

National Research Tomsk Polytechnic University, Tomsk Polytechnic University

Математическое моделирование в приборных системах

Профессор отделения ОКД,
д.ф.-м.н., профессор
В.А.Перминов

- **Тема 2: Классификация моделей.**

- Целью моделирования являются получение, обработка, представление и использование информации об объектах, которые взаимодействуют между собой и внешней средой; а модель здесь выступает как средство познания свойств и закономерности поведения объекта.

- Основные моменты моделирования:
- а) уточнение предметной области исследования, ее структуризация на задачи:
- б) выбор параметров и критериев оценки эффективности системы:
- в) подбор нужных математических моделей:

- г) уточнение деталей и целей анализа системы:
- д) синтезирование математических моделей, обеспечивающих достижение поставленных целей.

- Теорией моделирования является раздел науки, изучающий способы исследования свойств объектов-оригиналов, на основе замещения их другими объектами-моделями. В основе теории моделирования лежит теория подобия.

- Все модели можно разделить на два класса:
- вещественные,
- идеальные.
- **Вещественные** модели можно разделить на:
- натурные,
- физические,
- математические.
- Идеальные модели можно разделить на:
- наглядные,
- знаковые,
- математические.

Вещественные натурные модели - это реальные объекты, процессы и системы, над которыми выполняются эксперименты научные, технические и производственные.

Физические модели - это макеты, муляжи, воспроизводящие физические свойства оригиналов (кинематические, динамические, гидравлические, тепловые, электрические, световые модели).

- Вещественные математические - это аналоговые, структурные, геометрические, графические, цифровые и кибернетические модели.

- **Идеальные модели**– это абстрактные образы замещаемых объектов.
- Два типа идеального моделирования:
- *интуитивное;*
- *знаковое.*
- **+Интуитивное моделирование**
используется для отражения окружающего мира и предсказания его реакций. Как оно осуществляется – неизвестно.

- **Знаковое моделирование**– это использование в качестве моделей знаков/символов: схемы; графики; чертежи; тексты на разных языках, включая математические формулы и теории, и т.п.
- *Математическое моделирование*– наиболее важный вид знакового моделирования.

- **Аналоговыми моделями** называют системы, имеющие физическую природу, отличающуюся от оригинала, но сходные с оригиналом процессы функционирования. Для создания аналоговой модели требуется наличие математического описания изучаемой системы. В качестве аналоговых моделей используются механические, гидравлические, пневматические и электрические системы.

- Аналоговое моделирование использует при исследовании приборных средств на уровне логических элементов и электрических цепей, а так же на системном уровне, когда функционирование системы описывается, например, дифференциальными или алгебраическими уравнениями.

- .

Используя эти соотношения, по зависимостям, обнаруженным в модели, строятся зависимости, справедливые для реального объекта. При физическом моделировании может использоваться не одна, а несколько моделей. Недостатками аналогового моделирования являются сложность создания модели и низкая точность результатов

- Математическое моделирование - это средство изучения реального объекта, процесса или системы путем их замены математической моделью, более удобной для экспериментального исследования с помощью ЭВМ.

Математическая модель является приближенным представлением реальных объектов, процессов или систем, выраженным в математических терминах и сохраняющим существенные черты оригинала.

Математические модели в количественной форме, с помощью логико-математических конструкций, описывают основные свойства объекта, процесса или системы, его параметры, внутренние и внешние связи.

- В общем случае математическая модель реального объекта, процесса или системы представляется в виде системы функционалов
- $\Phi_i (X, Y, Z, t) = 0$,
- где X - вектор входных переменных,
 $X = [x_1, x_2, x_3, \dots, x_N]^t$,
- Y - вектор выходных переменных, $Y = [y_1, y_2, y_3, \dots, y_N]^t$,
- Z - вектор внешних воздействий, $Z = [z_1, z_2, z_3, \dots, z_N]^t$,
- t - координата времени.

- Построение математической модели заключается:
- в определении связей между теми или иными процессами и явлениями,
- создании математического аппарата, позволяющего выразить количественно и качественно связь между теми или иными процессами и явлениями, между интересующими специалиста физическими величинами, и факторами, влияющими на конечный результат.

- По принципам построения математические модели разделяют на:
 - аналитические;
 - имитационные.
- В аналитических моделях процессы функционирования реальных объектов, процессов или систем записываются в виде явных функциональных зависимостей.

- Аналитическая модель разделяется на типы в зависимости от математической проблемы:
- уравнения (алгебраические, трансцендентные, дифференциальные, интегральные),
- аппроксимационные задачи (интерполяция, экстраполяция, численное интегрирование и дифференцирование),
- задачи оптимизации,
- стохастические проблемы

- В зависимости от характера исследуемых реальных процессов и систем математические модели могут быть:
- детерминированные,
- стохастические.

В детерминированных моделях предполагается отсутствие всяких случайных воздействий, элементы модели (переменные, математические связи) достаточно точно установленные, поведение системы можно точно определить. При построении детерминированных моделей чаще всего используются алгебраические уравнения, интегральные уравнения, матричная алгебра.

- Стохастическая модель учитывает случайный характер процессов в исследуемых объектах и системах, который описывается методами теории вероятности и математической статистики.

- По виду входной информации модели разделяются на:
 - непрерывные,
 - дискретные.

- Если информация и параметры являются непрерывными, а математические связи устойчивы, то модель - непрерывная. И наоборот, если информация и параметры - дискретны, а связи неустойчивы, то и математическая модель - дискретная.

- По поведению моделей во времени они разделяются на:
- статические,
- динамические.
- Статические модели описывают поведение объекта, процесса или системы в какой-либо момент времени. Динамические модели отражают поведение объекта, процесса или системы во времени.

Имитационная модель(ИМ) – это совокупность описания системы и внешних воздействий, алгоритмов функционирования системы или правил изменения состояния системы под влиянием внешних и внутренних возмущений. Эти алгоритмы и правила не дают возможности использования имеющихся математических методов аналитического и численного решения, но позволяют имитировать процесс функционирования системы и производить вычисления интересующих характеристик.

Имитационные модели могут быть созданы для гораздо более широкого класса объектов и процессов, чем аналитические и численные. Поскольку для реализации имитационных моделей используются вычислительные системы и средства, средствами формализованного описания ИМ служат универсальные и специальные алгоритмические языки. ИМ в наибольшей степени подходят для исследований на системном уровне.

- **Имитационная модель** — логико-математическое описание объекта, которое может быть использовано для экспериментирования на компьютере в целях проектирования, анализа и оценки функционирования объекта.

- ИМИТАЦИОННАЯ МОДЕЛЬ [simulator] — математическая модель изучаемой системы, предназначенная для использования в процессе машинной имитации. Она является по существу программой для компьютера, а эксперимент над ней состоит в наблюдении за результатами расчетов по этой программе при различных задаваемых значениях вводимых экзогенных переменных.

- Имитационная модель является динамической моделью в том смысле, что в ней присутствует время — когда проигрывается серия вариантов развития исследуемого объекта. С другой стороны, имитационная модель, как правило, является адаптивной моделью, т.к. совершенствуется, уточняется в процессе использования.



Пример

Battelle — крупнейшая в мире некоммерческая независимая исследовательская организация. Она управляет лабораториями для Министерства энергетики и Министерства внутренней безопасности США, а также международной ядерной лабораторией в Великобритании.

В одном из проектов задачей Battelle было найти быстрые и эффективные методы реагирования на чрезвычайные ситуации различного характера.

В частности, исследователи должны были определить эффективность мер реагирования в случае **детонации ядерного устройства**, сравнив **планы эвакуации населения** и его укрытия на месте ЧС.

Цель таких мер реагирования – уменьшить полученные населением дозы радиации.

Моделирование чрезвычайной ситуации представляет множество трудностей.

Нужно учитывать физические последствия ЧС, многие возможные сценарии и географические векторы угроз.

К тому же, нужно учитывать тот факт, что меры по урегулированию ЧС редко проводятся в жизнь так, как запланировано, т.к. в их проведение вмешивается человеческий фактор.

Исследователи использовали имитационное моделирование, т.к. оно дает возможность рассмотреть множество случайных сценариев.

Детерминистские же модели, напротив, не могут учесть непредсказуемость реакции людей и не дают возможности протестировать разные сценарии.

- Специалисты Battelle выбрали для задачи имитационного моделирования инструмент AnyLogic, т.к. он использовался в других проектах организации, в том числе:

- **Здравоохранение:** управление ресурсами и моделирование работы больниц, контроль распространения инфекций.
- **Экономическое развитие и прогнозирование** развития отраслевых кластеров.
- **Логистика и техобслуживание** парка транспортных средств.
- **Национальная безопасность** и меры реагирования на ЧС.

Возможности моделирования в AnyLogic позволили исследователям Battelle отразить в модели основную динамику чрезвычайной ситуации. Неопределенность и непрогнозируемость – главное, что нужно учитывать при моделировании поведения людей в таких случаях.

Модель включала дорожную сеть, транспортные средства, водителей и чрезвычайные происшествия. Дорожная сеть была построена на ГИС-карте, куда были загружены данные местной дорожной организации (ограничения скорости, пропускная способность). По мере развития чрезвычайной ситуации в дорожной сети посредством динамических событий происходили изменения – например, затопления дорог и разрушения мостов.

Физические ограничения техники были заданы. Для моделирования поведения водителей использовались данные из исследований – была отражена разница между нормальным поведением и действиями во время всеобщей эвакуации. Движение автомобилей было связано с состоянием водителей – машины останавливались, когда водители теряли возможность управлять ими.

Специалисты использовали динамические контуры последствий аварии, чтобы измерить их последствия по регионам. Контуры обновлялись в реальном времени на основе прогнозов погоды, рельефа и т.д. Различные наборы контуров давали данные по почти любому варианту последствий ЧС (уровни наводнения, развития пожаров, распределения загрязнения/выпадения радиоактивных осадков и др.).

В сценарии с детонацией ядерного устройства использовались два набора контуров: радиус взрывной волны (огненный шар и зона с высоким избыточным давлением) и распределение радиоактивных осадков (уровни загрязнения в воздухе и на земле от разных типов элементов).

- Имитационная модель, сравнив сценарии немедленной эвакуации и укрытия на месте аварии, показала, что при больших взрывах укрытие снижает полученные дозы радиации и количество случаев сильного радиационного заражения.
- Также модель предоставила данные для проведения дальнейших исследований по выбору лучших стратегий реагирования на ЧС.

Специалисты Battelle реализовали в одной модели разные типы, отвечающих за реагирование на ЧС, разные сценарии реагирования и возможность изменять компоненты модели (выбирать разные локации для одного сценария или разные сценарии для одной локации). Это позволило найти практические решения для реагирования на разные типы кризисов и природных чрезвычайных ситуаций.

Особенности построения математических моделей

- Для построения математической модели необходимо:
- тщательно проанализировать реальный объект или процесс;
- выделить его наиболее существенные черты и свойства;
- определить переменные, т.е. параметры, значения которых влияют на основные черты и свойства объекта;

- описать зависимость основных свойств объекта, процесса или системы от значения переменных с помощью логико-математических соотношений (уравнения, равенства, неравенства, логико-математические конструкции);
- выделить внутренние связи объекта, процесса или системы с помощью ограничений, уравнений, равенств, неравенств, логико-математических конструкций;
- определить внешние связи и описать их с помощью ограничений, уравнений, равенств, неравенств, логико-математических конструкций.

Математическое моделирование, кроме исследования объекта, процесса или системы и составления их математического описания, также включает:

- построение алгоритма, моделирующего поведение объекта, процесса или системы;
- проверка адекватности модели и объекта, процесса или системы на основе вычислительного и натурального эксперимента;
- корректировка модели;
- использование модели

Математическое описание исследуемых процессов и систем зависит от:

- природы реального процесса или системы и составляется на основе законов физики, химии, механики, термодинамики, гидродинамики, электротехники, теории пластичности, теории упругости и т.д.
- требуемой достоверности и точности изучения и исследования реальных процессов и систем.

На этапе выбора математической модели устанавливаются:

1. линейность и нелинейность объекта, процесса или системы,
2. динамичность или статичность,
3. стационарность или нестационарность,
4. степень детерминированности исследуемого объекта или процесса.

При математическом моделировании сознательно отвлекаются от конкретной физической природы объектов, процессов или систем и, в основном, сосредотачиваются на изучении количественных зависимостей между величинами, описывающими эти процессы.

Математическая модель никогда не бывает полностью тождественна рассматриваемому объекту, процессу или системе.

Основанная на упрощении, идеализации она является приближенным описанием объекта.

Результаты, полученные при анализе модели, носят приближенный характер.

Их точность определяется степенью адекватности (соответствия) модели и объекта.

Построение математической модели обычно начинается с построения и анализа простейшей, наиболее грубой математической модели рассматриваемого объекта, процесса или системы.

В дальнейшем, в случае необходимости, модель уточняется, делается ее соответствие объекту более полным.

- Чем выше требования к точности результатов решения задачи, тем больше необходимость учитывать при построении математической модели особенности изучаемого объекта, процесса или системы.
- Однако, здесь важно во время остановиться, так как сложная математическая модель может превратиться в трудно разрешимую задачу.

- Наиболее просто строится модель, когда хорошо известны законы, определяющие поведение и свойства объекта, процесса или системы, и имеется большой практический опыт их применения.

- Построение математической модели в прикладных задачах – один из наиболее сложных и ответственных этапов работы. Построить модель – значит решить проблему более, чем наполовину.

- Трудность данного этапа состоит в том, что он требует соединения математических и специальных знаний.
- Поэтому очень важно, чтобы при решении прикладных задач математики обладали специальными знаниями об объекте, а их партнеры, специалисты, – определенной математической культурой, опытом исследования в своей области, знанием ЭВМ и программирования.

Компьютерное моделирование и вычислительный эксперимент. Решение математических моделей

- Компьютерное моделирование как новый метод научных исследований основывается на:
 - построении математических моделей для описания изучаемых процессов;
 - использовании новейших вычислительных машин, обладающих высоким быстродействием (миллионы операций в секунду) и способных вести диалог с человеком.

Суть компьютерного моделирования состоит в следующем:

1. на основе математической модели с помощью ЭВМ проводится серия вычислительных экспериментов, т.е. исследуются свойства объектов или процессов,
2. находятся их оптимальные параметры и режимы работы,
3. уточняется модель.

Вычислительный эксперимент позволяет

- заменить дорогостоящий натурный эксперимент расчетами на ЭВМ,
- в короткие сроки и без значительных материальных затрат осуществить исследование большого числа вариантов проектируемого объекта или процесса для различных режимов его эксплуатации,
- это значительно сокращает сроки разработки сложных систем и их внедрение в производство

- Для проверки адекватности математической модели и реального объекта, процесса или системы результаты исследований на ЭВМ сравниваются с результатами эксперимента на опытном натурном образце.
- Результаты проверки используются для корректировки математической модели или решается вопрос о применимости построенной математической модели к проектированию либо исследованию заданных объектов, процессов или систем.

- Компьютерное моделирование и вычислительный эксперимент позволяют свести исследование "нематематического" объекта к решению математической задачи.

- Этим самым открывается возможность использования для его изучения хорошо разработанного математического аппарата в сочетании с мощной вычислительной техникой.
- На этом основано применение математики и ЭВМ для познания законов реального мира и их использования на практике.

- Итак, мы построили математическую модель рассматриваемого объекта, процесса или системы, т.е. представили прикладную задачу как математическую. После этого наступает второй этап решения прикладной задачи – поиск или разработка метода решения сформулированной математической задачи. Метод должен быть удобным для его реализации на ЭВМ, обеспечивать необходимое качество решения.

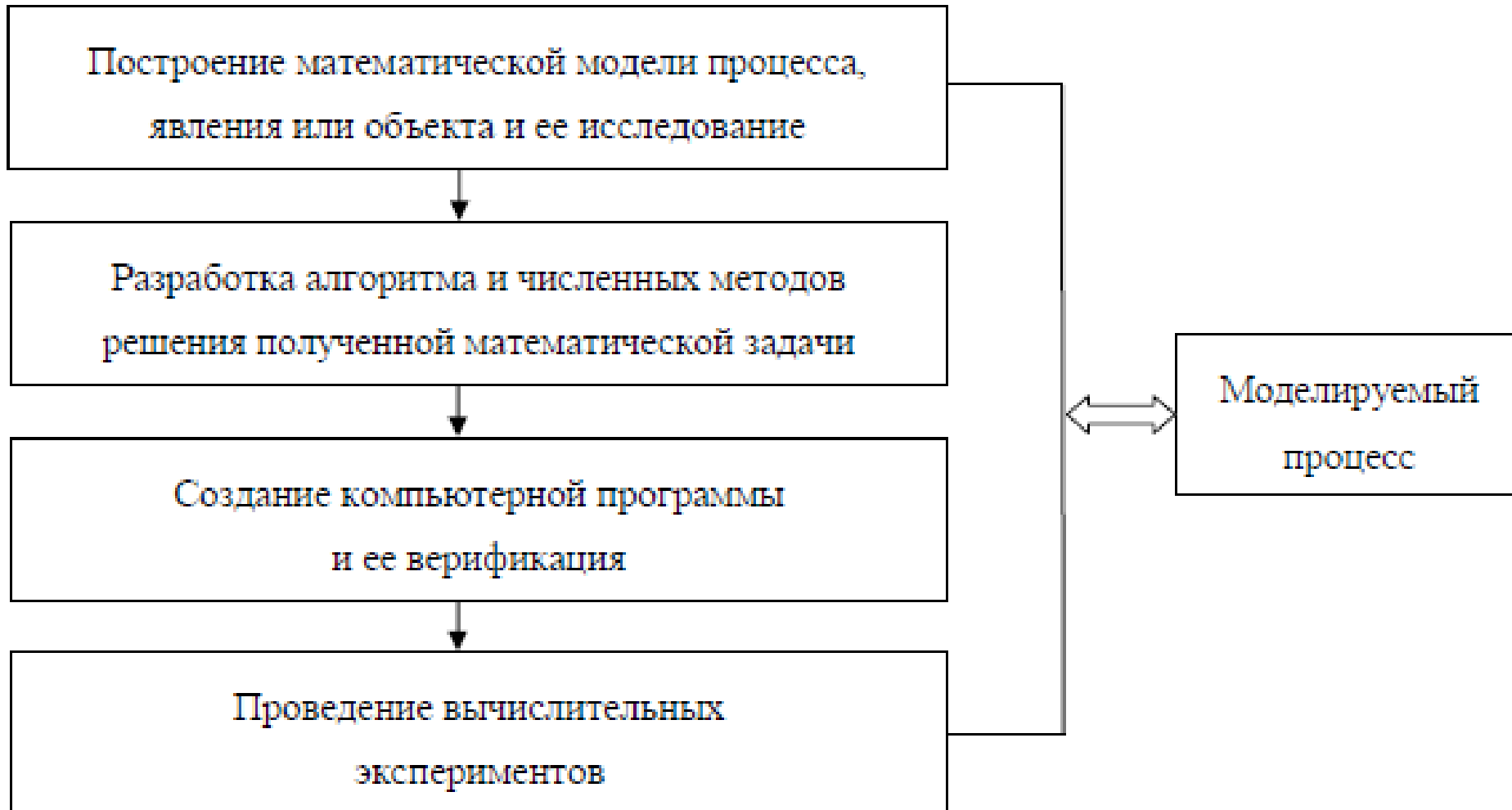
- Все методы решения математических задач можно разделить на 2 группы:
- точные методы решения задач;
- численные методы решения задач.
- В точных методах решения математических задач ответ удастся получить в виде формул.

- Для большинства задач, встречающихся на практике, точные методы решения или неизвестны, или дают очень громоздкие формулы.
- Однако, они не всегда являются необходимыми.
- Прикладную задачу можно считать практически решенной, если мы сумеем ее решить с нужной степенью точности.

- Для решения таких задач разработаны численные методы, в которых решение сложных математических задач сводится к последовательному выполнению большого числа простых арифметических операций.
- Непосредственная разработка численных методов относится к вычислительной математике.

- Таким образом, в прикладных задачах и при применении точных методов решения, и при применении численных методов решения результаты вычислений носят приближенный характер.
- Важно только добиться того, чтобы ошибки укладывались в рамки требуемой точности

Алгоритм решения задач методом математического моделирования



- **Приборные системы** являются сложными техническими объектами, предназначенными для:
- различного рода измерений; контроля, диагностики и управления техническими объектами и технологическими процессами; контроля окружающей природной и техногенной среды.
- Под **системой** будем понимать совокупность элементов, отношений, ограничений и связей между ними.

- Под исследуемым объектом понимается реально существующий объект различной природы или некоторый природный либо технологический процесс.
- Целью исследования объекта, в нашем случае **системы**, является изучение и описание свойств и связей составляющих элементов и откликов на внешние воздействия.

- Принципиально идея замены системы (оригинала) его математической моделью (образом) носит крайне общий характер, поэтому практически невозможно чётко определить функции математической модели.

- Можно сформулировать ряд групп функций (целевых ожиданий), предъявляемых к математическим моделям, которые могут быть переведены и в разряд требований.
- Целевые ожидания от математических моделей:

- 1) упорядочение и структурирование нечётких, невнятных и, зачастую, противоречивых понятий и взаимных отношений;
- 2) создание специализированных метаязыков, заменяющих традиционные языки, для использования в описании сложных систем с целью компактификации представлений, которые включают в себя совокупность принятых обозначений, блоков, уравнений, отношений и ограничений;

- 3) возможность трансформации математической модели системы в имитационную модель или симулятор с целью анализа влияния случайностей на поведение системы и адаптации субъекта измерения и контроля, использующего систему, к изменению внешних обстоятельств;
- 4) прогнозирование поведения оригиналов на основе анализа их образов (математических моделей) с ответом на вопрос или их совокупность – «что будет, если...»;

- 5) возможность проведения вычислительных экспериментов, имитируя реальные системы, в случаях невозможности или нецелесообразности проведения натурных экспериментов.

Тема 2: Математические модели приборных систем

- Математическое моделирование приборных систем является обязательной составляющей в процессе их проектирования на стадии научного обоснования их физической и технической реализуемости.
Экспериментальные же исследования, как правило, являются затратными по времени и материальным ресурсам, причём могут завести в тупик.

- ***2.1 Основные подходы к построению математических моделей систем***
- *Математическая схема* – это звено при переходе от содержательного к формализованному описанию процесса функционирования системы с учётом воздействия внешней среды, т.е. имеет место цепочка: описательная модель – математическая схема – имитационная модель.

- Каждая конкретная система S характеризуется набором свойств, под которыми понимаются величины, отображающие поведение моделируемого объекта (реальной системы) и учитываются условия её функционирования во взаимодействии с внешней средой (системой) E .
- Полнота моделирования регулируется, в основном, выбором границ «Система S – среда E ».

- Математическую модель (ММ) объекта моделирования, т.е. системы S можно представить в виде множества величин, описывающих процесс функционирования реальной системы и образующих в общем случае следующие подмножества:

- совокупность входных воздействий на S
 $x_i \in X, i=1, \dots, n_x;$
- совокупность воздействий внешней среды
 $v_l \in V, l=1, \dots, n_v;$
- совокупность внутренних (собственных) параметров системы $h_k \in H, k=1, \dots, n_h;$
- совокупность выходных характеристик системы $y_j \in H, j=1, \dots, n_y.$

- В перечисленных множествах можно выделить управляемые и неуправляемые величины.
- В общем случае непересекающиеся множества X , V , H , Y содержат как детерминированные так и стохастические составляющие.

- Входные воздействия E и внутренние параметры S являются *независимыми (экзогенными) переменными* $\bar{X}(t); \bar{V}(t); \bar{H}(t)$
- Выходные характеристики – *зависимые (эндогенные) переменные* $\bar{Y}(t)$
- Процесс функционирования S описывается оператором F_S :
- $$\bar{Y}(t) = F_S(\bar{X}(t), \bar{V}(t), \bar{H}(t), t) \tag{2.1}$$

- Где $\bar{Y}(t)$ - выходная траектория,
- F_S – закон функционирования S .
- F_S может быть функция, функционал, логические условия, алгоритм, таблица или словесное описание правил.

•

Алгоритм функционирования A_S — метод получения выходных характеристик $\bar{Y}(t)$ с учётом входных воздействий $\bar{X}(t); \bar{V}(t); \bar{H}(t)$. Очевидно, один и тот же F_S может быть реализован различными способами, т.е. с помощью множества различных A_S . ¶

- Соотношение (2.1) является математическим описанием поведения объекта S моделирования во времени t , т.е. отражает его *динамические свойства*. (2.1) – это динамическая модель системы S . Для статических условий ММ есть отображения X, V, H в Y :

$$\bar{Y} = f(\bar{X}, \bar{V}, \bar{H}) \quad (2.2)$$

- Соотношения (2.1), (2.2) могут быть заданы формулами, таблицами и т.д. Также соотношения в ряде случаев могут быть получены через свойства системы в конкретные моменты времени, называемые *состояниями*.
- Состояния системы S характеризуются векторами:

•

$\bar{Z}^l = (\bar{z}_1^l, \dots, \bar{z}_k^l)$ и $\bar{Z}^{ll} = (\bar{z}_1^{ll}, \dots, \bar{z}_k^{ll})$, где $z_1^l = z_1(t^l), \dots, z_k^l = z_k(t^l)$ в

МОМЕНТ $t^l \in (t_0, T)$; $z_1^{ll} = z_1(t^{ll}), \dots, z_k^{ll} = z_k(t^{ll})$ в МОМЕНТ

$t^{ll} \in (t_0, T)$ и т.д., $k=1, \dots, n_z$ ¶

$Z_1(t), Z_2(t), \dots, Z_k(t)$ — это координаты

точки в k -мерном фазовом пространстве.

Каждой реализации процесса будет соответствовать некоторая *фазовая траектория*. ¶

Совокупность всех возможных значений состояний $\{\bar{z}\}$ называется *пространством состояний* объекта моделирования Z , причём $z_k^o \in Z$. ¶

Состояние системы S в интервале времени $t_0 < t \leq T$ полностью определяется начальными условиями $\bar{z}^0 = (z_1^0, \dots, z_k^0)$, где $z_1^0 = z_1(t_0), \dots$; входными воздействиями $\bar{X}(t)$, внутренними параметрами $\bar{H}(t)$ и воздействиями внешней среды $\bar{V}(t)$, которые имели место за промежуток времени $t^* - t_0$ с помощью 2-х векторных уравнений:

$$\bar{Z}(t) = \Phi(\bar{Z}^0, \bar{X}, \bar{V}, \bar{h}, t); \quad (2.3)$$

$$\bar{Y}(t) = F(\bar{Z}, t) \quad (2.4)$$

иначе: $\bar{Y}(t) = F(\Phi(\bar{Z}^0, \bar{X}, \bar{V}, \bar{h}, t)).$ (2.5)

Время в модели S может рассматриваться на интервале моделирования (t_0, T) как непрерывное, так и дискретное, т.е. квантованное на отрезке длиной Δt .

Таким образом, под ММ объекта понимаем конечное множество переменных $\{\bar{X}, \bar{Z}, \bar{h}\}$ вместе с математическими связями между ними и характеристиками \bar{Y} .

Моделирование называется *детерминированным*, если операторы F, Φ – детерминированные. Детерминированное моделирование – частный случай стохастического моделирования. В практике моделирования объектов в области системного анализа на первичных этапах исследования рациональнее использовать типовые математические схемы: дифференциальные уравнения, конечные и вероятностные автоматы, СМО и т.д.

В стохастических моделях (при учёте случайного фактора) для представления систем с дискретным временем используются вероятностные автоматы, а для представления систем с непрерывным временем – системы массового обслуживания (СМО). Большое практическое значение при исследовании сложных индивидуальных управленческих систем, к которым относятся АСУ, имеют так называемые агрегативные модели.

Агрегативные модели (системы) позволяют описать широкий круг объектов исследования с отображением системного характера этих объектов. Именно при агрегативном описании сложный объект расчленяется на конечное число частей (подсистем), сохраняя при этом связи, обеспечивая взаимодействие частей.

2.2 Непрерывно-детерминированные модели (Д-схемы)

Если в модели системы не учитывается воздействие случайных факторов, а операторы переходов и выходов непрерывны (это означает, что малые изменения входных воздействий приводят к такого же порядка малым изменениям выходного воздействия и состояния системы), то состояния системы и выхода соответственно могут быть представлены в виде дифференциальных уравнений:

$$\frac{dz}{dt} = \mathbf{h}(z(t), x(t), t); \quad (2.6)$$

$$\mathbf{y} = \mathbf{g}(z(t), t), \quad (2.7)$$

где \mathbf{h} , \mathbf{g} – вектор-функции состояний и выходов соответственно; \mathbf{x} , \mathbf{z} , \mathbf{y} – векторы входных воздействий, состояний и выходных воздействий соответственно.

Дифференциальными уравнениями называются такие уравнения, в которых неизвестными будут функции одной переменной или нескольких переменных, причём в уравнение входят не только их функции, но их производные различных порядков. Если неизвестные – функции многих переменных, то уравнения называются – *уравнения в частных производных*. Если неизвестные функции одной независимой переменной, то имеют место *обыкновенные дифференциальные уравнения*.

Математическое соотношение для детерминированных систем в общем виде:

$$\bar{y}' = \bar{f}(\bar{y}, t), \quad \bar{y}(t_0) = \bar{y}_0.$$

В случае линейности таких систем, когда операторы переходов и выходов обладают свойствами однородности и аддитивности, вид уравнений (2.6) и (2.7) упрощается, что дает возможность аналитического решения или исследования известными методами с помощью вычислительных машин.

- Построение математических моделей непрерывных линейных детерминированных систем в виде дифференциальных уравнений используется при анализе функционирования элементов и электрических цепей приборных систем.

2.3 Дискретно-детерминированные модели (F -схемы)

Дискретно-детерминированные модели (ДДМ) являются предметом рассмотрения теории автоматов (ТА) – раздела теоретической кибернетики, изучающей устройства, перерабатывающие дискретную информацию и меняющие свои внутренние состояния лишь в допустимые моменты времени.

Конечный автомат (КА) имеет множество внутренних состояний и входных сигналов, являющихся конечными множествами. Автомат задаётся F -схемой:

$$F = \langle z, x, y, \varphi, \psi, z_0 \rangle, \quad (2.8)$$

где z, x, y – соответственно конечные множества входных, выходных сигналов (алфавитов) и конечное множество внутренних состояний (алфавита). $z_0 \in Z$ – начальное состояние; $\varphi(z, x)$ – функция переходов; $\psi(z, x)$ – функция выхода.

Автомат функционирует в дискретном автоматном времени, моментами которого являются *такты*, т.е. примыкающие друг к другу равные интервалы времени, каждому из которых соответствуют постоянные значения входного, выходного сигнала и внутреннего состояния. Абстрактный автомат имеет один входной и один выходной каналы.

В момент t , будучи в состоянии $z(t)$, автомат способен воспринять сигнал $x(t)$ и выдать сигнал $y(t) = \psi[z(t), x(t)]$, переходя в состояние $z(t + 1) = \varphi[z(t), x(t)]$, $z(t) \in Z$; $y(t) \in Y$; $x(t) \in X$. Абстрактный КА в начальном состоянии z_0 , принимая сигналы $x(0), x(1), x(2), \dots$ (входное слово), выдаёт сигналы $y(0), y(1), y(2), \dots$ (выходное слово).

1) F' -автомат 1-ого рода (автомат Миля), функционирующий по схеме:

$$z(t + 1) = \varphi[z(t), z(t)], t = 0, 1, 2, \dots \quad (2.9)$$

$$y(t) = \psi[z(t), x(t)], t = 0, 1, 2, \dots; \quad (2.10)$$

2) F -автомат 2-ого рода:

$$z(t + 1) = \varphi[z(t), z(t)], t = 0, 1, 2, \dots \quad (2.11)$$

$$y(t) = \psi[z(t), x(t - 1)], t = 1, 2, 3, \dots; \quad (2.12)$$

3) F -автомат 2-ого рода, для которого функция выходов не зависит от входной переменной $x(t)$ (автомат Мура):

$$z(t + 1) = \varphi[z(t), z(t)], t = 0, 1, 2, \dots \quad (2.13)$$

$$y(t) = \psi[z(t)], t = 0, 1, 2, \dots; \quad (2.14)$$

Уравнения (2.9-2.14), полностью задающие F -автомат, являются частным случаем уравнения:

$$\bar{z}(t) = \Phi(\bar{z}_0, \bar{x}, \bar{v}, \bar{h}, t) \quad (2.15)$$

где \bar{z} – вектор состояний, \bar{x} – вектор независимых входных переменных, \bar{v} – вектор воздействий внешней среды, \bar{h} – вектор собственных внутренних параметров системы, \bar{z}_0 – вектор начального состояния, t – время; и уравнения

$$\bar{y}(t) = F(\bar{z}, t), \quad (2.16)$$

когда система S – детерминированная, и на её вход поступает дискретный сигнал x .

По числу состояний конечные автоматы бывают *с памятью* и *без памяти*. Автоматы с памятью имеют более одного состояния, а автоматы без памяти (комбинационные или логические схемы) обладают лишь одним состоянием.

Согласно (2.10) работа комбинационной схемы заключается в том, что она ставит в соответствие каждому входному сигналу $x(t)$ определённый выходной сигнал $y(t)$, т.е. реализует логическую функцию вида:

$$y(t) = \psi[x(t)], t = 0, 1, 2, \dots$$

Эта функция называется *булевой*, если алфавиты X и Y , которым принадлежат значения сигналов x и y , состоят из 2-х букв.

По характеру отсчёта времени (дискретному) F -автоматы делятся на синхронные и асинхронные. В *синхронных автоматах* моменты времени, в которые автомат «считывает» входные сигналы, определяются принудительно синхронизирующими сигналами. Реакция автомата на каждое значение входного сигнала заканчивается за один такт синхронизации.

Асинхронный F-автомат считывает входной сигнал непрерывно, и поэтому, реагируя на достаточно длинный входной сигнал постоянной величины x , он может, как это следует из (2.8)-(2.14), несколько раз изменить своё состояние, выдавая соответствующее число выходных сигналов, пока не перейдёт в устойчивое.

Для задания F -автомата необходимо описать все элементы множества $F = \langle z, x, y, \varphi, \psi, z_0 \rangle$, т.е. входной, внутренний и выходной алфавиты, а также функции переходов и выходов. Для задания работы F -автоматов наиболее часто используются табличный, графический и матричный способ.

В табличном способе задания используется таблицы переходов и выходов, строки которых соответствуют входным сигналам автомата, а столбцы – его состояниям. При этом обычно 1-ый столбец слева соответствует начальному состоянию z_0 . На пересечении i -ой строки и j -ого столбца таблицы переходов помещается соответствующее значение $\varphi(z_k, x_i)$ функции переходов, а в таблице выходов – $\psi(z_k, x_i)$ функции выходов.

Для F -автомата Мура обе таблицы можно совместить, получив *отмеченную* таблицу переходов, в которой над каждым состоянием z_k автомата, обозначающим столбец таблицы, стоит соответствующий этому состоянию, согласно (2.14), выходной сигнал $\psi(z_i)$.

Описание работы F -автомата Миля таблицами переходов φ и выходов ψ иллюстрируется таблицей 2.1, а описание F -автомата Мура – таблицей переходов 2.2.

Таблица 2.1.

x_i	z_k			
	z_0	z_1	...	z_k
Переходы				
x_1	$\varphi(z_0, x_1)$	$\varphi(z_1, x_1)$...	$\varphi(z_k, x_1)$
x_2	$\varphi(z_0, x_2)$	$\varphi(z_1, x_2)$...	$\varphi(z_k, x_2)$
.....				
x_n	$\varphi(z_0, x_n)$	$\varphi(z_1, x_n)$...	$\varphi(z_k, x_n)$
Выходы				
x_1	$\psi(z_0, x_1)$	$\psi(z_1, x_1)$...	$\psi(z_k, x_1)$
.....				
x_n	$\psi(z_0, x_n)$	$\psi(z_1, x_n)$...	$\psi(z_k, x_n)$

Таблица 2.2.

x_i	$\psi(z_k)$			
	$\psi(z_0)$	$\psi(z_1)$...	$\psi(z_k)$
	z_0	z_1	...	z_k
x_1	$\varphi(z_0, x_1)$	$\varphi(z_1, x_1)$...	$\varphi(z_k, x_1)$
x_2	$\varphi(z_0, x_2)$	$\varphi(z_1, x_2)$...	$\varphi(z_k, x_2)$
.....				
x_n	$\varphi(z_0, x_n)$	$\varphi(z_1, x_n)$...	$\varphi(z_k, x_n)$

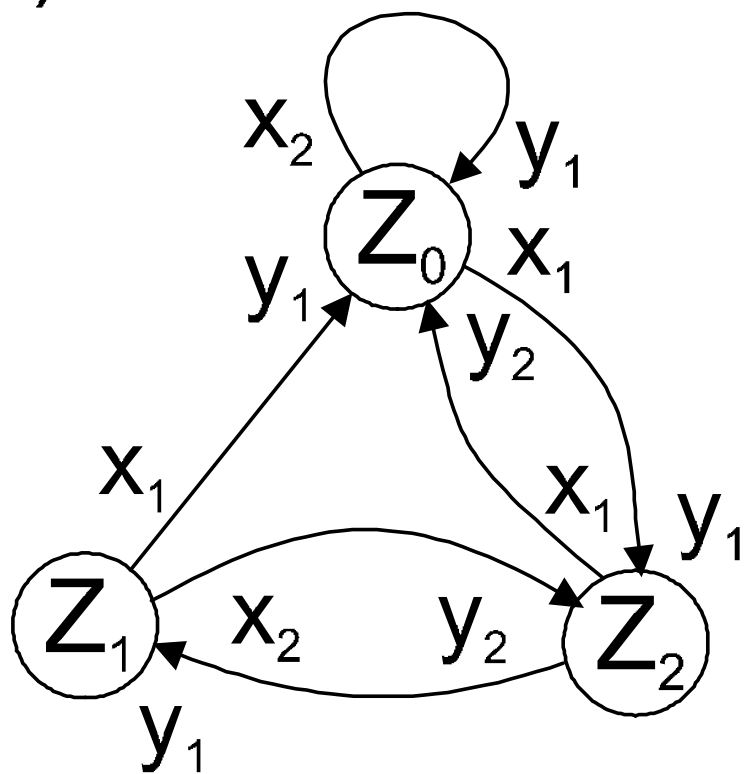
При другом способе задания конечного автомата используется понятие направленного графа. Граф автомата представляет собой набор вершин, соответствующих различным состояниям автомата и соединяющих вершин дуг графа, соответствующих тем или иным переходам автомата. Если входной сигнал x_k вызывает переход из состояния z_i в состояние z_j , то на графе автомата дуга, соединяющая вершину z_i с вершиной z_j , обозначается x_k .

- Для того чтобы задать функцию переходов, дуги графа необходимо отметить соответствующими выходными сигналами. Для автоматов Миля эта разметка производится так: если входной сигнал x_k действует на состояние z_i , то получается дуга, исходящая из z_i и помеченная x_k ;

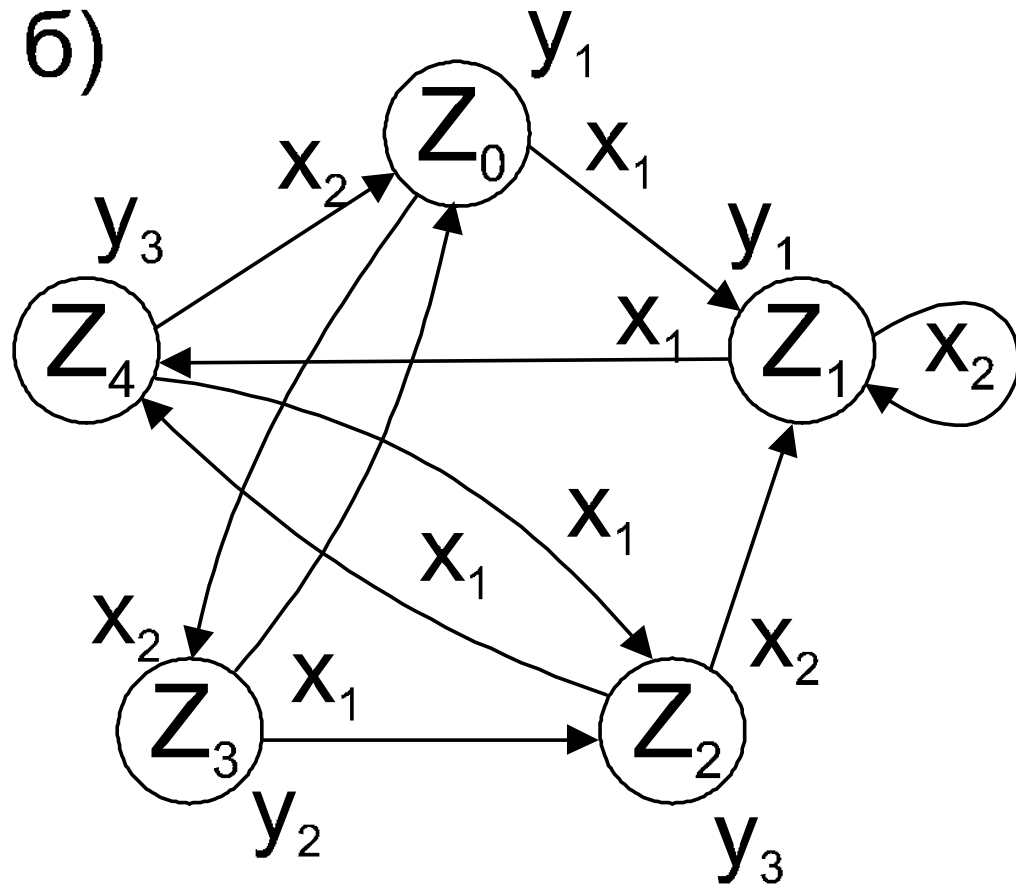
- эту дугу дополнительно отмечают выходным сигналом $y = \psi(z_i, x_k)$. Для автомата Мура аналогичная разметка графа такова: если входной сигнал x_k , действуя на некоторое состояние автомата, вызывает переход в состояние z_j , то дугу, направленную в z_j и помеченную x_k , дополнительно отмечают выходным сигналом $y = \psi(z_j, x_k)$.

На рис. 2.1 приведены заданные ранее таблицами F -автоматы Миля F_1 и Мура F_2 соответственно.

а)



б)



При решении задач моделирования часто более удобной формой является матричное задание конечного автомата. При этом матрица соединений автомата есть квадратная матрица $C = \|c_{ij}\|$, строки которой соответствуют исходным состояниям, а столбцы – состояниям перехода.

Элемент $c_{ij} = x_k/y_s$ в случае автомата Миля соответствует входному сигналу x_k , вызывающему переход из состояния z_i в состояние z_j , и выходному сигналу y_s , выдаваемому при этом переходе. Для автомата Миля F_1 , рассмотренного выше, матрица соединений имеет вид:

$$\mathbf{C}_1 = \begin{pmatrix} x_2 / y_1 & - & x_1 / y_1 \\ x_1 / y_1 & - & x_2 / y_2 \\ x_1 / y_2 & x_2 / y_1 & - \end{pmatrix}$$

- Если переход из состояния z_i в состояние z_j происходит под действием нескольких сигналов, элемент матрицы c_{ij} представляет собой множество пар «вход/выход» для этого перехода, соединённых знаком дизъюнкции.
- Для F -автомата Мура элемент c_{ij} равен множеству входных сигналов на переходе $(z_i \rightarrow z_j)$, а выход описывается вектором выходов:

$$y = \begin{pmatrix} \psi(z_0) \\ \psi(z_1) \\ \dots \\ \psi(z_k) \end{pmatrix}$$

i -ая компонента которого выходной сигнал, отмечающий состояние z_i .

- Для детерминированных автоматов переходы однозначны. Это означает, что в графе F -автомата из любой вершины не могут выходить 2 и более ребра, отмеченные одним и тем же входным сигналом. Аналогично этому в матрице соединений автомата C в каждой строке любой входной сигнал не должен встречаться более одного раза.

- Для F -автомата состояние z_k называется *устойчивым*, если для любого входа $x_i \in X$, для которого $\varphi(z_k, x_i) = z_k$ имеет место $\psi(z_k, x_i) = y_k$. Таким образом, F -автомат называется *асинхронным*, если каждое его состояние $z_k \in Z$ устойчиво.

- На практике всегда автоматы являются асинхронными, а устойчивость их состояний обеспечивается тем или иным способом, например, введением сигналов синхронизации. На уровне абстрактной теории удобно оперировать синхронными конечными автоматами.

- Если в таблице переходов асинхронного автомата некоторое состояние z_k стоит на пересечении строки x_s и столбца z_s ($s \neq k$), то это состояние z_k обязательно должно встретиться в этой же строке в столбце z_k .

- С помощью F -схем описываются узлы и элементы ЭВМ, устройства контроля, регулирования и управления, системы временной и пространственной коммутации в технике обмена информацией. Широта применения F -схем не означает их универсальность. Этот подход непригоден для описания процессов принятия решений, процессов в динамических системах с наличием переходных процессов и стохастических элементов.

Потоки систем

- Одним из наиболее важных классов систем являются те системы, изменение состояния которых происходят в случайные моменты времени. Изменения системы логично назвать **событием** и ассоциировать наступление события с моментом времени.

- Под ***поток*** ***событий*** в дальнейшем мы будем понимать последовательность событий, происходящих один за другим в случайные моменты времени.

$\Delta t \rightarrow 0$

- **Определение 1.** Интенсивностью потока событий называется математическое ожидание числа событий в единицу времени.
- **Определение 2.** Параметром потока событий $\lambda(t)$ называется предел отношения вероятности появления хотя бы одного события за промежуток времени Δt к данному промежутку времени при $\Delta t \rightarrow 0$:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P_1(t, t + \Delta t)}{\Delta t}$$

3.1 Простейший поток событий

- Простейшим потоком событий называется такой поток, при котором время возникновения события удовлетворяет одновременно условиям **стационарности**, **отсутствия последействия** и **ординарности**.

- **Стационарность** случайного процесса – означает, что на любом промежутке времени Δt вероятность возникновения n событий зависит только от n и от величины промежутка Δt , но не изменяется от сдвига Δt по оси времени.

- **Отсутствие последствия** означает, что вероятность наступления n событий в течение промежутка времени Δt не зависит от того, сколько уже было событий, и как они распределялись до этого промежутка.

- **Ординарность потока** событий означает невозможность появления в один и тот момент времени.

- Если элементы сложной системы функционируют одновременно,
- события, связанные с ними, имеют мгновенный характер,
- событие, связанное с одним элементом, ассоциируется со всей системой,
- изменение характеристик элементов системы не происходит, и
- процесс функционирования стабилизирован, то **поток событий, связанных с элементами и всей системой в целом, можно считать простейшим**

Свойства простейшего потока:

- а) Случайные события, образующие простейший поток, распределены по закону Пуассона, то есть вероятность $P_n(t)$ возникновения в течение времени t ровно n событий находится по формуле

$$P_n(t) = \frac{(\lambda \cdot t)^n}{n!} \cdot e^{-\lambda t}$$

- где λ - параметр потока.

- б) Закон распределения промежутков времени между соседними событиями является показательным с параметром λ

$$f(t) = \lambda \cdot e^{-\lambda t}$$

- в) Плотность распределения промежутков времени от начала потока до k –го события определяется выражением:

$$f_k(t_k) = \lambda \cdot \frac{(\lambda \cdot t_k)^{k-1}}{(k-1)!} \cdot e^{-\lambda t_k}$$

формула описывает гамма–распределение

- г) Если промежуток времени, распределенный по показательному закону, уже длится некоторое время τ , то это никак не влияет на закон распределения оставшейся части промежутка: он будет таким же, как и закон распределения всего промежутка.

- д) Сумма большого числа простейших потоков в течение времени образует также простейший поток с интенсивностью Λ , равной сумме интенсивностей составных потоков в течение того же времени t
- ,
- где λ_i – интенсивность составных потоков, N – число простейших потоков.

- е) Интенсивность простейшего потока совпадает с его параметром $\mu = \lambda$.

3.2 Нестационарный пуассоновский поток

- **Нестационарным пуассоновским потоком** называется поток нестационарный, но удовлетворяющий одновременно условиям **отсутствия последствия и ординарности**.

- Если события, связанные с элементами сложной системы, носят мгновенный характер, событие, связанное с любым элементом, ассоциируется с событием, связанным с системой в целом, и изменение элементов отсутствует, то поток событий в течение длительного времени функционирования системы является **нестационарным пуассоновским**

Основной характеристикой нестационарного пуассоновского потока является мгновенная интенсивность $\mu(t)$.

Мгновенной интенсивностью потока называется предел отношения среднего числа событий на участке Δt к длительности этого временного участка, когда $\Delta t \rightarrow 0$

$$\mu(t) = \lim_{\Delta t \rightarrow 0} \frac{m(t + \Delta t) - m(t)}{\Delta t} = m'(t)$$

где $m(t)$ - математическое ожидание числа событий на участке времени от нуля до t .

- Нестационарный пуассоновский поток обладает следующими свойствами:
- а) Случайные события, образующие поток, распределены по Пуассону с параметром, зависящим от длины участка, в течение которого наблюдается поток, и от его расположения на оси времени

$$P_n(\tau, t_0) = \frac{a^n}{n!} \cdot e^{-a}$$

- Где $P_n(\tau, t_0)$ – вероятность появления n событий на участке $(t_0, t_0 + \tau)$, a – математическое ожидание числа событий на участке $(t_0, t_0 + \tau)$, определяемое выражением

- $$a = \int_{t_0}^{t_0 + \tau} \lambda(t) dt$$

- б) Закон распределения промежутков времени между соседними событиями определяется выражением

$$f_{t_0}(t) = \lambda(t_0 + t) \cdot \exp\left(-\int_{t_0}^{t_0+t} \lambda(t) dt\right)$$

- t_0 – время появления первого из соседних событий.

Закон распределения промежутков времени между соседними событиями в нестационарном пуассоновском потоке не является показательным, зависит от вида функции $\lambda(t)$ и параметра t_0 .

- в) Мгновенная интенсивность ординарного потока без последствия совпадает с мгновенным значением его параметра
- $\mu(t) = \lambda(t)$.

3.3 Поток с ограниченным последствием (поток Пальма)

- Поток с ограниченным последствием(или потоком Пальма) называется ординарный поток, при котором промежуток времени между последовательными событиями являются независимыми случайными величинами.

Свойства потока Пальма

- а) Плотность распределения промежутков времени в потоке Эрланга k –го порядка имеет вид

$$f_k(t) = \lambda \cdot \frac{(\lambda \cdot t)^k}{k!} \cdot e^{-\lambda \cdot t}$$

- Этот закон распределения называется законом Эрланга k –го порядка.

б) Параметр потока и плотность распределения промежутка времени от начала потока до первого события связаны между собой следующим уравнением в преобразованиях Лапласа:

$$\lambda(s) = \frac{f_1(s)}{1 - f_1(s)}$$

где f_1 – плотность распределения. Для того, чтобы найти параметр потока надо применить обратное преобразование Лапласа.

- в) При любом законе распределения $f_1(t)$ параметр потока $\lambda(t)$ при $t \rightarrow \infty$
- имеет предел, равный величине, обратной математическому ожиданию промежутков времени между событиями

$$\lim_{t \rightarrow \infty} \lambda(t) = \frac{1}{T}$$

- г) Если суммарный поток состоит из большого числа независимых потоков малой интенсивности, каждый из которых удовлетворяет условиям стационарности и ординарности, то он весьма близок к простейшему.

- Для сложных систем справедлив экспоненциальный закон потока событий.
- Если независимые потоки являются ординарными, то параметр суммарного потока $\lambda_c(t)$ равен сумме параметров независимых потоков

$$\lambda_c(t) = \sum_{i=1}^n \lambda_i(t)$$

- При этом независимые потоки могут быть с последствием и нестационарные.

3.4 Законы распределения времени между событиями

- Время между соседними событиями является непрерывной случайной величиной. Эта случайная величина с вероятностной точки зрения будет полностью определена, если известна плотность распределения.
- Это объясняется тем, что одна из количественных характеристик потока событий, а именно частота событий, является статистическим выражением для плотности распределения времени между соседними событиями.

- Случайные величины в зависимости от их физического смысла могут иметь различные законы распределения. На практике, время между событиями в сложных системах подчиняется многим законам распределения, среди которых следует выделить:

- 1) экспоненциальный (показательный);
- 2) Рэляя;
- 3) нормальный;
- 4) гамма–распределение;
- 5) Вэйбулла.

- Экспоненциальный закон. При экспоненциальном законе распределения времени возникновения событий интенсивность событий $\lambda(t) = \lambda = const$.
Зависимости между основными количественными характеристиками потока событий будут выражены формулами:

$$P(t) = \exp\left(-\int_0^t \lambda(t) dt\right) = \exp(-\lambda \cdot t)$$

$$Q(t) = 1 - \exp(-\lambda \cdot t)$$

$$a(t) = \lambda \cdot \exp(-\lambda \cdot t)$$

$$T = \int_0^{\infty} \exp(-\lambda \cdot t) dt = -\frac{1}{\lambda} \cdot \exp(-\lambda \cdot t) \Big|_0^{\infty} = \frac{1}{\lambda}$$

- $P(t)$ – вероятность того, что событие не наступит,
- $Q(t)$ – вероятность события,
- $a(t)$ – частота событий,
- T – среднее время наступления событий.
Средняя частота событий равна интенсивности событий – $\omega(t)=\lambda(t)=\lambda$.
- Среднее время между соседними событиями равно среднему времени наступления первого события – $t_{cp} = T$

- Для экспоненциального закона распределения времени возникновения событий средняя частота событий превращается в интенсивность событий, а среднее время между соседними событиями – в среднее время до первого события.

- Основными характеристиками потока являются:
- 1. $P(t)$ – вероятность того, что событие не произойдет;
- 2. $Q(t)$ – вероятность события;
- 3. $a(t)$ – частота событий;
- 4. T – среднее время до первого события.

- В случае экспоненциального закона эти характеристики пригодны для оценки характеристик, как простейших элементов, так и сложных систем, работающей в режиме смены элементов.
- Смысл среднего времени до первого события, при $t=T$ вероятность работы до первого события будет иметь значение

$$P(T) = \exp(-\lambda \cdot T) = \exp\left(-\frac{1}{\lambda} \cdot \pi\right) = \frac{1}{e} \approx 0,37$$

При экспоненциальном законе надежности среднее время до первого события – время, в течение которого вероятность наступления события уменьшается в e раз.

- Дисперсия времени возникновения событий D_t для экспоненциального закона равна $D_t = T^2$. На практике используется среднее квадратичное отклонение, равное квадратному корню из дисперсии

- $$\sigma_t = \sqrt{D_t} = T$$

- При экспоненциальном законе квадратичное отклонение времени возникновения времени между событиями равно среднему времени возникновения первого события. Это свойство на практике может быть использовано для проверки истинности статистической гипотезы о существовании экспоненциального закона надежности.

- Для этого из статистических данных о потоке событий определяется среднее время между событиями и среднее квадратичное отклонение. Если они близки, то это может служить доказательством правдоподобия принятой гипотезы.

- **Закон Рэля.**
- При распределении времени возникновения событий по закону Рэля частота событий определяется выражением

$$a(t) = \frac{t}{\sigma^2} \cdot \exp\left(-\frac{t^2}{2 \cdot \sigma^2}\right)$$

- , где σ – параметр распределения Рэля.
- Вероятность безотказной работы, интенсивность событий и среднее время до первого события будут выражаться следующими формулами:

- Вероятность безотказной работы, интенсивность событий и среднее время до первого события будут выражаться следующими формулами:

$$P(t) = 1 - \int_0^t a(t) dt = \exp\left(-\frac{t^2}{2 \cdot \sigma^2}\right),$$

$$\lambda(t) = \frac{a(t)}{P(t)} = \frac{t}{\sigma^2},$$

$$T = \int_0^{\infty} P(t) dt = \sqrt{\frac{\pi}{2}} \cdot \sigma$$

- При распределении времени распределения событий по закону Рэлея происходит интенсивное изменение системы, и события не удовлетворяют условиям стационарного случайного процесса. Вероятность наступления первого события с течением времени уменьшается значительно интенсивнее, чем при экспоненциальном законе

Нормальное распределение

- Длительность функционирования системы не может быть отрицательной. Поэтому количественные характеристики параметра событий имеет смысл рассматривать только при усеченном нормальном законе распределения времени между событиями.

- Частота событий в этом случае определяется выражением

$$a(t) = C \cdot \exp\left(-\frac{(t - T_1)^2}{2 \cdot \sigma^2}\right)$$

- где T_1, σ^2 – среднее значение и дисперсия времени между событиями в нормальном законе, C – постоянная усеченного нормального распределения

$$C = \frac{\sqrt{\frac{2}{\pi}}}{\sigma \cdot \left(1 + \Phi \left(\frac{T_1}{\sigma \cdot \sqrt{2}} \right) \right)}$$

Φ – интеграл вероятностей

$$\Phi = \frac{2}{\sqrt{\pi}} \cdot \int_0^x \exp(-t^2) dt$$

Вероятность безотказной работы

$$P(t) = \frac{1 - \Phi \left(\frac{t - T_1}{\sigma \cdot \sqrt{2}} \right)}{1 + \Phi \left(\frac{t - T_1}{\sigma \cdot \sqrt{2}} \right)}$$

- Формула для вычисления интенсивности событий выглядит следующим образом

$$\lambda(t) = \frac{a(t)}{P(t)} = \frac{\sqrt{\frac{2}{\pi}} \cdot \exp\left(-\frac{(t - T_1)^2}{2 \cdot \sigma^2}\right)}{\sigma \cdot \left[1 - \Phi\left(\frac{t - T_1}{\sigma \cdot \sqrt{2}}\right)\right]}$$

Среднее время T находится по формуле :

$$T = T_1 + C \cdot \sigma^2 \cdot \exp\left(-\frac{T_1^2}{2 \cdot \sigma^2}\right)$$

Гамма–распределение с параметрами

$$\lambda_0, k \text{ .}$$

- При этом распределении частота событий выражается формулой:

$$a(t) = \lambda_0 \cdot \frac{(\lambda_0 \cdot t)^{k-1}}{(k-1)!} \cdot \exp(-\lambda_0 \cdot t)$$

- При целом и положительном k вероятность безотказной работы, интенсивность износа и среднее время безотказной работы выражается следующими формулами:

$$P(t) = 1 - \int_0^t a(t) dt = \exp(-\lambda_0 \cdot t) \cdot \sum_{i=0}^{k-1} \frac{(\lambda_0 \cdot t)^i}{i!}$$

$$\lambda(t) = \frac{a(t)}{P(t)} = \frac{\lambda_0 \cdot (\lambda_0 \cdot t)^{k-1}}{(k-1)! \cdot \sum_{i=0}^{k-1} \frac{(\lambda_0 \cdot t)^i}{i!}}$$

$$T = \int_0^{\infty} P(t) dt = \frac{k}{\lambda_0}$$

Распределение Вейбулла с параметрами λ_0, k .

При этом распределении частота событий задается выражением :

$$a(t) = \lambda_0 \cdot k \cdot t^{k-1} \cdot \exp(-\lambda_0 \cdot t \cdot k)$$

- Для распределения Вейбулла основные количественные характеристики надежности выражаются следующими формулами:

$$P(t) = 1 - \int_0^t a(t) dt = \exp(-\lambda_0 \cdot t \cdot k) \quad \lambda(t) = \frac{a(t)}{P(t)} = \lambda_0 \cdot k \cdot t^{k-1} \quad T = \int_0^{\infty} P(t) dt = \frac{\Gamma\left(\frac{1}{k+1}\right)}{\lambda_0^{\frac{1}{k}}}$$

Суперпозиция распределений.

- Рассмотренные законы распределения времени событий могут в большинстве случаев характеризовать потоки событий для системы лишь на ограниченных участках времени ее работы.

- В связи с этим, для оценки потока событий сложной системы на длительном временном участке ее функционирования целесообразно использовать суперпозицию рассмотренных законов распределения времени между событиями.

- В качестве примера рассмотрим суперпозицию двух экспоненциальных законов :

$$a(t) = C_1 \cdot \lambda_1 \cdot \exp(-\lambda_1 \cdot t) + C_2 \cdot \lambda_2 \cdot \exp(-\lambda_2 \cdot t)$$

$$\lambda_1 < \lambda_2 \quad C_1, C_2$$

- -постоянные, определяемые из условия:

$$C_1 + C_2 = 1$$

- и зависящие от соотношения параметров $\lambda_1; \lambda_2$
Для рассматриваемого случая основные количественные характеристики надежности будут иметь следующий вид:

$$P(t) = 1 - \int_0^t a(t) dt = C_1 \exp(-\lambda_1 \cdot t) + C_2 \exp(-\lambda_2 \cdot t)$$

$$\lambda(t) = \frac{a(t)}{P(t)} = \frac{C_1 \cdot \lambda_1 \cdot \exp(-\lambda_1 \cdot t) + C_2 \cdot \lambda_2 \cdot \exp(-\lambda_2 \cdot t)}{C_1 \exp(-\lambda_1 \cdot t) + C_2 \exp(-\lambda_2 \cdot t)}$$

$$T = \int_0^{\infty} P(t) dt = \frac{C_1}{\lambda_1} + \frac{C_2}{\lambda_2}$$

Лекция 4. Имитационное моделирование в приборных системах

4.1 Понятие имитационного моделирования

Когда говорят об имитационном моделировании, то предполагают использование некоторой структурной схемы, математического обеспечения, а также вычислительного устройства для создания соответствующей программы. Отсюда следует определение имитационного моделирования

- ***Имитационной моделью*** называют абстрактную динамическую модель, реализованную, как правило, на ЭВМ, и воспроизводящую в рамках установленных ограничений поведение оригинала в хронологическом порядке.

- Имитационные модели не способны формировать решение в таком виде, как в аналитических моделях, а служат лишь средством для анализа поведения системы (оригинала) в условиях, которые определяются экспериментатором. По сути, имитационное моделирование является экспериментальной и прикладной методологией, которая имеет следующие цели:

- 1) описать поведение системы;
- 2) построить теории и гипотезы, которые могут объяснить наблюдаемое поведение;
- 3) использовать данные теории для предсказания будущего поведения системы.

- *Имитационная система* реализуется на ЭВМ и позволяет исследовать имитационную модель M , задаваемую в виде определенной совокупности отдельных блочных моделей и связей между ними в их взаимодействии в пространстве и времени при реализации какого-либо процесса.
- Можно выделить три основные группы блоков:

- блоки, характеризующие моделируемый процесс функционирования системы S ;
- блоки, отражающие внешнюю среду E и ее воздействие на реализуемый процесс;
- блоки, играющие служебную вспомогательную роль, обеспечивая взаимодействие первых двух, а также выполняющие дополнительные функции по получению и обработке результатов моделирования.

Кроме того, имитационная система характеризуется набором переменных, с помощью которых удастся управлять изучаемым процессом, и набором начальных условий, когда можно изменять условия (план) проведения машинного эксперимента.

- **Математическое обеспечение**
имитационной системы — совокупность математических соотношений, описывающих поведение реального объекта, совокупность алгоритмов, обеспечивающих как подготовку (ввод исходных данных), так и работу с моделью (имитация, вывод, обработка результатов).

Программное обеспечение — совокупность программ: планирования эксперимента, имитационной модели, проведения эксперимента, обработки и интерпретации результатов, синхронизации процессов в модели (псевдопараллельное выполнение процессов в модели).

Информационное обеспечение — средства и технология организации и реорганизации базы данных моделирования, методы логической и физической организации массивов, формы документов, описывающих процесс моделирования и его результаты.

Техническое обеспечение — средства вычислительной техники, связи и обмена между оператором и сетью ЭВМ, ввода и вывода информации, управления проведением эксперимента.

Эргономическое обеспечение — совокупность научных и прикладных методик и методов, а также нормативно-технических и организационно-методических документов, создающих оптимальные условия для высокопроизводительной деятельности человека во взаимодействии с моделирующим комплексом.

2.2 Преимущества и недостатки имитационного моделирования

- Имитационное моделирование целесообразно применять при наличии условий:
- 1) Не существует законченной математической постановки данной задачи (например, модель многофазных, многоканальных систем массового обслуживания).

- 2) Аналитические методы имеются, но очень сложны и трудоемки, а имитационное моделирование дает более простой способ решения.
- 3) Аналитические решения имеются, но их реализация невозможна из-за недостаточной подготовки имеющегося персонала. В этом случае сопоставляются затраты на работу с имитационным моделированием и затраты на приглашение специалистов со стороны.

- 4) Кроме оценки определенных параметров необходимо осуществлять наблюдение за ходом процесса в течение определенного периода.
- 5) Имитационное моделирование может быть единственным возможным вследствие трудности постановки эксперимента и наблюдения явлений в реальных условиях (наблюдение за поведением космических кораблей).
- 6) Может понадобиться сжатие шкалы времени (как замедление, так и ускорение; например, исследование проблемы развития городов).

- Имитационное моделирование являются непревзойденным средством создания средств обучения в виде тренажеров, симуляторов и т.д.
- С помощью имитационного моделирования можно разыграть реальные процессы и ситуации, которые помогут исследователю понять и прочувствовать проблему, что стимулирует процесс поиска нововведений.

- Благодаря этому, порядка 30% из всех используемых на практике моделей являются имитационными моделями.

Основные ***достоинства*** имитационного моделирования при исследовании сложных систем:

- возможность исследовать особенности процесса функционирования системы S в любых условиях;
- существенное сокращение за счет применения ЭВМ продолжительности испытаний по сравнению с натурным экспериментом;

- возможность использовать результаты натурных испытаний реальной системы или ее частей для проведения имитационного моделирования;
- гибкость варьирования структуры, алгоритмов и параметров моделируемой системы при поиске оптимального варианта системы;
- для сложных систем — это единственный практически реализуемый метод исследования процесса функционирования систем.

- Вместе с тем имитационное моделирование обладает рядом недостатков.

- Во-первых, имитационное моделирование представляет собой весьма дорогостоящий процесс, требующий существенных затрат временных ресурсов и привлечения высококвалифицированных специалистов.

- Во-вторых, в процессе моделирования не представляется возможным получить точный результат.
- При этом оценка точности может быть выполнена путем анализа чувствительности модели к изменению определенных параметров.

- В-третьих, имитационное моделирование в действительности не отражает полного положения вещей. Данный факт необходимо учитывать при анализе исследуемого объекта (процесса).

В 4-х необходимость многократно воспроизводить имитационный эксперимент, варьируя исходные данные задачи, для полного анализа характеристик процесса функционирования систем и поиска оптимального варианта;

- 5. большие затраты машинного времени.

2.3 Процесс имитационного моделирования

- Построение имитационной модели, так же, как и любое исследование, требует проведения работ по следующим этапам.
- 1) Определение границ модели.
- 2) Разработка концептуальной модели.
- 3) Подготовка исходных данных.

- 4) Создание концептуальной модели в виде диаграммы.
- 5) Трансляция модели.
- 6) Оценка адекватности модели.

- 7) Планирование машинных экспериментов.
 - а) Стратегическое планирование.
 - б) Tактическое планирование.

- 8) Моделирование – проведение эксперимента.
- 9) Анализ (интерпретация) результатов.
- 10) Документирование и реализация.

Границы системы определяются таким образом, чтобы охватить те компоненты, взаимодействие которых определяет важные стороны поведения системы.

При этом система должна быть способна сама генерировать любую ситуацию, любые затруднения, которые, возможно, потребуются проанализировать.

- ***Разработка концептуальной схемы объекта (системы)*** является одним из самых важных этапов исследования. На этом этапе осуществляется *формализация системы*, т.е. переход от реального объекта к некоторой логической схеме (абстракции).

Такая формализация начинается со словесного описания реальности в системе принятых терминов и формальных понятий. Здесь «приводятся сведения о природе и параметрах (характеристиках) элементарных явлений исследуемой системы, о виде и степени взаимодействия между ними, о месте и значении каждого элементарного явления в общем процессе функционирования системы». Завершается формализация построением общей схемы процессов, подлежащих исследованию.

В дальнейшем полученная схема уточняется и дополняется в соответствии с тем уровнем детализации, который определяется (стратифицируется) постановкой задачи.

Действия, которые позволяют представить модель в виде совокупности частей (подсистем, элементов), называют ***декомпозицией*** системы.

Составные части модели должны обеспечивать сохранение целостности системы, с одной стороны, а с другой – достижение поставленных целей моделирования.

- Процесс построения концептуальной схемы системы завершается структуризацией (указанием и общим описанием связей между выделенными элементами системы), а также укрупненным описанием динамики функционирования системы и ее возможных состояний.
- От того, как будет построена концептуальная схема имитационной модели, зависит результат исследования.

Следующий, не менее важный этап имитационного моделирования – ***подготовка исходных данных***. В некоторых случаях он проходит параллельно с построением концептуальной схемы. Фактически на данном этапе формируется информационное пространство системы. Здесь выявляются количественные характеристики (параметры) функционирования системы и ее элементов, численные значения которых составят исходные данные для моделирования

Когда подготовлены исходные данные и концептуальная схема модели, последняя ***оформляется в виде диаграммы***, состоящей из стандартных блоков. Это технический этап, благодаря которому схема модели становится доступной для понимания широкому кругу специалистов, владеющих соответствующей методикой.

На этапе *трансляции модели* осуществляется преобразование диаграммы модели в отдельную компьютерную программу или сценарий специализированной системы моделирования.

В современных версиях таких систем такой этап выполняется автоматически, благодаря наличию визуальных средств построения моделей.

Оценка адекватности полученной модели осуществляется путем ее экспертизы и проигрывания на тестовых данных.

На данном этапе модель проверяется на корректность, т.е. на соответствие реальному объекту в рамках поставленной задачи (границ системы).

Когда модель не адекватна, то она подвергается исправлениям и корректировкам до приемлемого уровня степени уверенности, с которой можно судить о корректности выводов, касающихся реальной системы.

Добившись адекватности модели, исследователи осуществляют ***стратегическое и тактическое планирование эксперимента.***

Когда говорят о стратегическом планировании, то предусматривают схему получения желаемых результатов с помощью имитационной модели.

На тактическом уровне планируют способ проведения каждой серии испытаний, предусмотренных планом эксперимента.

- ***На этапе экспериментирования*** осуществляется проигрывание запланированных сценариев с целью получения желаемого результата.
- После того, как получены результаты моделирования, наступает важный этап исследования – ***интерпретация результатов.***

- По полученным выходным данным эксперимента строятся выводы о поведении исследуемой системы. При этом очень важно не пропустить эффект двойного прочтения одних и тех же результатов.
- В этом случае следует дорабатывать модель.
- На этапе интерпретации результатов также дается заключение о полезности или бесполезности модели.

В случае полезности модели осуществляется ее ***реализация***, т.е. практическое использование.

На данном этапе выполняются вспомогательные действия, такие как регистрация хода осуществления исследования и его результатов, документирование процесса создания и использования модели.

2.4 Концепция универсальной модели системы имитационного моделирования

Основные трудности при построении имитационных моделей вызывают математический и программный этапы.

В большинстве случаев математическое и программное обеспечение в разных моделях содержат одинаковые и (или) похожие блоки. Поэтому при создании моделей на разработку необходимых модулей не рационально расходуются время.

Подход (концепция) универсальной системы моделирования, который основывается на автоматизации процедур, позволяет не только сэкономить время разработки и реализации имитационной модели, но и сделать процесс самого моделирования более простым и доступным.

Тем самым сокращается вероятность возникновения ошибок в ходе создания моделей из-за недостаточного знания языковых средств, невнимательности в работе с большими объемами информации и т.д.

- Концепция универсальной системы имитационного моделирования (УСИМ) основывается на трех принципах:
 - 1) простота;
 - 2) модульность;
 - 3) универсальность.

- ***Принцип простоты*** заключается в минимуме необходимых знаний пользователя о системе моделирования, и, как следствие в минимуме трудозатрат.

- ***Принцип модульности*** предусматривает наличие в УСИМ общих для различных классов систем моделей – готовых модулей. При необходимости они подсоединяются к разрабатываемой модели, уточняются и дополняются.

Как показывает анализ задач, возникающих в сфере системного анализа и исследования операций, автоматизации проектирования и исследования сложных систем, основной формой задания реальных объектов является описание элементов системы (модулей УСИМ) в виде конечных автоматов, а связей между элементами – многоуровневых схем сопряжения.

Заключительным принципом построения УСИМ является ***принцип универсальности***. Он определяется способностью УСИМ охватить сложные многоуровневые объекты произвольной структуры, элементы которых являются динамическими системами в широком понимании.

Полностью универсальную систему имитационного моделирования построить, естественно, невозможно.

УСИМ должна осуществлять, по крайней мере, следующую совокупность операций:

- 1) ввод и формирование области исходных данных;
- 2) визуальное представление элементов моделируемой системы и схем сопряжения;
- 3) имитация модуля и взаимодействия элементов системы;
- 4) обработка результатов моделирования;
- 5) управление моделью.

- Исходя из этого, основным назначением элементов концептуальной схемы УСИМ будет являться:
- 1) для интерфейса пользователя – обеспечение представления состава и структуры модели реальной системы посредством модулей и возможных схем сопряжения УСИМ;
- 2) для библиотек стандартных процедур имитации и модулей – реализация модельного представления исследуемой системы;

- 3) для библиотек стандартных процедур обработки результатов – представление и предварительный анализ результатов моделирования в виде, удобном для восприятия и дальнейшего анализа;
- 4) для ядра УСИМ – обеспечение взаимодействия элементов УСИМ и управление процессом моделирования.

- Применение УСИМ позволяет сократить количество этапов разработки, при этом время реализации имитационной модели существенно сокращается.

Этапы построения универсальной модели:

- 1) разработка концептуальной модели;
- 2) подготовка исходных данных;
- 3) реализация концептуальной модели в виде диаграммы;

- 4) планирование машинных экспериментов;
- 5) моделирование;
- 6) анализ результатов.

2.5 Датчики псевдослучайных чисел

- Когда ставится задача имитационного моделирования некоторой сложной системы, зачастую в ней предполагается наличие случайностей.
- Например, в описании системы финансовых потоков в торговой компании встречается случайный фактор – спрос на продаваемую продукцию. Если мы рассматриваем производство болтов для некоторой сложной металлической конструкции, то их размер также может служить случайной величиной.

Поэтому в имитационной модели должен присутствовать блок, отвечающий за моделирование случайных величин.

Такие блоки должны выдавать последовательность случайных значений, которые должны быть распределены по заданному закону.

В качестве «истинных» генераторов случайных чисел используют различные физические объекты, изменение состояния которых можно зарегистрировать существующими приборами.

Например, из урановой руды вылетают частицы, которые можно зарегистрировать. Интервал времени между регистрациями частиц можно принять как случайную величину. Мы же будем говорить о генераторах псевдослучайных чисел, которые можно реализовать на ЭВМ.

Датчиком псевдослучайных чисел называют алгоритм (программный модуль), который позволяет генерировать последовательность разных элементов, значения которых находятся в заданном диапазоне.

Датчики псевдослучайных чисел имеются в любой системе для математических вычислений, в языках высокого уровня и в некоторых офисных приложениях, например, EXCEL

Нужно четко понимать, что любой генератор дает не случайную, а детерминированную последовательность чисел.

В этой последовательности чисел можно выделить максимальный период, в котором все значения чисел не повторяются.

- ***Оценка качества***
- Оценка качества генератора псевдослучайных чисел осуществляется по двум критериям:
- 1) Равномерность распределения.
- 2) Независимость испытаний.

Равномерность распределения можно трактовать как симметричность расстановки дискретных точек, посещаемых последовательностью.

Независимость также можно рассматривать как определенную симметрию, но уже в пространстве размерности больше единицы.

- ***Проверка генераторов псевдослучайных чисел***
- Как мы уже обсуждали выше, оценка качества генератора псевдослучайных чисел осуществляется по двум критериям:
 - 1) Равномерность распределения.
 - 2) Независимость испытаний.

- Тестирование генератора производится по следующему
- принципу:
- 1) с помощью генератора вырабатывается последовательность чисел заданной длины;
- 2) вычисляется некоторый функционал от полученной выборки;

- 3) процесс повторяется некоторое количество раз и строится эмпирическое распределение функционала;
- 4) эмпирическое распределение с помощью стандартной техники гипотез сравнивается с тем, которое функционал должен иметь, если считать выборку извлеченной из действительно равномерно распределенной случайной последовательности.

К традиционным критериям проверки равномерности распределения псевдослучайных чисел относятся критерии Пирсона, Колмогорова-Смирнова, а также частотный тест, который является наиболее наглядным и интуитивно понятным

При моделировании необходимо обеспечить максимальную эффективность модели системы.

- *Эффективность* обычно определяется как некоторая разность между какими-то показателями ценности результатов, полученных при эксплуатации модели, и теми затратами, которые были вложены в ее разработку и создание.

Эффективность имитационного моделирования может оцениваться рядом критериев:

- точностью и достоверностью результатов моделирования;
- временем построения и работы с моделью M ;
- затратой машинных ресурсов (время и память);
- стоимостью разработки и эксплуатации модели.

- Наилучшей оценкой эффективности является сравнение полученных результатов с реальными исследованиями.
- С помощью статистического подхода с определенной степенью точности (в зависимости от числа реализаций машинного эксперимента) получают усредненные характеристики поведения системы.

Суммарные затраты машинного времени складываются из времени по

- вводу и выводу по каждому алгоритму моделирования,
- времени на проведение вычислительных операций, с учетом обращения к оперативной памяти и внешним устройствам,
- а также сложности каждого моделирующего алгоритма и планирования экспериментов.

Метод ИМ заключается в создании логико-аналитической (математической модели системы и внешних воздействий), имитации функционирования системы, т.е. в определении временных изменений состояния системы под влиянием внешних воздействий и в получении выборок значений выходных параметров, по которым определяются их основные вероятностные характеристики. Данное определение справедливо для стохастических систем.

- При исследовании детерминированных систем отпадает необходимость изучения выборок значений выходных параметров.
- Модель системы со структурным принципом управления представляет собой совокупность моделей элементов и их функциональные взаимосвязи.

Модель элемента (агрегата, обслуживающего прибора) - это, в первую очередь, набор правил (алгоритмов) поведения устройства по отношению к выходным воздействиям (заявкам) и правил изменений состояний элемента.

Элемент отображает функциональное устройство на том или ином уровне детализации. В простейшем случае устройство может находиться в работоспособном состоянии или в состоянии отказа.

В работоспособном состоянии устройство может быть занято, например, выполнением операции по обслуживанию заявки или быть свободным.

К правилам поведения устройства относятся:

- правила выборки заявок из очереди;
- реакция устройства на поступление заявки, когда устройство занято или к нему имеется очередь заявок;
- реакция устройства на возникновение отказа в процессе обслуживания заявки и некоторые другие.

- **Имитационное моделирование (ИМ)** — это метод исследования, который основан на том, что анализируемая динамическая система заменяется имитатором и с ним производятся эксперименты для получения об изучаемой системе. Роль имитатора зачастую выполняет программа ЭВМ.

Основная идея метода ИМ состоит в следующем :

Пусть необходимо определить функцию распределения случайной величины y .

Допустим, что искомая величина y может быть представлена в виде зависимости:

$y=f(a,b,\dots,z)$ где a,b,c,\dots,z - случайные величины с известными функциями распределения.

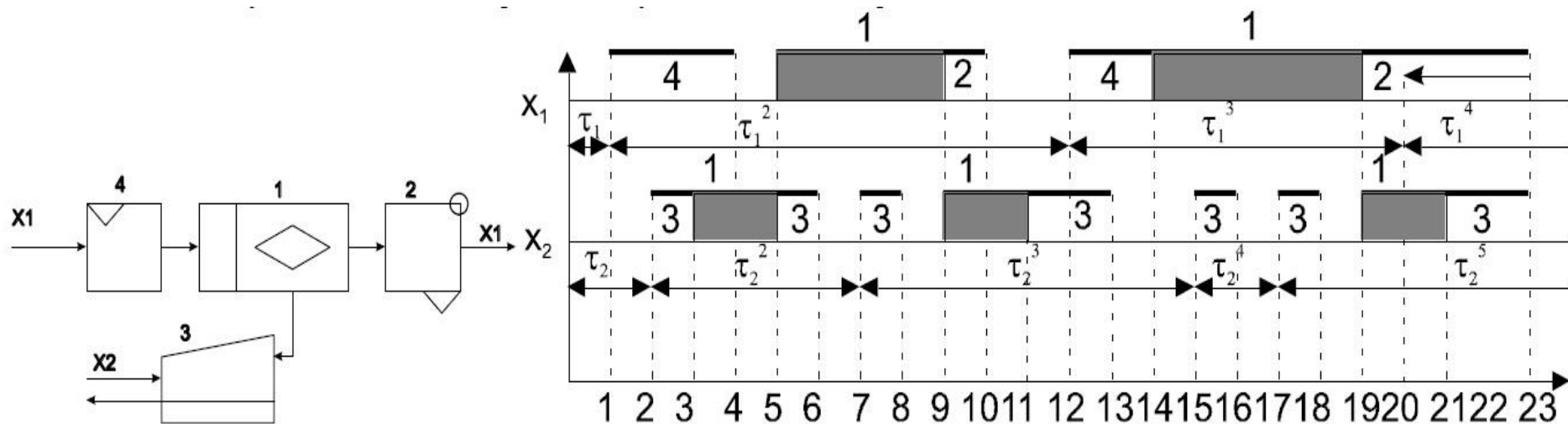
Для решения задач такого вида применяется следующий алгоритм :

- 1) по каждой из величин a, b, \dots, z производится случайное испытание, в результате каждого определяется некоторое конкретное значение случайной величины a_i, b_i, \dots, z_i

- 2) используя найденные величины, определяется одно частное значение y_i по выше приведённой зависимости;
- 3) предыдущие операции повторяются N раз, в результате чего определяется N значений случайной величины y ;
- 4) на основании N значений величины находится её эмпирическая функция распределения.

Предположим, исследуется вычислительная система (ВС), состоящая из процессора - 1 с основной памятью, устройство ввода - 4, АЦПУ - 2 и дисплея - 3 (Рис.)

Рис. Упрощённая схема моделируемой системы.



Через устройство 4 поступает поток заданий X1. Процессор обрабатывает задания и результаты выдаёт на АЦПУ 2. Одновременно с этим ВС используется, например, как информационно-справочная система. Оператор-пользователь, работающий за дисплеем, посылает в систему запросы X2, которые обрабатываются процессором и ответы выводятся на экран дисплея.

Процессор работает в 2-х программном режиме: в одном разделе обрабатываются задания X_1 , в другом, с более высоким относительным приоритетом запросы X_2 . Представим данную ВС в упрощённом варианте в виде стохастической сети из 4-х СМО. Потоки заданий и запросы будем называть потоками заявок. Считаем потоки X_1 и X_2 независимыми.

Известны ф.р. периодов следования заявок τ_1 и τ_2 и длительность обслуживания T_{1K} , T_{2K} заявок в k -ом устройстве.

Требуется определить времена загрузки каждого устройства и времена реакции по каждому из потоков.

Вначале определяется момент поступления в систему 1-ой заявки потока X_1 по результатам случайного испытания в соответствии с ф.р. периода следования заявок. На рис. 2 это момент времени $t_1=0+\tau_{11}$ (здесь и далее второй индекс обозначает порядковый номер заявки данного потока). То же самое делается для потока X_2 . На рис.2 момент поступления 1-ой заявки потока X_2 $t_2=0+\tau_{21}$.

- Затем находится минимальное время, т.е. наиболее раннее событие. В примере это время t_1 . Для 1-ой заявки потока X_1 определяется время обслуживания устройством ввода перфокарт T_{114} методом случайного испытания и отмечается момент окончания обслуживания $t_4 = t_1 + T_{114}$.

На рис. показан переход устройства 4 в состояние "занято". Одновременно определяется момент поступления следующей заявки потока X1: $t_{12} = t_1 + \tau_{12}$. Следующее минимальное время это момент поступления заявки потока X2 - t_2 . Для этой заявки находится время обслуживания на дисплее T_{123} и отслеживается время окончания обслуживания $t_3 = t_2 + T_{123}$.

Определяется момент поступления второй заявки потока X_2 : $t_7 = t_2 + \tau_{22}$. Снова выбирается минимальное время — это t_3 . В этот момент заявка потока X_2 начинает обрабатываться процессором. По результату случайного испытания определяется время её обслуживания T_{121} и отмечается момент $t_5 = t_3 + T_{121}$ окончания обслуживания.

- Следующее минимальное время t_4 - момент завершения обслуживания заявки потока X_1 устройством 4. С этого момента заявка может начать обрабатываться процессором, но он занят обслуживанием потока X_2 . Тогда заявка потока X_1 переходит в состояние ожидания, становится в очередь.

- В следующий момент времени t_5 освобождается процессор. С этого момента процессор начинает обрабатывать заявку потока X_1 , а заявка потока X_2 переходит на обслуживание дисплеем, т.е. ответ на запрос пользователя передаётся из основной памяти в буферный накопитель дисплея. Далее определяются соответствующие времена обслуживания: T_{111} и T_{123} и отмечаются моменты времени $t_9 = t_5 + T_{111}$ и $t_6 = t_5 + T_{123}$.

- В момент t_6 полностью завершается обработка первой заявки потока X2. По разности времени t_6 и t_2 вычисляется время реакции по этой заявке $u_{12} = t_6 - t_2$. Следующий минимальный момент t_7 - это наступление 2-ой заявки потока X2. Определяет время поступления очередной заявки этого потока $t_{15} = t_7 + \tau_{23}$.

- Затем вычисляется время обслуживания 2-ой заявки на дисплее T223 и отмечается момент $t_8 = t_7 + T_{223}$, после чего заявка становится в очередь, т.к. процессор занят. Эта заявка поступит на обслуживание в процессор только после его освобождения в момент t_9 . В этот момент заявка потока X1 начинает обслуживаться в АЦПУ.

- Определяются времена обслуживания T_{221} и T_{112} по результатам случайных испытаний и отмечаются моменты окончания обслуживания $t_{11} = t_9 + T_{223}$ и $t_{10} = t_9 + T_{112}$. В момент времени t_{10} завершается полное обслуживание 1-ой заявки потока X_1 . Разность между этим моментом и моментом времени t_1 даёт 1-ое значение времени реакции по потоку X_1 $u_{11} = t_{10} - t_1$.

- Указанные процедуры выполняются до истечения времени моделирования. В результате получается некоторое количество (выборка) случайных значений времени реакции (u_1) и (u_2) по 1-ому и 2-ому потокам. По этим значениям могут быть определены эмпирические функции распределения и вычислены количественные вероятностные характеристики времени реакции. В процессе моделирования можно суммировать продолжительности занятости каждого устройства обслуживанием всех потоков.

- Например, на рис. 2 занятость процессора 1 выделена заштрихованными ступеньками. Если результаты суммирования разделить на время моделирования, то получатся коэффициенты загрузки устройств.
- Можно определить время ожидания заявок в очереди, обслуженных системой, среднюю и максимальную длину очереди заявок к каждому устройству, требуемая ёмкость памяти и др.

- Имитация даёт возможность учесть надёжные характеристики ВС. В частности, если известны времена наработки на отказ и восстановления всех входящих в систему устройств, то определяются моменты возникновения отказов
- устройств в период моделирования и моменты восстановления.

- Если устройство отказало, то возможны решения:
- снятие заявки без возврата;
- помещение заявки в очередь и дообслуживание после восстановления;
- поступление на повторное обслуживание из очереди;

В качестве событий в примере выделены:

- поступление заявки в систему;
- освобождение элемента после обслуживания заявки;
- завершения моделирования;
- возникновение отказа устройств (другие-типы
- завершение восстановления устройств событий)

- Процесс имитации развивался с использованием управляющих последовательностей, определяемых по функциям распределения вероятностей исходных данных путём проведения случайных испытаний.

В качестве управляющих последовательностей использовались в примере последовательности значений периодов следования заявок по каждому i -ому потоку $\{\tau_i\}$ и длительности обслуживания заявок i -ого потока устройством $\{T_{ik}\}$.

Моменты наступления будущих событий определялись по простым рекуррентным соотношениям. Эта особенность даёт возможность построить простой циклический алгоритм моделирования, который сводится к следующим действиям:

- 1) определяется событие с минимальным временем — наиболее раннее событие;
- 2) модельному времени присваивается значение времени наступления наиболее раннего события;
- 3) определяется тип наступившего события;

- 4) в зависимости от типа события предпринимаются действия, направленные на загрузку устройств и продвижение заявок в соответствии с алгоритмом их обработки, и вычисляются моменты наступления будущих событий; эти действия называют реакцией модели на события;
- 5) перечисленные действия повторяются до истечения времени моделирования.

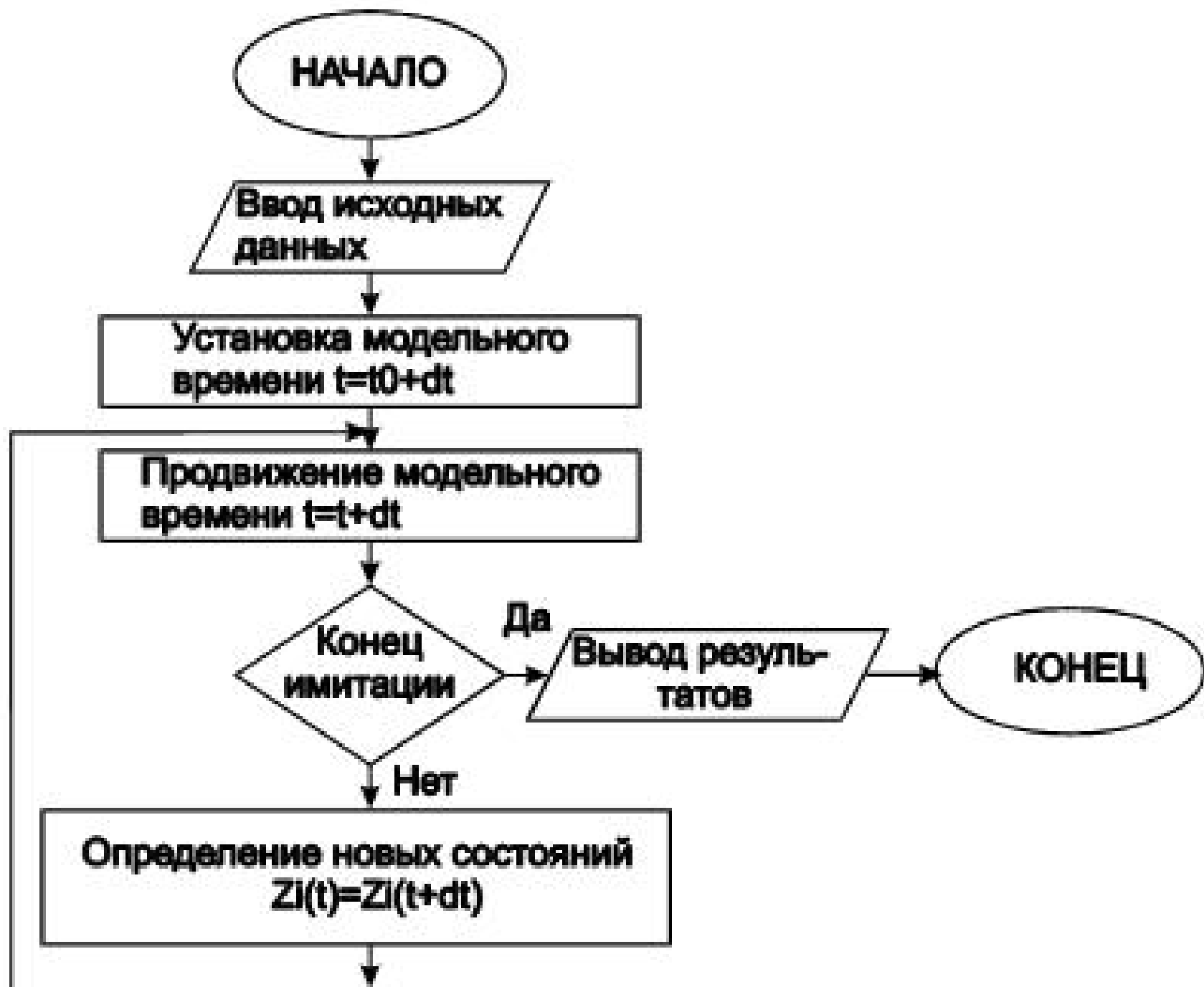
- В процессе моделирования производится измерение и статистическая обработка значений выходных характеристик. Обобщённая схема алгоритма моделирования по принципу особых состояний приведена на рисунке



Алгоритм моделирования

- 1) Устанавливаются начальные значения (состояния) $t=0$
- 2) Задается приращение модельного времени
- 3) На основе вектора текущих состояний элементов модели и нового времени мы рассчитываем новые значения этих состояний
- 4) Если не превышено время моделирования, то три предыдущих этапа повторяются.

- Моделирующий алгоритм, основанный на применении dt применим для более широкого круга систем, чем алгоритм, построенный по принципу особых состояний. Однако при его реализации возникают проблемы определения величины dt . Для моделирования ВС на системном уровне в основном используются принцип особых состояний.



5. Сети Петри

Современные вычислительные системы невозможно представить без устройств, обеспечивающих обмен информацией между отдельными компьютерами, т. е. без локальных или корпоративных сетей передачи данных.

Отличительной особенностью процессов в любой сети является одновременное прохождение информационных и служебных потоков данных по ее различным фрагментам.

Для описания логики работы сетевой модели могут быть использованы различные средства: либо русский язык (устный или письменный), либо традиционные схемы алгоритмов, либо какие-то другие инструментальные средства.

Однако описание в виде блок-схемы алгоритма модели даже простой сети может оказаться затруднительным прежде всего потому, что такие схемы слабо приспособлены для описания параллельных процессов.

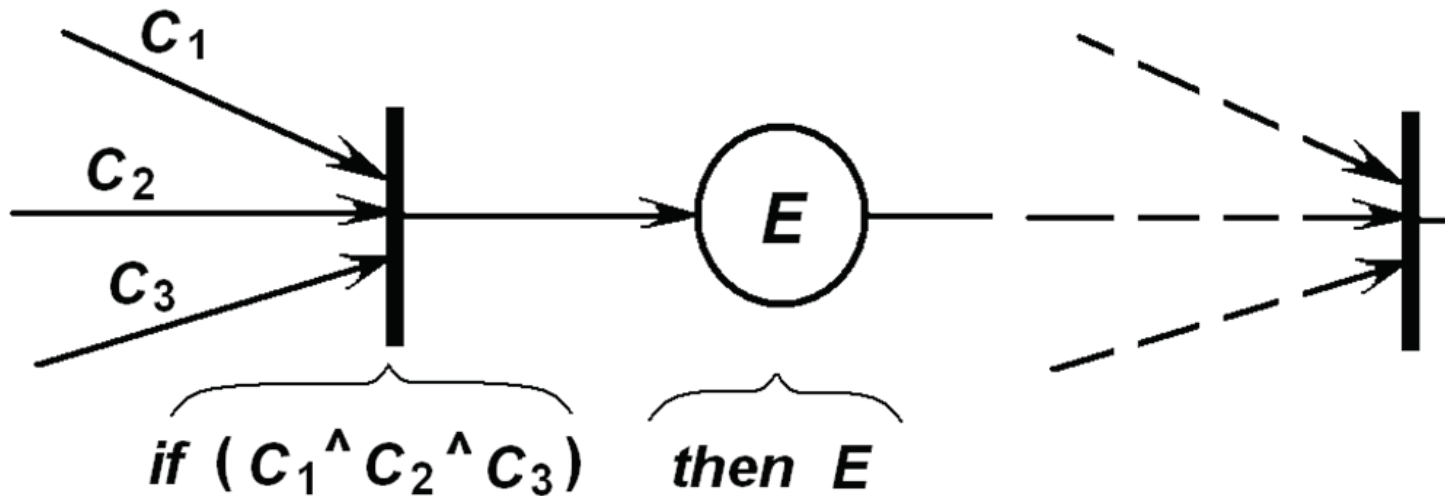
Именно поэтому представляется необходимым изучить распространенные средств описания параллельных процессов — **сетями Петри**

5.1. Примитивные сети Петри

Одно из основных достоинств аппарата сетей Петри (сокращенно N-сети) заключается в том, что они могут быть представлены как в графической форме (это обеспечивает наглядность), так и в аналитической (это позволяет автоматизировать процесс их анализа).

При графической интерпретации сеть Петри представляет собой граф особого вида, состоящий из вершин двух типов — позиций P и переходов T , соединенных ориентированными ветвями (множества I и O , характеризующие начала и концы ветвей),

Каждая ветвь может связывать лишь разнотипные вершины (позицию с переходом или переход с позицией). Вершины-позиции P обозначаются точками или кружками, вершины-переходы T — черточками или узкими прямоугольниками



С содержательной точки зрения переходы соответствуют событиям, присущим исследуемой системе, а позиции — условиям их возникновения. Таким образом, совокупность переходов, позиций и ветвей графа позволяет наглядно описать причинно-следственные связи, присущие системе, но в статике (см.Рис.)

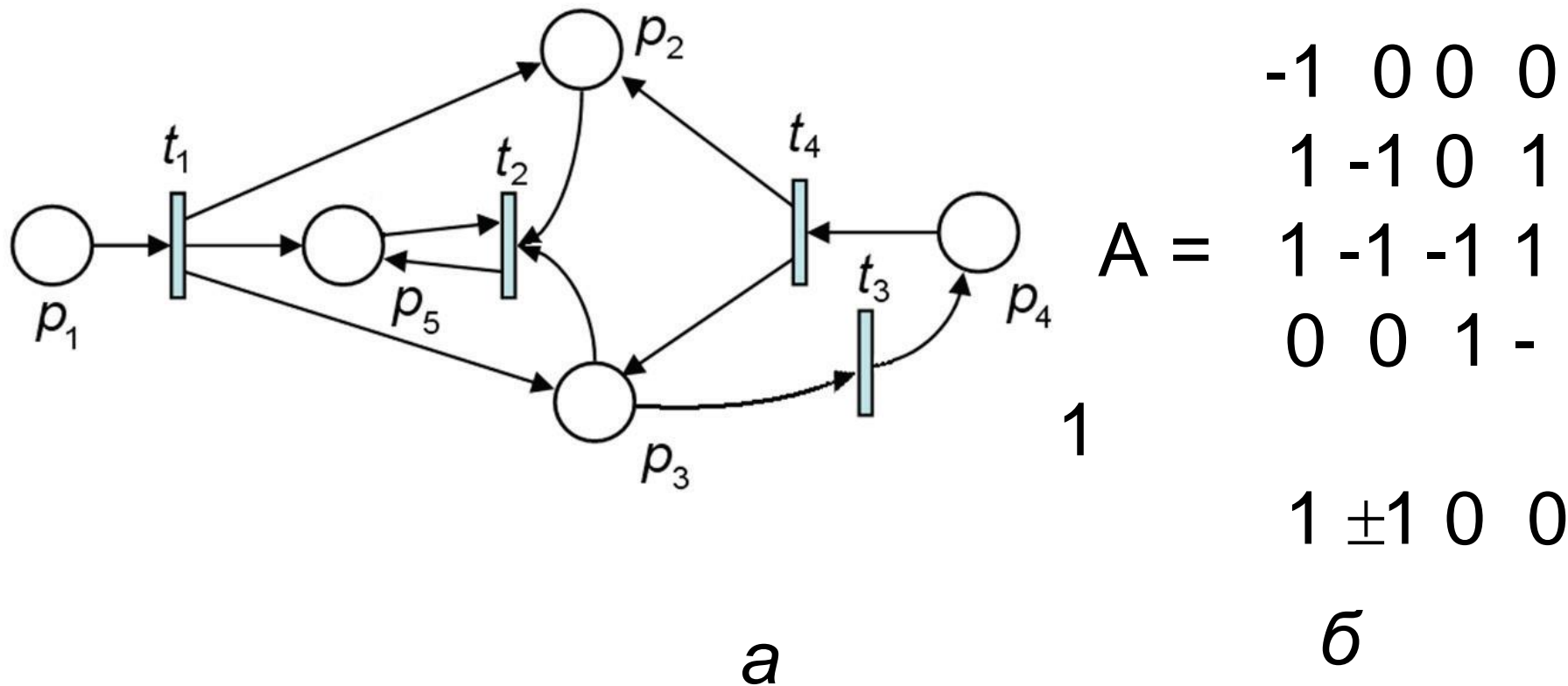


Рис. Классический пример графа N -сети — *a*
и его матрица инциденций — *b*

Формально простая сеть Петри
(N -схема) задается четверкой вида:

$$N = \langle P, T, I, O \rangle,$$

где P — конечное непустое множество
символов-позиций;

T — конечное непустое множество
символов-переходов;

I — входная функция (прямая
функция инцидентности);

O — выходная функция (обратная
функция инцидентности).

Таким образом, входная функция I отображает переход t_j во множество входных позиций

$p_i \in I(t_j)$, а выходная функция O отображает переход t_j во множество выходных позиций $p_i \in O(t_j)$. Для каждого перехода $t_j \in T$, $j=1, \dots, m$ можно определить множество входных позиций перехода $I(t_j)$ и выходных позиций перехода $O(t_j)$

$$I(t_j) = \{p_i \in P \mid I(p_i, t_j) = 1\},$$

$$O(t_j) = \{p_i \in P \mid O(t_j, p_i) = 1\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}, \quad n = |P|, \quad m = |T| \quad (*)$$

Входная и выходная функции сети Петри (I и O) позволяют описать любую сеть с помощью двух матриц размером $n \times m$ (матриц входных и выходных позиций), имеющих следующую структуру (состоят только из положительных единиц и нулей)

$$I = \begin{matrix} & t_1 & t_2 & \cdots & t_m \\ p_1 & \left[\begin{array}{cccc} 1 & 0 & \cdots & 1 \end{array} \right. \\ p_2 & & 0 & 1 & \cdots & 0 \\ p_3 & & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_n & & 1 & 0 & \cdots & 1 \end{matrix},$$

$$O = \begin{matrix} & t_1 & t_2 & \cdots & t_m \\ p_1 & \left[\begin{array}{cccc} 0 & 0 & \cdots & 1 \end{array} \right. \\ p_2 & & 1 & 1 & \cdots & 1 \\ p_3 & & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_n & & 1 & 0 & \cdots & 0 \end{matrix}.$$

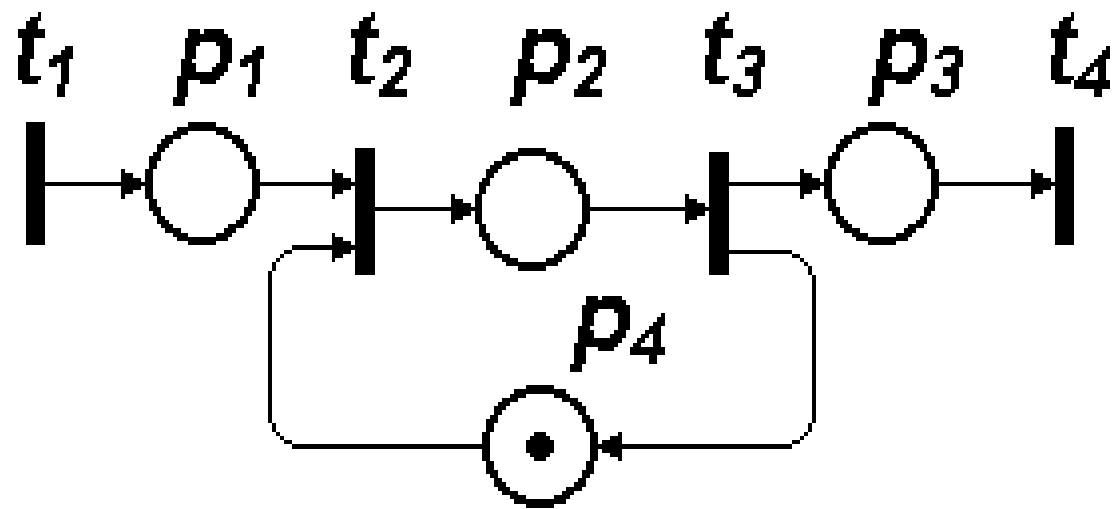
Если ориентацию ветвей сделать одинаковой, например, положительным направлением (t_j, p_i) считать выход из перехода t_j и вход в позицию p_i , то соотношения (*) можно представить как алгебраическую запись матрицы инциденций (см. рис. б), т. е. объединить обе матрицы позиций в одну:

$$A_{ij} = \begin{cases} -I(p_i, t_j) \cup O(t_j, p_i), & I(p_i, t_j) \times O(t_j, p_i) = 0; \\ \pm 1, & I(p_i, t_j) \times O(t_j, p_i) \neq 0. \end{cases} \quad (**)$$

Первая строка в выражении (**)
соответствует одной ветви,
связывающей i -ю позицию и j -й
переход, вторая строка — двум
встречно направленным ветвям (петля
из двух ветвей, как, например, на рис.,
а между позицией p_5 и переходом t_2).

Пример 1

Требуется описать с помощью сети Петри работу группы пользователей на единственной рабочей станции WS при заданных характеристиках потока запросов на пользование WS и характеристиках поступающих задач. Сеть Петри представлена на рисунке:



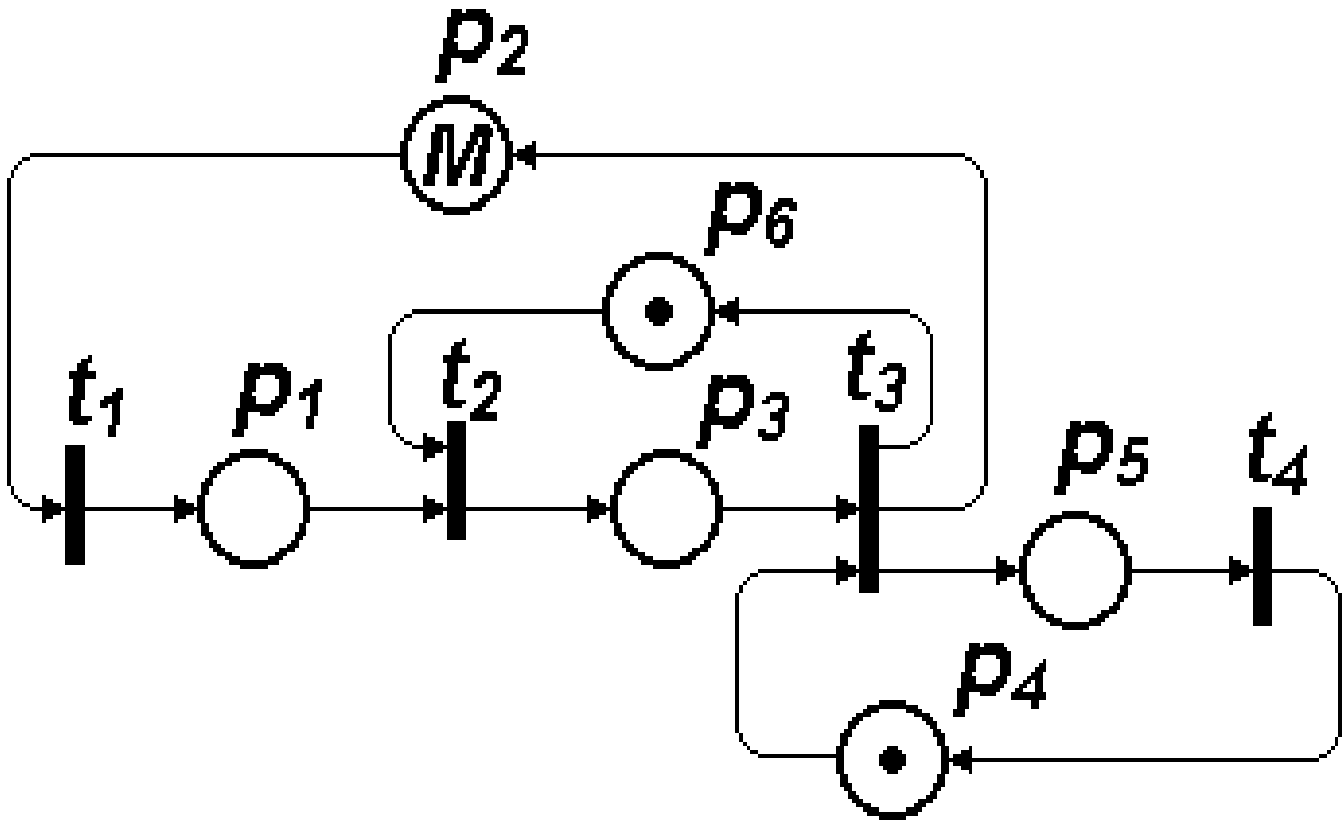
Здесь переходы связаны со следующими событиями: t_1 — поступление запроса на использование WS , t_2 — занятие станции, t_3 — освобождение станции, t_4 — выход обслуженной заявки; позиция p_4 используется для отображения состояния WS : если в p_4 имеется метка, то WS свободна и пришедшая заявка вызывает срабатывание перехода t_2 ; пока эта заявка не будет обслужена, метки в p_4 не будет, следовательно, пришедшие в позицию p_1 запросы вынуждены ожидать срабатывания перехода t_4

Пример.2. Требуется описать с помощью сети Петри процессы возникновения и устранения неисправностей в некоторой технической системе, состоящей из множества однотипных блоков; в запасе имеется один исправный блок; известны статистические данные об интенсивностях возникновения отказов и длительностях таких операций, как поиск неисправностей, замена и ремонт отказавшего блока.

Поиск и замену отказавшего блока производит одна бригада, а ремонт замененного блока — другая бригада. Сеть Петри показана на рис. (сл.слайд). Отметим, что при числе меток в позиции, равном M , можно в ней не ставить точек, а записать в позиции значение M .

В примере значение M в позиции P_2 соответствует числу имеющихся в системе блоков. Переходы отображают следующие события: t_1 — отказ блока, t_2 — поиск неисправного блока, t_3 — его замена, t_4 — окончание ремонта.

Графически сеть Петри изображается в виде двудольного ориентированного мультиграфа , представляющего собой совокупность позиций и переходов (см. рис.). Как видно из этого рисунка, граф N-схемы имеет два типа вершин: позиции и переходы, изображаемые Р и Т соответственно.



Ориентированные ветви $a_{ji} = (t_j, p_i)$ соединяют позиции и переходы, причем каждая ветвь направлена от элемента одного множества (позиции или перехода) к элементу другого множества (переходу или позиции). Граф N-схемы является мультиграфом, так как он допускает существование кратных ветвей от одной вершины к другой.

Приведенные выше представления N-схемы в виде множеств, графа и матриц могут использоваться только для отображения статических состояний моделируемой системы (взаимосвязи событий и условий), описывая ее, по существу, в форме структурной схемы.

Для отражения в модели динамики функционирования моделируемой системы вводят еще один вид объектов сети Петри — так называемые фишки или метки m_i позиций $p_i \in P, i = 1, \dots, n$.

Процесс присвоения неких абстрактных объектов — меток — позициям N-схемы называется маркировкой и обозначается символом M, а ее результат — положение меток в позициях сети — называется разметкой сети.

Количество меток,
соответствующее каждой позиции,
может меняться. При графическом
задании N -схемы разметка
отображается помещением внутри
вершин- позиций
соответствующего числа точек
(когда количество точек велико,
ставят цифры).

Важной особенностью моделей процесса функционирования систем с использованием типовых N -схем является простота построения иерархических конструкций модели. С одной стороны, каждая N -схема может рассматриваться как макропереход или макропозиция модели более высокого уровня.

С другой стороны, переход или позиция *N*-схемы могут детализироваться в форме отдельной подсети для более углубленного исследования процессов в моделируемой системе (рис. Сл.слайд).

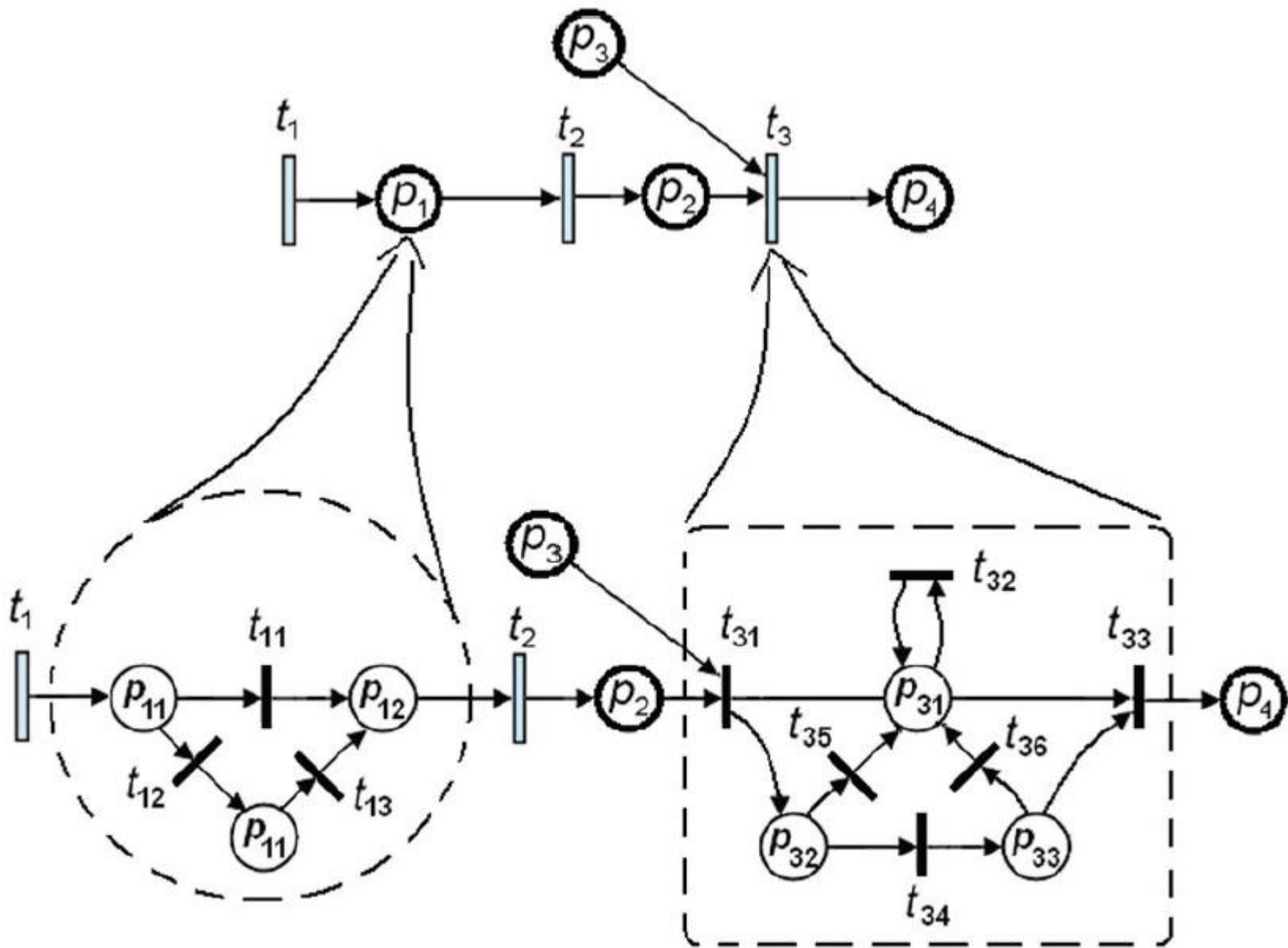


Рис. Иерархические расширения сети Петри

Отсюда вытекает возможность эффективного использования N -схем для моделирования параллельных и конкурирующих процессов в различных системах. Речь идет о так называемых вложенных сетях Петри (*Nested Petri Nets* — *NPN*).

Появление указанной разновидности сетей Петри связано с желанием исследователей иметь инструмент для адекватного и удобного представления систем со сложной иерархической и мультиагентной структурой.

Вложенные сети Петри представляют собой расширение стандартного формализма сетей Петри, в котором метки, представляющие локальные ресурсы в позициях системной сети, сами могут быть сложными объектами с сетевой структурой и моделироваться сетями Петри нижнего уровня — их мы также будем называть *спутниковыми* сетями EN_i .

Структурно такая сеть состоит из системной N -сети и набора сетей-меток (спутников). При этом между некоторыми переходами системной сети и переходами сетей-меток может быть установлена связь, разрешающая только их совместное срабатывание. Такие переходы называются *помеченными*.

Достоинства применения сетей Петри для изучения процессов в сложных системах заключаются в следующем:

- позволяют моделировать параллельные процессы всех возможных типов с учетом возможных конфликтов между ними (универсальность);

- обладают наглядностью и обеспечивают возможность автоматизированного анализа (эффективность);
- позволяют переходить от одного уровня детализации описания системы к другому за счет раскрытия или закрытия позиций и переходов (иерархичность).

Вместе с тем сети Петри имеют ряд недостатков, ограничивающих их возможности. Основной недостаток — время срабатывания перехода считается равным нулю, что не позволяет исследовать с помощью простых сетей Петри временные характеристики моделируемых систем.

Хотя типовые N -схемы на основе обычных размеченных сетей Петри пригодны для описания в моделируемой системе событий произвольной длительности, но в этом случае модель отражает только порядок наступления событий в исследуемой системе.

Для отражения временных параметров процесса функционирования моделируемой системы на базе N -схем следует использовать различные расширения аппарата сетей Петри: временные сети, E -сети, сети Мерлина и т. п.

N-сети с мгновенным выполнением переходов называются *примитивными*. В них возникновение двух событий одновременно невозможно.

Непримитивными называются такие события, длительность которых отлична от нуля. Любое непримитивное событие может быть представлено в виде двух примитивных событий, определяющих начало и конец непримитивного события и состояния, когда это непримитивное событие происходит.

Одним из самых известных расширений сетей Петри являются так называемые *раскрашенные* сети Петри, которые применяются для моделирования временных процессов.

Особенность этой методологии состоит в том, что она моделирует системы в терминах «разнотипные условия — события», что позволяет исследовать динамику работы сложной системы. Модели на основе *CPN*-схем («событийные модели») в настоящее время встраиваются в универсальные системы моделирования информационных систем, таких как ARJS, IDEF и другие.

Для автономного использования методологии CPN-схем разработан специальный язык моделирования *Coloured Petri Net Modeling Language (CPN ML)* [17] и созданы соответствующие программные средства. С возможностями системы (*CPN/Tools*), работающей под *Windows*, можно ознакомиться на сайте <http://www.daimi.au.dk/CPnets/CPN2000>.

Методология *CPN* близка к структурным методам моделирования систем, однако в отличие от многих из них она базируется на хорошо разработанном математическом аппарате и поэтому допускает проведение аналитических и имитационных исследований.

Модель динамической системы в методологии CPN представляет собой сеть — двудольный ориентированный мультиграф. Сеть, как и ранее, содержит узлы двух видов — позиции и переходы, связанные ветвями. При этом позиции моделируют наличие в системе определенных ресурсов, необходимых для наступления события.

Различные виды ресурсов условно обозначаются разными «цветами», что объясняет название методологии.

Переходы моделируют сами события. Наступление какого-либо события в системе изменяет условия, что может привести к новым событиям. Ветви моделируют причинно-следственные связи в системе, а развернутая во времени последовательность условий и событий во времени образует модель динамики системы.

При анализе сети Петри основное внимание уделяется, как правило, решению трех проблем:

- *Проблема достижимости.* В сети Петри с начальной разметкой M_0 требуется определить, достижима ли принципиально из M_0 некоторая разметка M_k . С точки зрения исследования моделируемой системы эта проблема интерпретируется как проблема достижимости (реализуемости) некоторого состояния системы.

Оценка живости переходов сети. Под живостью перехода T_j понимают возможность его срабатывания в данной сети при начальной разметке M_0 . Анализ модели на свойство живости позволяет выявить невозможные состояния в моделируемой системе (например, неисполняемые ветви в программе).

Оценка безопасности сети. Безопасной является такая сеть Петри, в которой ни при каких условиях не может появиться более одной метки в каждой из позиций. Для исследуемой системы это означает возможность функционирования ее в стационарном режиме. На основе анализа данного свойства могут быть, например, определены требования к буферным накопителям в системе.

Пример. Моделирование процесса интерактивного обучения.

Применим рассмотренную выше теорию к моделированию процесса обучения.

Рассмотрим модель прохождения под управлением автоматизированной обучающей системы (АОС)

интерактивного курса, содержащего ряд учебных модулей, которые должны быть изучены

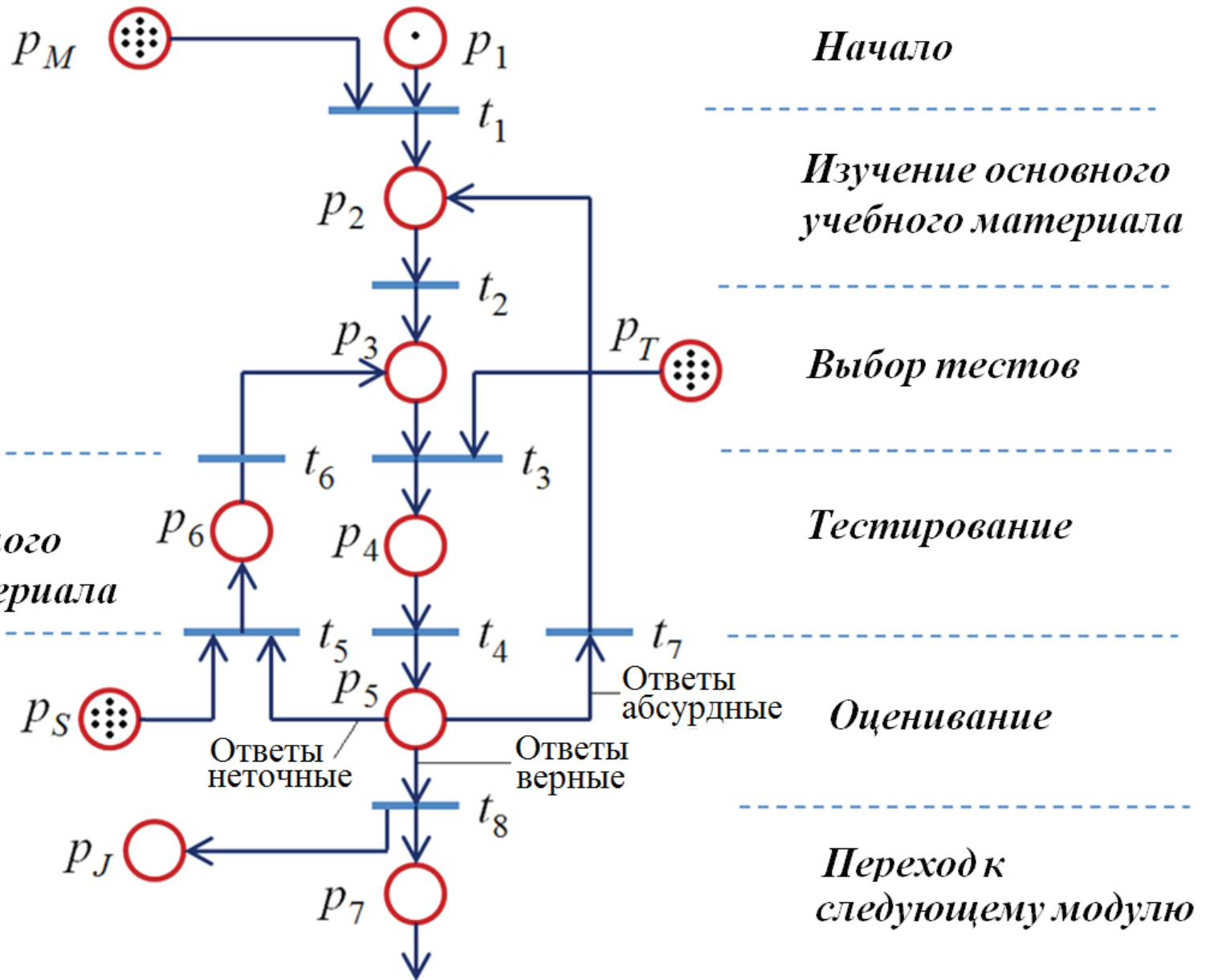
в последовательном порядке.

Процесс прохождения студентом учебного модуля заключается в следующем. Из базы учебных модулей извлекается очередная порция теоретического материала, которую предлагается освоить обучаемому. После того как обучаемый окончил изучение этого материала, система приступает к тестированию.

Из базы тестов выбирается тестовый материал и предъявляется обучаемому, который готовит и вводит в систему ответы на тестовые задания. Эти ответы анализируются системой оценивания, которая принимает решение:

1. ответы верные: в этом случае изучение данного модуля завершается и возможен переход к следующему модулю;
2. · ответы неточные: в этом случае обучаемый должен изучить дополнительный материал и затем пройти повторное тестирование;
3. · ответы абсурдные: в этом случае обучаемый должен изучить материал модуля с самого начала.

Построение модели проводится на основе принципа от простого к сложному для того, чтобы продемонстрировать моделирующие возможности сетей Петри и в то же время облегчить понимание используемых формализмов.



Сеть Петри, моделирующая изучение отдельного модуля

Во вложенных сетях Петри фишки, помечающие позиции, рассматриваются как объекты, имеющие самостоятельное поведение, которое описывается также некоторыми сетями Петри. Название «вложенные сети» указывает на то, что элементы сетей в них сами являются сетями, подобно тому, как в системе вложенных множеств элементами некоторого множества могут быть другие множества.

Примером приложения указанной выше методики может быть разработка функциональной модели информационного портала сетями Петри, апробированная созданием на ее основе нескольких реальных порталов научно-образовательной тематики.

Теория сетей Петри развивается в нескольких направлениях:

- 1). разработка математических основ,
- 2). структурная теория сетей,
- 3). различные приложения (параллельное программирование,
- 4). дискретные динамические системы и т. д.).

МОДЕЛИРОВАНИЯ ФИЗИЧЕСКИХ ПРОЦЕССОВ В ПЕРСПЕКТИВНЫХ ПОЛУПРОВОДНИКОВЫХ ПРИБОРАХ С УЧЕТОМ РАДИАЦИОННОГО ВОЗДЕЙСТВИЯ

Начало исследований в области радиационной физики полупроводниковых приборов приходится на 60-е года XX века и обусловлено: активным внедрением дискретных элементов в военные и космические системы, успехами в радиационной физике твердого тела, появлением лабораторных источников проникающих излучений и другими факторами.

В это время появляются первые аналитические модели реакции полупроводниковых приборов на радиационное воздействие. Однако такого рода модели принципиально не могут учесть нелинейные эффекты, связанные с высоким уровнем инжекции и фотовозбуждения неравновесных носителей заряда в полупроводниковых структурах.

Строгий расчет переходных ионизационных процессов при высоких концентрациях неравновесных носителей заряда можно провести только на основе численного моделирования. Впервые данная задача была решена в 1967г. на основе диффузионно-дрейфового приближения. что ограничивает его применимость при высоких уровнях инжекции или фотовозбуждения.

Однако итерационный алгоритм, основанный на поочередном решении разностных аналогов уравнения Пуассона и уравнений непрерывности для электронов и дырок, имеет низкую скорость сходимости вдали от состояния равновесия, что ограничивает его применимость при высоких уровнях инжекции или фотовозбуждения.

Для преодоления этого недостатка был разработан ряд моделей, носящих обобщенное название квазигидродинамические. В указанной работе для учета разогревающего воздействия потока квантов высоких энергий в уравнения баланса энергии введено дополнительное слагаемое, учитывающее данный эффект

Позднее, в систему уравнений квазигидродинамической модели были добавлены уравнения баланса импульса электронного и дырочного газа, позволяющие учесть дополнительное рассеяние носителей заряда на кластерах радиационных дефектов. Однако для численного решения по-прежнему был использован алгоритм Гуммеля.

В дальнейшем пространственно-дискретизированная диффузионнодрейфовая модель была представлена как система дифференциально-алгебраических уравнений. Данный подход позволил применить к решению задачи переноса носителей заряда жестко-устойчивые методы Гира высокого порядка точности.

Подход на основе метода прямых к решению задач моделирования физических процессов в полупроводниковых приборах представляется эффективным в связи с использованием при интегрировании возникающих систем дифференциально-алгебраических уравнений стандартных пакетов математических программ, предусматривающих:

- автоматический выбор шага интегрирования и порядка точности метода при заданной относительной ошибке;
- возможность использования стандартных программ для решения возникающих линейных уравнений с разреженной матрицей коэффициентов, что особенно ценно при решении многомерных задач;
- достаточно широкие возможности по изменению параметров алгоритма, что позволяет пользователю наилучшим образом подстраивать программу под конкретную задачу.

Анализ вышеизложенных тенденций развития численного моделирования переноса носителей заряда в полупроводниковых приборах при радиационном воздействии показывает, что объем вычислений увеличивается с уменьшением размеров рабочей области элемента. Это обусловлено как увеличением числа уравнений модели, так и уменьшением временного шага интегрирования системы уравнений переноса при переходе от диффузионно- дрейфового к квазигидродинамическому приближению.

Последнее связано с тем, что характерное время изменения концентраций электронов и дырок в диффузионно-дрейфовой модели определяется их временами жизни, а характерное время изменения их средних энергий (характеристика электронно-дырочного газа, не учитываемая в диффузионно-дрейфовой модели) в квазигидродинамической модели определяется временами релаксации энергии и импульса.

Учитывая, что времена жизни носителей заряда лежат в пределах $10E-3...10E-9$ с, а времена релаксации энергии и импульса электронно-дырочного газа – $10E-12...10E-13$ с, характерное уменьшение временного шага интегрирования составляет как минимум 4 порядка. Данное положение подтверждается путем анализа результатов численного моделирования.

Столь резкое увеличение объема вычислений требует применения радикальных мер повышения производительности, основным из которых в настоящее время является распараллеливание численных алгоритмов.

Использование стандартных библиотек программ решения систем дифференциальных, нелинейных и линейных алгебраических уравнений, многие из которых содержат эффективные параллельные алгоритмы, определяет применение данного подхода в области высокопроизводительных вычислений на супер-ЭВМ в качестве базового.

Физико-топологическое моделирование важно как для понимания физических процессов, протекающих в полупроводниковых приборах, в том числе при радиационном воздействии, так и для оптимизации их конструкции с целью достижения наилучших характеристик. Именно моделирование отражает эволюцию представлений о том, как работает транзистор при различных его размерах/.

Для микронных размеров диодов и транзисторов, когда длина свободного пробега была гораздо меньше длины области пространственного заряда p-n перехода, базы биполярного транзистора и канала полевого транзистора, используется представление о том, что в рабочей области полупроводникового прибора протекает заряженная жидкость, и для описания такого ламинарного течения применяются гидродинамические модели: **диффузионно-дрейфовая, электротепловая и квазигидродинамическая**

Для субмикронных размеров элементов, когда длина свободного пробега носителей заряда уже оказывается сопоставимой с характерным масштабом изменения потенциала и напряженности электрического поля, более реалистичными оказываются представления о частицах, пролетающих через активную область прибора. Для описания переноса носителей заряда в этом случае используется кинетическое **уравнение Больцмана.**

Отметим, что семейство гидродинамических моделей переноса носителей заряда может быть формально получено из кинетического уравнения Больцмана с использованием ряда упрощающих предположений.

Границей применимости кинетического уравнения **Больцмана** и метода **Монте-Карло** как процедуры его решения является рабочая область прибора, меньшая длины волны де Бройля для электронов и дырок. В этом случае активная область прибора ведет себя как волновод, для корректного описания которого необходимо использовать уравнение **Шредингера**. Такой полупроводниковый прибор становится квантовым, хотя на внешних электродах измеряются классические значения токов и напряжений.

Мы ограничимся рассмотрением гидродинамических моделей переноса носителей заряда и методов их распараллеливания. Так отмечалось, существует большое количество вариантов формулировок систем уравнений, описывающих перенос носителей заряда в гидродинамическом приближении. Каждая модель содержит уравнение Пуассона и набор уравнений непрерывности (баланса).

Наиболее полное описание переноса в гидродинамическом приближении (так называемая «полная» квазигидродинамическая модель) включает в себя: уравнение Пуассона, уравнения непрерывности для электронов и дырок, уравнения баланса энергии электронов, дырок и кристаллической решетки, уравнения баланса импульса электронов и дырок и уравнения плотностей токов и потоков энергии электронов и дырок:

$$\operatorname{div}(\epsilon \nabla \phi) = -\frac{q}{\epsilon_0}(p - n + N_d - N_a), \quad (1)$$

$$\frac{\partial n}{\partial t} = \frac{1}{q} \operatorname{div}(j_n) - R + G, \quad (2)$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \operatorname{div}(j_p) - R + G, \quad (3)$$

$$\frac{\partial(W_n n)}{\partial t} = \operatorname{div}(S_n) + j_n E - \frac{n(W_n - W_0)}{\tau_{W_n}(W_n)} - RW_n + GW_e, \quad (4)$$

$$\frac{\partial(W_p p)}{\partial t} = \text{div}(S_p) + j_p E - \frac{p(W_p - W_0)}{\tau_{W_p}(W_p)} - RW_p + GW_h, \quad (5)$$

$$c\rho \frac{\partial T}{\partial t} = \text{div}(k\nabla T) + \frac{n(W_n - W_0)}{\tau_{W_n}(W_n)} + \frac{p(W_p - W_0)}{\tau_{W_p}(W_p)} + R(W_g + W_n + W_p), \quad (6)$$

$$\frac{\partial(m_e(W_n)v_n)}{\partial t} = -m_e(W_n)v_n \text{div}(v_n) + q\nabla\phi - \frac{m_e(W_e)v_n}{\tau_{pn}(W_n)}, \quad (7)$$

$$\frac{\partial(m_h(W_p)v_p)}{\partial t} = -m_h(W_p)v_p \text{div}(v_p) - q\nabla\phi - \frac{m_h(W_e)v_p}{\tau_{pp}(W_p)}, \quad (8)$$

$$j_n = -qnv_n + qD_n\nabla n, \quad (9)$$

$$j_p = qp v_p - qD_p\nabla p, \quad (10)$$

$$S_n = -W_n n v_n + D_n\nabla(W_n n), \quad (11)$$

$$S_p = W_p p v_p - D_p\nabla(W_p p), \quad (12)$$

- где φ – потенциал электрического поля, n – концентрация электронов; p – концентрация дырок, T – температура кристаллической решетки, Wn – средняя энергия электронов, Wp – средняя энергия дырок, U_n – дрейфовая скорость электронов, U_p – дрейфовая скорость дырок, j_n – плотность электронного тока, j_p – плотность дырочного тока, S_n – плотность потока энергии электронов, S_p – плотность потока энергии дырок, Dn – коэффициент диффузии электронов, Dp – коэффициент диффузии дырок, τ_{Wn} – время релаксации энергии электронов

τ_{Wp} – время релаксации энергии дырок, W_e – средняя энергия радиационно-генерируемых электронов, W_h – средняя энергия радиационно-генерируемых дырок, N_d – концентрация положительно заряженных ионов примеси, N_a – концентрация отрицательно заряженных ионов примеси, ϵ_0 – диэлектрическая проницаемость вакуума, ϵ – диэлектрическая проницаемость полупроводника, R – коэффициент рекомбинации, G – коэффициент генерации, k – теплопроводность, c – удельная теплоемкость,

ρ – плотность, m_e – эффективная масса электронов, m_h – эффективная масса дырок, τ_{np} – время релаксации импульса электронов, τ_{pp} – время релаксации импульса дырок.

- В «усеченной» квазигидродинамической модели (модели локальной температуры) уравнения баланса импульсов электронов (7) и дырок (8) отсутствуют. В связи с этим средняя дрейфовая скорость электронов и дырок становятся функцией их средней энергии $u_n(W_n)$ и $u_p(W_p)$, соответственно.

В электротепловой модели электрофизические параметры полупроводника становятся функциями напряженности электрического поля и температуры кристаллической решетки, а уравнения баланса энергии электронов (4) и дырок (5) исключаются. В диффузионно-дрейфовой модели моделирование переноса носителей заряда проводится при фиксированной температуре полупроводникового прибора и уравнение теплопроводности (6) также исключается.

Для учета воздействия проникающих излучений при моделировании переноса носителей заряда в уравнения непрерывности электронов (2) и дырок (3) вводится слагаемое, отвечающее за генерацию неравновесных носителей заряда, в уравнения баланса энергии добавлено слагаемое, отвечающее за разогрев электронно-дырочного газа, а электрофизические параметры полупроводника зависят от потоков дефектообразующих излучений.

Численные методы решения задачи переноса носителей заряда в полупроводниковых приборах при воздействии проникающих излучений

При построении численных алгоритмов удобно пользоваться системой уравнений переноса носителей заряда, приведенной к безразмерному виду. Это позволяет не только уменьшить мантиссы обрабатываемых чисел до приемлемых величин, пригодных для цифровой обработки, что особо актуально для вычислительных архитектур, использующих числа одинарной точности

Например, технология параллельных вычислений CUDA корпорации Nvidia, но также избавиться от некоторых постоянных размерных коэффициентов в уравнениях модели.

Искомými функциями при решении системы уравнений переноса носителей заряда являются зависимости от координат и времени:

- в диффузионно-дрейфовом приближении:
потенциала электрического поля и концентрации электронов и дырок;
- – в электротепловом приближении:
потенциала электрического поля,
концентрации электронов и дырок и температуры кристаллической решетки;

– в квазигидродинамическом приближении:
потенциала электрического поля,
концентрации электронов и дырок,
температуры кристаллической решетки,
средней энергии электронов и дырок
при заданных зависимостях
электрофизических параметров
полупроводника, начальных и граничных
условиях.

Множество исходных функций системы уравнений переноса носителей заряда принято называть базисом переменных. В диффузионно-дрейфовом приближении распространены три базиса:

- $\{\varphi, n, p\}$ – потенциал электрического поля, концентрации электронов и дырок;
- $\{\varphi, \varphi_n, \varphi_p\}$ – потенциал электрического поля, квазиуровни Ферми для электронов и дырок;
- $\{\varphi, \Phi_n, \Phi_p\}$ – потенциал электрического поля, экспоненты квазиуровней Ферми для электронов и дырок.

В электротепловом и квазигидродинамическом приближениях принято использовать базисы $\{\varphi, n, p, T\}$ и $\{\varphi, n, p, T, Wn, Wp\}$, $\{\varphi, n, p, T, Wn, Wp, un, up\}$, соответственно.

Сведение системы уравнений переноса носителей заряда к дифференциально-алгебраической системе уравнений

Любая система дифференциальных уравнений в частных производных параболического и эллиптического типов, к числу которых относится семейство систем уравнений переноса носителей заряда в полупроводниках, может быть сведена к системе дифференциально-алгебраических уравнений вида

$$M \frac{du}{dt} = f(u, t) \quad (13)$$

путем дискретизации пространственной части уравнений в частных производных. В выражении (13) u – вектор нормированных искомым переменных, M – матрица массы, f – векторная функция нормированных пространственно-дискретизированных правых частей системы уравнений переноса носителей заряда, t – нормированное время.

В настоящее время применяются 3 метода пространственной дискретизации:

- конечных объемов (FVM);
- конечных элементов (FEM);
- конечных разностей (FDM).

- Дискретизация определяется формой записи системы дифференциальных уравнений в частных производных: для дифференциальной формы записи дискретизация осуществляется при помощи конечно-разностных методов; интегральная форма записи дискретизируется при помощи метода конечных объемов, а вариационная форма – при помощи метода конечных элементов.

Методы дискретизации, их преимущества и недостатки, широко рассмотрены в литературе. Применительно к одномерной квазигидродинамической модели переноса носителей заряда, рассматриваемой ниже, явным преимуществом будет обладать метод конечных разностей.

Основной акцент сделаем на дискретизации плотностей токов и потоков энергии носителей заряда. Поэтому без ущерба общности обсуждения дальнейших результатов исключим из системы (1)-(12) уравнения баланса импульса электронов и дырок, а также уравнение теплопроводности кристаллической решетки.

На отрезке $[0, 1]$ введем неравномерную пространственную сетку с узлами $x = \{x_i\}$, $i = 0, 1, \dots, N - 1$, согласованную с распределением легирующей примеси: вблизи границ раздела и контактов полупроводникового прибора шаг сетки делаем более мелким. Определим значения базисных функций $\{\varphi, n, p, W_n, W_p\}$ в следующем виде: $u = \{u_i\}$, $u_{5i} = \varphi_i$,
 $u_{5i+1} = n_i$, $u_{5i+2} = p_i$, $u_{5i+3} = (W_n)_i$,
 $u_{5i+4} = (W_p)_i$

Тогда векторная функция нормированных пространственно - дискретизированных правых частей системы уравнений переноса носителей заряда в квазигидродинамическом приближении записывается в виде

$$(j_n)_{i+1/2} = \begin{cases} \frac{u_{5i+8} - u_{5i+3} + u_{5i} - u_{5i+5}}{x_{i+1} - x_i} \frac{((\mu_n)_{i+1} u_{5i+6} - (\mu_n)_i u_{5i+1}) \times h_n(u_{5i}, u_{5i+3}, u_{5i+5}, u_{5i+8})}{1 - h_n(u_{5i}, u_{5i+3}, u_{5i+5}, u_{5i+8})} & h_n(\dots) \neq 1 \\ \frac{((\mu_n)_{i+1} u_{5i+6} u_{5i+8} - (\mu_n)_i u_{5i+1} u_{5i+3})}{x_{i+1} - x_i} & h_n(\dots) = 1 \end{cases},$$

$$(j_p)_{i+1/2} = \begin{cases} \frac{u_{5i} - u_{5i+5} - u_{5i+9} + u_{5i+4}}{x_{i+1} - x_i} \frac{((\mu_p)_{i+1} u_{5i+7} - (\mu_p)_i u_{5i+2}) \times h_p(u_{5i}, u_{5i+4}, u_{5i+5}, u_{5i+9})}{1 - h_p(u_{5i}, u_{5i+4}, u_{5i+5}, u_{5i+9})} & h_p(\dots) \neq 1 \\ \frac{((\mu_p)_i u_{5i+2} u_{5i+4} - (\mu_p)_{i+1} u_{5i+7} u_{5i+9})}{x_{i+1} - x_i} & h_p(\dots) = 1 \end{cases},$$

$$(S_n)_{i+1/2} = \begin{cases} \frac{u_{5i+8} - u_{5i+3} + u_{5i} - u_{5i+5}}{x_{i+1} - x_i} \frac{((\mu_n)_{i+1} u_{5i+6} u_{5i+8} - (\mu_n)_i u_{5i+1} u_{5i+3}) \times h_n(u_{5i}, u_{5i+3}, u_{5i+5}, u_{5i+8})}{1 - h_n(u_{5i}, u_{5i+3}, u_{5i+5}, u_{5i+8})} & h_n(\dots) \neq 1 \\ \frac{(\mu_n)_{i+1} u_{5i+6} u_{5i+8}^2 - (\mu_n)_i u_{5i+1} u_{5i+3}^2}{x_{i+1} - x_i} & h_n(\dots) = 1 \end{cases},$$

$$(S_p)_{i+1/2} = \begin{cases} \frac{u_{5i} - u_{5i+5} - u_{5i+9} + u_{5i+4}}{x_{i+1} - x_i} \frac{((\mu_p)_{i+1} u_{5i+7} u_{5i+9} - (\mu_p)_i u_{5i+2} u_{5i+4}) \times h_p(u_{5i}, u_{5i+4}, u_{5i+5}, u_{5i+9})}{1 - h_p(u_{5i}, u_{5i+4}, u_{5i+5}, u_{5i+9})} & h_p(\dots) \neq 1 \\ \frac{(\mu_p)_i u_{5i+2} u_{5i+4}^2 - (\mu_p)_{i+1} u_{5i+7} u_{5i+9}^2}{x_{i+1} - x_i} & h_p(\dots) = 1 \end{cases},$$

$$h_n(u_{5i}, u_{5i+3}, u_{5i+5}, u_{5i+8}) = \exp\left(-2 \frac{u_{5i+8} - u_{5i+3} + u_{5i} - u_{5i+5}}{u_{5i+8} + u_{5i+3}}\right),$$

$$h_p(u_{5i}, u_{5i+4}, u_{5i+5}, u_{5i+9}) = \exp\left(2 \frac{u_{5i} - u_{5i+5} - u_{5i+9} + u_{5i+4}}{u_{5i+9} + u_{5i+4}}\right).$$

- Отметим, что если положить равенство средних энергий электронов и
- дырок тепловой энергии носителей заряда в каждом узле расчетной сетки
- $U_{5i+3} \equiv U_{5i+4} \equiv 1$ и, следовательно, постоянные их подвижности $\mu_i \equiv \mu_{i+1} \equiv \mu_{i+1/2}$, то выражения для плотностей токов переходят в традиционный для диффузионно-дрейфовой модели вид

$$(j_n)_{i+1/2} = \begin{cases} (\mu_n)_{i+1/2} \frac{u_{5i} - u_{5i+5}}{x_{i+1} - x_i} \frac{(u_{5i+6} - u_{5i+1} \exp(u_{5i+5} - u_{5i}))}{1 - \exp(u_{5i+5} - u_{5i})} & u_{5i+5} \neq u_{5i} \\ (\mu_n)_{i+1/2} \frac{u_{5i+6} - u_{5i+1}}{x_{i+1} - x_i} & u_{5i+5} = u_{5i} \end{cases},$$

$$(j_p)_{i+1/2} = \begin{cases} (\mu_p)_{i+1/2} \frac{u_{5i} - u_{5i+5}}{x_{i+1} - x_i} \frac{(u_{5i+7} - u_{5i+2} \exp(u_{5i} - u_{5i+5}))}{1 - \exp(u_{5i} - u_{5i+5})} & u_{5i+5} \neq u_{5i} \\ (\mu_p)_{i+1/2} \frac{u_{5i+2} - u_{5i+7}}{x_{i+1} - x_i} & u_{5i+5} = u_{5i} \end{cases}.$$

Методы решения системы дифференциально - алгебраических уравнений

Любая гидродинамическая модель переноса носителей заряда в полупроводнике включает уравнение Пуассона, определяющее алгебраическую компоненту дискретизированной задачи (13). Матрица массы дифференциально-алгебраических уравнений является вырожденной, что определяет «бесконечную жесткость» задачи (13).

При построении разностных схем для жестких систем предъявляются повышенные требования к устойчивости решения – явные схемы для решения жестких задач требуют очень мелкого шага интегрирования и поэтому практически неприменимы.

Неявные итерационные схемы

Среди неявных схем широко распространены методы Рунге-Кутты (IRK – implicit Runge-Kutta methods). Не все схемы Рунге-Кутты подходят для решения жестких задач. В литературе дан всеобъемлющий обзор методов Рунге-Кутты, пригодных для решения жестких задач, выделены методы, являющиеся жестко точными, то есть сохраняющими при решении систем дифференциально-алгебраических уравнений тот же порядок точности, что и для систем дифференциальных уравнений.

Любой неявный метод Рунге-Кутты для перехода на новый временной слой требует решения системы нелинейных алгебраических уравнений при помощи итераций ньютоновского типа. Для s -стадийного неявного метода Рунге-Кутты минимальное число возникающих нелинейных систем s соответствует диагонально неявным методам (DIRK – diagonal implicit Runge-Kutta methods). Именно они чаще всего и используются на практике

Неявные многошаговые методы (формулы дифференцирования назад) положены в основу популярных программ Гира.

Коэффициенты многошаговых методов подбираются так, чтобы q -шаговый метод имел точность $O(\tau^q)$. Однако можно показать, что неявные многошаговые методы с $q > 2$ теряют свойство безусловной устойчивости. При $q > 6$ неявные многошаговые методы становятся абсолютно неустойчивыми.

Этим неявные многошаговые методы сильно уступают неявным методам Кунге-Кутты. Тем не менее, неявные многошаговые методы первого и второго порядков (неявный метод Эйлера и неявное правило трапеций) реализованы в системе автоматизированного

проектирования изделий микроэлектроники
TCAD Sentaurus фирмы Synopsys для решения
задачи переноса носителей заряда. Также на
базе методов Гира реализован комплекс
программ оценки радиационного
воздействия на изделия микроэлектроники
DIODE-2D разработки НИЯУ МИФИ

Безитерационные схемы

Наличие итераций сильно усложняет использование неявных методов Рунге-Кутты и многошаговых методов, так как к проблемам устойчивости добавляется проблема сходимости итерационного процесса при решении систем нелинейных алгебраических уравнений. Альтернатива, которая обходит эту трудность – методы типа Розенброка (ROS) и Розенброка-Ваннера (ROW)

Формально эти схемы неявные, но итераций в них не возникает и число арифметических действий для перехода на новый временной слой фиксировано и заранее известно (как в явных схемах). За это безусловное преимущество эти схемы получили название явно-неявных или полунеявных. Формулы перехода на новый временной слой однопараметрического семейства одностадийных схем Розенброка имеют вид

$$u = u + \tau \text{Re}k,$$

$$(M - \alpha \tau f'_u(u, t))k = f(u, t + 0.5\tau), \quad (17)$$

$$f'_u \equiv \frac{\partial f}{\partial u} \text{ — матрица Якоби,}$$

τ — шаг по времени, u — решение на текущем временном слое, u' — решение на новом временном слое, α — числовой параметр, определяющий свойства схемы

При $\alpha = 0$ это явная схема, имеющая точность $O(\tau)$. Этот вариант схемы практически непригоден для расчета жестких задач. При $\alpha = 0,5$ получается известная схема «с полусуммой». Она имеет точность $O(\tau^2)$ и безусловно устойчива. При $\alpha = 1$ имеем неявный метод Эйлера

Помимо безусловной устойчивости он имеет хорошее качественное поведение численного решения (за счет L1-устойчивости). Однако неявный метод Эйлера имеет невысокую точность $O(\tau)$, что препятствует его применению.

Описанные выше схемы вещественны. Однако существует одна комплексная схема этого семейства которая обладает уникальными свойствами [точность $O(\tau^2)$], L2-устойчивость и, соответственно, безусловная устойчивость. Эта схема обладает высокой надежностью и пригодна для расчетов с сильной жесткостью. В литературе ее принято называть одностадийной схемой Розенброка с комплексным коэффициентом (CROS).

Лекция . Интегральные уравнения и преобразования в приборных системах

Интегральными уравнениями называются уравнения, в которых неизвестная функция входит под знак интеграла.

Интегральные уравнения делят на **линейные** и **нелинейные уравнения**.

В линейных интегральных уравнениях над неизвестной функцией осуществляется исключительно линейное преобразование, а в нелинейных интегральных уравнениях это преобразование нелинейно.

1 Лине́йные интегральные уравнения

- Пусть $f(t)$ – известная функция; $\varphi(t)$ – неизвестная искомая функция.
- Классическая запись линейного интегрального уравнения имеет вид

- $$\varphi(t) = \lambda \int_a^b k(t, S)\varphi(S)dS + f(t) \quad (1)$$

где λ – параметр интегрального уравнения, который определяет семейство решений интегральных уравнений; $k(t,S)$ – ядро интегрального уравнения. Функция $f(t)$ задана на интервале $a \leq t \leq b$.

Двухмерная функция $k(t,S)$ – ядро интегрального уравнения , определена на следующей области

$$\begin{cases} a \leq t \leq b \\ a \leq S \leq b \end{cases}$$

5.2 Общий вид интегральных уравнений Фредгольма 1-го и 2-го рода

Интегральное уравнение Фредгольма 1-го 2-го рода описывается следующим образом

$$\int_a^b k(t, S)\varphi(S)dS = f(t) \quad (2)$$

$$\varphi(t) = \lambda \int_a^b k(t, S)\varphi(S)dS + f(t) \quad (3)$$

Границы области определения a и b могут быть конечными или бесконечными

Решение уравнений (2), (3) существует, если функция $f(t)$ непрерывна на интервале (a, b) и удовлетворяет соотношению

$$\int_a^b |f(t)|^2 dt < +\infty,$$

а функция $k(t, S)$, являющаяся ядром интегрального уравнения, непрерывна по своим аргументам в области

$$\begin{cases} a \leq t \leq b \\ a \leq S \leq b \end{cases}$$

- Представляет интерес случай $f(t)=0$.

Это интегральное уравнение называется однородным

$$\varphi(t) + \lambda \int_a^b k(t, S) \varphi(S) dS = 0 \quad (4)$$

5.3 Интегральные уравнения Вольтерры 1-го и 2-го рода

- Интегральные уравнения Вольтерры отличаются от интегральных уравнений Фредгольма переменными границами интегрирования. Интегральное уравнение Вольтерры 1-го рода имеет вид

$$\int_a^t k(t, S)\varphi(S)dS = f(t) \quad (5)$$

Интегральное уравнение Вольтерры 2-го рода записывается следующим образом

$$\varphi(t) + \lambda \int_a^t k(t, S) \varphi(S) dS = f(t) \quad (6)$$

В случае $f(t)=0$ уравнение (6) превращается в однородное уравнение Вольтерры

Уравнение Вольтерры, строго говоря, являются обобщением уравнений Фредгольма.

Указанный вывод следует из возможности модификации интегрального ядра в уравнении Вольтерры, в результате которой оно (уравнение) превращается в интегральное уравнение Фредгольма.

Предлагаемая модификация ядра имеет вид

$$H(t, S) = \begin{cases} k(t, S), & S \leq t \\ 0 & , \quad S > t \end{cases}$$

В результате модификации ядра уравнение (6) трансформируется в интегральное уравнение

$$\varphi(t) = \lambda \int_a^b H(t, S) \varphi(S) dS + f(t)$$

полученное уравнение имеет вид (3) и удовлетворяет всем необходимым требованиям, поэтому является интегральным уравнением Фредгольма 2-го рода.

5.4 Виды нелинейных интегральных уравнений

Существует огромное количество нелинейных интегральных уравнений, вытекающих из практических приложений. К таким интегральным уравнениям относится, например, интегральное уравнение Урысона с наиболее общим видом

$$\varphi(t) = \int_a^b k[t, S, \varphi(S)] dS \quad (7)$$

- Таким образом если неизвестная функция входит внутрь ядра, то такие интегральные уравнения называются **уравнениями Урысона**

Уравнение Гаммерштейна имеет вид, схожий с уравнением Фредгольма,

$$\varphi(t) = \int_a^b k(t, S) F[S, \varphi(S)] dS \quad (8)$$

Неизвестная функция является одним из аргументов нелинейного оператора

$$F[S, \varphi(S)]$$

Представление уравнений Ляпунова – Лихтенштейна является ещё более сложным, чем упомянутые выше уравнения. У них ярко выраженная нелинейность, и они имеют два ядра и зависят от двух параметров

$$\varphi(t) = f(t) + \lambda \int_a^b k_1(t, S) \varphi(S) dS + \mu \int_a^b \int_a^b k_2(t, S, z) \varphi(S) \varphi(z) dS dz \quad (9)$$

Обобщённые нелинейные интегральные уравнения Вольтерра описываются выражением.

$$\varphi(t) = \int_a^t k[t, S, \varphi(S)] dS \quad (10)$$

- **5.5 Практически важные примеры**

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{jxy} f(y) dy \quad (10)$$

- Это интегральное уравнение Фредгольма 1-го рода с ядром $k(x, y) = \frac{e^{ixy}}{\sqrt{2\pi}}$

Решение интегрального уравнения имеет

вид:

$$f(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ixy} g(x) dx$$

Эти уравнения представляют **обратное и прямое интегральные преобразования Фурье** и демонстрируют крайне важный подход к решению некоторых классов интегральных уравнений.

$$2 \quad \frac{dx(t)}{dt} = F[t, x(t)]$$

Проинтегрируем обе части этого выражения по t :

$$x(t) = x_0 + \int_a^t F[t, x(t)] dt$$

$$x(a) = x_0$$

3. Общая задача решения линейного дифференциального уравнения n -го порядка

$$\frac{d^n x}{dt^n} + a_1(t) \frac{d^{n-1} x}{dt^{n-1}} + \dots + a_n(t)x(t) = F(t) \quad ,$$

$$x(a) = c_0;$$

$$x'(a) = c_1;$$

...

$$x^{(n-1)}(a) = c_{n-1}.$$

Рассмотрим случай $n=2$

$$\frac{d^2 x}{dt^2} + a_1(t) \frac{dx}{dt} + a_2(t)x(t) = F(t)$$

$$\frac{d^2 x}{dt^2} = \varphi(t)$$

$$\frac{dx}{dt} = \int_0^t \varphi(S) dS + C_1$$

$$\int_0^t dt \int_0^t dt \dots \int_0^t f(t) dt = \frac{1}{(n-1)!} \int_0^t (t-S)^{n-1} f(S) dS$$

$$x(t) = \int_0^t (t-S) \varphi(S) dS + C_1 t + C_0$$

- Подставим в исходное дифференциальное уравнение:

$$\varphi(t) + \int_0^t [a_1(t) + a_2(t)(t - S)]\varphi(S)dS = F(t) - C_1a_1(t) - C_1ta_2(t) - C_0a_2(t)$$

- Обозначим:

$$k(t, S) = -[a_1(t) + a_2(t)(t - S)]$$

$$f(t) = F(t) - C_1a_1(t) - C_1ta_2(t) - C_0a_2(t)$$

- Интегральное уравнение Вольтерра 2-го рода

$$\varphi(t) = \int_0^t k(t, S)\varphi(S) + f(t)$$

5.6 Интегральное уравнение типа свёртки

- Во многих случаях ядро $k(t, S) = k(t - S)$ (зависит от разности аргументов), тогда уравнение Вольтерра называется интегральным уравнением типа свертки.

5.7 Интегральное уравнение Абеля

$$f(x) = \int_0^x \frac{\varphi(S)}{\sqrt{x-S}} dS$$

- *5.8 Интегрально-дифференциальные уравнения*
- Если неизвестная функция входит и под знак производной и под знак интеграла, то такое уравнение называется интегро-дифференциальное (ИДУ).

6 Решение интегральных уравнений методами интегральных преобразований

- Лекция посвящена вопросам, связанным с интегральными уравнениями, возникающими в рентгеновской вычислительной томографии.

Тысячелетия у человечества существовала мечта выяснить внутреннее содержание объекта, ответить на вопрос – а что же там внутри? Первые ответы на указанный вопрос были получены в 1895 году Вильгельмом Конрадом Рентгеном, который открыл рентгеновское излучение, по скромности названное им X-лучами.

Рентгеновское излучение на тот момент времени оказалось единственным видом излучения, для которого анализируемые объекты исследования являются почти прозрачными. Для изображений объектов в рентгеновских лучах характерно наложение теней одних структурных фрагментов на другие, упомянутое наложение при большом количестве фрагментов устранить невозможно.

В истории человечества существует много примеров гениальных предвидений учёных, которые на кончике пера разрабатывали теории, которые в будущем переворачивали мир и в практической реализации становились известны и необходимы человечеству в целом и каждому конкретному человеку в отдельности. К числу таких примеров относится постановка и решение одной из задач интегральной геометрии в 1917 году великим австрийским математиком Иоганном Радоном.

Отметим, что раздел математики под наименованием «интегральная математика» появился лишь в середине 50-ых годов XX-го века. Радон метод восстановления (реконструкции) многомерных функций по их интегральным характеристикам.

Постановку и решение теоретической задачи от их практического воплощения разделило более сорока лет. В 1963 году американский физик Алан Кормак решил задачу восстановления сечения по проекциям, причём сделал это способом, который отличался от преобразования Радона. Следует отметить, что подходы Радона и Кормака к решению анализируемой задачи являются близкими.

Прошло ещё шесть лет и в 1969 году под руководством английского инженера-физика Годфри Ньюболда Хаунсфилда из фирмы «EMI Ltd.» был создан первый компьютерный рентгеновский томограф. Испытания первого томографа прошли в 1971 году. В 1979 году Алан Кормак и Годфри Хаунсфилд были удостоены Нобелевской премии по физиологии и медицине «за разработку компьютерной томографии».

В нашей стране наряду с термином «компьютерная томография» используют более правильный, с физической точки зрения, термин – «рентгеновская вычислительная томография», впрочем, более точно – «рентгеновская трансмиссионная вычислительная томография». Определение «трансмиссионная» подчёркивает, что в основе метода лежит перенос (трансмиссия) рентгеновских фотонов.

До появления компьютеров в медицине использовалась классическая томография. Поясним её идею. Пусть необходимо получить изображение объекта в плоскости В (смотри рисунок 1). Для этого плоский регистратор рентгеновского излучения, пример – рентгеновская плёнка, помещается в плоскости С, а источник рентгеновского излучения помещается в плоскости С, параллельной плоскостям А и В

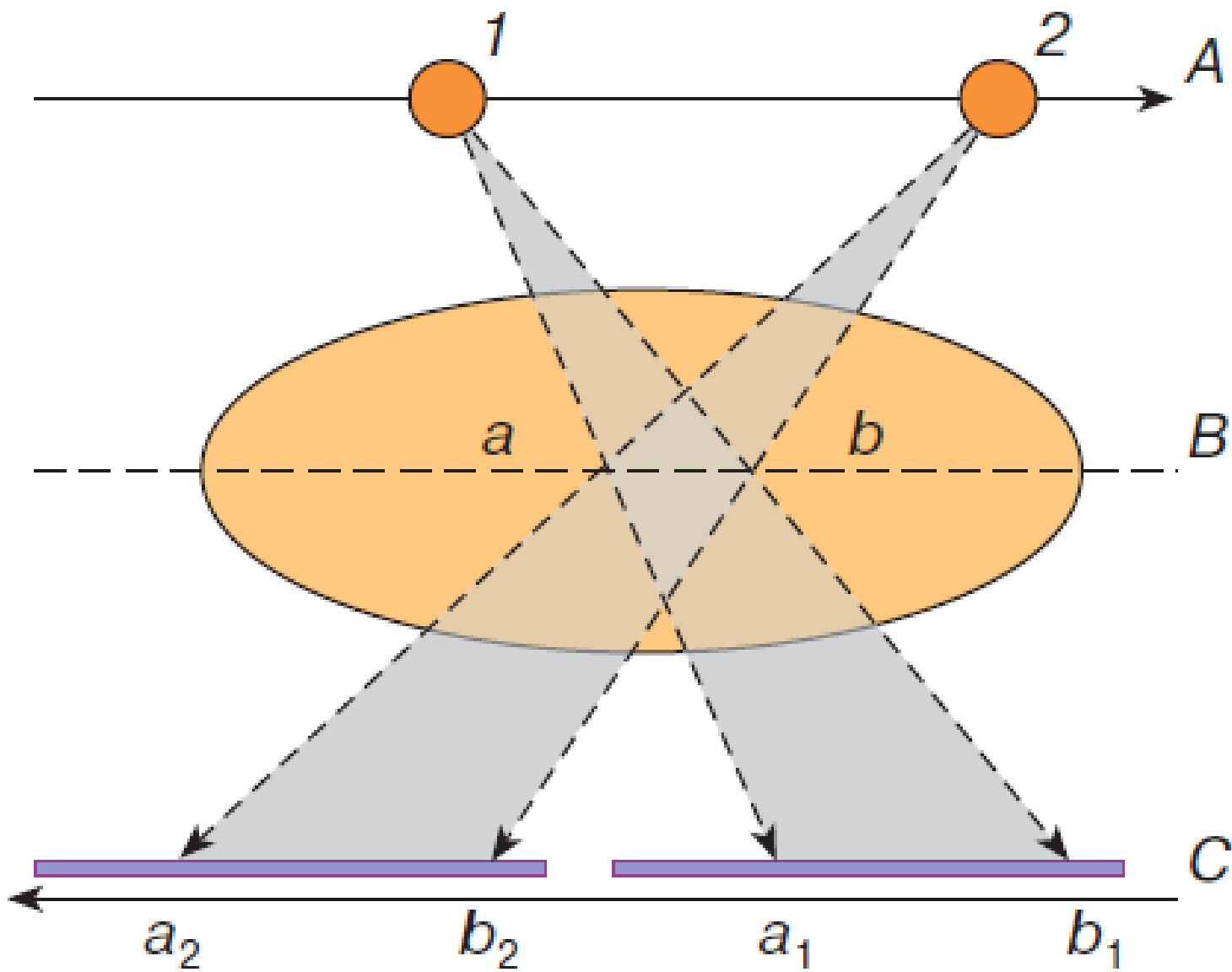


Рисунок 1 – Схема классической томографии

Регистратор и источник рентгеновского излучения перемещаются в противоположных направлениях с одинаковой скоростью. Из геометрических соображений следует, что множество точек, состоящее из точек пересечения осей источника рентгеновского излучения, принадлежит плоскости В. Поэтому изображение плоскости В на регистраторе будет неподвижным.

Очевидно, что все точки, лежащие вне плоскости V , будут отображаться в разных местах конечного изображения.

Следовательно, изображение плоскости будет более чёткое, а изображения остальных сечений объекта размазываются за счёт движения, создавая искажения изображений, в том числе и сечения V .

Описанный метод отличается исключительной простотой и доступен для понимания, но он является лишь частичным решением задачи исследования внутренней структуры объекта контроля, так как изображение сечения В остаётся затенённым другими слоями исследуемого объекта.

В научной литературе часто используется подход «от сложного к простому», особенно это характерно для обсуждения теоретических принципов рентгеновской вычислительной томографии. Но мы будем придерживаться традиционного подхода – «от простого к сложному».

Для этого исследуем осесимметричные объекты. Для таких объектов контроля формирования исходных проекций и восстановление изображений связывают с распределением Абеля.

6.1 Прямое и обратное преобразования Абеля

Итак, объект контроля имеет осевую симметрию. Пусть этот объект является ограниченным, тогда он ограничен произвольной цилиндрической поверхностью. В рентгеновской вычислительной томографии восстанавливается изображение поперечного сечения объекта, в нашем случае оно ограничено окружностью радиуса R , смотри рисунок 2.

Свяжем с сечением ортогональную систему координат XOY . Под изображением сечения понимается распределение некоторого параметра, характеризующего материал объекта, по сечению. В качестве указанной характеристики, как правило, выступают линейный коэффициент ослабления излучения или плотность.

- Рассмотрим распределение линейного коэффициента ослабления $\mu(x, y)$, $r^2 = x^2 + y^2 \leq R^2$. Приведённое соотношение означает, что для рассматриваемого случая нас интересует радиальное распределение линейного коэффициента ослабления $\mu(r)$.

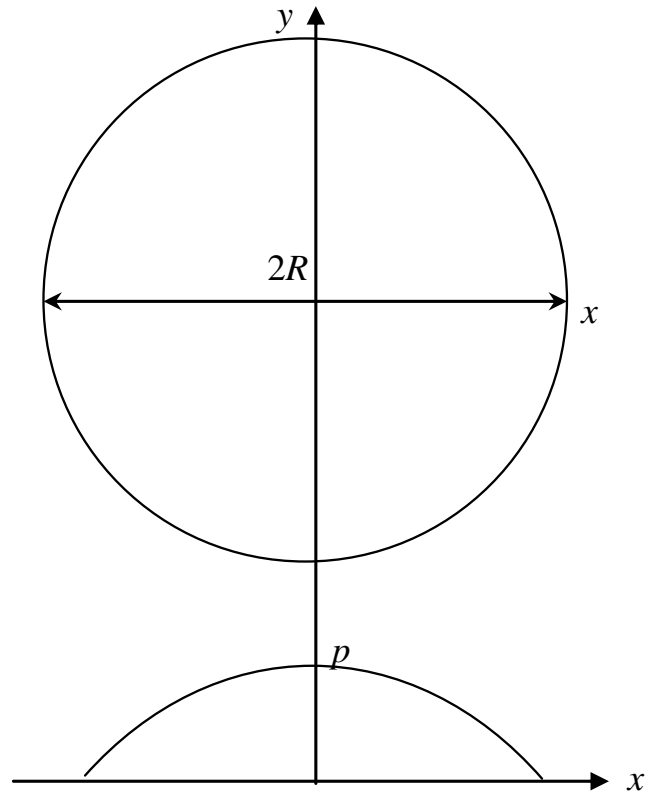


Рисунок 2 – Объект контроля и единичная радиографическая проекция

В рентгеновской вычислительной томографии измеряемой физической величиной (радиометрическим сигналом) является либо количество зарегистрированных фотонов N , либо величина поглощённой энергии гамма- или рентгеновского излучения I .

Для простоты будем рассматривать моноэнергетический источник излучения. Для анализируемого объекта в качестве проекции $p(x)$ будем рассматривать распределение толщины объекта контроля в длинах свободного пробега

$$p(x) = -\ln \frac{I(x)}{I_0} \quad (1)$$

здесь I_0 – величина радиометрического сигнала без объекта контроля при прочих одинаковых характеристиках процесса формирования проекции.

Выражение для описания радиального распределения $I(r)$ связано с радиальным распределением $\mu(r)$ и имеет следующий вид

$$I(r) = I_0 e^{-\int_{-\sqrt{R^2-r^2}}^{\sqrt{R^2-r^2}} \mu(\sqrt{r^2+y^2}) dy} \quad (2)$$

Из выражения (2) найдём проекцию $P(r)$ по формуле (2) с учётом симметрии объекта

$$p(r) = 2 \int_0^{\sqrt{R^2 - r^2}} \mu(\sqrt{r^2 + y^2}) dy \quad (3)$$

После несложных преобразований получим более удобную запись выражения (3)

$$p(r) = 2 \int_r^R \frac{\mu(z)z}{\sqrt{z^2 - r^2}} dz \quad (4)$$

- Полученная формула и есть классическая трактовка прямого преобразования Абеля.
- Пусть в уравнении (4) известна проекция $p(r)$, а неизвестная функция – радиальное распределение $\mu(r)$. Тогда их связь описывается выражением

- $$\mu(r) = -\frac{1}{\pi} \int_r^R \frac{p'(z)}{\sqrt{z^2 - r^2}} dz \quad (5)$$

- Формула (5) представляет собой обратное преобразование Абеля, называемое также инверсией Абеля.

Моделирование работы систем автоматического управления

- Для исследования работы любой системы автоматического управления (САУ) необходимо иметь ее математическое описание (математическую модель).

- Математическое описание системы начинается с разбиения ее на отдельные элементы (звенья) направленного действия.
- В каждом из которых сигнал со входа передается на выход, при этом выполняется определенное преобразование передаваемого сигнала

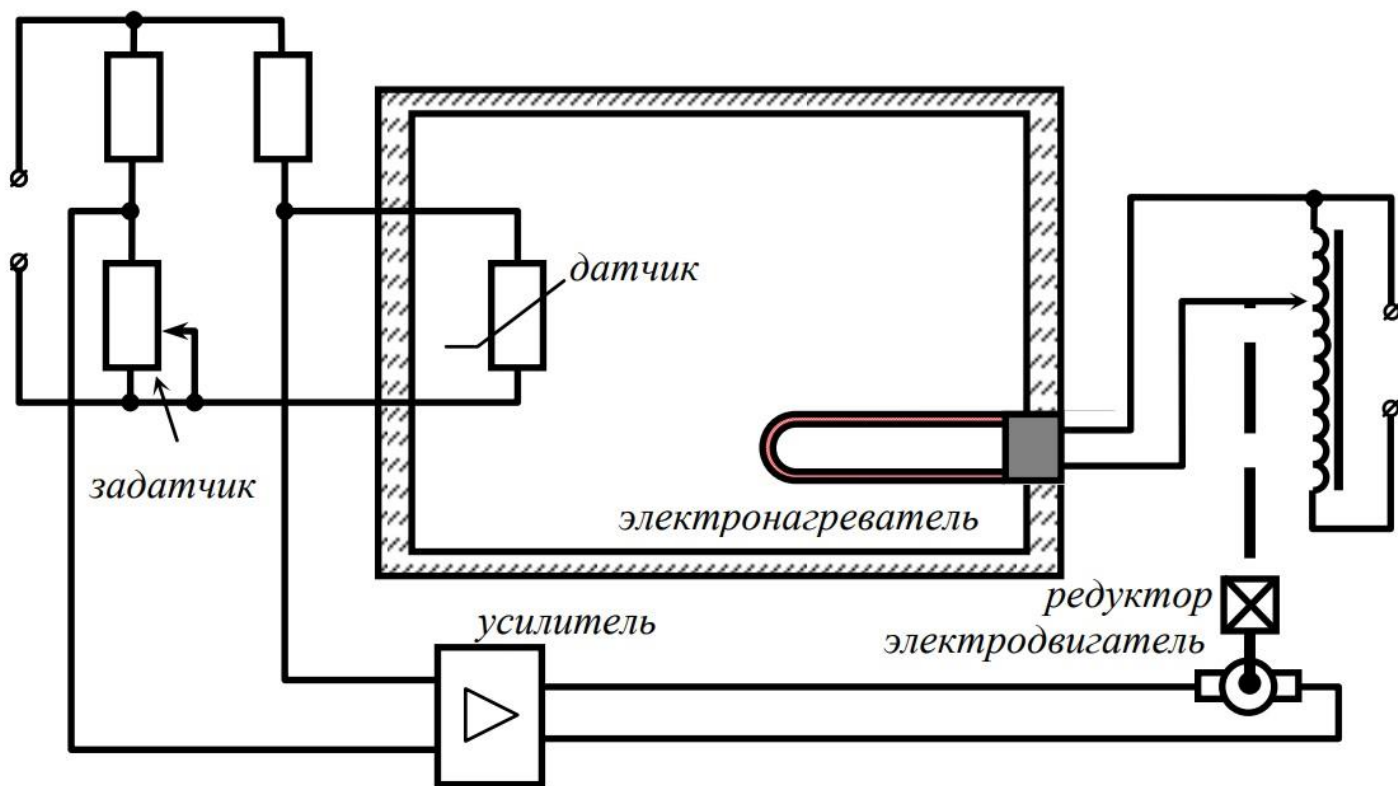


Рис. 1. Упрощенное схематическое изображение системы автоматического регулирования температуры в помещении

- **Объектом управления** данной САУ является помещение.
- **Управляемая величина** – температура воздуха θ в помещении. Для ее измерения в помещении имеется датчик (термосопротивление)

- Для обогрева воздуха в помещении установлен электронагреватель.

Управляющим воздействием в данном случае является электрическая мощность электронагревателя P , которая может изменяться с помощью подвижного контакта автотрансформатора, перемещением которого управляет электродвигатель постоянного тока

- **Возмущающим воздействием** на объект управления является изменение наружной температуры воздуха θ_n .
- Для названных элементов необходимо получить их **математическое описание**, т. е. соответствующие математические формулы, описывающие их работу

- Одним из способов получения **математического описания** элемента (системы) является теоретический (исходя из законов физики):
 - закон сохранения энергии,
 - массы,
 - импульса, законы
 - Кирхгофа и т. д.

- Полученные таким образом математические модели описывают
- внутренние связи в объекте и, как правило, наиболее точны.
- Получим теоретическим путем математическое описание для некоторых
- элементов из рассматриваемой нами САР температуры в помещении

- Значение температуры воздуха в помещении θ зависит от двух величин:
- мощности работающего электронагревателя P
- и температуры наружного воздуха θ_n .
- Влиянием остальных факторов пренебрегаем.

В процессе нагрева помещения одна часть тепловой мощности электронагревателя P_1 расходуется на увеличение температуры в помещении, а другая часть P_2 рассеивается наружу путем теплопередачи через стены.

Соответствующие математические формулы имеют следующий вид

$$P_1 = c \cdot \frac{d\theta}{dt}, \quad (1)$$

$$P_2 = \frac{S \cdot \kappa}{l} \cdot (\theta - \theta_{\text{н}}), \quad (2)$$

где c [Дж/°С] – теплоемкость помещения;

S [м²] – площадь стен;

κ [Вт/(м·°С)] – коэффициент теплопроводности стен (зависит от материала стен);

l [м] – толщина стен.

- Преобразуем эту формулу таким образом, чтобы все члены, содержащие выходную величину (температура θ), располагались в левой части уравнения, а входные величины (P и θ_n) – в правой части, получим в итоге математическое описание рассматриваемого объекта управления:

$$T \cdot \frac{d\theta}{dt} + \theta = K \cdot P + \theta_{\text{н}}, \quad (3)$$

где $T = \frac{c \cdot l}{S \cdot \kappa}$ [с] – постоянная времени;

$K = \frac{l}{S \cdot \kappa}$ [°С/ Вт] – коэффициент усиления (преобразования).

- 2. Датчик – термосопротивление – представляет собой металлический элемент (медный, платиновый, никелевый), имеющий при температуре 0°C некоторое определенное сопротивление R_0 .
- Как известно из курса физики, электрическое сопротивление металла при температурах до 600 °C почти линейно возрастает с увеличением температуры:

$$R = R_0 \cdot (1 + A \cdot \theta_{\text{д}}), \quad (4)$$

$$R = R_0 \cdot (1 + A \cdot \theta_{\text{д}}), \quad (4)$$

где R_0 [Ом] – значение сопротивления датчика при температуре 0 °С;
 $\theta_{\text{д}}$ [°С] – температура чувствительного элемента датчика;
 A [°С⁻¹] – постоянный коэффициент, значение которого зависит от металла, из которого изготовлен датчик (например, для датчика Pt100 $A = 3,9 \cdot 10^{-3}$ °С⁻¹).

Если такой датчик поместить в среду с некоторой температурой θ , отличной от θ_d , то будет происходить теплообмен между чувствительным элементом датчика и окружающей средой, в результате чего температура датчика будет стремиться к температуре среды, при этом, соответственно, будет меняться и его сопротивление (рис. 2).

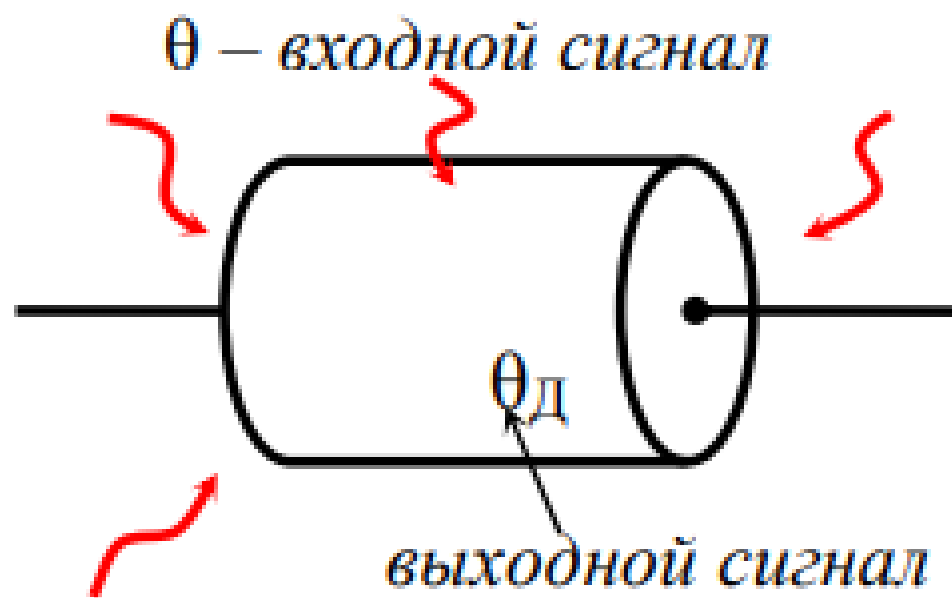


Рис. 2. Термосопротивление

Математически этот процесс будет описываться следующими формулами:

$$dQ = \kappa_{\text{д}} \cdot (\theta - \theta_{\text{д}}) \cdot dt, \quad (5)$$

$$dQ = c_{\text{д}} \cdot d\theta_{\text{д}}, \quad (6)$$

где dQ [Дж] – количество теплоты, переданное от окружающей среды датчику за малый промежуток времени dt ;

$\kappa_{\text{д}}$ [Дж/(°С·с)] – некоторый постоянный коэффициент, характеризующий теплопроводность чувствительного элемента датчика;

о

$c_{\text{д}}$ [Дж/°С] – теплоемкость чувствительного элемента датчика (зависит от его массы и удельной теплоемкости металла).

Приравняв выражения (5) и (6), а также выразив из формулы (4) температуру θ_d :

$$\theta_d = \frac{1}{R_0 \cdot A} \cdot (R - R_0),$$

получим

$$T_d \cdot \frac{dR_d}{dt} + R_d = K_d \cdot \theta, \quad (7)$$

где $R_d = R - R_0$ – выходной сигнал датчика;

$T_d = \frac{C_d}{K_d}$ [с] – постоянная времени датчика;

$K_d = R_0 \cdot A$ [Ом/°С] – коэффициент преобразования (усиления)

3. Устройство сравнения – измерительный резистивный мост из четырех сопротивлений R_1, R_2, R_3, R_d . Входными сигналами являются величины сопротивления задатчика R_1 и сопротивления датчика R_d ; выходной сигнал – напряжение разбаланса моста U_m (рис. 3, а). Используя законы Кирхгофа, получим

$$U_2 = i_2 \cdot R_{\text{д}} = \frac{U_{\text{п}}}{(R_2 + R_{\text{д}})} \cdot R_{\text{д}},$$

$$U_{\text{м}} = \frac{U_{\text{п}}}{(R_1 + R_{\text{зАд}})} \cdot R_{\text{зАд}} - \frac{U_{\text{п}}}{(R_2 + R_{\text{д}})} \cdot R_{\text{д}}, \quad (8)$$

где $U_{\text{п}}$ [В] – некоторое постоянное напряжение питания моста.

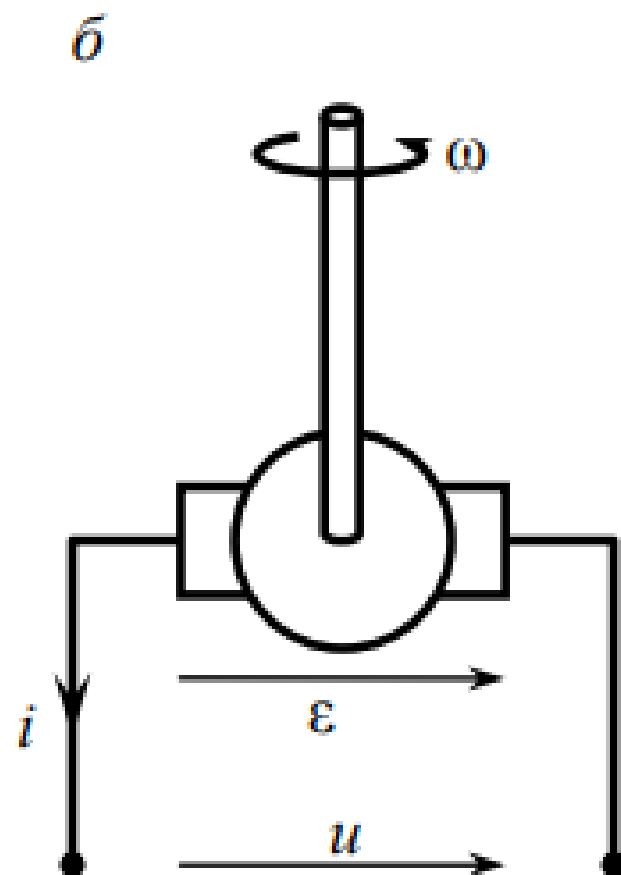
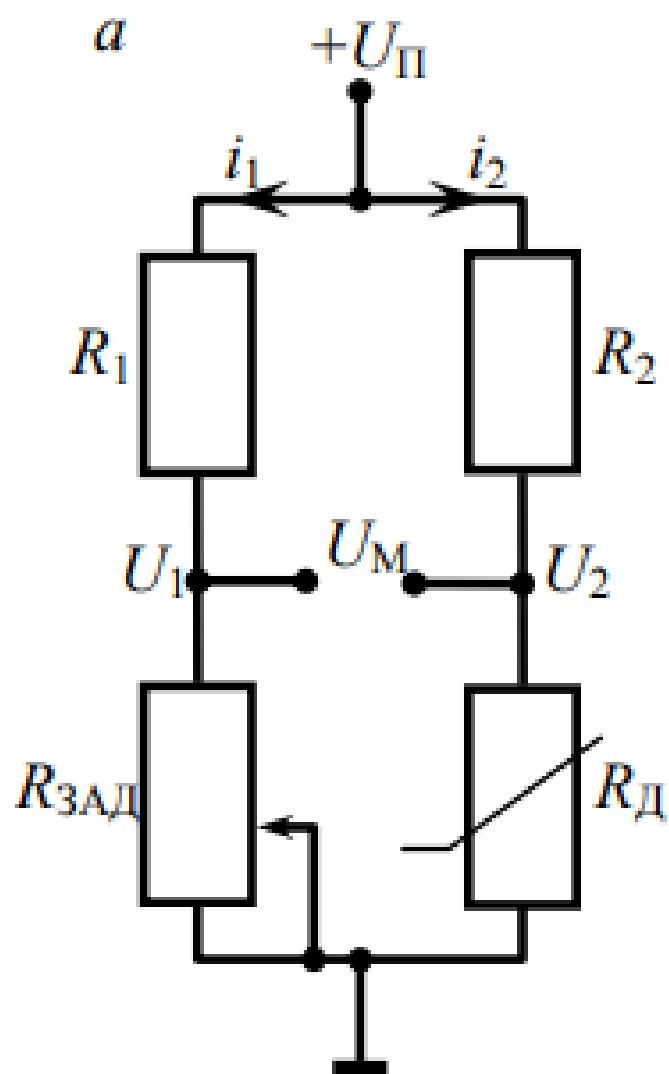


Рис. 3. Измерительный резистивный мост (*a*)
и двигатель постоянного тока (*б*)

4. Двигатель постоянного тока. Для получения его математического описания теоретическим путем необходимо использовать законы механики и электротехники.

Входным сигналом данного объекта является напряжение якоря $u(t)$; выходным сигналом – угол поворота вала двигателя $\phi(t)$.

Вал двигателя начинает вращаться, когда приложено напряжение питания.

Если напряжение питания не меняется ($u(t) = \text{const}$), то угловая скорость вращения вала двигателя остается постоянной ($\omega(t) = \text{const}$), при этом угол поворота вала двигателя будет равномерно увеличиваться.

Угол поворота вала двигателя $\phi(t)$ связан с его угловой скоростью вращения интегральным выражением:

$$\phi(t) = \int_0^t \omega(\tau) \cdot d\tau . \quad (9)$$

Чем больше напряжение, тем быстрее вращается вал. Уравнение динамики вращательного движения вала имеет следующий вид:

$$J \cdot \frac{d\omega(t)}{dt} = M(t) - M_H(t), \quad (10)$$

где $M(t)$ [Н·м] – вращающий момент – электромагнитный момент двигателя;

$M_H(t)$ [Н·м] – момент нагрузки;

J [кг·м²] – суммарный момент инерции якоря и нагрузки.

электромагнитный момент двигателя определяется по формуле:

$$M(t) = C_M \cdot \Phi \cdot i(t), \quad (11)$$

где C_M – безразмерный коэффициент;

Φ [Вб] – магнитный поток, создаваемый обмоткой возбуждения;

$i(t)$ [А] – ток якоря, который может быть найден из уравнения:

$$u(t) = e(t) + R \cdot i(t), \quad (12)$$

где $e(t)$ [В] – электродвижущая сила (ЭДС) якоря;

R [Ом] – сопротивление якорной цепи. В свою

очередь, ЭДС рассчитывается через магнитный

поток и частоту вращения:

$$e(t) = C_{\omega} \cdot \Phi \cdot \omega(t), \quad (13)$$

где C_{ω} – безразмерный коэффициент.

Таким образом, из уравнений (9)–(13) можно получить следующее дифференциальное уравнение, представляющее собой математическое описание электродвигателя постоянного тока:

$$T \cdot \frac{d^2 \varphi}{dt^2} + \frac{d\varphi}{dt} = K_1 \cdot u(t) - K_2 \cdot M_H(t), \quad (14)$$

$$T = \frac{J \cdot R}{C_M \cdot C_\omega \cdot \Phi^2};$$

$$K_1 = \frac{R}{C_M \cdot \Phi};$$

$$K_2 = \frac{R}{C_M \cdot C_\omega \cdot \Phi^2}.$$

По математическому описанию элементы и системы автоматики делятся на линейные и нелинейные.

Линейными элементами (звеньями) называются элементы автоматики, математически описываемые линейными дифференциальными уравнениями.

Линейными системами автоматического управления называются системы, состоящие только из линейных элементов. В общем случае для линейного элемента (или системы) взаимосвязь входного сигнала u и выходного сигнала y математически описывается дифференциальным уравнением вида:

$$\begin{aligned} a_0 \cdot \frac{d^n y}{dt^n} + a_1 \cdot \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_{n-1} \cdot \frac{dy}{dt} + a_n \cdot y = \\ = b_0 \cdot \frac{d^m u}{dt^m} + b_1 \cdot \frac{d^{m-1} u}{dt^{m-1}} + \dots + a_m \cdot u, \end{aligned} \quad (15)$$

где $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_n$ – некоторые постоянные коэффициенты,
 n – порядок системы,
 $m < n$.

- Так, например, уравнение для термосопротивления (7) получается из общего уравнения (15), если в (15) принять следующие обозначения: $n = 1$; $a_0 = T_d$; $a_1 = 1$; $m = 0$; $b_0 = K_d$. Нелинейными элементами называются элементы, описываемые нелинейными уравнениями.

- Система, содержащая хотя бы один нелинейный элемент, является нелинейной.

Из приведенных выше примеров математического описания элементов систем автоматики к линейным можно отнести математические модели помещения (3), термосопротивления (7) и электродвигателя постоянного тока (14). Математическое описание измерительного резистивного моста является нелинейным, так как в выражении (8) входные величины R_1 и R_d входят в знаменатели слагаемых дробей.

Режимы работы и характеристики элементов и систем автоматики

- В теории автоматического управления принято выделять два режима работы элементов и систем автоматики: **статический и динамический. Статическим (установившимся)** называется такой режим работы, при котором входная и выходная величины данного элемента неизменны во времени, т. е. $x(t) = \text{const}$ и $y(t) = \text{const}$.

Статической характеристикой звена (системы) называется зависимость между постоянным входным воздействием и постоянной выходной величиной звена (системы) в установившемся режиме (после окончания переходного процесса).

Например, рассмотрим датчик температуры (термосопротивление), помещенный в среду с некоторой неизменной во времени температурой θ ($\theta(t) = \text{const}$). Очевидно, что после того, как температура чувствительного элемента датчика станет равна температуре среды θ , сопротивление датчика также установится, т. е. примет некоторое неизменное во времени значение $R_d(t) = \text{const}$ тогда $dR_d/dt=0$

- В результате математическое выражение (7) примет следующий вид:
- $R_d = K_d Q$.
- Математическая зависимость (16) представляет собой статическую характеристику термосопротивления. Динамическим режимом работы (переходным процессом) называется режим работы элементов (систем) автоматики при изменяющихся во времени входном воздействии и выходной величине, т. е. при $x = x(t)$ и $y = y(t)$.

- .Динамическим режимом работы (переходным процессом) называется режим работы элементов (систем) автоматики при изменяющихся во времени входном воздействии и выходной величине, т. е. при $x = x(t)$ и $y = y(t)$.

- Если датчик температуры с некоторой начальной температурой
- чувствительного элемента θ_0 и соответствующим начальным значением сопротивления $R_{д0}$ поместить в среду с температурой θ_1 , отличной от θ_0 , то начнет происходить переходный процесс:

температура чувствительного элемента датчика за счет теплообмена с окружающей средой начнет постепенно приближаться к значению температуры среды θ_1 , при этом его сопротивление также будет изменяться в соответствии с формулой (7).

- Имеет место изменение
- входного воздействия на датчик температуры: до момента времени
- $t = 0$ температура среды (и самого датчика) была равна θ_0 , а начиная с момента $t = 0$, температура среды стала θ_1 :

$$\theta(t) = \begin{cases} \theta_0, & t < 0; \\ \theta_1, & t \geq 0. \end{cases}$$

Тогда решение дифференциального уравнения (7) имеет вид:

$$R_{\text{д}}(t) = R_{\text{д}0} \cdot e^{-\frac{t}{T_{\text{д}}}} + K_{\text{д}} \cdot \theta_1 \cdot \left(1 - e^{-\frac{t}{T_{\text{д}}}}\right). \quad (17)$$

Соответствующий график изменения во времени сопротивления датчика температуры приведен на рис. 4. После окончания переходного процесса датчик придет к статическому режиму работы.

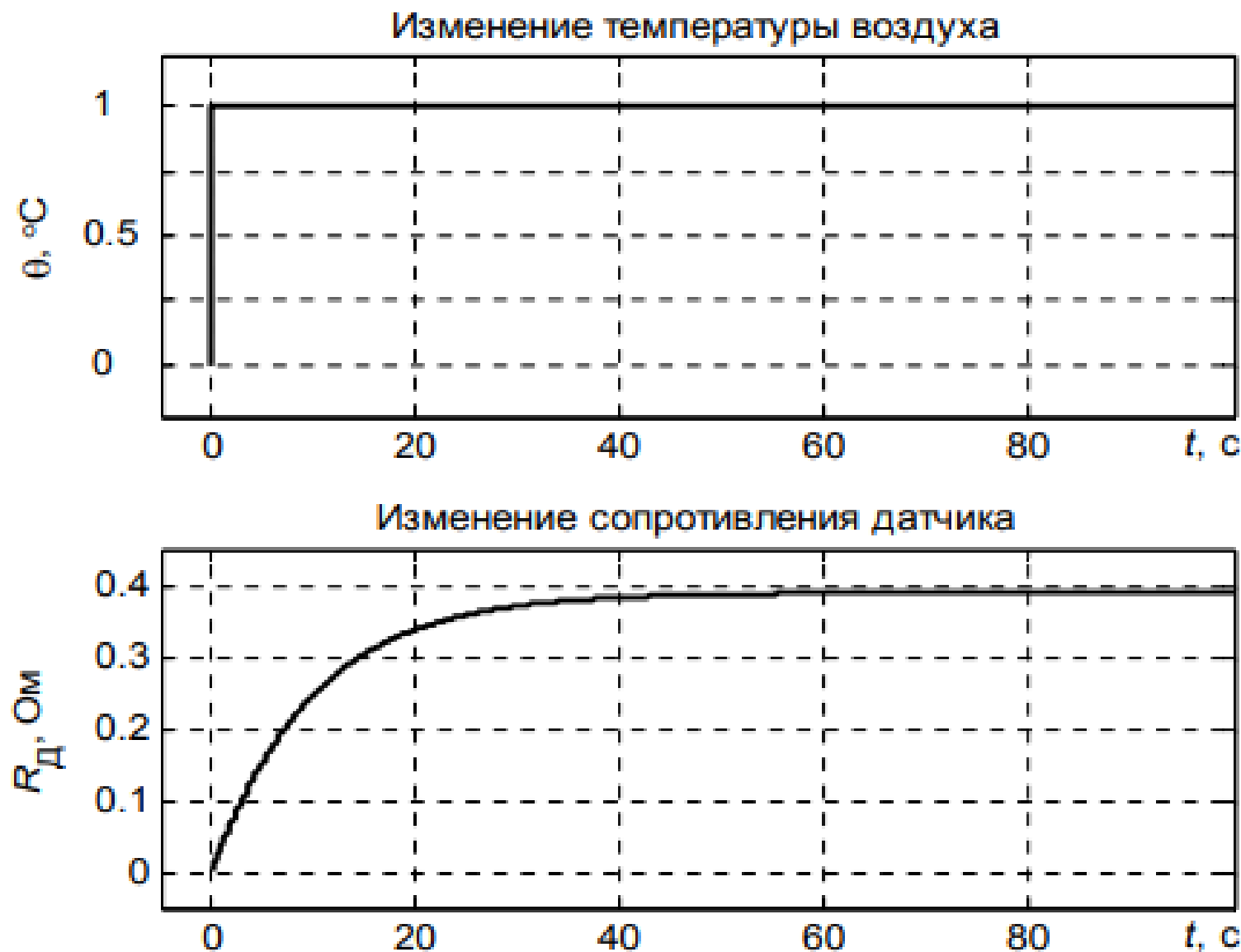


Рис. 4. Изменение сигнала датчика температуры Pt100 при ступенчатом входном воздействии (изменении температуры воздуха)

- После окончания переходного процесса датчик придет к статическому режиму работы. Поведение элементов (систем) автоматики в реальных условиях представляет собой сочетание переходного и установившегося режимов и, очевидно, существенно зависит от формы входного сигнала $x = x(t)$.

Изменение во времени внешних воздействий (например, изменение температуры наружного воздуха в течение суток) могут быть весьма различны и заранее неизвестны. В этом случае при математическом описании элементов (систем) автоматики возникают трудности принципиального характера, т. к. заранее неизвестны законы изменения внешних воздействий, что затрудняет анализ динамики и статики САУ.

- Для устранения этого затруднения при анализе работы элементов (систем) автоматики принято использовать так называемые типовые воздействия, которые являются математическими моделями часто встречающихся реальных воздействий. В качестве типовых рассматривают следующие воздействия:

1. Ступенчатое воздействие – это воздействие, которое мгновенно возрастает от нуля до некоторого значения и далее остается постоянным:

$$x(t) = \begin{cases} 0, & \text{при } t < 0, \\ a, & \text{при } t \geq 0, \end{cases} \quad (18)$$

где a – величина ступенчатого воздействия.

При $a = 1$ имеет место единичное ступенчатое воздействие, обозначаемое следующим образом:

$$1(t) = \begin{cases} 0, & \text{при } t < 0, \\ 1, & \text{при } t \geq 0. \end{cases} \quad (19)$$

Реакция элемента (системы) на единичное ступенчатое воздействие называется переходной функцией элемента (системы). Так, на рис. 4 изменение температуры воздуха представляет собой единичное ступенчатое воздействие на датчик температуры, а изменение сопротивления датчика – переходную функцию. Переходная функция обычно обозначается $h(t)$.

Следует заметить, что ступенчатый сигнал легко получить на практике. Так, например, включение в некоторый момент времени в помещении электронагревателя на некоторую постоянную мощность можно рассматривать как ступенчатое управляющее воздействие на объект управления (помещение).

Подача постоянного электрического напряжения на вход электродвигателя постоянного тока путем замыкания идеального электрического ключа также представляет собой ступенчатое воздействие. Поэтому переходную функцию элемента (системы) можно получить экспериментально

2. Импульсное воздействие (дельта-функция):

$$x(t) = \delta(t) = \begin{cases} \infty, & \text{при } t = 0, \\ 0, & \text{при } t \neq 0, \end{cases} \quad (20)$$

при этом согласно определению:

$$\int_{-\infty}^{+\infty} \delta(t) \cdot dt = 1. \quad (21)$$

Это идеальный сигнал, который равен нулю во всех точках, кроме $t = 0$, где он уходит в бесконечность, причем его площадь (интеграл по всей оси времени) равна единице.

В качестве физической интерпретации импульсного воздействия можно привести следующий пример. Пусть частица некоторой массы, движущаяся вдоль прямой, в некоторый момент времени $t = 0$ при соударении с другой частицей изменяет свою скорость с v_1 до v_2 .

- Тогда зависимость от времени ускорения частицы будет описываться следующим выражением:

$$a(t) \approx \begin{cases} 0, & \text{при } t < 0, \\ \frac{v_2 - v_1}{\Delta t}, & \text{при } 0 \leq t \leq \Delta t, \\ 0, & \text{при } t > \Delta t. \end{cases}$$

- Если считать, что длительность соударения $\Delta t \rightarrow 0$ (соударение происходит почти мгновенно), то получается, что

$$a(0) = \lim_{\Delta t \rightarrow 0} \frac{v_2 - v_1}{\Delta t} = \infty,$$

т. е. изменение во времени ускорения частицы приблизительно описывается дельта-функцией (20): $a(t) \approx \delta(t)$. Дельта-функция связана с единичным ступенчатым воздействием выражением

$$\delta(t) = l'(t), \quad (22)$$

т. е. дельта-функция равна производной от ступенчатого воздействия. Реакция элемента (системы) на единичный импульс (дельта-функцию) называется импульсной характеристикой и обозначается $w(t)$. Импульсная характеристика линейного элемента (системы) $w(t)$ связана с его переходной функцией выражением:

$$w(t) = h'(t). \quad (23)$$

Гармоническое воздействие

$$x(t) = A_x \cdot \sin(\omega \cdot t), \quad (24)$$

или

$$x(t) = A_x \cdot \cos(\omega \cdot t), \quad (25)$$

где A_x – амплитуда; ω [рад/с] – круговая частота.

Если на линейное звено (систему) подать гармоническое воздействие (24), то после окончания переходного процесса на выходе звена установится также гармонический сигнал:

$$y(t) = A_y \cdot \sin(\omega \cdot t + \varphi), \quad (26)$$

- где A_y – амплитуда выходного сигнала; ϕ [рад] – разность фаз между выходным и входным сигналами. На рис. 5, а показано линейное звено, представляющее собой электрическую схему из последовательно соединенных активного сопротивления R и конденсатора емкостью C . Математическое описание такого звена имеет следующий вид:

$$T \cdot \frac{du_{\text{ВЫХ}}}{dt} + u_{\text{ВЫХ}} = K \cdot u_{\text{ВХ}},$$

где $T = R \cdot C$;

$$K = 1.$$

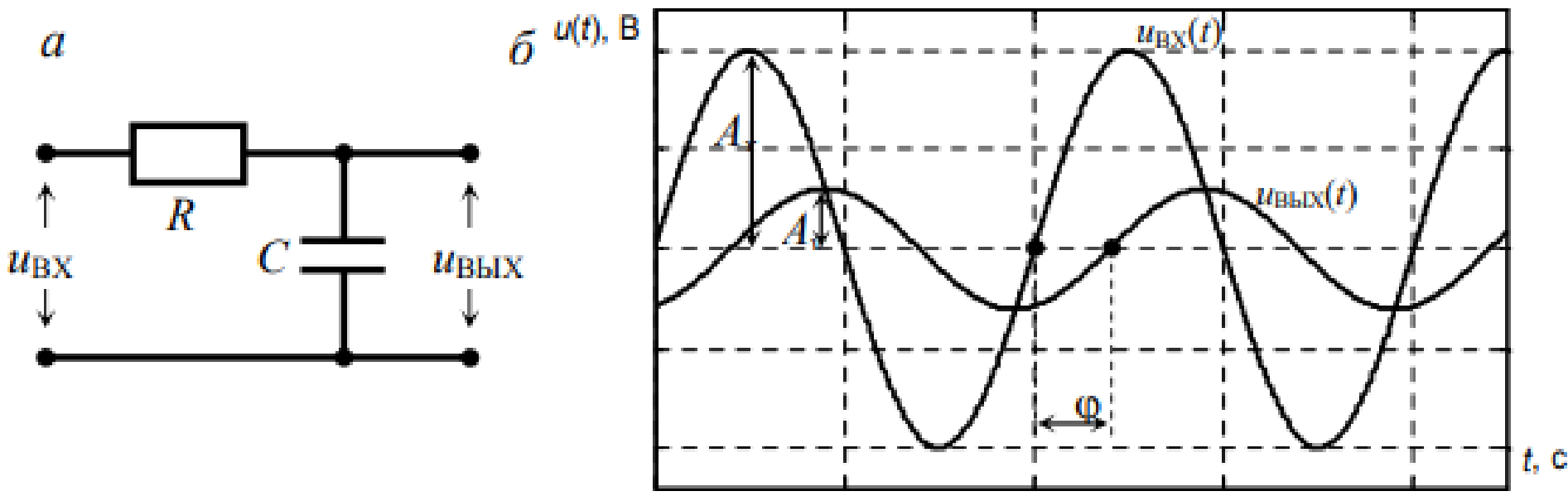


Рис. 5. При подаче гармонического сигнала на вход линейного динамического звена (а) выходной сигнал также гармонический (б)

На рис. 5 б показано входное гармоническое воздействие – изменение напряжения $u_{вх}(t)$ на входе RC-звена, и соответствующий выходной сигнал – изменение напряжения на конденсаторе. Физические свойства линейных звеньев (систем) таковы, что частота входного и выходного сигналов одинакова. Амплитуда выходного сигнала при постоянной амплитуде входного сигнала и разность фаз между выходным и входным сигналами зависят от частоты.

Амплитудной частотной характеристикой (АЧХ) называется зависимость отношения амплитуд выходного и входного гармонических сигналов от частоты в установившемся режиме:

$$A(\omega) = \frac{A_y}{A_x}. \quad (27)$$

Фазовой частотной характеристикой (ФЧХ) называется зависимость разности фаз $\phi(\omega)$ между выходным и входным сигналами от частоты в установившемся режиме.

Переходная функция и импульсная характеристика относятся к временным характеристикам звена, так как они показывают зависимость от времени реакции звена на ступенчатое или импульсное входное воздействие.

- АЧХ и ФЧХ называются частотными характеристиками звена и определяют реакцию звена на гармоническое входное воздействие в зависимости от его частоты.

В теории автоматического управления математическое описание линейных элементов, объектов или систем принято выражать не в виде дифференциальных уравнений (как, например, дифференциальные уравнения (3), (7), (14)), которые часто могут иметь достаточно сложный вид и еще более сложное решение, а в виде передаточной функции W

Передаточная функция элемента (объекта) W связывает его входной и выходной сигналы x и y при нулевых начальных условиях (т. е. считается, что в начальный момент времени $t = 0$ значения входного и выходного сигналов равны нулю: $x(t = 0) = 0, y(t = 0) = 0$).

- Формально передаточная функция объекта определяется как отношение преобразования Лапласа его выходного сигнала к преобразованию Лапласа входного сигнала:

- $W(s) = Y(s)/X(s)$ (28)

где $s = \sigma + i\omega$ – комплексная переменная, в которой $\sigma = 0$ для всех реализуемых на практике физических сигналов, а ω имеет физический смысл частоты гармонического сигнала.

- В качестве примера рассмотрим еще раз гармонический сигнал вида (25). Известно, что гармонический сигнал с помощью формулы Эйлера может быть выражен в комплексной форме:

$$x = A_x e^{i\omega t}$$

- A_x – амплитуда сигнала; ω [рад./с] – его круговая частота;

Перепишем уравнение $x = A_x e^{st}$ (30)

- где $s = i\omega$

- Для гармонического сигнала вида (30) математическая операция взятия производной эквивалентна (т. е. может быть заменена) умножению самого сигнала на комплексную переменную $s = i^* \omega$.
Производная сигнала вида (30) по времени равна

- $$\frac{dx}{dt} = A_x e^{st} s = s A_x e^{st} = s x \quad (31)$$

- Вторая производная сигнала по времени в этом случае будет равна

- $$\frac{d^2x}{dt^2} = sA_x e^{st} s = s^2 A_x e^{st} = s^2 x \quad (32)$$

- Допустим, что имеется некоторый элемент, который математически описывается дифференциальным уравнением

- $$T \frac{dy}{dt} + y = Kx \quad (33)$$

- Пусть на вход этого элемента подается гармонический сигнал x вида (30). Так как дифференциальное уравнение (33) линейное, а значит, и сам элемент является линейным, то выходной сигнал y также будет гармоническим. Поэтому для нахождения связи между входным и выходным сигналами не обязательно решать уравнение (33), а можно заменить в нем операцию взятия производной умножением на комплексную переменную $s = i \omega$

$$Tsy+y=Kx, \quad (Ts+1)y=Kx$$

$$W(s) = \frac{y}{x} = \frac{K}{Ts + 1} \quad (34)$$

Известно, что практически любой изменяющийся во времени физический сигнал (например, изменение электрического напряжения, тока, скорости, давления, уровня жидкости, температуры и т. д.) можно представить в виде суммы гармонических сигналов разной частоты и амплитуды

- Поэтому описанный подход, при котором операция дифференцирования заменяется умножением на комплексную переменную $s = i * \omega$, справедлив не только для гармонических, а для всех физически реализуемых сигналов (т. е. для всех сигналов любой формы, встречающихся на практике). Описанный подход был предложен французскими учеными П. С. Лапласом и Ш. Фурье.