЭРДУ на низкоорбитальных космических аппаратах при уровнях солнечной активности  $F_0=250$  Вт/м<sup>2</sup>Гц и  $F_0=75$  Вт/м<sup>2</sup>Гц

### 3.2.2. Постановка задачи многокритериального синтеза

В указанной области применения ЭРДУ задачей совместной оптимизации будем называть задачу отыскания проектных параметров  $\bar{p} \in P$  и функций управления  $(\bar{u}(t, x, z), \bar{x}(t, z))$  из множества допустимых  $D = U(p, z) \times X(p, z)$ , обеспечивающих реализацию диапазона динамических маневров  $Z$  при максимальном (минимальном) значении заданного критерия эффективности  $\mu$ :

$$\bar{\mu} = \max_{u(t,x) \in U(p), p \in P} \mu(z, p, x(t), u(t, x)) \quad (3.14)$$

Проектный облик КА характеризуется набором функциональных модулей, образующих конструктивно-компоновочную схему. В рамках выбранной схемы задается вектор проектных параметров  $\bar{p} = (p_1, p_2, \dots, p_l)^T \in P$ . Такими параметрами являются: а) параметры, определяющие массовые характеристики КА, б) параметры, определяющие геометрические характеристики КА, в) параметры, определяющие энергетические характеристики КА:

$$\bar{p} = (p_D, p_{\Pi})^T,$$

где  $\bar{p}_D = (p_D^1, p_D^2, \dots, p_D^k)^T$  - вектор проектных параметров, непосредственно связанных с динамическими процессами,  $\bar{p}_{\Pi} = (p_{\Pi}^1, p_{\Pi}^2, \dots, p_{\Pi}^j)$  - вектор проектных параметров, изменение которых не оказывает прямого влияния на динамику движения,  $k + j = l$ . Минимальный вектор проектных параметров КА содержит: баллистический коэффициент  $\sigma_{КА}$ ,  $F_T$  - сила тяги,  $N_{ЭРД}$  - потребляемая мощность,  $n$  - количество ЭРД в ЭРДУ (целое число).

В основу математической модели движения положено уравнение возмущённого движения ИСЗ в векторной форме, которое имеет вид:

$$\frac{d\bar{V}}{dt} = -\mu \cdot \frac{\bar{r}}{r^3} + \bar{f}_B, \quad (3.15)$$

где  $\bar{r}$  - радиус-вектор (соединяет центр Земли и центр масс КА),  $\mu$  - гравитационный параметр Земли,  $\bar{f}_B$  - равнодействующая всех возмущающих ускорений. Выберем в качестве математической модели

движения систему уравнений в оскулирующей системе координат, тогда рабочая орбита ИСЗ описывается параметрами  $\{A, e, i, \omega, \Omega\}$ .

Для поддержания рабочей орбиты ИСЗ можно использовать различные альтернативы типов корректирующих двигательных установок (КДУ) или их комбинации:

- жидкостные реактивные двигатели (ЖРД);
- использование ЖРД совместно с ЭРДУ;
- использование ЭРДУ.

Наиболее эффективные варианты будем отбирать согласно следующим критериям предпочтения:

- $T_{\text{сущ}} \rightarrow \max$  (срок активного существования КА);
- $M_{\text{рт}} \rightarrow \min$  (масса рабочего тела, необходимого для коррекции).

При этом необходимо контролировать выполнение ограничений:

- $N_{\text{эрдю}} \leq N_{\text{доп}}$  (потребляемая ЭРДУ мощность не превышает допустимой величины);
- $M_{\text{эрдю}} \leq M_{\text{доп}}$  (масса ЭРДУ не более допустимой величины).

Таким образом, задача выбора проектных характеристик ЭРДУ и циклограмм коррекции орбиты формулируется следующим образом:

Выбрать основные проектные характеристики и баллистические схемы коррекции орбит КА с ЭРДУ, получить гарантирующие оценки затрат рабочего тела на поддержание орбит обеспечивающие экстремум критериев предпочтения при выполнении ограничений. Для решения этой задачи требуется создать методы выбора основных проектных характеристик энергодвигательного модуля, создать программное обеспечение для решения задачи и показать на примере реальных КА эффективность предложенных методов и моделей.

### ***3.2.3. Процедура совместной оптимизации законов управления и проектных параметров КА***

Процедура решения задачи оптимизации в описанной постановке сводится к следующему.

1. Формируется общая задача совместной оптимизации космического аппарата, предназначенного для множества динамических

маневров  $Z$ . Вводится проектная модель КА, а также модель распределения масс по отдельным компонентам КА.

2. Из общей задачи оптимизации выделяется её динамическая часть (выбор режимов управления движением). Динамическая задача решается последовательно, с применением моделей, обладающих различной степенью полноты, детализации и точности.

3. Решается задача оптимизации циклограмм коррекций совместно с синтезом проектных параметров. В результате выбираются в первом приближении основные параметры, характеризующие проектный облик КА, а также энергетика маневра и режимы управления.

4. Исследуется влияние неопределенных факторов, как на результаты решения динамической задачи, так и на значения проектных параметров. В случае необходимости определяются параметры, обеспечивающие минимум максимального проигрыша в критерии оптимальности, то есть реализуется гарантирующий подход.

Проектную модель КА с двигателями малой тяги будем строить на основе уравнения баланса масс его компонентов и систем на стартовой орбите. Анализ работ в области оптимизации КА с ЭРД малой тяги позволяет записать это уравнение в следующем виде:

$$m_0 = m_{\Pi} + m_{\text{Э}} + m_{\text{Д}} + m_{\text{РТ}} + m_{\text{СПХ}} + m_{\text{К}}, \quad (3.16)$$

где  $m_0$  - масса полезного груза;  $m_{\text{Э}}$  - масса энергоустановки, состоящая из источника и преобразователя энергии;  $m_{\text{Д}}$  - масса двигательной установки;  $m_{\text{РТ}}$  - масса рабочего тела, необходимая для управления КА;  $m_{\text{СПХ}}$  - масса систем подачи и хранения рабочего тела;  $m_{\text{К}}$  - масса конструктивных элементов.

Наиболее простыми и часто используемыми зависимостями массы отдельных компонентов КА от проектных параметров являются линейные зависимости масс от номинальной мощности энергоустановки  $N$  и тяги двигателей на стартовой орбите  $P$ :

$$m_{\text{Э}} = \alpha_{\text{Э}} \cdot N, \quad m_{\text{Д}} = \gamma_{\text{Д}} \cdot P, \quad m_{\text{К}} = \alpha_{\text{К}} \cdot N + \gamma_{\text{К}} \cdot P, \quad m_{\text{СПХ}} = k_{\text{СПХ}} \cdot m_{\text{РТ}}, \quad (3.17)$$

где  $\alpha_{\text{Э}}$  - удельная массовая характеристика энергоустановки (кг/Вт);  $\alpha_{\text{К}}$  - удельная масса конструкции по мощности (кг/Вт);  $\gamma_{\text{Д}}$  - удельная масса

двигательной установки (кг/Н);  $k_{СПХ}$  - отношение массы системы подачи и хранения рабочего тела к массе рабочего тела.

Реальная зависимость стартовой массы от проектных параметров аппарата имеет более сложный нелинейный вид и уравнение баланса масс значительно усложняется. В любом случае можно говорить о том, что стартовая масса аппарата является функцией тяги двигателей, скорости истечения рабочего тела, массы рабочего тела, требуемого на выполнение целевой задачи и ограничений, наложенных на фазовые координаты, законы управления и проектные параметры КА.

### 3.2.4. Учет неопределенности возмущающих факторов

В уравнениях движения искусственного спутника Земли воздействие внешних факторов выражается вектором возмущающих и управляющих ускорений  $\vec{f}_B$ .

$$\vec{f}_B = \vec{f}_g + \vec{f}_a + \vec{f}_m + \vec{f}_{упр} \quad (3.18)$$

где  $\vec{f}_g$  - вектор гравитационных возмущающих ускорений,  $\vec{f}_a$  - вектор аэродинамических возмущающих ускорений,  $\vec{f}_m$  - вектор электромагнитных возмущающих ускорений,  $\vec{f}_{упр}$  - вектор управляющих ускорений.

Если движение КА описывать в оскулирующих элементах уравнениями (3.1), то возмущающие и управляющие ускорения удобно проектировать на оси орбитальной системы координат:

$$\begin{aligned} S &= S_g + S_a + S_m + a_s \\ \vec{f}_B &= (S, T, W)^T, \quad T = T_g + T_a + T_m + a_T, \\ W &= W_g + W_a + W_m + a_T \end{aligned} \quad (3.19)$$

где  $S_g, T_g, W_g$  - проекции гравитационного возмущающего ускорения;  $S_a, T_a, W_a$  - проекции аэродинамического возмущающего ускорения;  $S_m, T_m, W_m$  - проекции электромагнитного возмущающего ускорения;  $a_T, a_T, a_T$  - проекции ускорения от тяги двигателя.



Проекция гравитационных возмущений от нецентральности поля сил Земного тяготения с учетом второй зональной гармоники находятся по формулам:

$$\begin{aligned}
 S_g &= 1,5 \cdot \frac{\mu}{r^2} \cdot J_2 \cdot \left(\frac{R_3}{r}\right)^2 \cdot (3 \cdot \sin^2 i \cdot \sin^2 u - 1); \\
 T_g &= -3 \cdot \frac{\mu}{r^2} \cdot J_2 \cdot \left(\frac{R_3}{r}\right)^2 \cdot \sin^2 i \cdot \sin u \cdot \cos u; \\
 W_g &= -3 \cdot \frac{\mu}{r^2} \cdot J_2 \cdot \left(\frac{R_3}{r}\right)^2 \cdot \sin i \cdot \cos i \cdot \sin u; \quad r = \frac{P}{R},
 \end{aligned}
 \tag{3.20}$$

где  $R_3$  - экваториальный радиус Земли;  $J_2$  - вторая гармоника.

Величину возмущающего аэродинамического ускорения в текущий момент времени можно определить по формуле:

$$\bar{a}_{амм} = \sigma \cdot \rho \cdot V \cdot \bar{V}, \tag{3.21}$$

где  $\sigma$  - баллистический коэффициент КА,  $\rho$  - плотность атмосферы,  $\bar{V}$  - скорость КА относительно потока атмосферы. Если считать, что атмосфера вращается вместе с Землей, проекции возмущающего аэродинамического ускорения на оси орбитальной системы координат имеют вид:

$$\begin{aligned}
 S_a &= -\sigma \cdot \rho(h) \cdot V \cdot V_r, \\
 T_a &= -\sigma \cdot \rho(h) \cdot V \cdot (V_T - \omega_3 \cdot r \cdot \cos i), \\
 W_a &= -\sigma \cdot \rho(h) \cdot V \cdot \omega_3 \cdot r \cdot \cos i \cdot \cos u,
 \end{aligned}
 \tag{3.22}$$

где  $V$  - модуль вектора скорости КА;  $V_r$  - радиальная составляющая вектора скорости;  $V_T$  - тангенциальная составляющая вектора скорости;  $\omega_3$  - угловая скорость вращения Земли вокруг своей оси,  $\sigma$  - баллистический коэффициент КА.

Особенностью движения искусственного спутника Земли (ИСЗ) на низких орбитах является неопределенность возмущающей силы аэродинамического сопротивления, обусловленная неточностью знания баллистического коэффициента спутника и плотности атмосферы. Для оценки возмущающего ускорения используется изотермическая модель плотности, где высота однородной атмосферы  $H$  изменяется в зависимости от уровня солнечной активности, а баллистический коэффициент в зависимости от текущего аргумента широты  $u$ .

Чтобы охарактеризовать зависимость среднего уровня плотности атмосферы от высоты можно использовать статическую модель плотности (рис. 3.15):

$$\rho_M = a_0 \cdot \exp \left[ a_1 - a_2 \cdot (H - a_3)^{\frac{1}{2}} \right], \quad (3.23)$$

где  $\rho_M$  - модельная плотность атмосферы (средний уровень),  $a_1, a_2, a_3, a_4$  - коэффициенты модели, используемые для расчета плотности атмосферы при различных значениях  $F_0$  (табулированные значения),  $F_0$  - фиксированное значение индекса солнечной активности  $F_{10.7}$  за рассматриваемый период времени,  $F_{10.7}$  - индекс солнечной активности равный плотности потока радиоизлучения на длине волны 10.7 см (частота 2800 МГц).

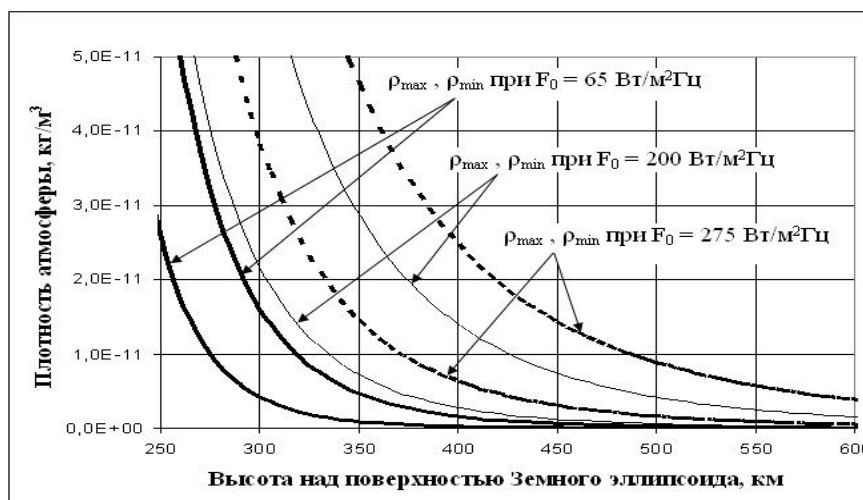


Рис. 3.15. Зависимость крайних состояний плотности атмосферы от высоты при различных уровнях солнечной активности по ГОСТ 25645.101-83

Плотность может существенно изменяться за сравнительно короткое время под действием внешних факторов. Как правило, состояние атмосферы характеризуют крайними уровнями плотности атмосферы на текущей высоте при различных уровнях солнечной активности. Величина плотности изменяется в указанных пределах, в течение исследуемого интервала времени. Если рассматривать более широкий диапазон времени,

то также следует учитывать полугодовые эффекты, геомагнитные возмущения и т.п.

В зависимости от плотности атмосферы, колеблется и средняя за виток величина реактивного ускорения  $a_{cp}$ , компенсирующего влияние аэродинамических возмущений, определяемая путем взятия интеграла возмущающего атмосферного ускорения на интервале, равном периоду обращения. Результаты расчета  $a_{cp}$  для  $\sigma = 0,002 \text{ м}^2/\text{кг}$  приведены в табл. 3.3. Видно, что разброс уровня компенсирующего ускорения достаточно велик.

Таблица 3.3. Средняя величина «компенсирующего» реактивного ускорения

Параметры орбиты		Величина «компенсирующего ускорения $a_{cp} \cdot 10^5$ , м/с <sup>2</sup>	
высота перигея $H_{\pi}$ , км	высота апогея, $H_{\alpha}$ , км	Минимальный уровень солнечной активности $F_0 = 75 \cdot 10^{-22} \frac{\text{Вт}}{\text{м}^2\text{Гц}}$	Максимальный уровень солнечной активности $F_0 = 250 \cdot 10^{-22} \frac{\text{Вт}}{\text{м}^2\text{Гц}}$
250	350	0.68	1.556
350	350	0.035	0.325

В качестве модели, характеризующей текущее состояние атмосферы, можно использовать так называемую “динамическую” модель плотности атмосферы:

$$\rho = \rho_M \cdot k_1 \cdot k_2 \cdot k_3 \cdot k_4 + \delta\rho, \quad (3.24)$$

где  $\rho_M$  - плотность ночной атмосферы;  $k_1$  - коэффициент, учитывающий влияние суточного эффекта;  $k_2$  - поправочный коэффициент, учитывающий полугодовой эффект,  $k_3$  - коэффициент, учитывающий отклонение среднесуточного индекса солнечной активности от его среднего значения за период;  $k_4$  - коэффициент, учитывающий корреляцию между плотностью атмосферы и геомагнитной возмущенностью;  $\delta\rho$  - случайные флуктуации плотности.

Данную модель плотности можно реализовать при имитационном моделировании возмущенного движения ИСЗ, в том числе и процесса

коррекции высоты орбиты, основанном на решении системы уравнений в огибающих элементах методом их численного интегрирования на ЭВМ.

В случае моделирования процессов движения КА по круговой орбите с наклоном, отличным от нуля, следует учитывать, что высота над поверхностью Земли не постоянна, а изменяется в зависимости от аргумента широты. С учётом сжатия земного эллипсоида высота КА над поверхностью Земли определяется выражением:

$$H = r - R_3 \cdot (1 - \varepsilon \cdot \sin^2 i \cdot \sin^2 u), \quad (3.25)$$

где  $r = H + R_3$  - текущий радиус орбиты,  $R_3 = 6371$  км - средний радиус земного эллипсоида,  $R_3 = 6378,245$  км - радиус экватора,  $\varepsilon \approx 0,0034$  - коэффициент сжатия земного эллипсоида,  $u$  - аргумент широты.

На рисунке 3.16 показано изменение высоты круговой орбиты высотой  $H = 350$  км в зависимости от аргумента широты КА в плоскости орбиты.

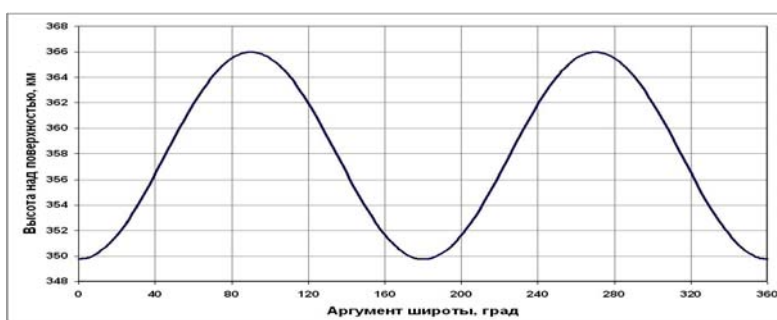


Рис. 3.16. Зависимость высоты орбиты КА от аргумента широты при наклоне орбиты  $i = 60^\circ$

Видно, что высота КА за виток изменяется, а, следовательно, изменяется плотность атмосферы на протяжении витка в следствии неферичности Земли.

### 3.2.5. Модель коррекции орбиты спутника Земли без учета ограничений по энергетике

Задачу поддержания орбиты формулируется как задача ликвидации вековых возмущений периода обращения  $T$ , возникающих из-за действия силы аэродинамического сопротивления. Потребуем, чтобы на отрезке

$[0, t_k]$  период обращения находился в диапазоне  $T_p + \delta T_{дон}$ , где  $T_p$  - расчетное значение периода обращения;  $T_{дон}$  - допустимое отклонение.

Весь интервал управления разобьем на  $N$  равных подынтервалов, каждый из которых состоит из  $m$  пассивных и  $n$  активных витков. На пассивных витках двигатель выключен, а на активных витках ЭРД создает постоянное по величине реактивное ускорение  $a_k$ . В результате за  $m+n$  витков период восстанавливается. Отметим, что на активных витках необходимо не только компенсировать влияние силы аэродинамического сопротивления, но и ликвидировать ошибки периода обращения, накопившиеся на пассивных витках (рис. 3.17).

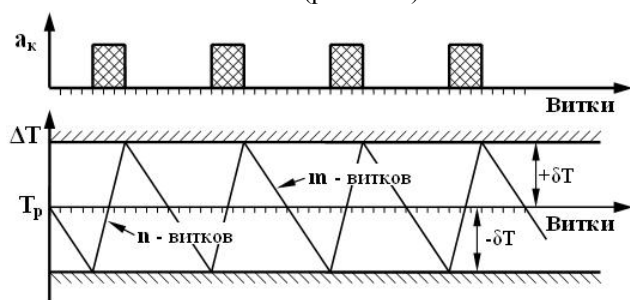


Рис. 3.17. Циклограмма включения ЭРД при поддержании периода обращения

Изменение периода обращения  $T$  на одном витке пропорционально величине баллистического коэффициента и зависит от плотности атмосферы и параметров орбиты. Для оценки падения периода обращения на одном витке  $\delta T(h_\pi, \rho)$  можно воспользоваться аналитическими аппроксимационными формулами. На рис. 3.18 приведены «предельные» зависимости:  $\delta T_n(h_\pi)$  – соответствует минимальной плотности атмосферы ( $\delta T_n(h_\pi) = \delta T(h_\pi, \rho_n)$ ),  $\delta T_g(h_\pi)$  – соответствует максимальной плотности атмосферы ( $\delta T_g(h_\pi) = \delta T(h_\pi, \rho_g)$ ).

Данные, приведенные на рис. 3.18, получены путем обработки результатов численного моделирования движения ИСЗ с использованием модели плотности по ГОСТ22721-77 и соответствуют баллистическому коэффициенту  $\sigma_x = 0,002 \text{ м}^2/\text{кг}$ . Аппроксимационные формулы имеют вид:  $\delta T_g(h_\pi) = \exp(\alpha_1 + \beta_1 \cdot h_\pi)$ ;  $\delta T_n(h_\pi) = \exp(\alpha_2 + \beta_2 \cdot h_\pi)$ , что соответствует

минимальной и максимальной плотности атмосферы,  $\alpha_1 = 2,34$ ,  $\alpha_2 = 3,53$ ,  $\beta_1 = -1,8125 \cdot 10^{-2} \text{ км}^{-1}$ ,  $\beta_2 = -3,1375 \cdot 10^{-2} \text{ км}^{-1}$  [25].

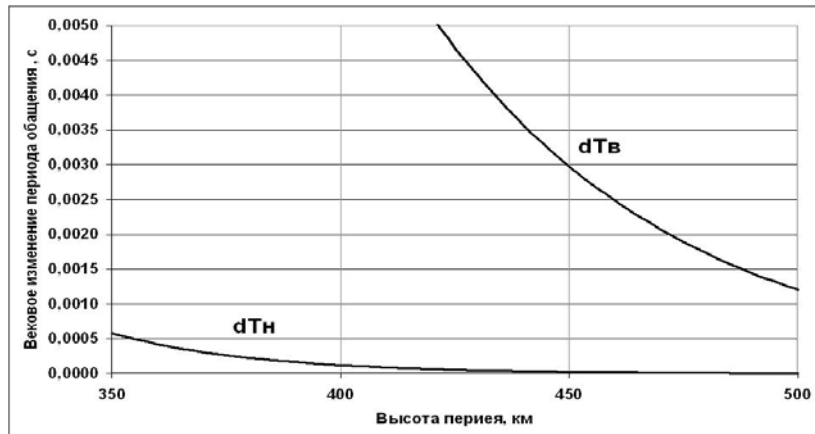


Рис. 3.18. Верхние и нижние границы изменения периода обращения на одном витке под действием атмосферы ( $\sigma_x = 0.002 \text{ м}^2/\text{кг}$ )

Для оценки изменения периода обращения за цикл коррекции ( $n$  активных витков) под действием постоянного ускорения  $a_k$  можно воспользоваться выражениями работы:

$$\Delta T_k = \frac{12 \cdot \pi^2 \cdot A^3}{\mu} \cdot \sqrt{\frac{A \cdot (1 - e^2)}{\mu}} \cdot a_k \cdot n, \quad (3.26)$$

$$\delta T_k = \alpha \cdot \frac{\Delta T_k}{n} = \alpha \frac{12 \pi^2 A^3}{\mu} \sqrt{\frac{A(1 - e^2)}{\mu}} \cdot a_k, \quad (3.27)$$

где  $\delta T_k$  - изменение периода обращения КА за виток (без учёта воздействия аэродинамического торможения),  $\alpha$  - относительное время работы ЭРД на витке ( $\alpha \in [0..1]$ ).

Модель задачи оптимизации управления движением спутника на низкой орбите достаточно сложна, а эллиптичность орбиты требует аккуратного описания «медленной» эволюции орбиты на больших интервалах времени. В этих задачах на первый план выходит стратегия гарантированного результата, как при выборе законов управления, так и при оптимизации параметров корректирующей двигательной установки.

### 3.2.6. Влияние ограничений по энергетике на формирование циклограмм коррекции орбиты

Для создания необходимых корректирующих ускорений требуется включение ЭРДУ на определенных интервалах времени. При своем функционировании ЭРД потребляют электроэнергию, при этом потребляемая мощность для включения пропорциональна силе тяги двигателя и скорости истечения рабочего тела, то есть напрямую зависит от характеристик двигателя.

В общем виде математические модели потребляемой тяги и потребляемой мощности записываются следующим образом:

$$F_T = M_{КА} \cdot a_K, \quad N = \frac{F_T \cdot C}{2 \cdot \eta} \quad (3.28)$$

где  $F_T$  - потребляемая сила тяги,  $N$  - потребляемая мощность для включения ЭРДУ,  $M_{КА}$  - масса КА,  $C$  - скорость истечения рабочего тела,  $\eta$  - КПД ЭРДУ,  $a_K$  - потребляемое корректирующее ускорение.

Для коррекции орбиты КА с массой  $M_{КА} \approx 6400$  кг и средним баллистическим коэффициентом  $\sigma_{cp} = 0,0036$  м<sup>2</sup>/кг зависимости потребляемой тяги и потребляемой мощности приведены на рис. 3.19.

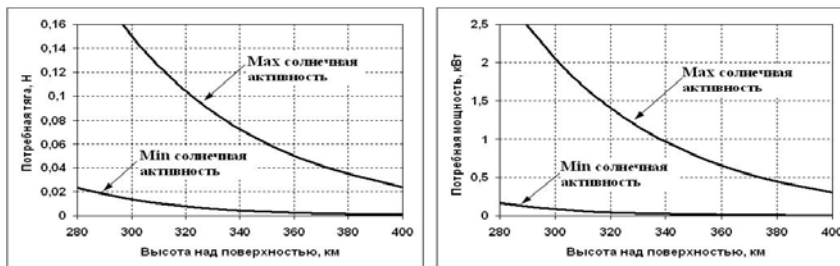


Рис. 3.19. Зависимость минимальной потребляемой тяги и потребляемой мощности ЭРДУ для осуществления коррекции от высоты перигея орбиты при крайних уровнях плотности атмосферы

По графикам видно, что для различных высот орбит в зависимости от уровня солнечной активности требуются различные уровни потребляемой тяги двигательной установки и потребляемой мощности энергетической установки. Причем эти значения могут отличаться на порядок (см. табл. 3.3).

При проектировании следует задавать сочетание характеристик питающей энергоустановки и ЭРДУ таким образом, чтобы располагаемой мощности было достаточно для работы полезной аппаратуры и включения ЭРДУ на заданном промежутке времени (рис. 3.20).



Рис. 3.20. Энергопотребление бортовых систем и ЭРДУ.

Обеспечение ЭРДУ необходимой мощностью возможно либо за счет изменения характеристик системы энергоснабжения КА (наращивание площади солнечных батарей, увеличение емкости аккумуляторных батарей), либо за счет резерва мощности для включения ЭРДУ. Резерв мощности можно обеспечивать за счет незначительного снижения интенсивности работы целевой аппаратуры.

Для КА с солнечной энергоустановкой предлагается следующая схема работы ЭРДУ на витке (рис. 3.21). Все энергопотребление бортовой аппаратуры (БА) за виток можно разделить на постоянное (для функционирования обслуживающих системы) и периодическое (для включений целевой аппаратуры (ЦА)). Включения ЦА производятся согласно алгоритму целевой работы, за исключением этих участков, на других возможно включение ЭРДУ для коррекции орбиты, при условии, что имеется необходимый резерв мощности.

Соответственно, если подключать к системе энергоснабжения КА электроракетные двигатели, потребляющие энергию разной мощности (разной тяги), то с увеличением мощности ЭРД будет уменьшаться время его включения. Также это время будет уменьшаться с повышением



интенсивности работы целевой аппаратуры.



Рис. 3.21. Схема работы ЭРДУ на витке

Дополнительную мощность для работы ЭРДУ можно найти за счёт уменьшения времени работы бортовой аппаратуры (отдельных её компонентов), затрачиваемой на съёмку объектов. Тогда ограничение по мощности выглядит следующим образом:

$$\frac{\sum N_{ЭРДУ} \cdot t_{\text{раб.ЭРДУнаВитке}}}{T_{\text{обр}}} = \alpha_{ЭРДУ} \cdot \sum N_{ЭРДУ}, \quad (3.29)$$

где  $\alpha_{ЭРДУ} = t_{\text{раб.ЭРДУнаВит}} / T$  - относительное время работы ЭРДУ за один виток;  $T$  - период обращения КА,  $t_{\text{раб.ЭРДУнаВитке}}$  - активное время работы ЭРДУ на одном витке.

При численном интегрировании с учетом ограничений по мощности систему уравнений (3.28) следует дополнить зависимостью располагаемой мощности от времени:

$$N_{\text{расп}} = f(t) \quad (3.30)$$

Для получения зависимости (3.30) КА с солнечной энергоустановкой требуется моделирование процесса энергопотребления бортовой аппаратуры. Функциональная схема солнечной энергоустановки показана на рисунке 3.22.

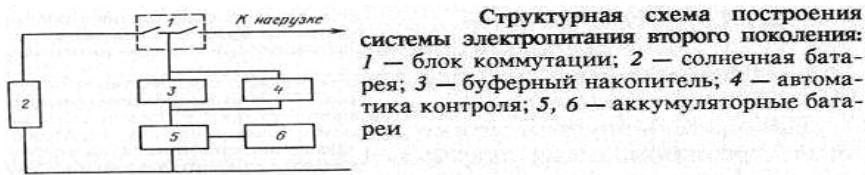


Рис. 3.22. Функциональная схема СЭУ

Следовательно, требуется использование математических моделей: освещенности КА, получения электроэнергии от солнечных батарей, циклов заряда-разряда аккумуляторных батарей КА.

### 3.2.7. Математическая модель коррекции орбиты низкоорбитального спутника Земли с учетом неопределенности плотности верхней атмосферы и ограничений по энергетике

При учете ограничений по энергетике величины корректирующего ускорения, баллистического коэффициента, оптимальных участков включения ЭРДУ, становятся зависимыми от проектных характеристик КА и ЭРДУ, то есть реализуется взаимосвязь проектных характеристик КА и законов его движения. Неопределенность плотности верхней атмосферы приводит к неоднозначности и появлению верхних и нижних границ области получаемых проектных параметров.

Структура режима управления орбитой на активном витке с учетом ограничений по энергетике показана на рис. 3.23.

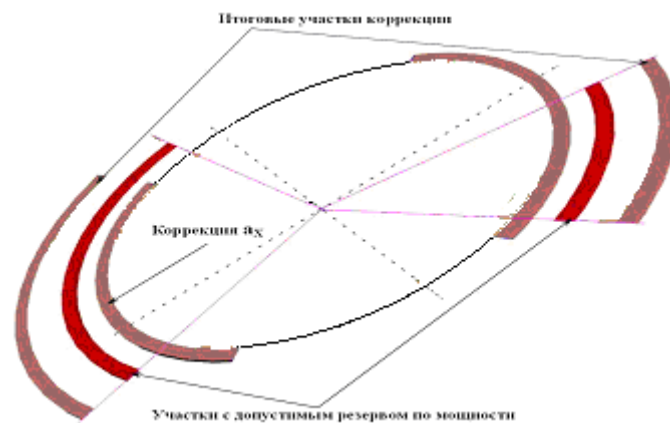


Рис. 3.23. Схема формирования циклограмм коррекции на активном витке с учетом ограничений по энергетике и неопределенности состояния плотности атмосферы

Математическая модель эволюции орбиты в поле земного сфероида, с учетом возмущающего влияния атмосферы и корректирующих включений ЭРДУ с оптимальной структурой управления, а также с учетом ограничений по мощности содержит уравнения

$$\frac{dA}{dt} = \frac{4a_x}{\pi} \sqrt{\frac{A^3}{\mu}} \left( \xi + \frac{\alpha - \pi}{2} \right) - 2\sigma\rho_{cp} \sqrt{\mu A} ,$$

$$\begin{aligned}
\frac{de}{dt} &= \frac{a_x}{\pi} \sqrt{\frac{A}{\mu}} \left[ -3e\left(\xi + \frac{\alpha - \pi}{2}\right) + 4\sin\left(\xi + \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\cos\eta - \right. \\
&\quad \left. - \frac{e}{2}\sin 2\left(\xi + \frac{\alpha}{2}\right)\cos\alpha\cos 2\eta \right] - 2e\sigma\rho_{ep}\sqrt{\frac{\mu}{A}}, \\
\frac{d\omega}{dt} &= \frac{a_x}{\pi \cdot e} \sqrt{\frac{A}{\mu}} \left[ 4\sin\left(\xi + \frac{\alpha}{2}\right)\cos\frac{\alpha}{2}\sin\eta - \frac{e}{2}\sin 2\left(\xi + \frac{\alpha}{2}\right)\cos\alpha\sin 2\eta \right] + \\
&\quad + \frac{\varepsilon(5\cos^2 i - 1)}{2\mu^{0,5}A^{3,5}}, \\
\frac{d\Delta u}{dt} &= \sqrt{\mu}(A^{-1,5} - A_K^{-1,5}),
\end{aligned} \tag{3.31}$$

в которых вместо средней плотности атмосферы используются ее текущие значения:

$$\begin{aligned}
\rho(t) &= K_0(t) \cdot K_1(t) \cdot K_2(t) \cdot K_3(t) \cdot a_0 \cdot \exp\left[ a_1 - a_2 \cdot ((r - 6371) - a_3)^{\frac{1}{2}} \right], \\
\rho(t) &\in [p_{\min}(t), p_{\max}(t)].
\end{aligned} \tag{3.32}$$

Текущее значение корректирующего ускорения рассчитывается согласно зависимостям:

$$\begin{aligned}
M_{KA} &= M_{II} + M_{OC} + M_{\mathcal{E}} + M_{\mathcal{E}PDY}, \\
a_K &= \begin{cases} \frac{\sum F_T}{M_{KA}}, & \delta T \geq \delta T_{дон} \\ 0, & \delta T < \delta T_{дон} \end{cases}, \\
\{\alpha, \xi\} &= f(N_{\mathcal{E}Y}), \quad T_{MOT} = f(\delta T), \quad M_{PT} = \frac{Q}{F_T} \cdot \sum T_{MOT}.
\end{aligned} \tag{3.33}$$

При численном моделировании процессов коррекции на основе описанных моделей можно проводить синтез проектных параметров ЭРДУ и алгоритмов управления для исследуемого КА ДЗЗ.

### 3.2.8. Определение гарантирующих затрат рабочего тела на поддержание орбиты

Для определения гарантирующих затрат рабочего тела (ксенона) на поддержание орбиты космического аппарата наблюдения и выбора

проектного варианта корректирующей электроракетной двигательной установки предлагается следующая методика:

- задаются параметры рабочей орбиты;
- задается множество масс КА на рабочей орбите  $M_{КА} = \{M_{\min}, \dots, M_{\max}\}$ ;
- задается множество значений среднего баллистического коэффициента КА  $\sigma_0 = \{\sigma_{0\min}, \dots, \sigma_{0\max}\}$ ;
- задается множество вариантов построения ЭРДУ (марка двигателя, количество одновременно работающих ЭРД);
- задаются характеристики системы энергопитания КА (среднесуточная мощность энергоустановки, средняя величина резерва мощности для ЭРДУ за сутки, площадь СБ);
- проводится моделирование процессов коррекции орбиты на множестве заданных проектных параметров при крайнеплотном состоянии атмосферы Земли с учетом ограничений по энергетике;
- по результатам моделирования определяются затраты рабочего тела на коррекцию орбиты за заданный период времени.

Применяя данную методику к расчету гарантирующих затрат рабочего тела для *американского спутника TacSat-2*, выведенному в 2006 году на орбиту с параметрами  $i = 40.01^\circ$ ,  $H_\alpha = 424.5$  км,  $H_\pi = 410$  км. Масса спутника около 400 кг. Энергопитание КА обеспечивают две солнечные батареи мощностью 550 Вт и литий-ионные аккумуляторы емкостью 30 А·ч. На TacSat-2 для компенсации снижения высоты орбиты из-за торможения в верхних слоях атмосферы используется ионный микродвигатель на основе эффекта Холла, разработанный лабораторией AFRL на основе электроракетного двигателя ВТН-200-ХЗ компании Busek. Удельный импульс двигателя – 1600 сек, тяга – 4...17 мН. Уровень энергопотребления в пределах 50...300 Вт при номинальной мощности электропитания 200 Вт. Масса двигателя диаметром 100 мм и длиной 105 мм – менее 1 кг. Предполагаемый срок активного существования КА «TacSat-2» составляет 5 лет. Характер эволюции высоты орбиты без проведения коррекции показан на рис. 3.24.

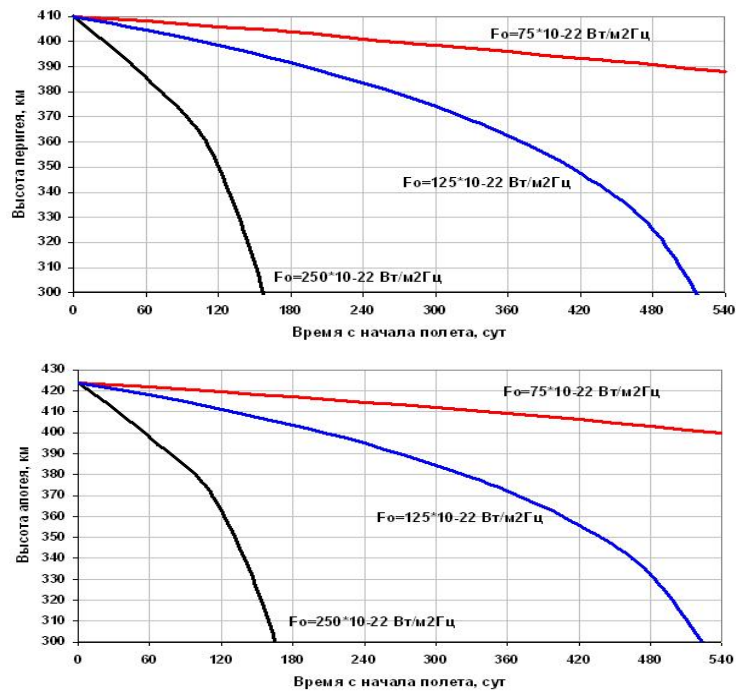


Рис. 3.24. Падение КА «TasSat-2» при различных состояниях плотности атмосферы

Из графиков видно, что высота перигея орбиты за год при минимальной солнечной активности уменьшается более чем на 10 км, при среднем состоянии атмосферы данная величина уже составляет около 50 км и при максимальной солнечной активности КА менее чем за год войдет в плотные слои. Следовательно, при расположении КА «TasSat-2» на данной орбите требуется производить периодическую коррекцию орбиты во время активного существования спутника.

Применим описанную методику для оценки возможности поддержания высоты орбиты с помощью ЭРДУ. На рис 3.25 – 3.27 показаны результаты расчетов в предположениях, что в среднем за виток резерв мощности составляет не более 50 Вт, коррекция производится только по трансверсали, одновременно может работать только один электрореактивный двигатель. Следует отметить, что при увеличении резерва мощности сокращается количество активных витков и время одной коррекции высоты орбиты.

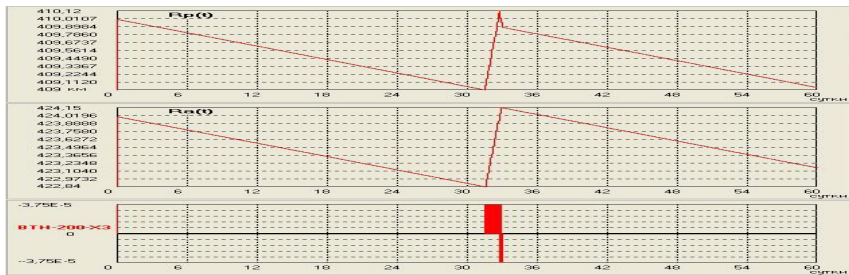


Рис. 3.25. Циклограмма коррекции высоты орбиты при минимальном состоянии плотности атмосферы ( $F_0=75 \cdot 10^{-22} \text{Вт/м}^2 \cdot \text{Гц}$ ) двигателем ВТН-200-Х3

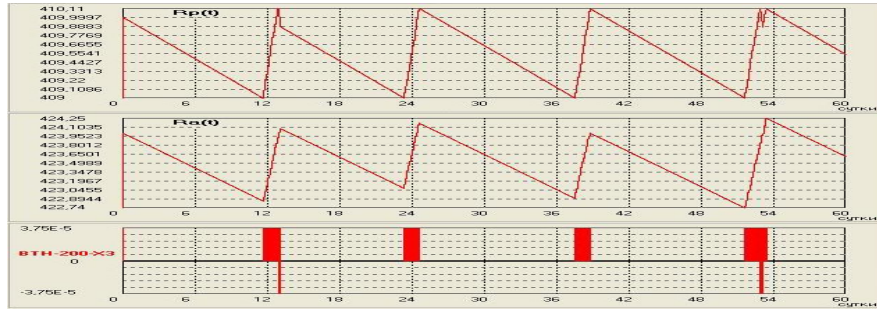


Рис. 3.26. Циклограмма коррекции высоты орбиты при среднем состоянии плотности атмосферы ( $F_0=125 \cdot 10^{-22} \text{Вт/м}^2 \cdot \text{Гц}$ ) двигателем ВТН-200-Х3

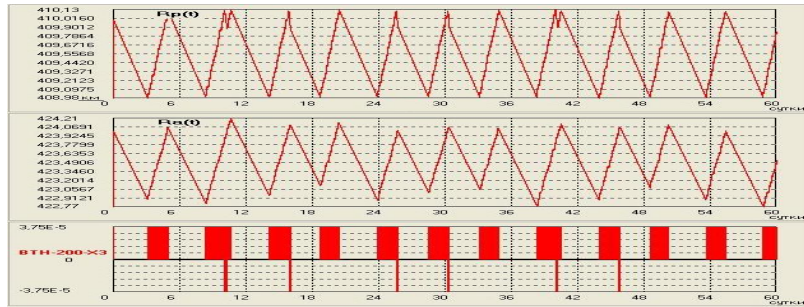


Рис. 3.27. Циклограмма коррекции высоты орбиты при максимальном состоянии плотности атмосферы ( $F_0=250 \cdot 10^{-22} \text{Вт/м}^2 \cdot \text{Гц}$ ) двигателем ВТН-200-Х3

Результаты моделирования позволяют построить области возможных сроков существования КА и соответствующих суммарных расходов рабочего тела (рис. 3.28).

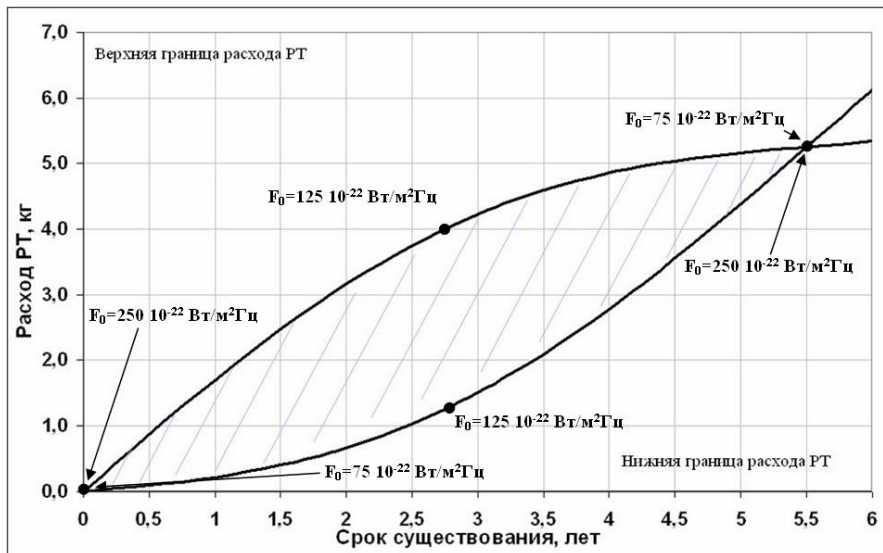
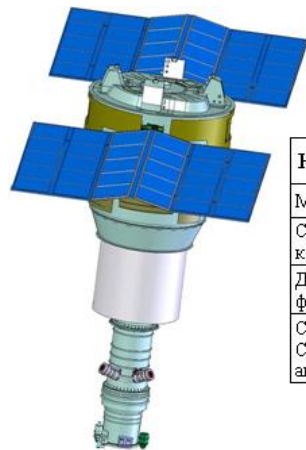


Рис. 3.28. Область возможных расходов рабочего тела и сроков существования КА ДЗЗ «TasSat-2» с двигателем ВТН-200-ХЗ

Анализируя данную область можно определить, что расход топлива за пять лет может изменяться от 4,5 до 5,2 кг, в зависимости от текущих условий.



Наименование параметра	Значение
Масса КА	≈6400 кг
Средний баллистический коэффициент	≈0,0035 м <sup>2</sup> /кг
Диапазон высот функционирования	300 ÷ 600 км
Среднесуточная мощность СЭП, выделяемая для целевой аппаратуры	≈400 Вт

Рис. 3.29. КА ДЗЗ класса «Ресурс»

Проведем оценку потребного для коррекции орбиты рабочего тела ЭРДУ для КА ДЗЗ класса «Ресурс», характеристики которого указаны на

рис. 3.29. Рассмотрим четыре марки электроракетных двигателей производства ОКБ «ФАКЕЛ», прошедших лётные испытания: СПД-50, СПД-60, СПД-70, СПД-100. Их характеристики показаны в табл. 3.4. Рассмотрим расчетный случай с околокруговыми орбитами функционирования с диапазоном высот 350...400 км. Диапазон коррекции высот  $dH$  подбирался таким образом, чтобы выполнялись условия  $M_{з.топ.ЭРДУ} \rightarrow \min$ ,  $\sum M_{ЭРДУ} \leq M_{доп}$ ,  $T_{сум} \rightarrow \max$ . Относительное время работы ЭРДУ на витке подбиралось таким образом, чтобы выполнялись условия:  $\sum N_{ЭРДУ} \leq N_{доп}$ ,  $P \rightarrow \max$ .

Таблица 3.4. Характеристики СПД производства ФГУП «ФАКЕЛ»

Двигатель	СПД-50	СПД-60	СПД-70	СПД-100
Тяга, мН	20	30	40	83
Удельный импульс, м/с	12500	12500	12500	16000
Мощность, кВт	0,35	0,517	0,593	1,221
Ресурс, ч	2250	2500	3100	7500
Масса, кг	0,8	1,2	1,5	3,5
Рабочее тело	Ксенон	Ксенон	Ксенон	Ксенон

При анализе результатов моделирования можно отметить, что наилучшие результаты дает использование для коррекции двигателя СПД-100, так как имеет наибольшую тягу. Но в большинстве случаев он не удовлетворяет ограничениям по энергетике, накладываемым энергетической установкой аппарата с учетом программы функционирования аппаратуры наблюдения. Поэтому для использования на аппарате класса «Ресурс», предпочтительнее использовать двигатели СПД-70. Оценку увеличения срока существования КА ДЗЗ класса «Ресурс» проведем при следующих условиях:

- ЭРДУ содержит два одновременно работающих СПД-70,
- большую полуось и эксцентриситет корректируем оптимальным образом,
- учитываем ограничения, накладываемые СЭП КА и программой работы аппаратуры наблюдения,
- КА рассчитан на три года активного существования.

Для определения потребного запаса ксенона, без учета использования ЖРД, проведем моделирование на интервале времени в



один год, для различных уровней солнечной активности, рабочих орбит с эксцентриситетом 0, 0.01, 0.183 и высотами перигея  $H_{\pi} \in [350;400]$  км. Результаты расчетов приведены в виде графиков на рис. 3.30 и в табл. 3.5.

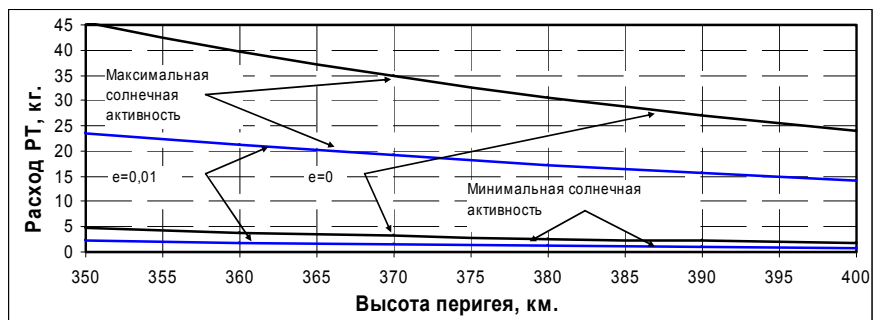


Рис. 3.30. Зависимости расхода рабочего тела в год для различных уровней солнечной активности

Эти результаты можно использовать в качестве исходных данных для проектирования системы коррекции орбиты на базе электроракетного двигателя.

Таблица 3.5. Расход рабочего тела за год полёта при крайних состояниях атмосферы

Высота перигея, км	Эксцентриситет e=0		Эксцентриситет e=0,01		Эксцентриситет e=0,0183	
	Расход РТ, кг (F max)	Расход РТ, кг (F min)	Расход РТ, кг (F max)	Расход РТ, кг (F min)	Расход РТ, кг (F max)	Расход РТ, кг (F min)
350	44,659	5,326	22,839	2,779	15,944	2,162
360	39,353	4,337	20,478	2,329	14,901	1,805
370	34,13	3,56	18,224	1,97	14,238	1,547
380	30,489	2,967	16,28	1,665	13,83	1,301
400	23,018	2,026	13,809	1,156	13,154	0,962

Как видно, наибольший потребный запас рабочего тела потребуется для поддержания круговой орбиты с высотой перигея 350 км. Для эллиптических орбит потребный запас составляет 22,8 кг ксенона. В расчёте на наихудший вариант примем запас рабочего тела равным 23 кг в год. Следовательно, на три года полёта потребуется 69 кг ксенона.

Примерный срок активного существования КА, с точки зрения запаса рабочего тела для ЭРДУ, составляет три года. При совместном использовании ЖРД и ЭРДУ срок активного существования может

увеличиться примерно в 1,5 раза.

При установке ЭРДУ на космический аппарат должны выполняться требования ограничения по массе и ограничения по габаритам. Основную долю в габаритные размеры ЭРДУ будет вносить баки хранения ксенона. Эти размеры будут определяться необходимым объёмом для топлива и формой бака.

Потребное количество двигателей определяется направлениями векторов ускорения для коррекции различных элементов орбиты и требуемыми характеристиками надёжности космического аппарата. Импульсами по трансверсальной оси корректируются большая полуось орбиты, эксцентриситет и аргумент перицентра. Импульсами по бинормали корректируются долгота восходящего узла и наклонение орбиты.

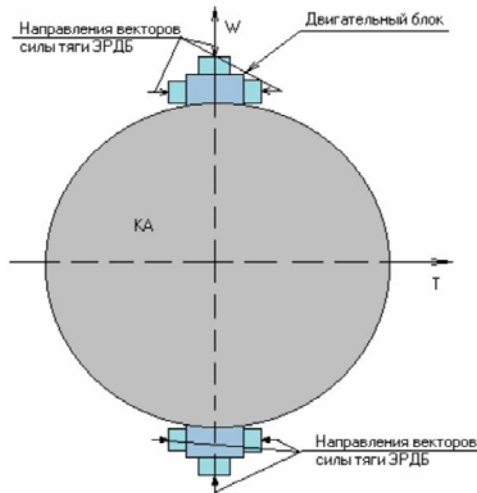


Рис. 3.31. Схема размещения двигательных блоков

Чтобы не нарушать режим работы аппаратуры наблюдения и обеспечить доступную максимальную энергию с солнечных батарей КА должен оставаться в ориентированном положении. Для этого разместим двигатели так, чтобы направления векторов тяги двигателей совпадали с направлениями осей (рис. 3.31). Кроме того, необходимо, чтобы линия действия силы тяги проходила через центр масс КА, для того чтобы не создавать крутящих моментов. Учитываются также ресурсные

характеристики КА.

Учитывая все выше сказанное, предлагается следующий вариант проектного облика космического аппарата класса «Ресурс» с дополнительной корректирующей электроракетной двигательной установкой на базе стационарных плазменных двигателей (см. рис.3.32).

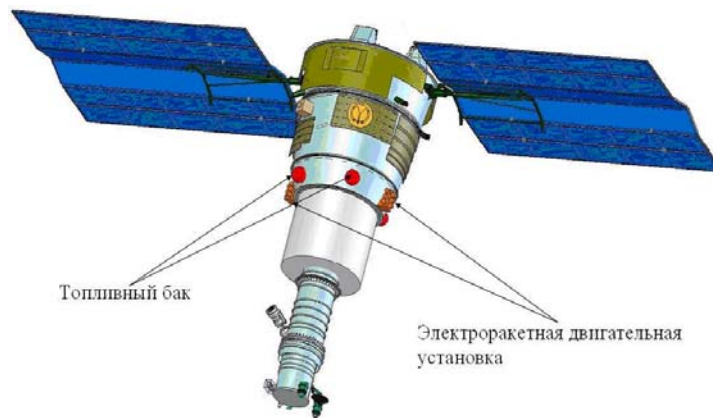


Рис. 3.32. Облик КА класса «Ресурс» с ЭРДУ

Для такой функциональной схемы ЭРДУ, суммарная масса ЭРДУ вычисляется по формуле:

$$\Sigma M_{ЭРДУ} = M_{МДБ} + M_{БХРТ} + M_{БПК} , \quad (3.34)$$

где  $M_{МДБ}$  - масса маршевого двигательного блока, включающего в себя двигатели СПД-70 и блоки газораспределения типа БГР-70,  $M_{БПК}$  - масса блока подачи ксенона (БПК), содержащего основную и резервную магистрали подачи, каждая из которых содержит регулятор давления (редуктор), клапаны, термодатчик, датчики высокого и низкого давления,  $M_{БХРТ}$  - масса блока хранения ксенона (БХРТ).

$$M_{МДБ} = n_1 \cdot M_{СПД-70} + n_2 \cdot M_{БГР} + \Delta M_{кр} , \quad (3.35)$$

где  $\Delta M_{кр}$  - масса крепёжных деталей,

$$M_{БХРТ} = M_{РТ} + n_{бак} \cdot M_{бак} + \delta M_n + \delta M_m + \delta M_k + M_{ред} ,$$

где  $M_{РТ}$  - масса ксенона,  $M_{бак}$  - масса конструкции одного топливного бака,  $\delta M_n$  - масса пироклапанов,  $\delta M_m$  - масса термодатчиков,  $\delta M_k$  - масса

элементов крепления,  $\delta M_{РЕД}$  - масса редуктора высокого давления.

Для КА класса «Ресурс», с учетом запаса рабочего тела на три года для коррекции околокруговой орбиты высотой порядка 350 км и дополнительным резервированием систем, общая масса ЭРДУ составит  $\sum M_{ЭРДУ} \approx 158 \text{ кг}$ , при массе  $M_{КА} \approx 6400 \text{ кг}$ .

Масса модернизированного космического аппарата будет определяться следующим выражением:

$$M_{КА}^2 = M_{КА}^1 + \sum M_{ЭРДУ} - \delta M_{ТОП.}^{ХДУ}, \quad (3.36)$$

где  $\delta M_{ТОП.}^{ХДУ}$  - часть массы топлива для химической двигательной установки, которую определяют как топливный резерв.

К примеру, у КА класса «Ресурс» максимальный запас химического топлива для маневров и коррекции – около 900 кг. При оснащении КА электроракетной двигательной установкой массой 150 кг запас топлива ХДУ можно сократить на такую же массу и более, при этом общая масса модернизированного КА не увеличивается.

### **3.3. Синтез Парето-оптимальных комбинированных схем выведения полезных нагрузок на геостационарную орбиту**

#### **3.3.1. Постановка задачи**

Выведение полезной нагрузки (ПН) на геостационарную орбиту (ГСО) может осуществляться тремя основными способами, задающими множество альтернатив:

1. «Импульсное» выведение с использованием традиционных химических двигателей. Достоинством данной схемы является малое время выведения ПН на ГСО и, как следствие, малое время пребывания КА в радиационных поясах Земли. Недостаток – сокращение массы ПН, выводимой на ГСО, из-за значительного возрастания массы топлива, необходимого для выполнения импульсных маневров.

2. Некомпланарный переход с низкой околоземной орбиты на ГСО с использованием электрореактивной двигательной установки (ЭРДУ). Достоинством указанной схемы является увеличение массы ПН, выводимой на рабочую орбиту, по сравнению с импульсным выведением.

Недостаток заключается в увеличении времени выведения и, следовательно, времени пребывания КА в радиационных поясах Земли.

3. Комбинированная схема выведения ПН на ГСО с использованием на первом этапе химического разгонного блока (ХРБ) для формирования промежуточной орбиты, а на втором этапе – бортовой ЭРДУ для доведения орбиты до целевой. Эта схема является наиболее универсальной, так как сочетает в себе достоинства первой и второй схем, а также позволяет варьировать соотношение массы ПН и времени выведения на ГСО, т.е. решить многокритериальную задачу оптимизации.

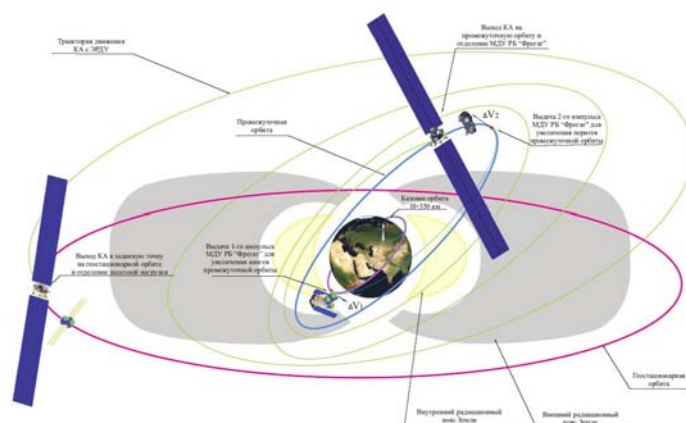


Рис. 3.33. Комбинированная схема выведения КА на ГСО

Будем рассматривать комбинированные схемы выведения, включающие в себя следующие этапы (рис. 3.33):

- выведение КА на базовую круговую орбиту высотой  $H_0=350$  км и наклоном  $i_0=51.5^\circ$  с помощью РН «Союз-2»;
- двухимпульсный компланарный переход КА на промежуточную эллиптическую орбиту за счет работы ХРБ (в данном случае маршевой двигательной установки (МДУ) разгонного блока (РБ) «Фрегат»), отделение МДУ РБ «Фрегат» и раскрытие панелей солнечных батарей.
- перелет КА с промежуточной орбиты на ГСО за счет работы ЭРДУ на основе стационарных плазменных двигателей СПД-140.

Уровень перегрузки, которую обеспечивает электрореактивный двигатель, весьма мал, поэтому продолжительность активного участка

велика, а траектория движения – многовитковая спираль с медленно изменяющимися оскулирующими элементами орбиты.

Для описания движения КА на этапе перехода с базовой орбиты на промежуточную используется компланарный двухимпульсный перелет между соосными эллиптическими орбитами [20].

Моделирование движения КА на этапе перехода с промежуточной на целевую орбиту производится на основе классической математической модели, представляющей собой систему дифференциальных уравнений первого порядка в оскулирующих элементах [21].

### 3.3.2. Проектная модель космической транспортной системы

Для выбора оптимальных проектных параметров КА представим стартовую массу КА как сумму масс отдельных систем:

$$M_0 = M_K^{XPB} + M_T^{XPB} + M_{\text{Э}}^{\text{ЭРДУ}} + M_{\text{Д}}^{\text{ЭРДУ}} + M_{\text{СПХ}}^{\text{ЭРДУ}} + M_T^{\text{ЭРДУ}} + M_K^{\text{ЭРДУ}} + M_{\text{ПН}}, \quad (3.37)$$

где  $M_0$  - начальная масса КА;  $M_K^{XPB}$  - масса конструкции ХРБ (сухая масса);  $M_T^{XPB}$  - масса топлива ХРБ;  $M_{\text{Э}}^{\text{ЭРДУ}}$  - масса источника и преобразователя энергии ЭРДУ;  $M_{\text{Д}}^{\text{ЭРДУ}}$  - масса двигательной установки ЭРДУ;  $M_{\text{СПХ}}^{\text{ЭРДУ}}$  - масса системы подачи и хранения рабочего тела ЭРДУ;  $M_T^{\text{ЭРДУ}}$  - масса рабочего тела ЭРДУ;  $M_K^{\text{ЭРДУ}}$  - масса корпуса КА, прочих элементов и систем КА с ЭРДУ;  $M_{\text{ПН}}$  - масса полезной нагрузки.

Масса ХРБ считается постоянной и равной сухой массе РБ «Фрегат» ( $M_K^{XPB} = 950$  кг). Массы отдельных компонентов КА с ЭРДУ зависят от проектных параметров. Обычно применяются следующие зависимости [19]:

$$M_{\text{Э}} = \alpha_{\text{ЭУ}} N, \quad M_{\text{Д}} = \gamma_{\text{ДУ}} (P + kP_{\text{УПР}}), \quad M_{\text{СПХ}} = \gamma_{\text{СПХ}} M_T, \\ M_K = \gamma'_K P + \gamma''_K N, \quad M_T^{\text{ЭРДУ}} = \frac{P}{c} T_M. \quad (3.38)$$

где  $P$  - тяга маршевых двигателей;  $P_{\text{УПР}}$  - тяга управляющих двигателей;  $N$  – мощность энергоустановки;  $T_M$  - моторное время (полное время работы ЭРДУ);  $\alpha_{\text{ЭУ}}$ ,  $\gamma_{\text{ДУ}}$ ,  $\gamma_{\text{СПХ}}$ ,  $\gamma'_K$ ,  $\gamma''_K$  - соответствующие удельные массовые

характеристики.

Мощность энергоустановки зависит от тяги двигателей и скорости истечения рабочего тела

$$N = \frac{Pc}{2} \frac{1 + \chi}{\eta_T \eta_{ПЭ}}, \quad (3.39)$$

где  $\chi = \frac{P_{УПР}}{P}$  - характеризует относительный расход массы управляющих двигателей,  $\eta_T$  - тяговый коэффициент полезного действия,  $\eta_{ПЭ}$  - КПД преобразователя энергии.

Если из уравнения баланса масс на начальной орбите (3.37) выразить массу полезной нагрузки, получим следующее выражение:

$$M_{ПН} = M_0 - M_K^{ХРБ} - M_T^{ХРБ} - M_{Э}^{ЭРДУ} - M_{Д}^{ЭРДУ} - M_{СПХ}^{ЭРДУ} - M_T^{ЭРДУ} - M_K^{ЭРДУ}. \quad (3.40)$$

Здесь  $M_0, M_K^{ХРБ} = fixe$ . Если количество двигателей одной и той же марки в составе ЭРДУ зафиксировать, то получим:  $M_{Э}^{ЭРДУ}, M_K^{ЭРДУ}, M_{Д}^{ЭРДУ} = fixe$ .

Из формул (3.38 - 3.40) видно, что масса системы подачи и хранения рабочего тела ЭРДУ зависит от массы рабочего тела ЭРДУ:  $M_{СПХ}^{ЭРДУ} = f(M_{РТ}^{ЭРДУ})$ . Тогда для массы полезной нагрузки можно записать:  $M_{ПН} = f(M_{РТ}^{ЭРДУ}, M_{РТ}^{ХРБ})$ . Масса топлива ХРБ равна потребной массе топлива, определяемого величинами импульсов для перевода КА на промежуточную орбиту. Чтобы масса полезной нагрузки на ГСО была максимальной, необходимо обеспечить минимум массы рабочего тела ЭРДУ. Масса рабочего тела ЭРДУ, в свою очередь, определяется временем перелета. В качестве второго критерия оптимальности будем рассматривать полное время перелета  $T_{\Sigma}$ .

$$T_{\Sigma} = T_{ХРБ} + T_{ЭРДУ}, \quad (3.41)$$

Поскольку сформулированная задача является многокритериальной, будем рассматривать максимизацию массы ПН  $M_{ПН} \rightarrow \max$ , выводимой на геостационарную орбиту, при фиксированном времени перелета:  $T_{\Sigma} = fixe$ , с последующим построением зависимости одного критерия от другого.

### 3.3.3. Модель управления космической транспортной системой

Законы управления вектором тяги КА с ЭРДУ реализуются посредством изменения углов ориентации вектора тяги (рис. 3.34) как в плоскости орбиты (угол  $\lambda$ ), так и в плоскости местного горизонта (угол  $\psi$ ).

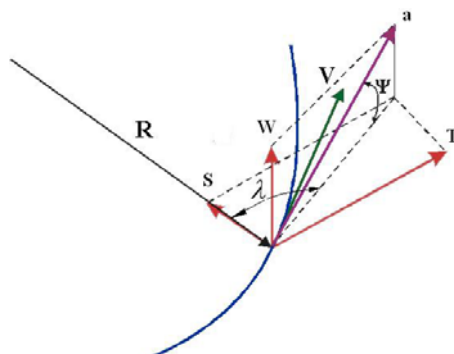


Рис. 3.34. К определению углов ориентации вектора тяги КА с ЭРДУ

Известны различные законы управления вектором тяги КА с ЭРДУ [21]. Главным образом, это законы, обеспечивающие изменение определенного элемента орбиты (большой полуоси, эксцентриситета, наклона) при постоянстве одного или двух других. В работе [21] рассматривались законы последовательного доведения элементов орбиты до заданных поочередно. Причем сначала производилось изменение большой полуоси или эксцентриситета, а затем – наклона, поскольку производить поворот плоскости орбиты выгодно на максимальном удалении от притягивающего центра. Такие законы управления предпочтительны в отношении их технической реализации, т.к. позволяют разделить каналы управления вектором тяги КА с ЭРДУ, однако не обеспечивают максимума массы ПН на ГСО.

Рассмотрим закон управления, обеспечивающий совместное изменение большой полуоси, эксцентриситета и наклона орбиты при движении КА с ЭРДУ.

В качестве критерия оптимальности динамической задачи был выбран квадратичный функционал, представляющий собой сумму квадратов невязок большой полуоси, эксцентриситета и наклона



орбиты, умноженных на соответствующие им весовые коэффициенты:

$$I = \sum_{j=1}^3 \alpha_j \mathcal{E}_j^2 = \alpha_1 \Delta A_k^2 + \alpha_2 \Delta e_k^2 + \alpha_3 \Delta i_k^2 \rightarrow \min, \quad (3.42)$$

$$\Delta A = A(t) - A_k, \quad \Delta e = e(t) - e_k \equiv e(t), \quad \Delta i = i(t) - i_k \equiv i(t).$$

Этот функционал легко приводится к интегральному виду с целью обеспечения максимальной скорости изменения критерия  $I$ :

$$\frac{dI}{dt} = 2\alpha_1 \left( \frac{A}{A_0} - \frac{A_k}{A_0} \right) \cdot \frac{1}{A_0} \frac{dA}{dt} + 2\alpha_2 e \frac{de}{dt} + 2\alpha_3 i \frac{di}{dt} \rightarrow \max, \quad (3.43)$$

В работе [19] получены аналитические выражения для углов ориентации вектора тяги КА с ЭРДУ в случае совместного управления классическим путем поиска максимума функционала по двум переменным

$$\frac{dI}{dt} = \frac{dI}{dt}(\lambda(t), \psi(t)) \rightarrow \max.$$

$$\begin{cases} \sin \lambda = \pm \frac{A_T}{\sqrt{A_T^2 + A_S^2}}, \\ \cos \lambda = \pm \frac{A_S}{\sqrt{A_T^2 + A_S^2}}; \end{cases} \quad \begin{cases} \sin \psi = \pm \frac{A_W}{\sqrt{A_T^2 + A_S^2 + A_W^2}}, \\ \cos \psi = \pm \frac{\sqrt{A_T^2 + A_S^2}}{\sqrt{A_T^2 + A_S^2 + A_W^2}}. \end{cases} \quad (3.44)$$

$$\text{Здесь } A_S = \alpha_1 \left( \frac{A}{A_0} - \frac{A_k}{A_0} \right) \cdot \frac{2A(1+e)}{A_0(1-e)} \cdot e \sin \vartheta + \alpha_2 e \cdot \sin \vartheta, \quad A_W = \alpha_3 i \cdot \frac{\cos u}{1 + e \cos \vartheta},$$

$$A_T = \alpha_1 \left( \frac{A}{A_0} - \frac{A_k}{A_0} \right) \cdot \frac{2A(1+e)}{A_0(1-e)} (1 + e \cos \vartheta) + \alpha_2 e \cdot \frac{e \cos^2 \vartheta + 2 \cos \vartheta + e}{1 + e \cos \vartheta}.$$

Можно рассмотреть два подхода к выбору весовых коэффициентов:

- итерационная процедура подбора весовых коэффициентов во внутреннем цикле путем оценки модулей средних производных большой полуоси, эксцентриситета и наклона орбиты за определенные интервалы времени ( $\Delta T = 1 \dots 4$  витка).

- частный случай:  $\alpha_1 = \alpha_2 = \alpha_3 \cong 0,333$ .

### 3.3.4. Результаты моделирования перелета на ГСО

На рис. 3.35 - 3.37 представлены зависимости большой полуоси, эксцентриситета и наклона орбиты при использовании совместного

управления для частного случая выбора весовых коэффициентов для исходных данных:  $A_0 = 22000$  км,  $e_0 = 0.273$ , 8 маршевых и 4 резервных двигателей СПД-140.

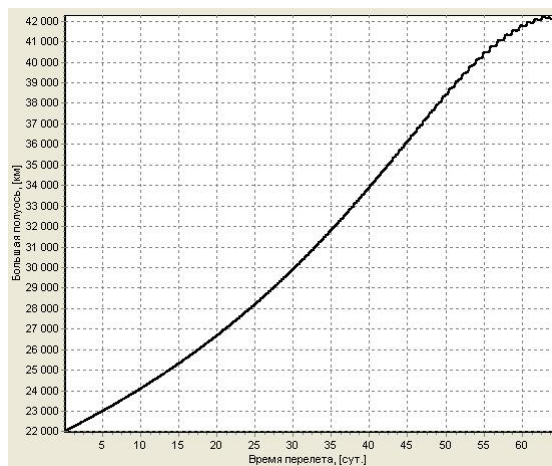


Рис. 3.35. Зависимость большой полуоси от времени перелета для локально-оптимальной схемы

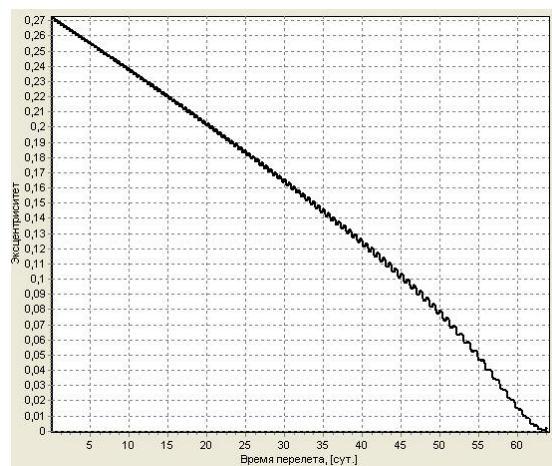


Рис. 3.36. Зависимость эксцентриситета от времени перелета для локально-оптимальной схемы

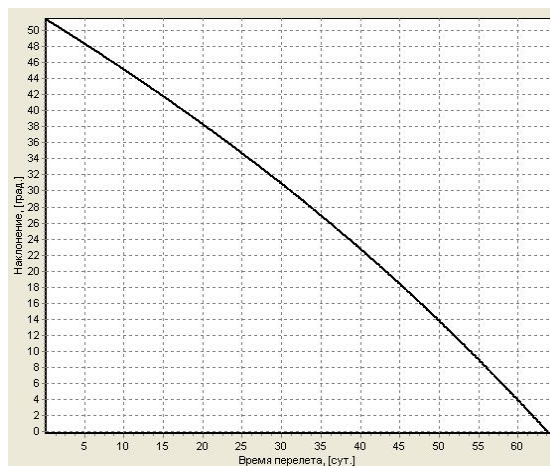


Рис. 3.37. Зависимость наклонения от времени перелета для локально-оптимальной схемы

На рис. 3.38 показаны результаты трехмерного моделирования траектории перелета на ГСО при использовании совместной схемы.

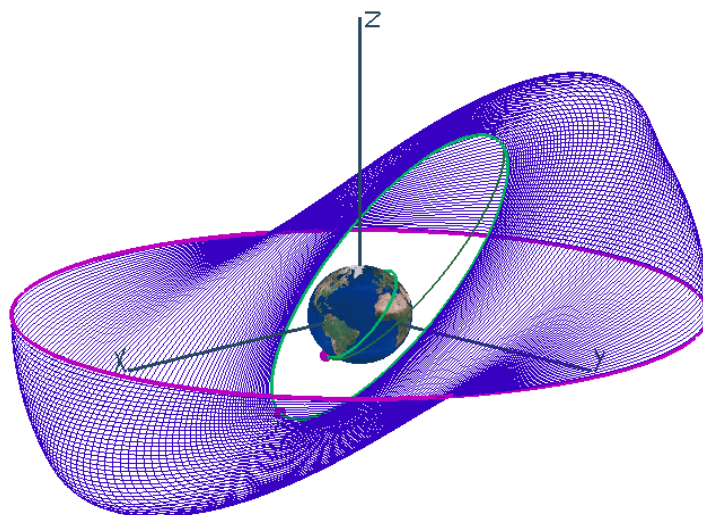


Рис. 3.37. Траектория перелета КА на ГСО

### 3.3.5. Построение множества Парето

Задача отыскания вариантов, обеспечивающих максимум массы ПН на целевой орбите за минимальное время выведения, является многокритериальной с векторным критерием  $W = \{T_{\Sigma}, \mu\}$ . Здесь

$T_{\Sigma} \rightarrow \min$  - критерий, определяющий полное время перелета;

$\mu = \frac{m_{ПН}}{m_{КА}} \rightarrow \max$  - критерий, определяющий массу полезной нагрузки,

отнесенную к массе КА на промежуточной орбите.

Одним из методов решения задачи многокритериальной оптимизации является выделение множества «эффективных» решений или метод Парето.

В работе [21] произведены расчеты комбинированных схем выведения для различных значений исходных данных, где, в частности, варьировались величины большой полуоси и эксцентриситета промежуточной орбиты. Построив точки в плоскости параметров «время перелета – относительная масса ПН» (рис. 3.38) и проведя процедуру сравнения, исключаем из исходного множества неконкурентоспособные решения по выбранным критериям.

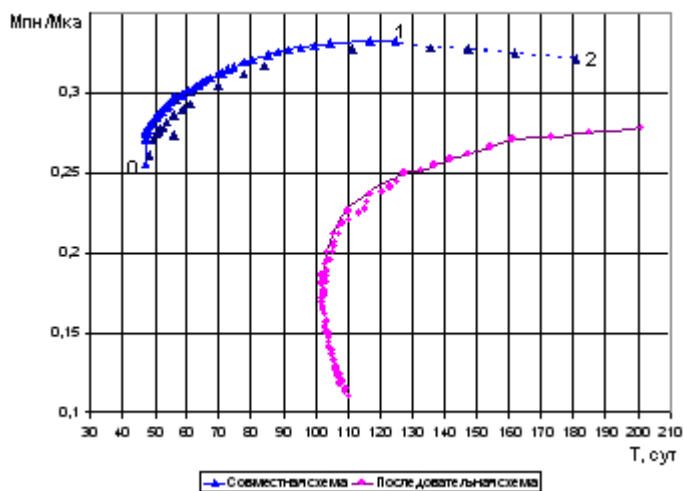


Рис. 3.38. Построение множества Парето

Таким образом, получаем более узкое множество решений, которое

является верхней границей исходного множества (на рис. 3.38 точки, принадлежащие множеству Парето, соединены сплошными линиями). Видно, что точки, принадлежащие Парето-множеству обладают следующим свойством: двигаясь вдоль этого множества, нельзя увеличить массу ПН на ГСО, не увеличив при этом и времени перелета.

Видно, что для кривой, соответствующей совместной схеме, в точке 1 существует слабый оптимум, а весь участок 0-1 образует множество неуправляемых решений, среди которых проектант может выбрать наиболее рациональное для поставленной задачи. Точки, принадлежащие участку 1-2, не образуют Парето-множество, поскольку здесь наблюдается ухудшение по массовому критерию с увеличением времени перелета. Этот рисунок подтверждает целесообразность совместной схемы как по массовому, так и по временному критериям.

На рис. 3.39 представлены множества «эффективных» решений для совместной схемы управления при различных значениях радиуса перигея промежуточной орбиты.

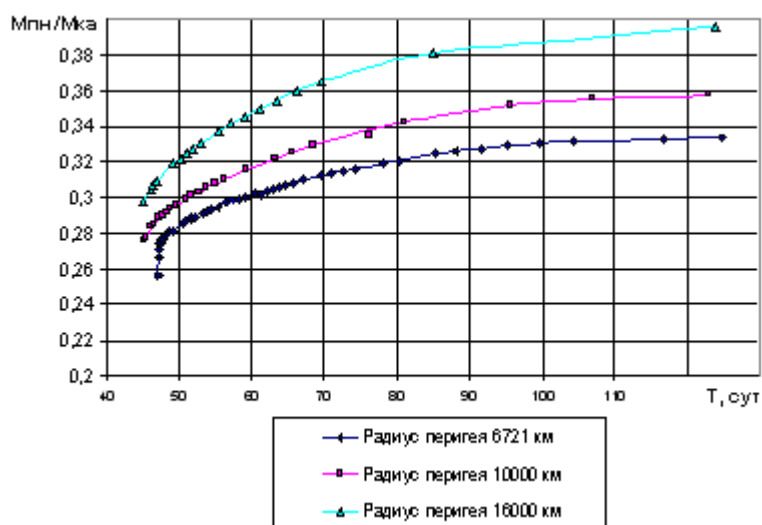


Рис. 3.39. Сравнительный анализ локально-оптимальных схем совместного управления элементами орбиты с различными значениями радиуса перигея

Из рис. 3.39 можно сделать вывод о том, что поднятие перигея до  $R_{\pi} = 16000$  км дает выигрыш как во времени выведения, так и в массе полезной нагрузки.

### 3.3.6. Формирование проектного облика МТА с использованием системы твердотельного моделирования Solid Works

В табл. 3.6 приведены результаты расчета проектных параметров для различных продолжительностей перелета. Во всех трех случаях в составе ЭРДУ использовались двенадцать двигателей СПД-140, восемь из которых являлись рабочими, а остальные – резервными.

Таблица 3.6. Результаты расчета проектных параметров для совместной схемы

$T^{\text{ЭРДУ}}$ , сут.	$T_{\min}^{\text{ЭРДУ}}=44,950$ сут.	$T_{\text{cp}}^{\text{ЭРДУ}}=53,028$ сут.	$T_{\max}^{\text{ЭРДУ}}=85,171$ сут.
$A_{np}$ , км	40500	27500	16500
$e_{np}$	0,605	0,418	0,030
$m_T^{XPB}$ , кг	3908,050	3590,115	2504,892
$V_{хар}^{XPB}$ , км/с	2,540	2,175	1,239
$m_{КА}^{np}$ , кг	2011,95	2329,885	3415,108
$V_{хар}^{\text{ЭРДУ}}$ , км/с	4,738	4,836	5,351
$m_{PT}^{\text{ЭРДУ}}$ , кг	511,608	603,637	969,415
$m_{ЭУ}^{\text{ЭРДУ}}$ , кг	84	84	84
$m_{СПХ}^{\text{ЭРДУ}}$ , кг	35,813	42,255	67,859
$m_K^{\text{ЭРДУ}}$ , кг	301,792	349,483	512,266
$m_{ДУ}^{\text{ЭРДУ}}$ , кг	84	84	84
$m_{ПН}^{\text{ЭРДУ}}$ , кг	598,737	770,602	1301,589
$a_0$ , мм/с <sup>2</sup>	1,113	0,961	0,656
$S_{CB}$ , м <sup>2</sup>	174	174	174

Обозначения:  $A_{np}$  - большая полуось промежуточной орбиты;  $e_{np}$  - эксцентриситет промежуточной орбиты;  $m_T^{XPB}$  - масса топлива ХРБ;  $V_{хар}^{XPB}$  - характеристическая скорость перелета КА с ХРБ;  $m_{КА}^{np}$  - масса КА на промежуточной орбите;  $V_{хар}^{\text{ЭРДУ}}$  - характеристическая скорость перелета КА с ЭРДУ;  $a_0$  - реактивное ускорение ЭРДУ;  $S_{CB}$  - площадь панелей солнечных батарей.

Результаты расчета проектных параметров позволяют сформировать

проектный облик космического аппарата с ЭРДУ (рис. 3.40) и составить упрощенную компоновку универсального транспортного модуля (рис. 3.41).

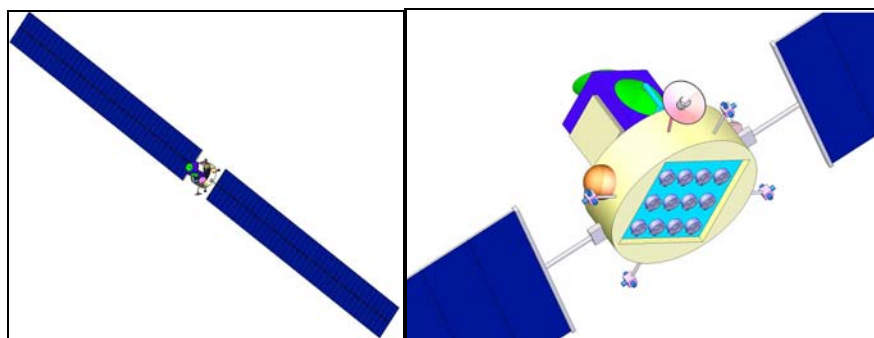


Рис. 3.40. Проектный облик космического аппарата с ЭРДУ

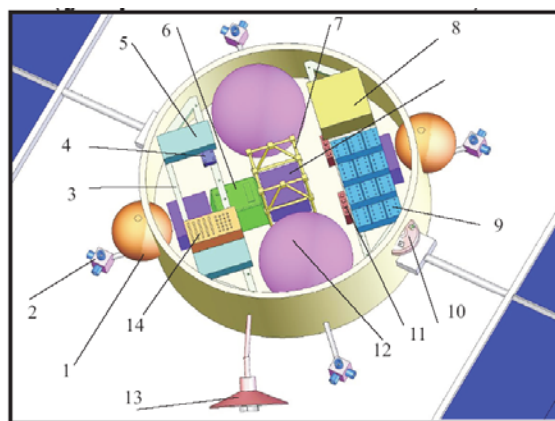


Рис. 3.41. Упрощенная компоновка универсального транспортного модуля: 1 – бак двигательной установки ориентации и стабилизации; 2 – двигательная установка ориентации и стабилизации; 3 – приборная рама; 4 – солнечный датчик; 5 – блок электроавтоматики; 6 – телеметрическая система; 7 – рама для установки ПН; 8 – бортовой комплекс управления; 9 – аккумуляторные батареи; 10 – астроприбор; 11 – комплекс ориентации и стабилизации; 12 – бак ЭРДУ; 13 – остронаправленная антенна; 14 – бортовой радиокомплекс.

Предполагается, что КА состоит из комбинированного транспортного модуля и полезной нагрузки. В качестве ПН может быть использован любой КА (как правило, спутник связи), выводимый на ГСО и удовлетворяющий требуемым массо-габаритным характеристикам.

Универсальный транспортный модуль имеет в своем составе многоцелевую космическую платформу, маршевую двигательную установку (ХРБ с ЖРД) и электрореактивную двигательную установку. Для этого использовался программный продукт Solid Works со специально разработанным интерфейсом между расчетным и графическим модулями.

На рис. 3.42 представлена блок-схема, содержащая все основные блоки для расчета комбинированной схемы выведения КА на геостационарную орбиту.

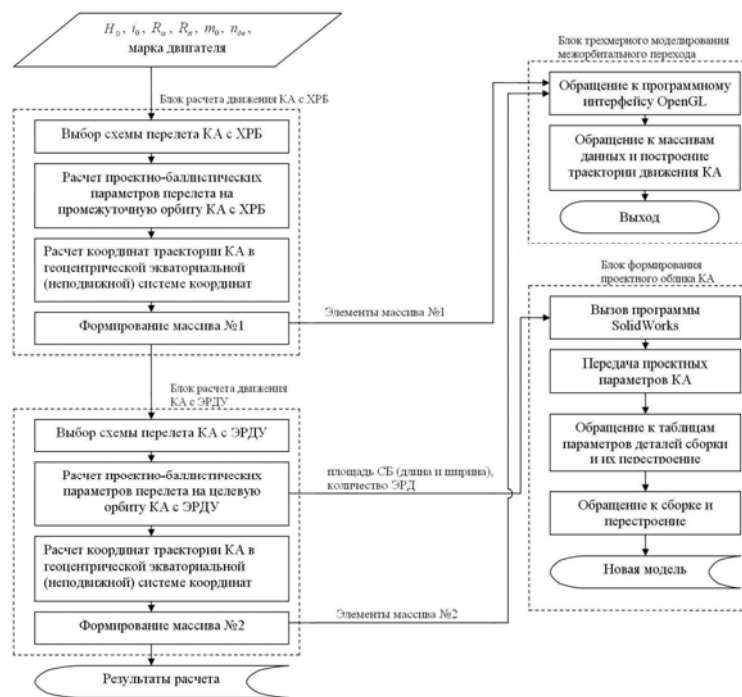


Рис. 3.42. Укрупненная блок-схема алгоритма оптимизации

Как уже говорилось выше, для выведения КА с ЭРДУ на промежуточную эллиптическую орбиту используется маршевая двигательная установка разгонного блока «Фрегат» (рис. 3.43).



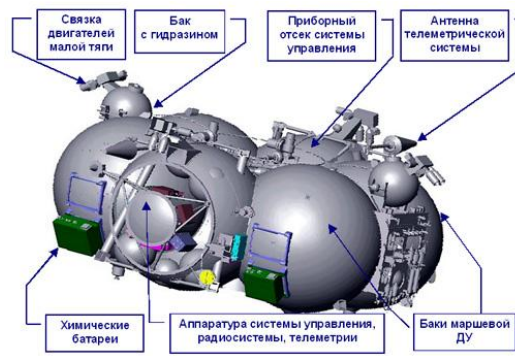


Рис. 3.43. Внешний вид разгонного блока «Фрегат»

### 3.4. Синтез баллистических схем межпланетных перелетов КА с ЭРДУ на основе многокритериальной оптимизации

#### 3.4.1. Постановка задачи проектно-баллистической оптимизации межпланетных перелетов

При решении задачи выбора оптимального состава средств, решающих задачу транспортировки ПГ в межпланетном пространстве, необходимо сформировать целевую функцию, наиболее полно отражающую качество системы межпланетного КА в целом. Рассмотрим случай, когда целью КА является доставка заданной полезной нагрузки в заданную точку пространства. При этом длительность перелета  $T_{\Sigma}$ , длительность пребывания в радиационных поясах  $T_{рад\Sigma}$ , стартовая масса аппарата  $M_0$  и стоимость миссии должны быть минимальными, а масса полезного груза должна быть максимальной.

$$\begin{aligned}
 T_{\Sigma} &= \{T_1, T_2, \dots, T_n\} = T_{\Sigma}(\bar{p}, \bar{b}) \rightarrow \min, \\
 T_{рад\Sigma} &= \{T_{рад1}, T_{рад2}, \dots, T_{раdn}\} = T_{рад\Sigma}(\bar{p}, \bar{b}) \rightarrow \min, \\
 M_{ПГ} &= \{M_{ПГ1}, M_{ПГ2}, \dots, M_{ПГn}\} = M_{ПГ}(\bar{p}, \bar{b}) \rightarrow \max, \\
 M_0 &= \{M_{01}, M_{02}, \dots, M_{0n}\} = M_0(\bar{p}, \bar{b}) \rightarrow \min, \\
 C_{\Sigma} &= \{C_{\Sigma1}, C_{\Sigma2}, \dots, C_{\Sigma n}\} = C_{\Sigma}(\bar{p}, \bar{b}) \rightarrow \min,
 \end{aligned}
 \tag{3.45}$$

где  $\bar{p}$  - вектор проектных параметров КА,  $\bar{b}$  - вектор параметров баллистической схемы перелета. Под *баллистической схемой* перелета будем понимать совокупность следующих параметров: положение небесных тел на момент старта миссии, определяющее оптимальную дату старта  $D_0$ ; длительность межпланетного перелета  $T$ , соотношение длительностей участков полета, и, наконец, при использовании гравитационных маневров их последовательность и длительности участков между маневрами  $T_i$ .

Длительность перелета ограничим:  $T_{\Sigma} \leq T_{\text{ПРЕД}}$ . Стартовую массу ограничиваем возможностями ракетоносителя и выбираем равной максимальной массе выводимой ракетоносителем на опорную орбиту  $M_0 = M_{PH \max}$ .

Для формализации постановки задачи оптимизации введем в рассмотрение:

- $\bar{x}(t) \in X$  - вектор фазовых координат аппарата, подчиняющийся граничным условиям, определяемым целью перелета и возможными ограничениями, где  $X$  - множество допустимых фазовых координат;
- $\bar{u}(t) \in U$  - вектор функций управления, на компоненты которого наложены ограничения, связанные с конструктивными особенностями аппарата и ДУ, где  $U$  - множество допустимых управлений;
- $\bar{p} \in P$  - вектор оптимизируемых проектных параметров КА, ограничен множеством допустимых проектных параметров  $P$ ;
- $\bar{b} \in B$  - вектор баллистических параметров, зависящий от цели миссии.

В зависимости от выбранной модели движения, *вектор фазовых координат* аппарата содержит  $\bar{x}(t) = (M, \bar{r}, \bar{V}, Rad, \bar{r}_i)^T$ , где

$M$  – текущая масса КА;

$\bar{r}$  – радиус-вектор КА относительно притягивающего центра в выбранной системе отсчета;

$\bar{V}$  – вектор скорости КА;

$Rad(\bar{r})$  – текущая интенсивность радиационного облучения КА;

$\bar{r}_i$  – радиус-вектор КА относительно других гравитирующих тел.

Вектор проектных параметров КА в зависимости от выбранной проектной модели КА может содержать компоненты  $\bar{p} = (a_0, c, \alpha_i, \gamma_i, \chi(t, \bar{r}))^T$ , где

$a_0$  – номинальное ускорение двигательной установки;

$c$  – скорость истечения рабочего тела;

$\alpha_i, \gamma_i$  – удельные массовые характеристики  $i$ -ого компонента КА по тяге и мощности;

$\chi(t, \bar{r})$  – коэффициент изменения тяги двигательной установки.

Состав вектора функций управления изменяется в зависимости от выбранной модели движения и критерия оптимальности  $\bar{u}(t) = (\lambda_1, \lambda_2, \delta)^T$ ,

где

$\lambda_1$  – угол, определяющий положение тяги в плоскости орбиты;

$\lambda_2$  – угол, определяющий положение тяги относительно плоскости орбиты;

$\delta$  – функция включения-выключения двигательной установки.

Состав вектора параметров баллистической схемы перелета зависит от цели и баллистической схемы миссии  $\bar{b} = (D_0, T_i, \Delta V_i)^T$ , где

$D_0$  – дата старта миссии;

$T_i$  – длительности участков перелета, определяющие положение планет назначения, финиша и промежуточных гравитационных маневров;

$\Delta V_i$  – гиперболические избытки скорости на границах участков движения, приращения скоростей в моменты гравитационных маневров или включений двигателей большой тяги.

При использовании критериев оптимальности (3.45) задача комплексной оптимизации миссии формулируется следующим образом. Определить вектора  $\bar{p} \in P$ ,  $\bar{u}(t) \in U$  и  $\bar{b} \in B$ , доставляющие экстремум критерию оптимальности обеспечивающие выполнение цели миссии  $\bar{x} \in X$ .

$$K_i = \underset{\bar{u}(t) \in U, \bar{p} \in P, \bar{b} \in B}{extr} K_i(\bar{x} \in X, \bar{u}(t), \bar{p}, \bar{b}). \quad (3.46)$$

### 3.4.2. Массовая и проектная модель межпланетного КА с малой тягой

Анализ работ в области оптимизации КА с ЭРД малой тяги [19, 27] позволяет выделить следующие элементы конструкции КА:

- 1) полезный груз, включая возвращаемую часть,
- 2) энергоустановка, состоящая из источника и преобразователя энергии,
- 3) двигательная установка, включающая маршевые и управляющие двигатели вместе с исполнительными органами,
- 4) рабочее тело, необходимое для прямого и обратного перелета с учетом расхода на управление,
- 5) система подачи и хранения рабочего тела (баки, трубопроводы и пр.),
- 6) корпус и конструкции КА.

Уравнение баланса масс на начальной орбите имеет вид (3.37). Наиболее простыми и часто используемыми зависимостями массы отдельных компонентов КА от проектных параметров являются линейные зависимости масс от номинальной мощности энергоустановки  $N_0$  и тяги двигателей на орбите Земли  $P_0$  [9, 18]:

$$M_{\text{Э}} = \alpha_{\text{Э}} \cdot N_0, \quad M_{\text{Д}} = \gamma_{\text{Д}} \cdot P_0, \quad (3.47)$$

$$M_{\text{КОН}} = \alpha_{\text{КОН}} \cdot N_0 + \gamma_{\text{КОН}} \cdot P_0.$$

где  $\alpha_{\text{Э}}$ ,  $\alpha_{\text{КОН}}$ ,  $\gamma_{\text{Д}}$ ,  $\gamma_{\text{КОН}}$  - соответствующие удельные массовые характеристики. Подставляя эти выражения в формулу (3.37) с учетом зависимости мощности от проектных параметров

$$N_0 = \frac{P_0 c_0}{2\eta_T \eta_{\text{ПЭ}}} \quad (3.48)$$

получим новый вид уравнения баланса масс на начальной орбите:

$$M_0 = M_{\text{гр}} + \alpha_{\text{ЭВ}} \cdot \frac{P_0 c_0}{2\eta_T \eta_{\text{ПЭ}}} + \gamma_{\text{Д}} \cdot P_0 + \frac{P_0}{c_0} \cdot T_{\mu}^*(T)(1 + k_{\text{СПХ}}) + \\ + \alpha_{\text{К}} \cdot \frac{P_0 c_0}{2\eta_T \eta_{\text{ПЭ}}} + \gamma_{\text{К}} \cdot P_0,$$

или

$$M_0 = M_{\text{гр}} + (\alpha_{\text{ЭВ}} + \alpha_{\text{К}}) \cdot \frac{P_0 c_0}{2\eta_T \eta_{\text{ПЭ}}} + (\gamma_{\text{Д}} + \gamma_{\text{К}}) \cdot P_0 +$$

$$+ \frac{P_0}{c_0} \cdot T_{\mu}^*(T)(1 + k_{СПХ}). \quad (3.49)$$

Уравнение (3.49) показывает, что в качестве основных проектных параметров в простейшем случае можно выбрать номинальную тягу двигателей и скорость истечения рабочего тела.

Для наиболее простой модели считается, что в условиях падения уровня энергии достигающей двигательной установки происходит падение только тяги двигателей, при неизменной скорости истечения рабочего тела. Считается, что КПД двигательной установки при этом остается неизменным. Однако для реальной двигательной установки существует оптимальный уровень тяги с наибольшим КПД, а при всех остальных значениях эффективность двигателей снижается.

Естественно не являются линейными и зависимости масс отдельных частей конструкции от проектных параметров аппарата. Дальнейшее уточнение проектной модели предполагает выбор реальных двигателей и энергоустановок среди нескольких альтернативных вариантов. Выбор двигателей и энергоустановки определяет массу конструкции. Дальнейшее решение проектной задачи оптимизации носит поворочный характер.

Таким образом, реальная зависимость стартовой массы от проектных параметров аппарата имеет более сложный нелинейный вид и уравнение баланса масс (3.49) значительно усложняется. В любом случае можно говорить о том, что стартовая масса аппарата является функцией тяги двигателей на орбите Земли, скорости истечения рабочего тела, массы рабочего тела требуемого на выполнение целевой задачи и ограничений, наложенных на фазовые координаты, законы управления и проектные параметры КА.

### **3.4.3. Модель оценки затрат на осуществление проекта**

Предварительная оценка стоимости создания и эксплуатации космического комплекса, осуществляющего полет по заданной траектории и миссию в целом, обычно включает затраты на его разработку, производство, испытание и наземное обслуживание.

Поскольку эта оценка делается для проведения сравнительного анализа различных вариантов миссий, то важны не абсолютные значения

стоимости того или иного варианта миссии, а их относительные показатели, дающие возможность сравнивать миссии между собой. Из всех составляющих стоимости миссии выделяются только те из них, которые вносят наиболее существенную разницу в сравнительную оценку. Исходя из существующих и разрабатываемых космических технологий, к таким составляющим космического комплекса относятся: ракета-носитель, разгонные блоки с химическими двигательными установками, электрореактивные двигательные установки, солнечные и ядерные энергоустановки.

### 3.4.4. Модели движения, используемые при оптимизации межпланетных перелетов с малой тягой

Для осуществления баллистической оптимизации миссии после разделения задачи требуется выбрать модель движения КА. На рис. 3.44 схематично представлена классификация **моделей движения** КА, использующихся для оптимизации межпланетных перелетов КА.



Рис. 3.44. Структурная схема моделей движения КА, использующихся для решения баллистической части задачи оптимизации межпланетных миссий

Отдельно выделена группа наиболее простых моделей, для которых

получено аналитическое решение, т.е. зависимости длительности маневра, траектории движения и критерия оптимальности от проектных параметров имеют явный вид и вычисляются по конечному числу аналитических формул. К этой группе относятся модели приближенно описывающие отдельные участки движения КА под действием простого, заранее известного управления. Например, это модели, описывающие маневры набора параболической скорости или торможения под действием трансверсального закона управления на планетоцентрических участках движения; гелиоцентрическое движение КА по логарифмической спирали; движение КА под действием традиционных двигателей большой тяги в импульсной постановке и некоторые другие.

Большая часть моделей движения, обеспечивающих более точное описание миссии, требует решения систем обыкновенных дифференциальных уравнений. Часть моделей этого уровня используют разделение траектории на участки движения в сфере действия одного притягивающего центра, другая группа наиболее сложных моделей использует описание движения в рамках задач трех и более гравитирующих тел.

При описании околопланетных участков движения уточнение модели достигается за счет учета нецентральной гравитационного поля притягивающего тела, влияния его атмосферы, притяжения спутников планеты и Солнца и других возмущающих ускорений. На выбор оптимальных программ управления влияют способы управления двигательной установкой и ограничения на ее перемещение, условия освещенности КА, прохождение радиационных поясов и другие факторы.

Гелиоцентрическое движение КА уточняется за счет учета эллиптичности и некомпланарности орбит планет, влияния солнечной радиации и температуры на эффективность работы энергодвигательной установки, ограничений на законы управления ДУ.

#### ***3.4.5. Результаты оптимизации межпланетной миссии к Марсу, осуществляемой КА с солнечным электроракетным двигателем***

Рассмотрена межпланетная миссия Земля-Марс фиксированной длительности 350 сут, осуществляемая КА с ЭРДУ и солнечным

источником энергии. Запуск КА предполагается производить при помощи РН «Союз - ФГ» из Байконура. Сухая масса КА равна 3370 кг, стартовая масса минимизируется. Все маневры КА (набор скорости и переход на отлетную траекторию, гелиоцентрический перелет и торможение при выходе на орбиту искусственного спутника Марса) предполагается производить при помощи ЭРДУ (рис. 3.45).

2. Выбирается программа управления на гелиоцентрической траектории.

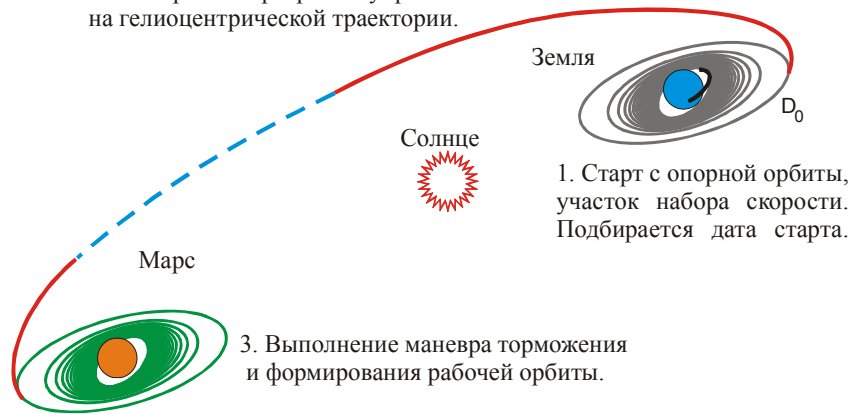


Рис. 3.45. Баллистическая схема межпланетной миссии Земля – Марс с использованием ЭРДУ с солнечной энергоустановкой

КА выводится на опорную круговую геоцентрическую орбиту высотой 500 км, наклонением 51.6 град. Для этой баллистической схемы проводится проектно-баллистическая оптимизация миссии. Для последовательной оптимизации миссии использовались следующие начальные значения вектора проектных параметров КА  $\bar{p} = (P_0 = 4Н, c = 7 \cdot 10^4 \text{ км/с})^T$  и стартовой массой КА в первом приближении  $M_0^0 = 5510 \text{ кг}$ .

Для этих проектных параметров проведено моделирование маневра набора скорости в сфере действия Земли. В табл. 3.7 приведены результаты этого моделирования для различных моделей планетоцентрического движения, в т.ч. координаты КА после выхода из сферы действия Земли в гелиоцентрической системе координат. Плоский маневр КА массой 5500 кг рассчитывался для начальной круговой орбиты с высотой орбиты 500 км и



постоянным наклоном орбиты 23.45 град, при пространственном маневре наклонение орбиты изменялось с 51.6 до 23.45 град.

Таблица 3.7. Результаты моделирования маневра набора параболической скорости в сфере действия Земли КА с солнечной энергоустановкой

Маневр и учетные возмущения	$T$ , сут	$M_{PT}$ , кг	Координаты КА в гелиоцентрической системе координат					
			$r$ , млн.км	$\varphi$ , град	$V_r$ , км/с	$V_\varphi$ , км/с	$i$ , град	$\Omega$ , град
Тангенциальный закон набора скорости, без учета возмущений ( $M 2.i$ )	106.1	550	150.859	346.3	-1.01	30.60	0.004	-0.16
Тангенциальный закон набора скорости, с учетом нахождения КА в тени, без учета возмущений	122.1	550	149.636	2.031	-0.41	28.41	0.007	0.065
Тангенциальный закон набора скорости, учитывается пребывание КА в тени и гравитационные возмущения от несферичности Земли	117.9	550	150.747	357.9	0.368	30.49	0.014	-0.17
Тангенциальный закон набора скорости, учитывается пребывание КА в тени, грав. возмущения от Луны, Солнца и несферичности Земли ( $M 3.i$ )	117.9	550	149.915	357.9	-1.44	29.44	0.030	-0.95
Набор параболической скорости с изменением наклона, возмущения не учитываются	108.3	560	150.617	348.4	0.372	29.07	0.004	-0.16
Набор параболической скорости с изменением наклона, учитывается пребывание КА в тени, грав. возмущения от Луны, Солнца и несферичности Земли ( $M 3.i$ )	126.5	560	149.006	6.307	-1.42	29.04	0.020	-0.03

Данные табл. 3.7 показывают, что длительность маневра набора скорости в сфере действия Земли слишком высока и не позволяет осуществить миссию за требуемые 350 сут.

Возможным выходом из ситуации является временное увеличение тяги двигателей в два раза, за счет уменьшения скорости истечения рабочего тела. В таблице 3.8 показаны проектно-баллистические параметры

маневра набора параболической скорости, полученные для такой схемы изменения проектных параметров.

В рамках модели *M 2.4* для круговых компланарных орбит планет и используя процедуру оптимизации, описанную в [28], получены основные проектно-баллистические характеристики миссии Земля – Марс, представленные в табл. 3.8.

Таблица 3.8. Основные проектно-баллистические характеристики миссии Земля – Марс, полученные в рамках модели *M 2.4* для круговых компланарных орбит планет

$i$	Маневр	$T_i$ , сут	$\varphi_i$ , град	$M(T_{i-1})$ , кг	$M(T_i)$ , кг	$M_{PT_i}$ , кг
1	Набор скорости у Земли	41.97	41.37	5510	4380	1130
2	Перелет Земля – Марс	248.58	173.97	4380	3560	820
3	Торможение у Марса	59.45	31.16	3560	3370	190
8	Суммарные параметры	350.0	246.49	5510	3370	2140

На рис. 3.46. показана траектория оптимального гелиоцентрического перелета Земля – Марс, а на рис. 3.47 этапы торможения в сфере действия Марса.

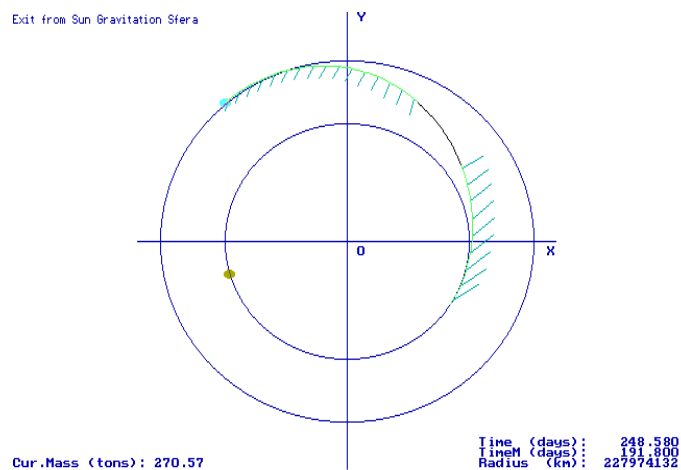


Рис. 3.46. Оптимальная траектория перелета Земля – Марс

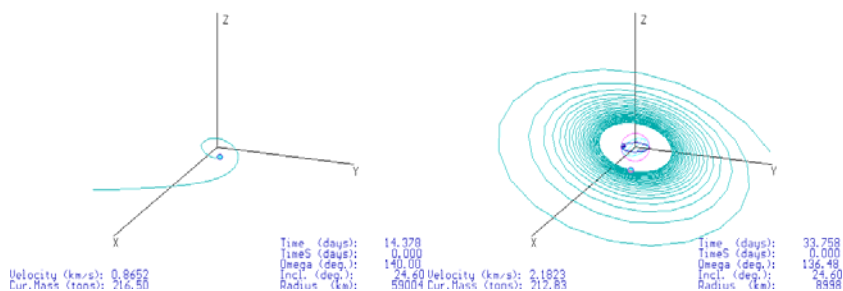


Рис. 3.47. Этапы торможения в сфере действия Марса

Как уже отмечалось, проектно-баллистические параметры миссии существенно изменяются при учете эллиптичности и некомпланарности орбит планет. На рис. 3.48 показаны зависимости расхода рабочего тела на оптимальной траектории перелета Земля-Мартс от оптимальных дат старта с учетом эллиптичности и некомпланарности орбит планет. Результаты получены для КА с солнечной энергодвигательной установкой, стартовой массой 5500 кг и фиксированной длительности перелета 175 сут.

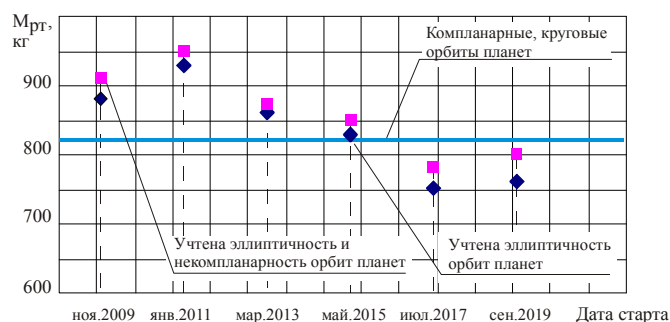


Рис. 3.48. Зависимость расхода рабочего тела в зависимости от дат старта для перелетов Земля-Мартс КА с солнечной энергодвигательной установкой с учетом эллиптичности и некомпланарности орбит. Получено по модели M i.5

Как отмечалось, при наборе скорости с помощью двигателя малой тяги, при достижении сферы действия Земли формируется небольшой гиперболический избыток скорости, величина которого зависит от проектных параметров КА. Используя описанную [29] методику ее оптимизации получены зависимости стартовой массы КА в зависимости от направления этого гиперболического избытка, показанные на рис. 3.49.

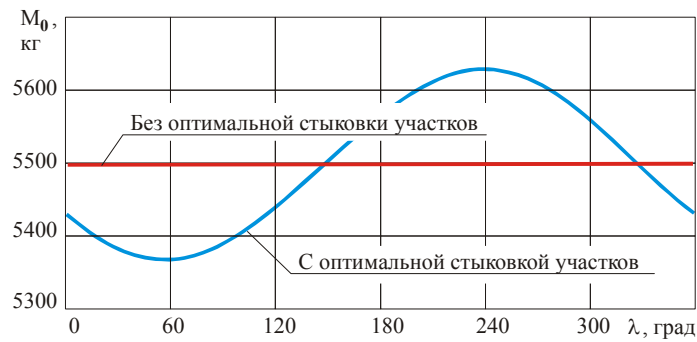


Рис. 3.49. Зависимость стартовой массы аппарата от направления гиперболического избытка скорости на границе сферы действия Земли

В табл. 3.9. представлены результаты итерационной оптимизации точек стыковки участков траектории вблизи сферы действия Земли.

С использованием описанных в предыдущих разделах моделей проектного облика и движения КА и методики комплексной оптимизации миссии, проведена оптимизация проектных параметров и даты старта проекта полета к Марсу малого межпланетного аппарата. В качестве начального приближения для решения задачи оптимизации использовались решения, полученные в рамках плоской модели движения (см. табл. 3.8).

Таблица 3.9. Изменение расхода рабочего тела на гео и гелиоцентрических участках в процессе оптимизации участка траектории вблизи сферы действия Земли

N итер.	Координаты КА в гелиоцентрической СК						Расход рабочего тела, кг		
	$r$ , млн.км	$\varphi$ , град	$V_r$ , км/с	$V_\varphi$ , км/с	$i$ , град	$\Omega$ , град	$M_{PT1}$	$M_{PT2}$	$M_{PT}$
0	152.345	280.0	-0.37	30.79	0.021	-1.41	1130	820	1950
1	152.057	280.9	0.48	30.28	0	-0.01	1160	614	1774
2	152.042	281.1	0.51	30.19	0	0	1164	618	1782
3	152.038	281.1	0.52	30.17	0	0	1163	617	1780

Для тех же масс полезной нагрузки и удельных массовых характеристик энергодвигательной установки были уточнены оптимальные проектные параметры КА. Из периода с 2009 года по 2019 год с учетом эллиптичности и некомпланарности орбит планет выбрана оптимальная дата старта экспедиции 24.07.2017 (см. рис. 3.48).

Для длительности экспедиции 350 сут и оптимальных проектных параметрах аппарата ( $P = 3.99$  Н,  $c = 69.97$  км/с) проведен сквозной расчет перелета Земля – Марс с оптимизацией точки стыковки участка набора скорости в сфере действия Земли и гелиоцентрического перелета и формированием рабочей ареоцентрической орбиты.

В процессе моделирования, учитывалось, что маневр набора параболической скорости в сфере действия Земли происходит с изменением наклона орбиты от 51.6 град (соответствует старту КА с космодрома Байконур) до 23.7 град. Маневр рассчитывался в рамках модели  $M 4.i$ , траектория набора параболической скорости с изменением наклона приведена на рис. 3.55.

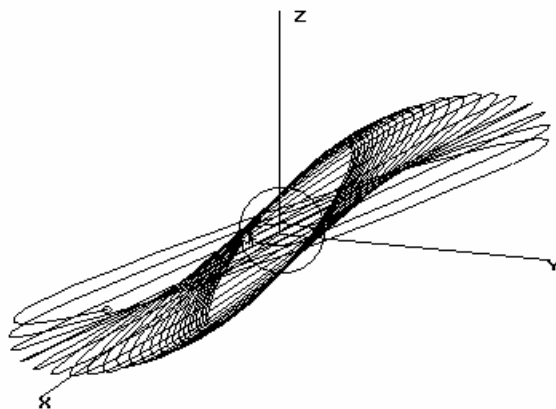


Рис. 3.50. Оптимальная траектория геоцентрического маневра набора скорости с изменением наклона орбиты

На рис. 3.51 отображена траектория гелиоцентрического перелета КА с учетом эллиптичности и некомпланарности орбит планет. На этапах торможения и формирования рабочей ареоцентрической орбиты

возможные погрешности подхода аппарата к Марсу не учитывались. На участке выхода из сферы действия Земли проводилась оптимизация направления планетоцентрической скорости КА.

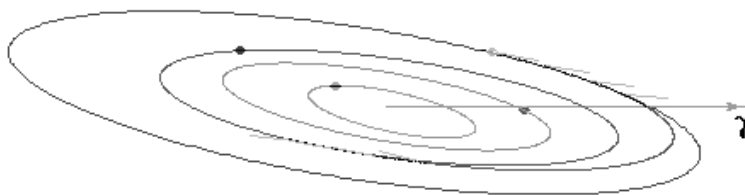


Рис. 3.51. Оптимальная траектория перелета Земля-Марс с учетом эллиптичности и некомпланарности орбит планет

Основные проектно-баллистические характеристики миссии, рассчитанной в рамках модели М 4.5, показаны в табл. 3.10. Значительное отклонение стартовой массы аппарата от массы, рассчитанной в рамках модели М 2.4, доказывает необходимость использования этой модели на этапе уточненной оценки проекта. Однако оптимальные проектные параметры аппарата и дата старта проекта изменяются слабо, что доказывает возможность использования результатов расчета по модели М 2.4 в качестве начального приближения для оценки оптимальных проектных и баллистических параметров системы.

Таблица 3.10. Основные проектно-баллистические характеристика миссии Земля – Марс, полученные в рамках модели М 2.4 с учетом эллиптичности и некомпланарности орбит планет

$i$	Маневр	Дата старта – финиша	$T_i$ , сут	$M(T_{i-1})$ , кг	$M(T_i)$ , кг	$M_{PTi}$ , кг
1	Набор скорости у Земли	24.07.2017 – 03.09.2017	40.5	5380	4117	1163
2	Перелет Земля – Марс	03.09.2019 – 07.05.2018	248.5	4117	3500	617
3	Торможение у Марса	07.05.2018 – 07.07.2018	60.0	3500	3370	130
8	Суммарные параметры	24.07.2017 – 07.07.2018	350.0	5380	3370	2010

### **3.4.6. Результаты выбора баллистической схемы межпланетного перелета на основе многокритериальной оптимизации**

Разработка любого проекта начинается с определения целей этого проекта, его основных характеристик и имеющихся для этого средств.

Рассмотрим пошагово разработку миссии: от цели - до основных проектно-баллистических характеристик КА на примере миссии исследования автоматическим КА поверхности Марса (спутник наблюдения).

Как отмечалось ранее, выбор любой баллистической схемы межпланетного перелета начинается с выбора РН. Анализ характеристик РН проводится по трем основным критериям: масса выводимой полезной нагрузки, стоимость вывода ПГ на опорную орбиту и надежность РН.

- При анализе характеристик РН сначала выбирается РН среднего класса "Союз-2", проводится проектно баллистический анализ возможных баллистических схем перелета. Если в результате расчетов масса ПН окажется меньше желаемой  $M_{ПГ}$ , то рассматривается следующий, - более тяжелый класс РН (Протон и т.п.).

- Затем определяется длительность экспедиции на основании ограничений ресурса элементов входящих в состав КА или соображений самого разработчика проекта. Например:  $T_{\Sigma} \leq T_1 = 600 = 600$  сут.

- Определяются максимально возможные затраты на осуществление проекта ( $C_1$ ).

- Определяется интервал старта, с учетом временных затрат на разработку и изготовление КА. Если рассматривать текущее время - то интервал старта находится в 2010-2011 гг.

- Рассматривается возможность использования ХРБ на участке разгона. (ХРБ "Фрегат").

- Рассматривается тип основной ДУ. Можно использовать уже имеющиеся ЭРД, надежные и прошедшие испытания СПД-100.

- Выбирается тип СБ. Наилучшим вариантом на сегодняшний день являются СБ с элементами на основе GaAs.

Когда определились с основными характеристиками и ограничениями, переходим к решению проектно-баллистической задачи

для различных вариантов баллистических схем. Выбор конкретной баллистической схемы перелета будет производиться с учетом ограничений ( $M_{ПГ}$ ,  $T_{\Sigma}$ ,  $C_1$ ) и значения основных критериев стоимости  $C_{\Sigma}$  длительности перелета  $T_{\Sigma}$ , массы полезной нагрузки  $M_{ПГ}$  и надежности успешного выполнения миссии  $P$ .

Рассмотрим некоторые баллистические схемы на примере перелета «Земля-Марс».

*Вариант №1.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью перспективного ХРБ.

*Вариант №2.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью перспективного ХРБ. Участок разгона предусматривает осуществление гравитационного маневра у Луны.

*Вариант №3.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью КДУ. Участок разгона формируется с помощью ХРБ «Фрегат», дальнейший перелет осуществляется с помощью блока ЭРДУ (СПД-100).

*Вариант №4.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью КДУ. Участок разгона формируется с помощью ХРБ «Фрегат», дальнейший перелет осуществляется с помощью блока ЭРДУ (СПД-100). Участок разгона предусматривает осуществление гравитационного маневра у Луны.

*Вариант №5.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью КДУ. Участок разгона разделяется на два, один из которых формируется с помощью ХРБ «Фрегат» до высокоэллиптической орбиты, доразгон и перелет осуществляется с помощью блока ЭРДУ (СПД-100).

*Вариант №6.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью ЭРДУ. Участок разгона оптимален по быстродействию.

*Вариант №7.* Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью ЭРДУ. Участок разгона оптимален по расходу РТ.



Вариант №8. Вывод КА на опорную орбиту осуществляется с помощью РН «Союз-2». Перелет осуществляется с помощью ЭРДУ. Участок разгона оптимален по расходу РТ. Участок разгона предусматривает осуществление гравитационного маневра у Луны.

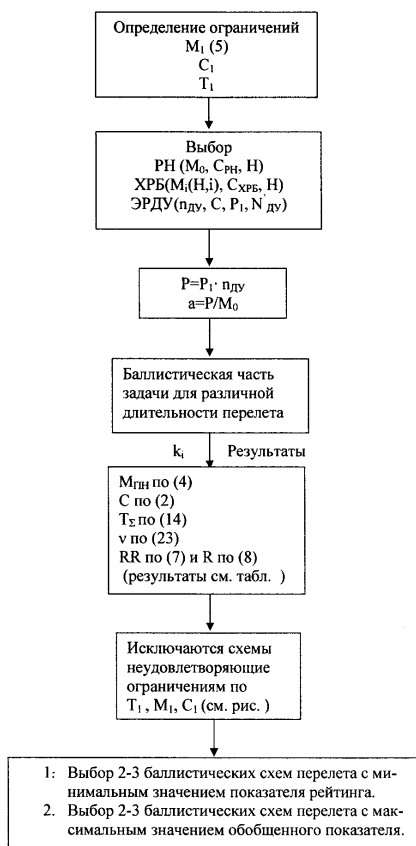


Рис. 3.52. Алгоритм проектирования баллистических схем межпланетного перелета КА с ЭРДУ

Процесс расчета и оптимизации основных параметров различных баллистических схем можно осуществлять с использованием алгоритма (рис. 3.52).

Для оценки преимущества перспективных проектов КА с ЭРДУ по сравнению с давно существующими проектами КА, у которых в качестве

ОДУ используется ХРБ, составляется сводная таблица оценки баллистических схем. Результаты моделирования представлены в табл. 3.11. Одновременно с баллистическими параметрами в сводную таблицу вносим коэффициенты, отражающие степень вклада каждого из показателей в составную «эффективность проекта»  $RR_j$ .

$$RR_j = k_1 \cdot RT_j + k_2 \cdot RM_j + k_3 \cdot RC_j + k_4 \cdot RP_j, \quad (3.50)$$

где  $k_i$  - весовые коэффициенты показателей эффективности ( $\sum k_i = 1$ , принимались равными:  $k_1 = 0.2$ ,  $k_2 = 0.3$ ,  $k_3 = 0.3$ ,  $k_4 = 0.2$ ),

$RT_j = \frac{\max_j T_j}{T_j}$  - показатель эффективности по длительности перелета,

$RM_j = \frac{M_{ПГj}}{\max_j M_{ПГj}}$  - показатель эффективности по массе ПГ,

$RC_j = \frac{\min_j C_j}{C_j}$  - по стоимости проекта,  $RP_j = \frac{P_j}{\max_j P_j}$  - показатель

успешного выполнения миссии. Показатель успешного выполнения  $j$  миссии равен вероятности ее успешного завершения:

$$P = P_1 \cdot P_2 \cdot P_3, \quad (3.51)$$

где  $P_1$  - вероятность успешного выведения КА ракетносителем на опорную орбиту или разгона,  $P_2$  - показатель успешной работы ЭРДУ,  $P_3$  - показатель успешной работы энергосистемы.

Если для набора параболической скорости используется химический разгонный блок (ХРБ), то

$$P_1 = P_{РН} \cdot P_{ХРБ}, \quad (3.52)$$

где  $P_{РН}$  и  $P_{ХРБ}$  - показатели надежности РН и ХРБ соответственно. Если разгон осуществляется без ХРБ, то  $P_{ХРБ} = 1$ .

$$P_2 = \left(1 - \frac{\tau}{3125}\right) \cdot \left(1 - \frac{\delta}{400}\right), \quad (3.53)$$

где  $\tau \in [0; 625]$  - длительность работы ЭРД, сут;  $\delta \in [0; 40]$  - количество включение – выключений ДУ.

$$P_3 = \begin{cases} 0.967 - 0.00033 \cdot S_{CB} & \text{- для СБ, } S_{CB} \in [50; 200] \text{ м}^2. \\ 0.871 - 0.0007 \cdot H & \text{- для ядерной, } P \in [30; 100] \text{ кВт} \end{cases}$$

Таблица 3.11. Результаты моделирования альтернативных баллистических схем перелета Земля-Марс

№ вар.	Даты старта и финиша	$T_{РАД}$ , сут	$T_{\Sigma}$ , сут	$R_1$ (по $T_{\Sigma}$ )	$RT$	$M_0$ , кг	$M_{ПГ}$ , кг	$R_2$ (по $\frac{M_{ПГ}}{M_0}$ )	$RM$
1	27.11.2011 – 13.09.2012	-	295	1	2.034	7250	214.2	17	0.11
2	28.11.2011 – 14.09.2012	-	295	1	2.034	7250	650.6	16	0.32
3	15.08.2009 – 22.09.2010	-	400	2	1.500	7250	1011	15	0.50
	01.08.2009 – 26.10.2010	-	450	3	1.333		1054	14	0.52
	16.08.2009 – 28.02.2011	-	600	6	1.000		1077	13	0.53
4	13.08.2009 – 20.09.2010	-	400	2	1.500	7250	1221	12	0.60
	02.08.2009 – 25.10.2010	-	450	3	1.333		1269	11	0.63
	17.08.2010 – 27.02.2012	-	600	6	1.000		0.180	10	0.65
5	03.08.2009 – 10.09.2011	-	400	2	1.500	7250	1507	9	0.75
	23.08.2009 – 15.10.2010	-	450	3	1.333		1528	8	0.76
	06.08.2010 – 21.10.2010	-	500	4	1.200		1554	7	0.77
6	29.02.2009 – 07.09.2010	52	550	5	1.091	7250	1822	6	0.90
	02.03.2009 – 22.10.2010	52	600	6	1.000		1901	5	0.94
7	29.02.2009 – 07.09.2010	65	550	5	1.091	7250	1905	4	0.94
	02.03.2010 – 22.10.2011	65	600	6	1.000		1942	3	0.96
8	01.03.2009 – 09.09.2010	65	550	5	1.091	7250	1985	2	0.98
	04.03.2010	65	600	6	1.000		2022	1	1.00

С использованием описанной методики оценивались затраты на выполнение миссии и вероятность ее успешного выполнения. Эти результаты и показатели общей эффективности рассматриваемых баллистических схем приведены в табл. 3.12.

Таблица 3.12. Результаты оценки затрат и общей эффективности миссий для альтернативных баллистических схем перелета Земля-Мартс

№ вар.	$C$ , млн. \$	$T_{\Sigma}$ , сут	$R_3$	$RC$	$P$	$R_4$	$RP$	$R$	$RR$
1	132.7	295	2	0.96	0.965	1	1.00	21	0.93
2	135.7	295	3	0.93	0.950	2	0.98	22	0.98
3	126.8	400	1	1.00	0.854	3	0.88	21	0.93
		450	1	1.00	0.850	5	0.88	23	0.90
		600	1	1.00	0.847	7	0.88	27	0.84
4	135.7	400	3	0.93	0.852	4	0.88	21	0.94
		450	3	0.93	0.849	6	0.88	23	0.91
		600	3	0.93	0.845	8	0.88	27	0.85
5	150.0	400	4	0.85	0.830	9	0.86	24	0.95
		450	4	0.85	0.820	10	0.85	25	0.92
		500	4	0.85	0.815	11	0.84	26	0.89
6	194.0	550	5	0.65	0.730	14	0.76	30	0.84
		600	5	0.65	0.710	15	0.74	31	0.82
7	201.0	550	6	0.63	0.740	12	0.77	27	0.84
		600	6	0.63	0.730	14	0.76	29	0.83
8	203.0	550	7	0.62	0.740	12	0.77	26	0.85
		600	7	0.62	0.737	13	0.76	27	0.84

Анализируя полученные результаты моделирования можно сделать следующие выводы:

- Доставить максимальную массу ПН на ареоцентрическую орбиту, можно используя 8-ю баллистическую схему. Однако это самая дорогая, сложная и длительная баллистическая схема.
- Самый быстрый перелет можно осуществить с помощью схем 1, 2 с ХРБ.
- Наименьшие затраты на миссию имеют схемы 7, 8 с КДУ.

При анализе результатов моделирования в таблице можно выделить приоритетные схемы: 1, 2, 3, 4 и 5. Оба метода сходятся в том, что данные схемы наиболее эффективны. После учета ограничений на минимальную массу полезной нагрузки  $M_{ПН} \geq 836$  кг, длительность перелета  $T_{\Sigma} = 600$  сут, максимальной стоимости миссии  $C = 180$  млн.\$, минимальной

надежности миссии  $P = 0.7$  (см. рис. 3.58, 3.59) для более глубокого анализа остались три баллистические схемы перелета: 3, 4, и 5. Наиболее перспективными являются эти схемы перелета при длительности 400 сут.

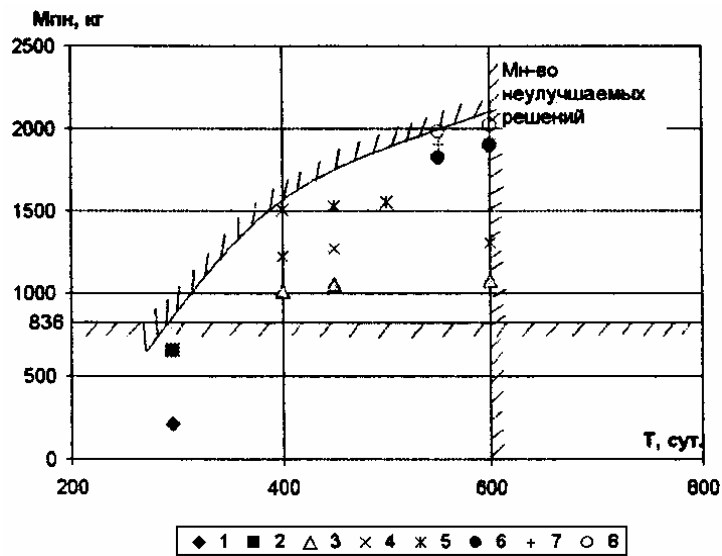


Рис. 3.58. Баллистические схемы перелета на поле ограничений: масса ПН – длительность перелета

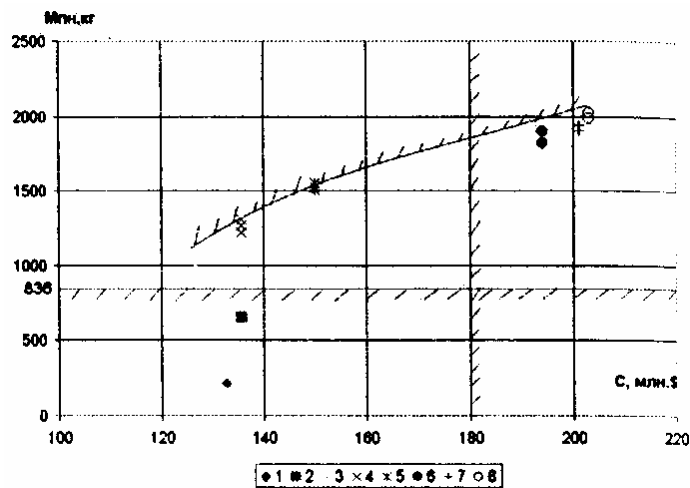


Рис. 3.59. Баллистические схемы на поле ограничений: масса ПН – затраты на осуществление проекта

Из таблицы несложно сделать вывод о значительном преимуществе схем с ЭРДУ перед схемами с ХРБ. Однако нельзя забывать о недостатках схем, когда на участке разгона используется ЭРДУ - это деградация элементов конструкции КА. В этом отношении схемы с КДУ имеют большое преимущество.

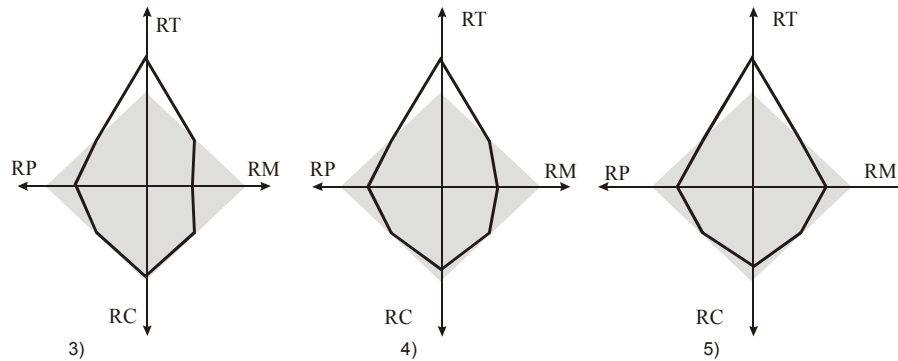


Рис. 3.60. Лепестковые диаграммы, характеризующая наиболее перспективные 3, 4, и 5 баллистических схем миссий, длительностью 400 сут

Полученные значения показателей, характеризующих баллистические схемы можно отобразить на лепестковой диаграмме, (рис. 3.60) которая дает наглядное представление о параметрах (критериях), входящих в систему и позволяет сделать выбор конкретной баллистической схемы.

## СПИСОК ЛИТЕРАТУРЫ

1. Лебедев А.А. Введение в анализ и синтез систем. - М.: Изд-во МАИ, 2001. – 352 с.
2. Авдеевский В.С., Успенский Г.Р. Космическая индустрия. – М. Машиностроение, 1989.
3. Аджимамудов Г.Г. Комплексное использование методов и средств дистанционной индикации для учета и контроля состояния окружающей среды. Ракетостроение, т.6. – М.: ВИНТИ, 1978, с. 89-288.
4. Аппазов Р.Ф., Сытин О.Г. Методы проектирования траекторий носителей и спутников Земли. – М.: Наука, 1987.
5. Ахундов В.М. Системный анализ в экономических исследованиях: Учебник. – М.: МАИ, 1987.
6. Военный энциклопедический словарь. – М.: Воениздат, 1986.
7. Вопросы математической теории надежности /Под ред. Б.В. Гнеденко. – М.: Радио и связь, 1983.
8. Гуткин Л.С. Проектирование радиосистем и радиоустройств: Учебное пособие для вузов. – М.: Радио и связь, 1986.
10. Гуцин В.Н. Проектирование искусственных спутников Земли: тексты лекций. – М.: Изд-во МАИ, 1999.
11. Гонин Г.Б. Космические съемки Земли. – М.: Наука, 1989.
12. Вентцель Е.С. Исследование операций. Задачи, принципы, методология.- М. Высш. шк., 2001.
13. Майзер Х., Эйджин Н., Тролл Р. Исследование операций: в 2-х томах. Пер. с англ./Под ред. Дж. Моудера, С.Элмаграби. -М.: Мир, 1981.
14. Зайченко Ю.П.. Исследование операций - Киев: Вища школа, 1979.
15. Волков И.К., Загоруйко Е.А. Исследование операций-М.: Изд-во МГТУ им. Н.Э.Баумана, 2002.
16. Вентцель Е.С. Теория вероятностей-М.: Высш. шк., 1999.
17. Катулев А.Н., Северцев Н.А. Исследование операций: принципы принятия решений и обеспечение безопасности. – М.: Физ.-мат. лит-ра, 2000.
18. Отчет о НИР: Транспортные модули на базе комбинации

современных химических двигателей и электроракетных двигательных установок для транспортных перевозок «орбита-орбита». НИИПМЭ, 1998.

19. Салмин В.В. Оптимизация космических перелетов с малой тягой: Проблемы совместного управления траекторным и угловым движением. – М. Машиностроение, 1987.

20. Сихарулидзе Ю.Г. Баллистика летательных аппаратов.– М.: Наука. Главная ред. физико-математической литературы, 1982.

21. Лебедев В.Н. Расчет движения космического аппарата с малой тягой. М.: ВЦ АН СССР, 1968.

22. Справочное руководство по небесной механике и астродинамике. Под ред. Г.Н.Дубошина. – М.: Наука, 1976.

23. Спутниковые системы мониторинга. Анализ, синтез и управление / В.В.Малышев, В.Т.Бобронников, О.П.Нестеренко, А.В.Федоров; Под ред. В.В.Мальшева. – М.: Изд-во МАИ, 2000.

24. Космические аппараты систем зондирования поверхности Земли: Математические модели повышения эффективности КА / А.В.Соллогуб, Г.П.Аншаков, В.В.Данилов; Под ред. Д.И.Козлова. – М. Машиностроение, 1993.

25. Салмин В.В., Соколов В.О. Приближенный расчет маневров формирования орбит спутника Земли с двигателем малой тяги // Космические исследования Т 27, №1, 1989 г.

26. Васильев В.В., Салмин В.В. Выбор универсальных параметров двигателя малой тяги, предназначенного для поддержания орбиты спутника Земли // Космические исследования Т 22, №6, 1984 г

27. В.В.Салмин, В.В.Васильев, С.А.Ишков, В.А.Романенко, О.Л.Старинова, В.О.Соколов, В.В.Юрин Приближенные методы расчета оптимальных перелетов космических аппаратов с двигателями малой тяги. Часть 1, 2. Вестник СГАУ, Самара, СГАУ, 2007, №1, 2.

28. В.В.Салмин, С.А.Ишков, О.Л.Старинова Методы решения вариационных задач механики космического полета с малой тягой. – Самара: Издательство Самарского научного центра РАН, 2006

29. В.В.Салмин, О.Л.Старинова Оптимизация межпланетных перелетов КА с двигателями малой тяги с учетом эллиптичности и некомпланарности орбит планет. Космические исследования .- 2001.-Т.39, N 1



Учебное издание

*Салмин Вадим Викторович,  
Кучеров Александр Степанович,  
Старинова Ольга Леонардовна,  
Прохоров Александр Георгиевич*

**МЕТОДЫ СИСТЕМНОГО АНАЛИЗА И ИССЛЕДОВАНИЯ ОПЕРАЦИЙ  
В ЗАДАЧАХ ПРОЕКТИРОВАНИЯ ЛЕТАТЕЛЬНЫХ АППАРАТОВ**

*Учебное пособие*

Редакторская обработка Н. С. К у п р и я н о в а  
Корректорская обработка Н. В. П р я д и л ь н и к о в а  
Доверстка В. С. Т е л е п о в а

Подписано в печать 27.12.07. Формат 60x84 1/16.

Бумага офсетная. Печать офсетная.

Печ. л. 17,0.

Тираж 120 экз. Заказ . ИП-143ж/2007

Самарский государственный  
аэрокосмический университет.  
443086 Самара, Московское шоссе, 34.

---

Изд-во Самарского государственного  
аэрокосмического университета.  
443086 Самара, Московское шоссе, 34.