

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ХАРЬКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ»**

**И. А. Чермных, А. Г. Журило,
Е. А. Краевская, И. Ю. Адашевская**

ГЕОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ В КОМПЬЮТЕРНОЙ ГРАФИКЕ

**УЧЕБНОЕ ПОСОБИЕ
ДЛЯ СТУДЕНТОВ
ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ**

ХАРЬКОВ 2017

УДК 004.92(075.8)

Ч 48

Рецензенты:

Л. Н. Куценко, д-р техн. наук, проф., Национальный Университет гражданской обороны Украины

В. Ф. Ткаченко, канд. техн. наук, профессор, Харьковский национальный университет радиоэлектроники

Черных И. А.

Ч 48 **Геометрическое моделирование в компьютерной графике : учеб. пособие / И. А. Черных, А. Г. Журило, Е. А. Краевская, И. Ю. Адашевская. – Харьков : «НТМТ», 2017. – 320 с. – На рус. яз.**

ISBN 978-617-578-279-8

Навчальний посібник призначено для студентів вищих навчальних закладів технічного напрямку, в тому числі іноземних студентів, аспірантів, а також для дизайнерів, ілюстраторів, інженерно-технічних працівників, які працюють з комп'ютерною графікою у різних її галузях. Посібник пропонує читачам поіншому розглянути використання матриць та векторів, перетворення кривих високих порядків та сплайнів і фракталів у геометричному моделюванні. Велика кількість ілюстрацій значно полегшує опанування запропонованою інформацією, дозволяє розглянути на прикладах найбільш складні питання.

У посібнику викладено теоретичні і практичні питання та існуючі методи геометричного моделювання точок, прямих, площин, поверхонь, перетворення координат, фракталів та сплайнів.

Запропонований матеріал може бути використаний при вивченні найважливіших питань геометричного моделювання, при курсовому й дипломному проектуванні.

Учебное пособие предназначено для студентов высших учебных заведений технического направления, в том числе для иностранных студентов, аспирантов, а также для дизайнеров, иллюстраторов, инженерно-технических работников, которые работают с компьютерной графикой в разных ее отраслях. Пособие предлагает читателям по-другому рассмотреть использование матриц и векторов, преобразования кривых высоких порядков, сплайнов и фракталов в геометрическом моделировании. Большое количество иллюстраций значительно облегчает усвоение предложенной информации, позволяет рассмотреть на примерах наиболее сложные вопросы.

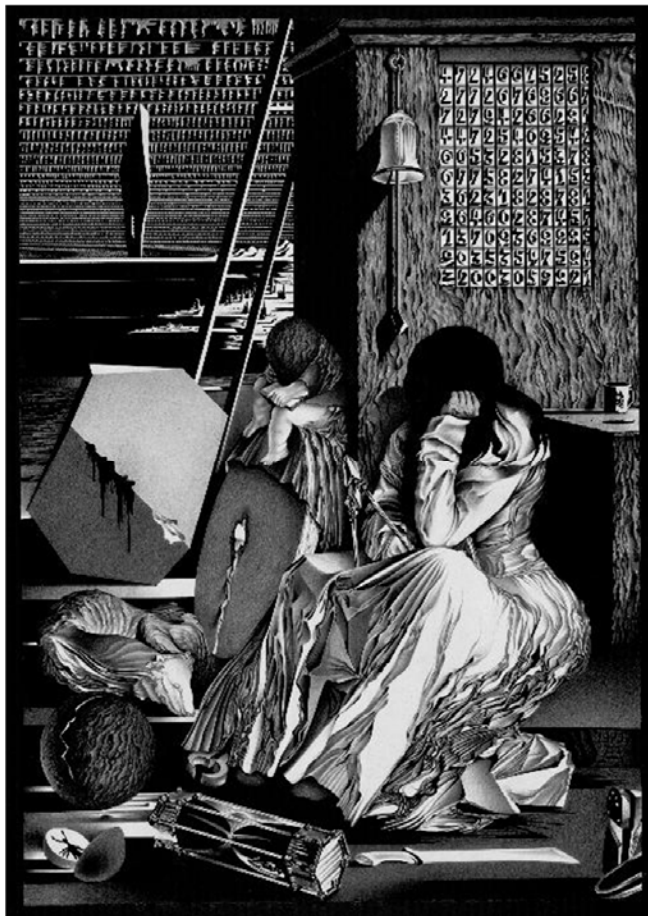
В пособии приведены теоретические и практические вопросы и существующие методы геометрического моделирования точек, прямых, плоскостей, поверхностей, преобразование координат, фракталов и сплайнов. Предложенный материал может быть использован при изучении важнейших вопросов геометрического моделирования, при курсовом и дипломном проектировании.

ISBN 978-617-578-279-8

УДК 004.92(075.8)

© И. А. Черных, А. Г. Журило,
Е. А. Краевская, И. Ю. Адашевская 2017

ВВЕДЕНИЕ. ГЕОМЕТРИЧЕСКИЕ МОДЕЛИ И МОДЕЛИРОВАНИЕ



*Вся природа выражается в символах
геометрического искусства*

И. Кеплер

*Прямоугольники, круги, квадраты,
Дома с бетона, стали и стекла,
Повсюду радио и автоматы,
И надо - всем победный знак числа.*

М. Драй-Хмара

Моделирование как метод научного познания

Вы приступаете к изучению нового курса «Теория геометрического моделирования». Зачем он Вам нужен? Чтобы ответить на этот вопрос, выясним, прежде всего, что такое геометрическое моделирование, чем оно занимается, какие вопросы решает, где используется.

Геометрическое моделирование – это взгляд на мир с геометрической точки зрения – т. е. описание геометрии отдельных объектов, их взаимного расположения, исследование различных геометрических свойств, конструирование объекта с заданными геометрическими свойствами. При этом в центре внимания и изучения находятся именно геометрические свойства, а остальные (цветовые, прочностные и т.п.) отсекаются.

Геометрический объект (ГО) – это и реально существующий физический объект, и задумка конструктора или дизайнера, определяющая собой (или удовлетворяющая) определённые геометрические параметры (ГП): форму, размеры (масштабность), положение в определенной системе отсчета.

Геометрический объект - реально существующая, физически воплощенная в определенном материале субстанция, определяющая собою (или удовлетворяющая) определенные геометрические параметры: форму, размеры (масштабность), положение в определенной системе отсчета. Геометрическими объектами можно назвать: камень, хрустальную вазу, футбольный мяч, вертолет, изотермическую поверхность уровня температуры t_i и многое другое.

Процессы генерации объектов и их моделей, описываемые ниже, должны привести нас к конечной цели - определению всех необходимых характеристик физических объектов и их моделей. Следовательно, процесс становится сам объектом изучения и моделирования. Объект изучения - процесс может быть физическим, химическим, расчетным (математическим), логическим, геометрическим, дискретизирующим, графическим, описательным (лингвистическим) и др.

Геометрическое моделирование базируется на:

- аналитической и дифференциальной геометрии,
- топологии;
- теории параметризации;
- векторной алгебре;
- теории матриц,
- булевой алгебре.

Геометрическое моделирование является математической основой всех графических пакетов от Автокада и Компаса до 3D MAX.

ГМ занимается описанием, исследованием и конструированием ГО, способами преобразования описаний к требуемому виду, визуализацией ГО и процессов.

[Французское *model* происходит (через итальянское *modello*) от латинского *modus* (мера, образец)]

Основные требования, которые должны выполняться при геометрическом моделировании:

- геометрическая модель (ГМ) не должна противоречить реальному объекту (правильность модели);
- возможность вычисления ряда геометрических величин (объемов, площадей, периметров и т. д.)
- допустимо конструирование модели целиком (мощность модели);
- предусмотрено использование различных функций (расчет ЧПУ (числовое программное управление), разработка серии изделий, расчет конструкций и т. п.).

ГМ может быть представлена в виде:

- лингвистической модели (ЛМ) – в виде текста;
- физической модели (ФМ) – объект, выполненный в масштабе;
- математической модели (ММ) – совокупность уравнений, неравенств и других ограничений и условий их разрешения (область изменения параметров формы и положения), которые в заданной системе координат определяют ГО единственным образом;
- цифровой модели (ЦМ) – это три массива координат X, Y, Z точек принадлежащих ГО, упорядоченных определённым образом и оптимизированных по критериям *min* количества точек и наперёд заданной точности описания объекта;
- графической модели (ГрМ) – его проекционный обратимый чертёж – обеспечивает решение задач визуализации исследуемого ГО или процесса.

Основные требования, которые должны выполняться при геометрическом моделировании:

- ГМ не должна противоречить реальному объекту (правильность модели);
- возможность вычисления ряда геометрических величин (объемов, площадей, периметров и т.д.)
- допустимо конструирование модели целиком (мощность модели);
- предусмотрено использование различных функций (расчет ЧПУ(числовое программное управление), разработка серии изделий, расчет конструкций и т.д.

Основные задачи, решаемые геометрическим моделированием

- конструирование линий, отвечающих наперед заданным условиям (прохождение кривой через определенные точки, степень гладкости линий, заданная кривизна в нужных точках и т.п.)
- построение эквидистантных кривых
- конструирование поверхностей по наперед заданным условиям.

Подобные задачи возникают как при конструировании и технологической подготовке так и при работе дизайнера.

Геометрические модели и моделирование

Функциональные модули геометрического моделирования призваны формировать оптимальную схему проектирования (генерации, создания) объекта или процесса, групп объектов или процессов и решения инженерно-технических задач. Критериями оптимальности здесь служат прозрачность алгоритмов и их универсальность относительно объектов или процессов одного вида. Удовлетворить критериям удастся благодаря установлению взаимозначимой связи между классификаторами объектов и процессов и существующими или вновь генерируемыми их геометрическими моделями, которые, в свою очередь, подразделяются на аппроксимационные, интерполяционные, конструктивные и кусочно-составные (обводы) в любых сочетаниях.

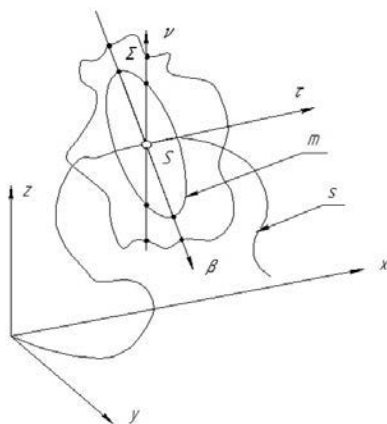


Рис. 1

Чтобы уяснить ситуации рассмотрим простой пример. Пусть требуется спроектировать и рассчитать трубчатую поверхность с пространственной линией центров (рис. 1). Для решения этой задачи можно использовать ряд геометрических моделей. Аппарат дифференциальной геометрии позволяет использовать триэдр Френе $\nu\tau\beta$ в качестве локальной системы координат, где в плоскости $\Sigma = \nu\tau\beta$ задана образующая окружность m , которая определяет искомую трубчатую поверхность при движении триэдра по направляющей линии центров S , определяемой единственным образом в декартовой системе координат $OXYZ$. Аппарат теории параметризации позволяет не прибегать к использованию направляющей линии S и триэдра Френе $\nu\tau\beta$. В системе координат $OXYZ$ положение плоскости Σ определяется единственным образом тремя параметрами. Задав определенным образом законы изменения каждого из параметров, определим любое положение плоскости Σ , а следовательно, и положение образующей m , таким образом, искомую трубчатую поверхность. Какая из этих двух (или еще каких-либо других) геометрических моделей лучше, может определить инженер-геометр на основании анализа алгоритмов, реализующих

генерацию трубчатой поверхности. Критериями оптимальности выбора наилучшей геометрической модели служат компактность (миниатюрность) алгоритма; удобство машинной реализации (простота формализации расчетов); максимальное соответствие результатов исходным инженерно-геометрическим требованиям. Примером удовлетворения последнему критерию может служить отдавание предпочтения аппарату дифференциальной геометрии, если в исходные данные (рис.0.1) включена линия тока S проектируемой трубчатой поверхности, и известно, что плоскость образующей окружности всегда нормальна к линии тока.

Мобильная (в смысле передачи ее от одного человека к другому) ГМ объекта или процесса может быть интерпретирована, скорее всего, их графическими или лингвистическими моделями или их сочетанием, например: схемой, рисунком, эскизом, чертежом и описанием алгоритма или инструкцией.

ГМ примитивных и достаточно простых объектов или процессов могут быть представлены только графическими или лингвистическими моделями. Например, чертежа прямоугольного треугольника; не требующий пояснений текст: "отрезок прямой длиной десять миллиметров" или "соединить произвольные точки A и B прямой или другой окружности" и др.

На определенном этапе идея конструктора, базирующаяся на прототипе, превращается в упорядоченную сумму условий, которым должен удовлетворять объект. Теперь можно сказать, что прототип, удовлетворяющий после корректирования его параметров всем наперед заданным условиям, является моделью объекта, визуализированной конструктором внутренним образом. Внутреннее "видение" объекта позволяет конструктору выделить определяющие объект геометрические условия в виде определителя, каркаса точек или линий и т.п. Наличие определяющих условий и опыта конструктора, знающего инженерную геометрию, позволяет сформировать модель геометрических преобразований, обеспечивающих генерацию объекта. Определяющие геометрические условия и геометрические преобразования, приведенные во взаимодействие, обуславливают возможность реализации процесса геометрического моделирования объекта, в результате чего генерируется собственно геометрическая модель объекта (ГМО). Геометрические определяющие условия и преобразования, визуализируемые внутренним образом, достаточно легко формализуются переводом ГМО в любой вид моделей низшего уровня (рис. 2).

Геометрической моделью геометрического объекта (ГМО) или геометрической моделью процесса (ГМП) назовем совокупность геометрических условий, определяющих объект или процесс с требуемой точностью, и геометрических преобразований, позволяющих конструктору визуализировать объект или процесс внутренним образом; формализовать в единственно возможном варианте, представление объекта или процесса в лингвистическом, физическом, математическом, цифровом, графическом или другом видах генерацией соответствующих алгоритмов. Например, ГМ внутренней поверхности проточной части тела корпуса турбины можно сформировать посредством реализации алгоритма расчета ортогональной сети образующих, одно семейство которых

представляет собою линии тока на поверхности и замкнутые линии, им ортогональные.

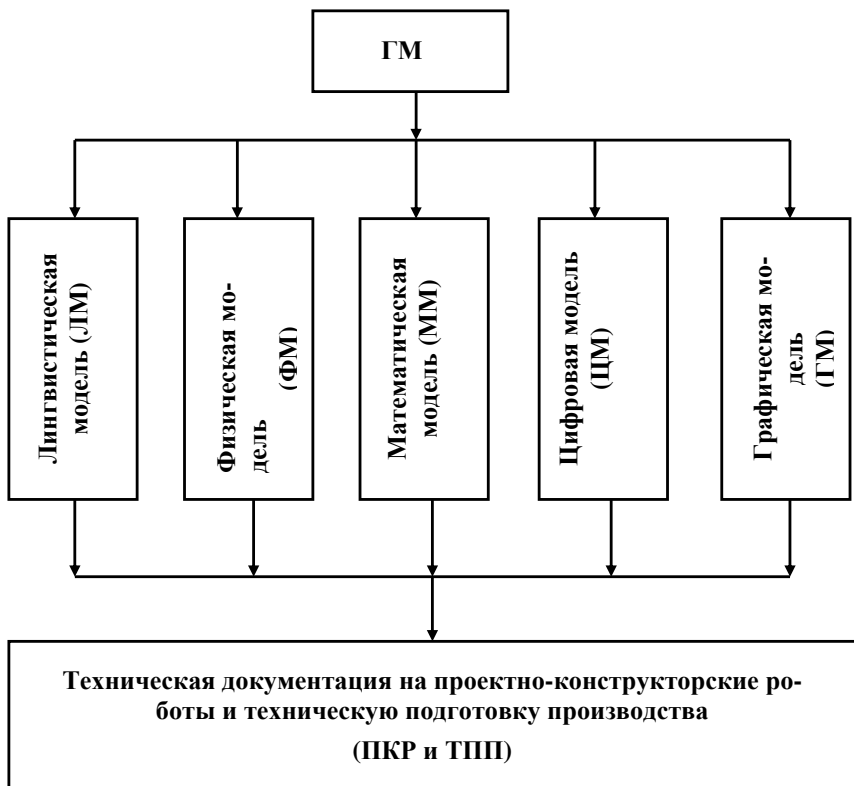


Рис. 2

ГМП течения воды в проточной части турбины можно сформировать посредством фиксации в пространстве плоских или пространственных линий тока частичек воды в трехмерном объеме проточной части.

Математические модели и моделирование

Математическое моделирование объекта сводится к математическому описанию его геометрической модели. Плохо структурируемые процессы успешно формируются посредством эвристического моделирования. Обратимся к примеру, рассмотренному выше (см. Рис.0.1). У трубчатой поверхности объектами математического моделирования могут быть: направляющая линия S центров; образующая окружность m ; зависимости, определяющие непрерыв-

ным образом изменение параметров положения плоскости Σ при расчете поверхности, и многие другие. В качестве математических моделей могут выступать сплайн-функции, кривые второго порядка, полиномиальные функции, специальные функции и т.д. Критериями оптимальности выбора наилучшей математической модели служат компактность (миниатюрность) алгоритма, точность получаемых решений, удобство машинной реализации (простота формализации расчетов).

Все математические модели (ММ) следует делить на универсальные (УММ) и специализированные (СММ). УММ применяются для описания сложных объектов, например кривой, S , а СММ - для описания широкого класса более простых объектов, например окружности m . Подробно изучаются вопросы, связанные с формированием и использованием УММ и СММ. В общем виде математическая модель объекта может выглядеть следующим образом:

$$\begin{cases} x = x(t, a, b, \dots) a = 0 \text{ нпу} A_1 \leq y \leq A_2; \\ y = y(t, c, d, \dots) e = 1 \text{ нпу} E_1 \leq y \leq E_2; \\ z = z(t, e, f, \dots); \\ a = a(y); \\ e = e(y); \\ \dots \dots \dots \end{cases}$$

Уравнения системы определяют собою метод математического моделирования объекта (МММО), а условия реализации метода определяются индивидуальными инженерно-геометрическими требованиями, предъявляемыми к данному объекту. Итак, уравнения системы вместе с условиями их разрешения представляют собой математическую модель объекта (ММО).

Параметры формы и положения извлекаются из системы уравнений, определяющей МММО. Опыт показывает, что их вид, количество и функциональное назначение и определяют индивидуальные свойства МММО, который, в свою очередь, создается под какую-то разновидность геометрических объектов. Модулирование же функций, ограничения, накладываемые на области изменения параметров, помогают "высечь" из множества $M10$, определяемых МШО, необходимую конкретную ММО.

Математической моделью (ММО) геометрического объекта назовем совокупность уравнений, неравенств и других ограничений и условий их разрешения (в том числе областей изменения всех параметров формы и положения), которые в заданной системе координат определяют объект единственным образом; каждым двум конкретным значениям координат точки объекта соответствует единственное значение третьей координаты точки.

Математической моделью процесса (ММП) назовем совокупность уравнений, неравенств и других ограничений и условий их разрешения (в том числе областей изменения всех параметров), которые в заданной системе координат определяют процесс единственным образом.

Графические модели и моделирование

Графической моделью (ГрМО) геометрического объекта назовем его проекционный обратимый чертеж. ГрМО обладает точностью, зависящей от способа ее исполнения, например: с точностью до чертежных и инструментально-приборных ошибок (ручное исполнение); с точностью до методов математического моделирования объектов и ошибок чертежно-графических автоматов.

Графическая модель (ГрМ) объекта или процесса формируется на базе их математической модели и алгоритма отображения.

Однако объекты и процессы высокой степени сложности потребуют при своем моделировании расширения методов деятельности пользователя и описания плохо обусловленных процессов с целью их последующей формализации. Современные эвристические методы, методы дискретной математики, алгебры логики и т.д. позволяют успешно справляться с решением этих задач, которыми изобилует инженерная многопроекционная машинная графика (например, построение сечения плоскостью сложной корпусной детали).

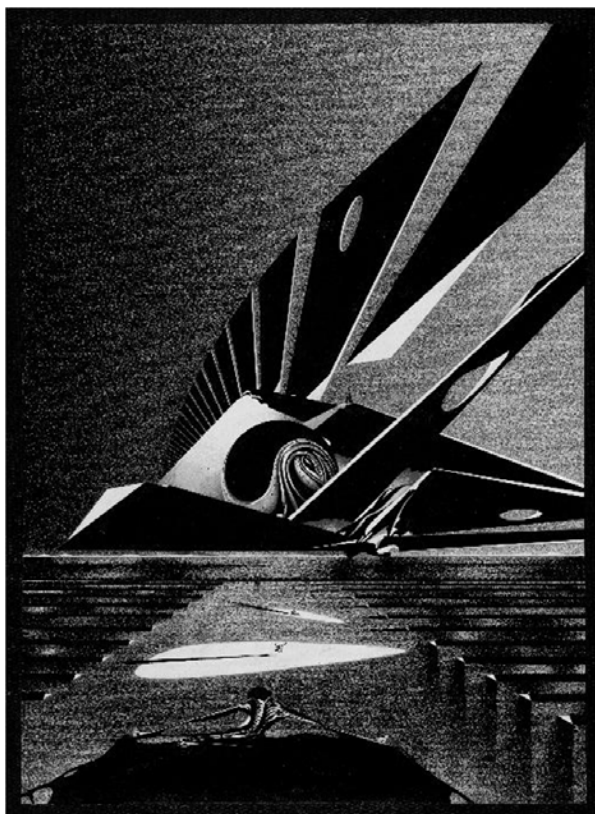
Известно, что главными задачами начертательной геометрии являются: построение проекционных изображений объектов, решение задач на проекционных чертежах и реконструкция этих решений в пространстве. Следовательно, ГрМ и ее роль могут быть различными.

Графической моделью интерпретирующей (ГрМИ) назовем проекционный чертеж, отображающий (графически интерпретирующий) результаты внутримашинного преобразования (переработки) информации, обеспечивающей решение задач визуализации исследуемого объекта или процесса. ГрМИ — результат визуализации информации об объекте или процессе.

Графической моделью конструктивной (ГрМК) назовем проекционный чертеж, отображающий (графически интерпретирующий) результаты внутримашинного преобразования (переработки) информации, обеспечивающей решение позиционных и метрических проектно-конструкторских и технологических задач на ГрМИ посредством реализации машинных геометрических и графических алгоритмов. ГрМК — результат визуализации решения инженерных задач на проекционном чертеже.

ГрМ, в особенности ГрМК, предъявляют высокие требования к математическому обеспечению базовых графических систем (БГС) и к собственно вычислительным средствам (ВС). Эти требования и должны определять все исходные требования к базовым графическим системам (ЕТС) и вычислительным средствам (ВС). Английская пословица гласит, что костюм должен подгоняться по человеку, а не наоборот. Следовательно, можно формулировать техническое задание к БГС, которое следует формировать для решения определенной группы инженерных задач определенного уровня сложности. Теперь можно говорить о соответствии или несоответствии существующих систем типа *Core sistem*, *GKS* и т.п. современным требованиям .

РАЗДЕЛ 1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ МАТЕМАТИКИ



Из всех языков, которыми мы обладаем, математика – единственный, не отягощенный предубеждениями, порожденными содержанием. Лишенная содержания и независимая от личного опыта, математика является единственным «космополитическим» языком человека. Она дает возможность соединять теории с очень различным содержанием, но близкие по логическим структурам. Она, таким образом, в высшей степени подходит для того, чтобы быть языком всей науки.

А. Рапопорт (1911-?)

1.1 Матрицы и операции над ними

1.1.1 Действия над матрицами.

Матрицей $m \times n$ называется прямоугольная таблица (массив элементов), имеющая m строк и n столбцов

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Матрицы обозначаются заглавными буквами A, B, \dots , а отдельные элементы матрицы строчными буквами a, b, \dots с различными нижними индексами. Например, элемент матрицы $B = [b_{ij}]$ —это элемент, расположенный на пересечении i -ой строки и j -ого столбца.

Матрица размерностью один на n называют «вектором-строкой», а матрица размерностью m на один – «вектором-столбцом».

Матрица $m \times m$ называется квадратной. Число ее строк и столбцов одинаково.

Квадратная матрица называется нулевой, если все ее элементы нули.

Квадратная матрица A называется диагональной, если все ее элементы, за исключением элементов главной диагонали a_{ij} (у которых $i=j$), равны нулю.

Диагональная матрица I называется единичной, если ее диагональные элементы равны единице.

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Две матрицы $A \equiv [a_{ij}]$ и $B \equiv [b_{ij}]$ одинакового размера $m \times n$ равны друг другу ($A=B$) в том и только в том случае, если $a_{ij} = b_{ij}$ для всех i и j .

Сложение - сумма двух матриц $A \equiv [a_{ij}]$ и $B \equiv [b_{ij}]$ одинакового размера $m \times n$ есть матрица размера $m \times n$

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

Масштабирование -произведение матрицы $A \equiv [a_{ij}]$ размера $m \times n$ на число (скаляр) α есть матрица $m \times n$

$$\alpha A = \alpha [a_{ij}] = [\alpha a_{ij}]$$

Из определений сложения и масштабирования непосредственно следуют свойства коммутативности, ассоциативности и дистрибутивности матриц одинаковой размерности:

$$\begin{aligned} A + B &= B + A \\ A + (B + C) &= (A + B) + C \\ (\alpha + \beta)(A + B) &= \alpha A + \alpha B + \beta A + \beta B \end{aligned}$$

Транспонирование матрицы M (обозначается M^T)-взаимная замена строк и столбцов матрицы M : ij -й элемент матрицы M^T равен ji -му элементу матрицы M .

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} A = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nm} \end{pmatrix}$$

Результатом транспонирования вектора-строки является вектор-столбец.

$$B = (b_1 b_2 \dots b_n) B^T = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

1.1.2. Умножение двух матриц

Произведение AB двух матриц A и B определено только в том случае, если эти матрицы согласованы, т.е. число столбцов m первой матрицы A равно числу строк n второй матрицы B . Например, если матрица A имеет размерность $m \times n$, а матрица B размерность $n \times r$, то произведение AB определено, а произведение BA нет ($r \neq m$).

Каждый элемент c_{ik} матрицы $C(m \times r)$ произведения матриц $A(m \times n)$ и $B(n \times r)$, $C=AB$ является скалярным произведением i -той строки матрицы A на j -тый столбец матрицы B

$$c_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$$

Некоторые полезные свойства умножения матриц. Пусть матрицы A, B, C согласуются должным образом. Тогда:

- $(AB)C = A(BC)$,
- $A(B + C) = AB + AC$.
- $(A + B)C = AC + BC$.
- $(AB)^T = A^T B^T$,
- $A(sB) = sAB$, где s число.

При формировании произведения двух матриц A и B имеет значение порядок следования этих матриц. Выражение AB – A умножается слева на B или B умножается справа на A . Если обе матрицы A и B квадратные одной и той же размерности, то оба произведения – AB и BA – определены, однако эти два произведения могут содержать различные элементы.

Если для двух матриц $AB = BA$, то эти две матрицы называются коммутативными.

Умножение вектора на матрицу.

В компьютерной графике часто встречается случай, когда матрица M умножается на вектор – столбец w , что обозначается Mw . Можно также умножить матрицу M слева на вектор-строку v – vM .

Скалярное и векторное произведения.

В ряде аналитических преобразований полезно записывать скалярное произведение $a \cdot b$ двух кортежей n -го порядка в форме произведения вектора на

матрицу. Для этого рассмотрим вектор b как матрицу- строку и транспонируем ее в матрицу- столбец b^T размерности n на 1. Тогда

$$a \cdot b = ab^T \quad \text{или} \quad a \cdot b = ba^T$$

Аналогично векторное произведение двух триад $a \times b$ можно записать в виде произведения

$$a \times b = (a_1 a_2 a_3) \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix}$$

Векторное произведение также является матрицей, умноженной справа на вектор- столбец a^T .

Внешнее, или тензорное произведение, записывается в виде:

$$a \otimes b = a^T b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} (b_1 b_2 b_3) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix},$$

откуда следует:

$$b \otimes a = (a \otimes b)^T \quad \text{и} \quad a(b \otimes c) = (ab).$$

1.1.3. Разбиение матрицы на блоки

Матрицу можно разбить на блоки из элементов и присвоить различным блокам свои имена. При этом, когда один блок находится под другим, то эти блоки то эти блоки должны иметь одинаковое число столбцов. Когда же два блока располагаются рядом, то они должны иметь одинаковое число строк.

Две блочные матрицы, разбиение которых было произведено одинаково (соответствующие блоки имеют одинаковую размерность), можно складывать поблочно. Для транспонирования блочных матриц следует транспонировать каждый блок отдельно, а затем транспонировать расположение блоков:

$$\begin{pmatrix} M_1 & \vdots & M_2 \\ \dots & \cdot & \dots \\ M_3 & \vdots & M_4 \end{pmatrix}^T = \begin{pmatrix} M_1^T & \vdots & M_3^T \\ \dots & \cdot & \dots \\ M_2^T & \vdots & M_4^T \end{pmatrix}$$

Две блочные матрицы можно умножать посредством обычного умножения их подматриц, если эти подматрицы согласованы:

$$\begin{pmatrix} M_1 & \vdots & M_2 \\ \dots & \cdot & \dots \\ M_3 & \vdots & M_4 \end{pmatrix} \begin{pmatrix} M_5 & \vdots & M_6 \\ \dots & \cdot & \dots \\ M_7 & \vdots & M_8 \end{pmatrix} = \begin{pmatrix} M_1 M_5 + M_2 M_7 & \vdots & M_1 M_6 + M_2 M_8 \\ \dots & \cdot & \dots \\ M_3 M_5 + M_4 M_7 & \vdots & M_3 M_6 + M_4 M_8 \end{pmatrix}.$$

1.1.4. Определитель матрицы

С каждой квадратной матрицей M связано число, которое называется ее определителем и обозначается $|M|$ или $\det M$. Определитель содержит информа-

цию о влиянии преобразования на площади и объемы геометрических объектов.

Определение. Каждому элементу m_{ij} квадратной матрицы M соответствует алгебраическое дополнение M_{ij} равное произведению $(-1)^{ij}$ на определитель матрицы, полученный из исходной матрицы M удалением (вычеркиванием) i -й строки и j -го столбца.

Чтобы вычислить определитель $|M|$ произвольной матрицы $n \times n$ необходимо:

- выбрать какую-либо строку матрицы M ,
- найти алгебраическое дополнение для каждого элемента этой строки,
- вычислить скалярное произведение этого ряда на n -кортеж, составленный из этих алгебраических дополнений.

Можно взять любой столбец матрицы M и использовать приведенный алгоритм.

Основные свойства определителей:

- $|M| = |M|^T$,
- Если две строки (два столбца) матрицы M одинаковы, то $|M| = 0$,
- Для двух квадратных матриц M и B $|MB| = |M||B|$,
- Если матрица B образована из матрицы M перестановкой двух строк (двух столбцов) матрицы M , то $|B| = -|M|$,
- Если матрица B образована из матрицы M умножением одной строки (одного столбца) матрицы M на число k , то $|B| = k|M|$,
- Если матрица B образована из матрицы M прибавлением число k произведения одной строки (одного столбца) матрицы M на какое либо другой ее строке, то $|B| = |M|$.

1.1.5. Обращение матриц

Квадратная матрица M размерностью $n \times n$ называется невырожденной, если ее определитель не равен нулю: $|M| \neq 0$. В этом случае матрица M имеет обратную матрицу M^{-1} , обладающую следующими свойствами:

- $MM^{-1} = M^{-1}M = I$,
- где I - единичная матрица размерностью $n \times n$
- $(AB)^{-1} = B^{-1}A^{-1}$.

Элементы обратной матрицы M^{-1} можно выразить через алгебраические дополнения матрицы M . Пусть A - матрица, обратная к матрице M , Тогда ij - м элементом матрица A будет число

$$a_{ij} = \frac{M_{ij}}{|M|}$$

т.е. алгебраическое дополнение элемента m_{ji} делится на определитель всей матрицы (при вычислении элемента a_{ij} используется алгебраическое дополнение m_{ji}).

Чтобы найти обратную матрицу необходимо:

• Создать промежуточную матрицу C из алгебраических дополнений

$$c_{ij} = M_{ij}.$$

• Найти определитель $|M|$ как скалярное произведение любой строки матрицы C на соответствующую строку матрицы M .

• Транспонировать матрицу C и получить C^T .

• Умножить каждый элемент матрицы C^T на $1/|M|$ и получить M^{-1} .

Обращение матриц используется при решении систем линейных алгебраических уравнений вида:

$$N \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

где заданы матрица $N \cdot [n \times n]$ и вектор-столбец b , требуется определить вектор x , который удовлетворял бы все n уравнений одновременно.

Если N - невырожденная матрица ($|N| \neq 0$), то решение имеет вид:

$$x = N^{-1}b.$$

Ортогональные матрицы.

Матрица M называется ортогональной, если $M^T = M^{-1} \Rightarrow MM^T = I$.

Следовательно, каждая строка матрицы M представляет собой вектор единичной длины, а строки взаимно ортогональны. (Это верно и для столбцов матрицы M).

Если матрица M имеет размерность 3×3 и разбита на три строки, то

$$M = \begin{pmatrix} a \\ \dots \\ b \\ \dots \\ c \end{pmatrix}$$

И каждая триада a, b, c имеет единичную длину, а $a \cdot b = a \cdot c = b \cdot c = 0$

1.2. Элементы векторной алгебры

1.2.1. Операции с векторами

Определение. Величины, которые полностью определяются заданием их числовых значений (например, длина, площадь, объем, масса и т. д.), называются скалярными.

Определение. Величины, для определения которых, кроме числового значения, необходимо знать еще и направление (например, сила, скорость, ускорение и т. д.), называются векторными.

Определение. Вектором называется направленный отрезок в пространстве или упорядоченная пара точек. Вектор изображается отрезком прямой, на

котором указано его направление, и обозначается \overline{AB} или a . Вектор, у которого начало и конец совпадают, называется нулевым и обозначается 0 .

Определение. Длина отрезка, изображающего вектор, называется длиной (модулем) вектора и обозначается так: $|\overline{AB}|, |a|$ или AB, a . Векторы, параллельные одной прямой или лежащие на одной прямой, называются коллинеарными.

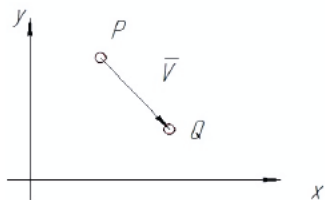


Рис.1.1

Полезно мыслить геометрически о векторе, как о перемещении от одной точки к другой. Вектор – это разность между двумя точками. Вектор $\overline{V} = Q - P \Rightarrow \overline{V}$ перемещает точку P для создания точки Q .

Тогда Q — это сумма $Q = P + \overline{V}$ точки и вектора, то есть точка: $P + \overline{V} = Q$ (рис.1.1)

Определение. Два вектора называются равными, если один из них может быть получен из другого путем параллельного переноса.



Рис.1.2

Определение. Векторы, расположенные в одной плоскости или параллельные одной и той же плоскости, называются компланарными.

Ориентацию вектора вдоль прямой указывает острие стрелки (рис.1.2). Нулевой вектор не имеет направления (ориентации).

Вектор может быть представлен списком его компонент; n -мерный вектор задается посредством n -кортежа (n -tuple) и записывается в форме матрицы – строки (*rowmatrix*) или матрицы – столбца (*colommatrix*)

$$\overline{W} = (w_1 \quad w_2 \quad w_n)$$

$$\overline{W} = \begin{pmatrix} w_1 \\ w_2 \\ w_n \end{pmatrix}$$

Произведением вектора на число $\lambda \in R$ называется вектор $\lambda a = \lambda a$, длина которого $|\lambda a|$ равна произведению $|\lambda|$ и $|a|$, т. е. $|\lambda a| = |\lambda| |a|$, а направление совпадает с направлением вектора a , если $\lambda > 0$, и противоположно ему, если $\lambda < 0$.

Перечислим свойства операции умножения вектора на число:

1. Для любого вектора a и чисел $\alpha, \beta \in R$ справедливо равенство $\alpha(\beta a) = (\alpha\beta)a$;

2. Если $a \neq 0$, то для любого вектора b , коллинеарного a , существует единственное число λ , удовлетворяющее равенству: $b = \lambda a$.

Вектор, длина которого равна единице, называется единичным вектором или ортом. Для любого вектора a справедливо равенство $a = |a|a_0$, где a_0 — единичный вектор.

1. Равенство. Два вектора считаются равными, если они имеют одинаковые величины, направления и ориентации.

Такие векторы можно представить в виде двух параллельных прямолинейных отрезков равной длины при условии, что эти отрезки вычерчиваются в одном масштабе. Заметим, что положение отрезков в пространстве (или на листе бумаги) не влияет на равенство.

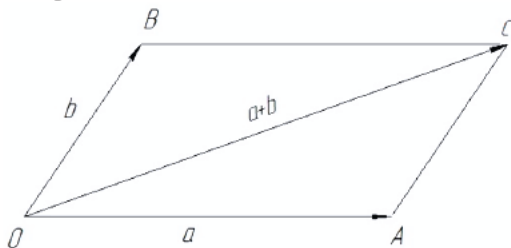


Рис. 1.3.

2. Сложение. Пусть даны два вектора a и b ;

чтобы найти сумму $a + b$ этих векторов, необходимо выполнить следующее.

Нарисуем A и B в одном и том же масштабе таким образом, чтобы начало B совпадало с концом A . Тогда отрезок, соединяющий начало A с концом B (рис. 1.3), будет вектором $A + B$ в том же масштабе, в котором представлены A и B . Если векторов больше двух, то чтобы найти сумму векторов a_1, a_2, \dots, a_n , нужно конец вектора a_1 совместить с началом вектора a_2 , конец вектора a_2 — с началом вектора a_3 , конец вектора a_3 — с началом a_4 и т. д., пока не дойдем до вектора a_n . Тогда суммой $a_1 + a_2 + \dots + a_n$ будет вектор, идущий из начала вектора a_1 в конец вектора a_n .



Рис. 1.4

3. Противоположный вектор. Вектор $(-a)$ определим как вектор, величина и направление которого такие же, что и у вектора a , но ориентация которого противоположна ориентации вектора a (рис. 1.4).

4. Вычитание. По определению $a - b = a + (-b)$ (рис. 1.5.)

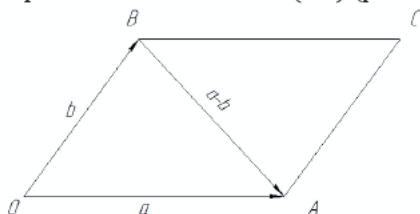


Рис. 1.5

5. Умножение на скаляр (масштабирование вектора).

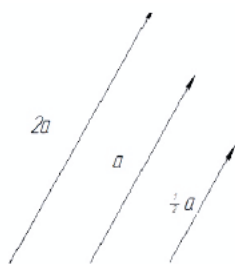


Рис. 1.6.

При умножении вектора \vec{a} на скаляр s каждый компонент вектора умножается на s $s\vec{a} = (sa_1, sa_2, \dots, sa_n)$. Если $s > 0$ то новый вектор увеличивается (или уменьшается) в s раз, не меняя своего направления. Если $s < 0$, то направление вектора изменяется на противоположное.

Параметр s - обычное число, или скаляр (рис.1.6)

6. Два вектора коллинеарны, если один из них пропорционален другому, т.е. $k\vec{a} + l\vec{b} = 0$. (1.1)

Следствия, вытекающие из приведенных выше определений

Пусть a, b, c - векторы, а λ, μ, ν - скаляры. Тогда (рис. 1.7):

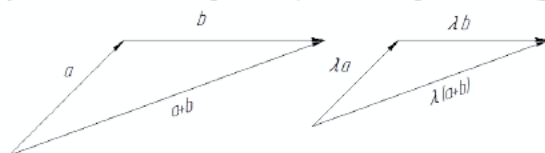


Рис. 1.7.

1. $a + b = b + a$
2. $a + (b + c) = (a + b)c$
3. $\lambda(\mu a) = (\lambda\mu)a$
4. $(\mu + \nu)a = \mu a + \nu a$
5. $\lambda(a + b) = \lambda a + \lambda b$

Проекцией вектора \overline{AB} на ось l называется длина отрезка CD этой оси, заключенного между проекциями на ось l точек A и B , взятая со знаком «+», если направление отрезка CD совпадает с направлением оси l , и со знаком «-», если эти направления противоположны. Обозначается это так:

$CD = np_1 \overline{AB}$. Проекция вектора на ось равна длине вектора, умноженной на косинус угла между ним и осью: $np_1 \overline{AB} = |\overline{AB}| \cos \alpha$.

Пример 1.

1. Точки K и L являются серединами сторон BC и CD параллелограмма $ABCD$. Полагая $\overline{AK} = k$ и $\overline{AL} = l$, выразить через k и l векторы \overline{BC} и \overline{DC} (рис. 1.8).

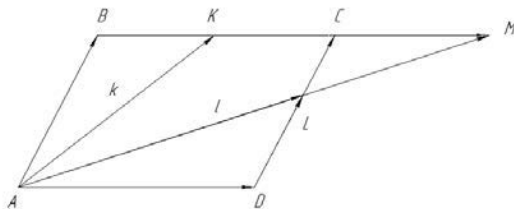


Рис. 1.8

Решение.

Из треугольника AKM получаем: $k + \overline{KM} + \overline{AM} = 2l, \overline{KM} = 2l - k$

$$\overline{BK} = \frac{1}{3}\overline{KM} = \frac{1}{3}(2l - k) = \frac{2l - k}{3}$$

$$\overline{BC} = 2\overline{BK} = \frac{2(2l - k)}{3} = \frac{4l - 2k}{3}$$

Из треугольника ADL имеем: $\overline{AD} + \overline{DL} = l\overline{DL} = l - \overline{AD} = \frac{2k - l}{3}$

$$\overline{DC} = 2\overline{DL}, \quad \overline{DC} = \frac{2(2k - l)}{3} = \frac{4k - 2l}{3}$$

1.2.2. Линейные комбинации векторов

Для того, чтобы составить линейную комбинацию двух векторов \vec{V} и \vec{W} одной и той же размерности нужно масштабировать каждый из них с помощью скаляров, например a и b , и сложить эти взвешенные величины для создания нового вектора $a\vec{V} + b\vec{W}$.

Например:

пусть $\vec{V}(3,4,-1)$, а $\vec{W}(-1,0,2)$, $a = 2$, $b = 6$. Тогда $2(3,4,-1) + 6(-1,0,2) = (0,8,10)$.

Определение 1. Линейной комбинацией m векторов $\vec{V}_1, \vec{V}_2, \dots, \vec{V}_m$ называется вектор

$$\vec{W} = a_1\vec{V}_1 + a_2\vec{V}_2 + \dots + a_m\vec{V}_m \quad (1.2)$$

где a_1, a_2, \dots, a_m - скаляры.

Линейные комбинации векторов применяются при представлении кривых и поверхностей с помощью сплайн - функций.

Определение 2. Линейная комбинация векторов называется аффинной, если

$$a_1 + a_2 + \dots + a_m = 1 \quad (1.3)$$

Коэффициенты линейной комбинации векторов \vec{a} и \vec{b} часто искусственно приводятся к виду, чтобы их сумма равнялась единице, умножая один вектор на некоторый скаляр t , а второй – на $(1 - t)$:

$$(1 - t)\vec{a} + t\vec{b} \text{ - а.л.к. (аффинная линейная комбинация)}$$

Определение 3. Аффинная линейная комбинация (а.л.к.) векторов называется выпуклой, если $a_1 + a_2 + \dots + a_m = 1$

$$a_i \geq 0 \text{ для } i = 1, 2, \dots, m.$$

Следовательно, все a_i должны находиться между нулем и единицей.

Совокупность коэффициентов a_1, a_2, \dots, a_n иногда называют разбиением единицы, понимая под этим, что некоторое количество «материала» разбивается на части.

Слайд— кривые фактически являются выпуклыми комбинациями некоторых векторов.

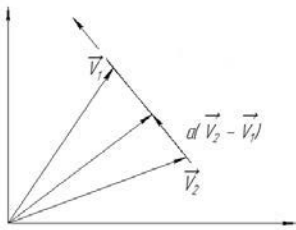


Рис.1.9.

В компьютерной графике оказывается полезным понятие «множество всех выпуклых комбинаций» некоторого набора векторов. Для двух векторов \vec{V}_1 и \vec{V}_2 выпуклая комбинация имеет вид:

$$\vec{V} = (1 - a)\vec{V}_1 + a\vec{V}_2 \Rightarrow a < 1$$

$$\vec{V} = \vec{V}_1 + a(\vec{V}_2 - \vec{V}_1)$$

На рис.1.9. показан вектор \vec{V} , равный сумме вектора \vec{V}_1 и некоторой части от $(\vec{V}_2 - \vec{V}_1)$, причем конец вектора \vec{V} расположен на прямой, соединяющей \vec{V}_1 и \vec{V}_2 .

На рис.1.10 показано множество всех выпуклых комбинаций трех векторов.

Выбраны два параметра $a_1 < 1$ и $a_2 < 1$, лежащих в промежутке между 0 и 1, и составляющих линейную комбинацию:

$g = a_1\vec{V}_1 + a_2\vec{V}_2 + (1 - a_1 - a_2)\vec{V}_3$ в которой сумма a_1 и a_2 не превышает 1: $a_1 + a_2 < 1$.

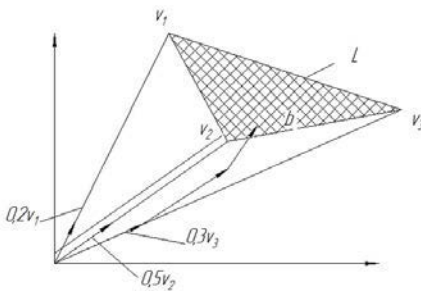


Рис. 1.10

Соответствующим выбором a_1 и a_2 можно выразить любой вектор, оканчивающийся внутри затемненного треугольника, и ни один из векторов, оканчивающийся вне этого треугольника, таким способом выражен быть не может.

1.2.3. Декартовы компоненты вектора

Определение. Векторы $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$ называются линейно зависимыми, если существуют числа $\lambda_1, \lambda_2, \dots, \lambda_n \in R$, одновременно не равные нулю, такие, что выполняется равенство

$$\lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2 + \dots + \lambda_n \bar{a}_n = 0.$$

В противном случае векторы $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$ называются линейно независимыми.

Приведем теоремы, устанавливающие условия линейной зависимости векторов на плоскости и в пространстве.

Теорема 1. Любые два коллинеарных вектора линейно зависимы, а два неколлинеарных вектора линейно независимы.

Теорема 2. Три компланарных вектора линейно зависимы, а три некопланарных вектора линейно независимы.

Базисом на плоскости (в R^2) называют два неколлинеарных вектора этой плоскости, взятых в определенном порядке. Если на плоскости выбран базис \bar{e}_1, \bar{e}_2 , то любой вектор \bar{a} этой плоскости можно разложить по векторам \bar{e}_1, \bar{e}_2 , и такое разложение единственно.

Базисом в пространстве (в R^3) называют три некопланарных вектора, взятых в определенном порядке. Если в пространстве R^3 выбран базис $\bar{e}_1, \bar{e}_2, \bar{e}_3$, то любой вектор \bar{a} можно разложить по векторам $\bar{e}_1, \bar{e}_2, \bar{e}_3$, и такое разложение единственно.

Если $\bar{e}_1, \bar{e}_2, \bar{e}_3$ — базис в R^3 и вектор $\bar{a} = \alpha_1 \bar{e}_1 + \alpha_2 \bar{e}_2 + \alpha_3 \bar{e}_3$, то числа $\alpha_1, \alpha_2, \alpha_3$ называют координатами вектора \bar{a} в базисе $\bar{e}_1, \bar{e}_2, \bar{e}_3$, и записывают: $\bar{a} = (\alpha_1, \alpha_2, \alpha_3)$.

При умножении вектора $\bar{a} = (\alpha_1, \alpha_2, \alpha_3)$ на число $\lambda \in R$ все его координаты умножаются на это число: $\lambda \bar{a} = (\lambda \alpha_1, \lambda \alpha_2, \lambda \alpha_3)$

При сложении (вычитании) векторов $\bar{a} = (\alpha_1, \alpha_2, \alpha_3)$ и $\bar{b} = (\beta_1, \beta_2, \beta_3)$ складываются (вычитаются) их соответствующие координаты: $a \pm b = (\alpha_1 \pm \beta_1, \alpha_2 \pm \beta_2, \alpha_3 \pm \beta_3)$.

Упорядоченная тройка некопланарных векторов $\bar{a}, \bar{b}, \bar{c}$ называется правой, если наблюдателю, находящемуся внутри телесного угла, образованного этими векторами, кратчайшие повороты от \bar{a} к \bar{b} и от \bar{b} к \bar{c} кажутся происходящими против хода часовой стрелки. В противном случае данная тройка называется левой.

Декартовой системой координат в пространстве называется совокупность фиксированной точки O и базиса $\bar{e}_1, \bar{e}_2, \bar{e}_3$. Точка O называется началом координат, а прямые, проходящие через начало координат в направлении базисных векторов, — осями координат. Прямая Ox называется осью абсцисс, прямая Oy — осью ординат, прямая Oz — осью аппликат. Плоскости, проходящие через оси координат, называются координатными плоскостями.

Вектор \overline{OM} называется радиусом-вектором точки M . Координаты радиус-вектора точки M по отношению к точке O называются координатами точки M в данной системе координат. Первая координата называется абсциссой, вторая — ординатой, третья — аппликатой.

Базис называется ортонормированным, если базисные векторы попарно ортогональны и длина каждого из них равна единице.

Декартова система координат с ортонормированным базисом $\bar{i} = (1, 0, 0), \bar{j} = (0, 1, 0), \bar{k} = (0, 0, 1)$ называется прямоугольной системой координат, а векторы $\bar{i}, \bar{j}, \bar{k}$ — ортами координатных осей.

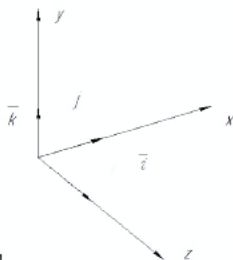


Рис. 1.11

Определив эти векторы, можно при их помощи представить любой вектор с помощью трех векторов, идущих вдоль осей координат (рис. 1.11). На чертеже эти векторы можно изобразить в виде трех ребер

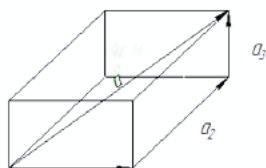


Рис. 1.12

прямоугольного параллелепипеда, диагональю которого является a (рис. 1.12). Три ребра изображенного на рисунке параллелепипеда образованы векторами \bar{a}_1, \bar{a}_2 и \bar{a}_3 . Если векторы \bar{a}_1, \bar{a}_2 и \bar{a}_3 расположить в ином порядке, путь от P до Q пройдет соответственно по иным ребрам параллелепипеда, но результирующий вектор \bar{a} не изменится. Таким образом, соотношение $\bar{a} = \bar{a}_1 + \bar{a}_2 + \bar{a}_3$ однозначно для любой заданной системы координат. Пусть $\bar{a}_1 = a_1\bar{i}, a_2 =$

$= a_2\bar{j}, a_3 = a_3\bar{k}$, где a_1, a_2, a_3 — длины векторов \bar{a}_1, \bar{a}_2 и \bar{a}_3 соответственно. Отсюда: $\bar{a} = a_1\bar{i} + a_2\bar{j} + a_3\bar{k}$, следовательно, (a_1, a_2, a_3) — декартовы компоненты вектора a для выбранной системы координат. Используя декартовы компоненты, можно установить следующие факты):

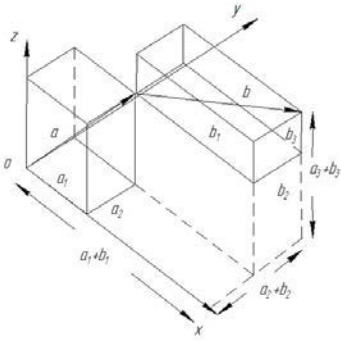


Рис.1.13

$$\begin{aligned} \bar{a} + \bar{b} &= (a_1 + b_1)\bar{i} + (a_2 + b_2)\bar{j} + (a_3 + b_3)\bar{k} \\ -\bar{a} &= -a_1\bar{i} - a_2\bar{j} - a_3\bar{k} \\ \lambda\bar{a} &= \lambda a_1\bar{i} + \lambda a_2\bar{j} + \lambda a_3\bar{k} \\ |\bar{a}| &= \sqrt{a_1^2 + a_2^2 + a_3^2} \end{aligned}$$

(согласно теореме Пифагора, рис.1.13.)

Компоненты единичного вектора дают косинусы углов, образованных этим вектором и осями Ox , Oy и Oz .

На рис. 1.14 видно, что $u = \overline{PU}/PU$. Разлагая \overline{PU} , мы получаем следующее выражение:

$$\bar{u} = \frac{\overline{PW}_i + \overline{PS}_j + \overline{PQ}_k}{PU} = \frac{PW}{PU}\bar{i} + \frac{PS}{PU}\bar{j} + \frac{PQ}{PU}\bar{k}$$

Следовательно, $\bar{u} = \cos \alpha \bar{i} + \cos \beta \bar{j} + \cos \gamma \bar{k}$,

где α, β, γ — углы, образованные вектором \bar{u} с положительными полуосями Ox, Oy и Oz . Если эти косинусы ассоциированы с некоторой прямой, то они называются *направляющими косинусами* этой прямой и обычно обозначаются через l, m и n .

Так направление вектора $\bar{a} = (a_x, a_y, a_z)$ определяется углами α, β, γ , образованными вектором \bar{a} с положительными направлениями осей Ox, Oy и Oz соответственно. Косинусы этих углов называются *направляющими косинусами вектора* \bar{a} и определяются по формулам:

$$\begin{aligned} l = \cos \alpha &= \frac{a_x}{a} = \frac{a_x}{\sqrt{a_x^2 + a_y^2 + a_z^2}} \\ m = \cos \beta &= \frac{a_y}{a} = \frac{a_y}{\sqrt{a_x^2 + a_y^2 + a_z^2}} \\ n = \cos \gamma &= \frac{a_z}{a} = \frac{a_z}{\sqrt{a_x^2 + a_y^2 + a_z^2}} \end{aligned}$$

Направляющие косинусы связаны соотношением:

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$$

Следовательно, $\bar{u} = l\bar{i} + m\bar{j} + n\bar{k}$. Любое множество чисел $(\lambda l, \lambda m, \lambda n)$, пропорциональных (l, m, n) , называется множеством *направляющих коэффициентов* прямой. Таким образом, компоненты любого вектора, параллельного некоторой прямой, являются направляющими коэффициентами этой прямой, и, если мы нормализуем вектор, мы получим направляющие косинусы.

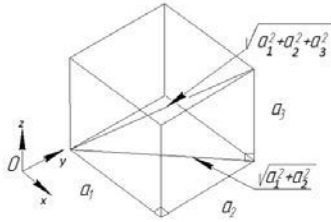


Рис. 1.15

Знание законов сложения, образования обратного вектора и умножения векторов на скаляр позволяет нам описывать положение отдельных точек и прямых в пространстве, определять средние точки и относительные положения прямых, а также устанавливать, пересекаются рассматриваемые прямые или нет.

Для того чтобы определить угол, образованный прямыми, установить наикратчайшее расстояние между скрещивающимися прямыми, получить уравнение плоскости и найти проекции точки на заданную плоскость, мы должны знать векторные эквиваленты тригонометрических соотношений.

Опираясь на векторы, мы можем столкнуться с вопросом разложения некоторого вектора на его составляющие, которые могут быть отличны от упомянутых ранее составляющих по осям x , y и z .

Любой вектор можно однозначным образом разложить на два взаимно перпендикулярных вектора, если задано направление одного из них. Предположим, что один из двух взаимно перпендикулярных векторов направлен вдоль вектора $\vec{b} = (\overline{AB})$. Опустим перпендикуляр из конца вектора \vec{a} на вектор \vec{b} (рис. 1.16) и обозначим его через \vec{c} . Тогда $\vec{a} = \lambda \vec{b} + \vec{c}$, где $\vec{b} \perp \vec{c}$ (рис. 1.17). Пусть $\angle CAB = \theta$; тогда $AB = AC \cos \theta = |\vec{a}| \cos \theta$, т. е. AB является компонентой AC вдоль вектора \vec{b} . Чтобы найти угол θ , рассмотрим треугольник ACD и применим правило косинусов (Рис. 1.18). Мы имеем соотношение $CD^2 = AC^2 + AD^2 - 2AC \cdot AD \cos \theta$, или $|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}||\vec{b}| \cos \theta$.

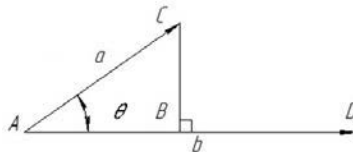


Рис. 1.16

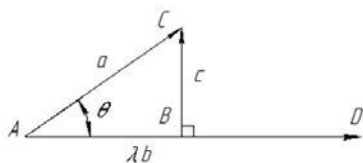


Рис. 1.17

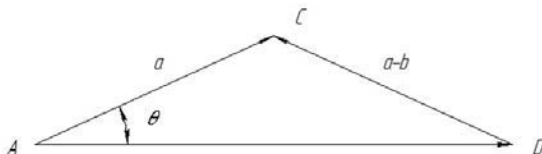


Рис. 1.18

Выражая эти модули через декартовы компоненты, мы получаем уравнение $(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 = (a_1^2 + a_2^2 + a_3^2) + (b_1^2 + b_2^2 + b_3^2) - 2|a||b| \cos \theta$

После некоторых упрощений находим, что угол θ , образованный двумя векторами a и b , определяется соотношением:

$$\cos \theta = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{|a||b|} \quad (1.5)$$

Компонента вектора \bar{a} вдоль вектора \bar{b} (или проекция a на b) определяется соотношением:

$$AB = |a| \cos \theta = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{|b|} \quad (1.6)$$

В частности, компонента вектора \bar{a} вдоль единичного вектора \bar{u} определяется формулой $a_1 u_1 + a_2 u_2 + a_3 u_3$.

Рассмотрим следующие примеры

1. Даны векторы: $\bar{a} = 3\bar{i} - 6\bar{j} - \bar{k}$, $\bar{b} = \bar{i} + 4\bar{j} - 5\bar{k}$, $\bar{c} = 3\bar{i} - 4\bar{j} + 12\bar{k}$.

Найти проекцию вектора $\bar{a} + \bar{b}$ на вектор \bar{c} .

Решение. Требуемая проекция $\text{пр}_c(\bar{a} + \bar{b}) = \frac{(\bar{a} + \bar{b}, \bar{c})}{|\bar{c}|}$.

Имеем: $\bar{a} = (3, -6, -1)$, $\bar{b} = (1, 4, -5)$, $\bar{c} = (3, -4, 12)$, $\bar{a} + \bar{b} = (4, -2, -6)$

Тогда

$$\text{пр}_c(\bar{a} + \bar{b}) = \frac{(\bar{a} + \bar{b}, \bar{c})}{|\bar{c}|} = \frac{4 \times 3 + (-2)(-4) + (-6) \times 12}{\sqrt{3^2 + (-4)^2 + 12^2}} = -52/13 = -4$$

2. Вектор \bar{x} , перпендикулярный к векторам $\bar{a} = (4, -2, -3)$ и $\bar{b} = (0, 1, 3)$, образует с осью Оутовой угол. Зная, что $|\bar{x}| = 26$, найти координаты вектора \bar{x} .

Решение. Обозначим координаты вектора \bar{x} через x_1, x_2, x_3 : $\bar{x} = (x_1, x_2, x_3)$. Так как $\bar{x} \perp a, \bar{x} \perp b$, то $(\bar{x}, \bar{a}) = 0, (\bar{x}, \bar{b}) = 0$. Переходя к координатам векторов \bar{x}, \bar{a} и \bar{b} , получаем систему уравнений:

$$\left. \begin{aligned} 4x_1 - 2x_2 - 3x_3 &= 0 \\ x_2 + 3x_3 &= 0 \end{aligned} \right\}$$

решив которую, имеем: $x_2 = -3x_3, 4x_1 = 2x_2 + 3x_3 = -3x_3, x_1 = -(3/4)x_3$

Следовательно, $x = (-(3/4)x_3, -3x_3, x_3)$. Так как $|x| = 26$, то $\sqrt{(9/16)x_3^2 + 9x_3^2 + x_3^2} = 26$

Тогда $\pm x_3 \times 13/4 = 26, x_3 = \pm 8$.

Так как по условию $\cos \beta < 0$, т. е. угол β - тупой, то $\cos \beta = x_2/26 = -3x_3/26 < 0$. Отсюда имеем: $-3x_3 < 0, x_3 > 0, x_3 = 8$. Тогда $x = (-6, -24, 8)$.

3. Даны векторы: $\bar{a} = 2\bar{i} - \bar{j} + 3\bar{k}, \bar{b} = \bar{i} - 3\bar{j} + 2\bar{k}, \bar{c} = 3\bar{i} + 2\bar{j} - 4\bar{k}$. Найти вектор \bar{x} , удовлетворяющий следующим условиям:

$$(\bar{x} \cdot \bar{a}) = -5, \quad (\bar{x} \cdot \bar{b}) = -11, \quad (\bar{x} \cdot \bar{c}) = -20.$$

Решение. Обозначим координаты вектора x через x_1, x_2, x_3 . Тогда

$$\left. \begin{aligned} (\bar{x} \cdot \bar{a}) &= -5 \\ (\bar{x} \cdot \bar{b}) &= -11 \\ (\bar{x} \cdot \bar{c}) &= 20 \end{aligned} \right\} \text{ или } \left. \begin{aligned} 2x_1 - x_2 + 3x_3 &= -5 \\ x_1 - 3x_2 + 2x_3 &= -11 \\ 3x_1 + 2x_2 - 4x_3 &= 20 \end{aligned} \right\}$$

Найдем решение полученной системы по формулам Крамера:

$$\Delta = \begin{vmatrix} 2 & -1 & 3 \\ 1 & -3 & 2 \\ 3 & 2 & -4 \end{vmatrix} = 20 - 6 + 6 + 27 - 4 - 8 = 39$$

$$\Delta_1 = \begin{vmatrix} -5 & -1 & 3 \\ -11 & -3 & 2 \\ 20 & 2 & -4 \end{vmatrix} = 60 - 40 - 66 + 180 + 44 + 20 = 78,$$

$$\Delta_2 = \begin{vmatrix} 2 & -5 & 3 \\ 1 & -11 & 2 \\ 3 & 20 & -4 \end{vmatrix} = 88 - 30 + 60 + 99 - 20 - 80 = 117,$$

$$\Delta_3 = \begin{vmatrix} 2 & -1 & -5 \\ 1 & -3 & -11 \\ 3 & 2 & 20 \end{vmatrix} = -120 + 33 - 10 - 45 + 20 + 44 = 78.$$

Тогда $x_1 = \Delta_1/\Delta = 78/39 = 2, x_2 = \Delta_2/\Delta = 117/39 = 3, x_3 = \Delta_3/\Delta = 78/39 = 2$, т. е. $\bar{x} = (2, 3, 2)$.

4. Даны векторы: $\bar{a} = (5, -1, 10)$, $\bar{b} = (3, 6, 4)$, $\bar{c} = (6, 0, 11)$, $\bar{d} = (2, -7, 6)$. Показать, что векторы $\bar{a}, \bar{b}, \bar{c}$ образуют базис, и найти координаты вектора \bar{d} в этом базисе.

Решение. Как известно, базисом в пространстве R^3 является любая упорядоченная система из трех линейно независимых векторов. Покажем, что векторы $\bar{a}, \bar{b}, \bar{c}$ линейно независимы, т. е. выполняется равенство: $\alpha_1 \bar{a} + \alpha_2 \bar{b} + \alpha_3 \bar{c} = 0$ только при условии, что все числа $\alpha_1, \alpha_2, \alpha_3$ одновременно равны нулю. Подставляя в это равенство координаты векторов $\bar{a}, \bar{b}, \bar{c}$, получаем:

$$\alpha_1(5\bar{e}_1 - \bar{e}_2 + 10\bar{e}_3) + \alpha_2(3\bar{e}_1 + 6\bar{e}_2 + 4\bar{e}_3) + \alpha_3(6\bar{e}_1 + 11\bar{e}_3) = 0$$

или $(5\alpha_1 + 3\alpha_2 + 6\alpha_3)\bar{e}_1 + (-\alpha_1 + 6\alpha_2)\bar{e}_2 + (10\alpha_1 + 4\alpha_2 + 11\alpha_3)\bar{e}_3 = 0$

Для того чтобы вектор, разложенный по базису $\bar{e}_1, \bar{e}_2, \bar{e}_3$, был равен нулевому вектору, его координаты должны равняться нулю, т. е.

$$\left. \begin{aligned} 5\alpha_1 + 3\alpha_2 + 6\alpha_3 &= 0 \\ -\alpha_1 + 6\alpha_2 &= 0 \\ 10\alpha_1 + 4\alpha_2 + 11\alpha_3 &= 0 \end{aligned} \right\}$$

Получим однородную систему трех линейных уравнений с тремя неизвестными $\alpha_1, \alpha_2, \alpha_3$. Такая система имеет нулевое решение ($\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$), если её определитель не равен нулю. Поскольку

$$\begin{vmatrix} 5 & 3 & 6 \\ -1 & 6 & 0 \\ 10 & 4 & 11 \end{vmatrix} = -21 \neq 0,$$

то векторы $\bar{a}, \bar{b}, \bar{c}$ линейно независимы. Следовательно, они образуют базис и вектор \bar{d} является линейной комбинацией векторов $\bar{a}, \bar{b}, \bar{c}$: $\bar{d} = \beta_1 \bar{a} + \beta_2 \bar{b} + \beta_3 \bar{c}$. Числа $\beta_1, \beta_2, \beta_3$ будут координатами вектора \bar{d} в базисе $\bar{a}, \bar{b}, \bar{c}$. Найдем их.

Воспользовавшись разложением $\bar{a}, \bar{b}, \bar{c}, \bar{d}$ в базисе $\bar{e}_1, \bar{e}_2, \bar{e}_3$, имеем:

$$\begin{aligned} & 2\bar{e}_1 - 7\bar{e}_2 + 6\bar{e}_3 = \\ & = \beta_1(5\bar{e}_1 - \bar{e}_2 + 10\bar{e}_3) + \beta_2(3\bar{e}_1 + 6\bar{e}_2 + 4\bar{e}_3) + \beta_3(6\bar{e}_1 + 11\bar{e}_3) \end{aligned}$$

Или

$$\begin{aligned} & 2\bar{e}_1 - 7\bar{e}_2 + 6\bar{e}_3 = \\ & = (5\beta_1 + 3\beta_2 + 6\beta_3)\bar{e}_1 + (-\beta_1 + 6\beta_2)\bar{e}_2 + (10\beta_1 + 4\beta_2 + 11\beta_3)\bar{e}_3 \end{aligned}$$

Из равенства векторов следует равенство их координат, поэтому получаем систему:

$$\left. \begin{aligned} 5\beta_1 + 3\beta_2 + 6\beta_3 &= 2 \\ -\beta_1 + 6\beta_2 &= -7 \\ 10\beta_1 + 4\beta_2 + 11\beta_3 &= 6 \end{aligned} \right\}$$

Решая ее по формулам Крамера $\beta_i = \Delta_i / \Delta$, $i = 1, 2, 3$, находим:

$$\Delta_1 = \begin{vmatrix} 2 & 2 & 6 \\ -7 & 6 & 0 \\ 6 & 4 & 11 \end{vmatrix} = -21, \Delta_2 = \begin{vmatrix} 5 & 2 & 6 \\ -1 & -7 & 0 \\ 10 & 6 & 11 \end{vmatrix} = 21, \Delta_3 = \begin{vmatrix} 5 & 3 & 2 \\ -1 & 6 & -7 \\ 10 & 4 & 6 \end{vmatrix} = 0$$

Следовательно,

$$\beta_1 = -21/-21 = 1, \beta_2 = 21/-21 = -1, \beta_3 = 0/-21 = 0,$$

т. е. $\vec{a} = (1, -1, 0)$.

5. Даны вершины треугольника ABC : $A(3, 2, -3)$, $B(5, 1, -1)$, $C(1, -2, 1)$. Вычислить его внешний угол при вершине A (рис.1.19).

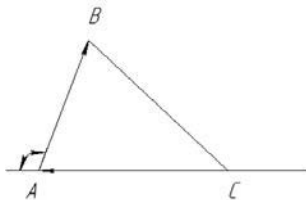


Рис. 1.19

Решение. Внешний угол при вершине A — это угол между векторами \overrightarrow{AB} и \overrightarrow{CA} . Так как $\overrightarrow{AB} = (2, -1, 2)$, $\overrightarrow{CA} = (2, 4, -4)$, то

$$\cos(\overrightarrow{AB}, \overrightarrow{CA}) = \frac{(\overrightarrow{AB}, \overrightarrow{CA})}{|\overrightarrow{AB}| \cdot |\overrightarrow{CA}|} = \frac{2 \times 2 + (-1) \times 4 + 2 \times (-4)}{\sqrt{2^2 + (-1)^2 + 2^2} \sqrt{(-2)^2 + (-4)^2 + 4^2}} = -4/9$$

Следовательно, $(\overrightarrow{AB}, \overrightarrow{CA}) = \arccos(-4/9) \approx 116^\circ 25'$

1.2.4. Скалярное произведение

Скалярным произведением двух векторов \vec{a} и \vec{b} называется число, равное произведению длин этих векторов на косинус угла φ между ними:

$$(\vec{a} \cdot \vec{b}) = |\vec{a}| |\vec{b}| \cos \varphi.$$

Обозначается скалярное произведение (\cdot) .

Если в прямоугольной системе координат векторы \vec{a} и \vec{b} заданы своими координатами: $\vec{a} = (a_x, a_y, a_z)$, $\vec{b} = (b_x, b_y, b_z)$, то $(\vec{a} \cdot \vec{b}) = a_x b_x + a_y b_y + a_z b_z$. Из определения скалярного произведения следует, что

$$\cos \varphi = \frac{(\vec{a} \cdot \vec{b})}{|\vec{a}| |\vec{b}|} = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2} \sqrt{b_x^2 + b_y^2 + b_z^2}} \quad (1.7)$$

С помощью последней формулы можно найти угол, который образуют векторы \vec{a} и \vec{b} .

Перечислим основные свойства скалярного произведения векторов:

1. $(\vec{a} \cdot \vec{a}) = |\vec{a}|^2 = \vec{a}^2$, отсюда $|\vec{a}| = \sqrt{(\vec{a} \cdot \vec{a})}$ - модуль вектора \vec{a}
2. $(\vec{a} \cdot \vec{b}) = 0$, если $\vec{a} = 0$, либо $\vec{b} = 0$, либо $\vec{a} \perp \vec{b}$

3. $(\bar{a} \cdot \bar{b}) = (\bar{b} \cdot \bar{a})$ - симметричность (коммутативность)
4. $(\lambda \bar{a} \cdot \bar{b}) = (\bar{a} \cdot \lambda \bar{b}) = \lambda(\bar{a} \cdot \bar{b}), \lambda \in R$;
5. $(\bar{a} \cdot \bar{b} + \bar{c}) = (\bar{a} \cdot \bar{b}) + (\bar{a} \cdot \bar{c})$; - линейность (дистрибутив)
6. $(\bar{a} \cdot \bar{b}) = |\bar{a}| \text{пр}_a \bar{b} = |\bar{a}| \text{пр}_b \bar{a}$, следовательно, $\text{пр}_b \bar{a} = \frac{(a \cdot b)}{|b|}$.

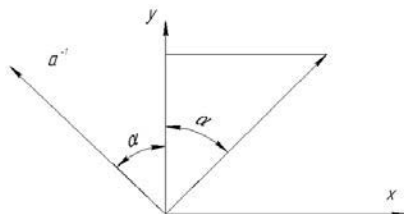


Рис. 1.20

Введем специальный символ для одного часто встречающегося в компьютерной графике вектора, ортогонального к заданному двумерному вектору a — a^\perp -(*perp*-перп).

Определение. Пусть $\bar{a}(a_x, a_y)$. Тогда

$$\bar{a}^\perp = (-a_y; a_x) \quad (1.8)$$

\bar{a}^\perp называется перпендикулярным против часовой стрелки к \bar{a} («перп» вектор \bar{a}^\perp).

В случае трех измерений не существует единственного вектора, перпендикулярного к заданному трехмерному вектору \bar{a} .

Можно рассматривать символ (\perp) «перп» как оператор, выполняющий над своим аргументом операцию «поворот против часовой стрелки на 90° ». В этом случае он обладает следующими свойствами:

- линейность $(a + b)^\perp = a^\perp + b^\perp$,
- для любого скаляра A : $(A \cdot a)^\perp = A a^\perp$
- $a^{\perp\perp} = (a^\perp)^\perp = -a$ два перпа осуществляют реверсирование (изменение направления на обратное).

Основные свойства перп скалярного произведения.

$$a^\perp \cdot b = a_x b_y - a_y b_x \quad (\text{величина перп скалярного произведения})$$

$$a^\perp \cdot a = 0, \quad (a^\perp \text{ перпендикулярен к } a)$$

$$|a^\perp|^2 = |a|^2, \quad (a^\perp \text{ и } a \text{ одинаковые по длине})$$

$$a^\perp \cdot b = -b^\perp \cdot a \quad (a^\perp - \text{антисимметричен})$$

$$a^\perp \cdot b = \begin{pmatrix} a_x & a_y \\ b_x & b_y \end{pmatrix} \quad (1.9)$$

Рассмотрим примеры.

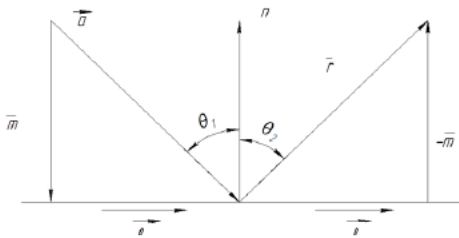


Рис. 1.21

Пример 1. Чтобы отобразить на экране дисплея отражение света от зеркальной поверхности или поведение соударяющихся частиц, необходимо найти направление, которое получает объект после отражения от поверхности.

На рис. 1.21 луч движется в направлении a , попадает на прямую l и отражается в направлении r . Вектор n перпендикулярен прямой. Угол $\theta_1 = \theta_2$.

Разложим вектор \bar{a} на две составляющие: $\bar{m} \parallel \bar{n}$ и $\bar{e} \perp \bar{n}$. Вследствие симметрии вектор \bar{r} имеет такой же компонент $\bar{e} \perp \bar{n}$, но противоположно направленный, следовательно $\bar{r} = \bar{e} - \bar{m}$. Из равенства $\bar{e} = \bar{a} - \bar{m}$ следует, что $\bar{r} = \bar{a} - 2\bar{m}$. Поскольку \bar{m} является ортогональной проекцией \bar{a} на \bar{n} , то:

$$\bar{m} = \frac{\bar{a} \cdot \bar{n}}{|\bar{n}|^2} \bar{n} = (a \cdot \hat{n}) \hat{n}$$

где \hat{n} — единичный вектор в направлении n .

Окончательно, для направления отраженного луча получаем

$$\bar{r} = \bar{a} - 2(\bar{a} \cdot \hat{n}) \hat{n} \quad (1.10)$$

Пример 2. Угол между векторами \bar{a} и \bar{b} равен 60° , $|\bar{a}| = 5$, $|\bar{b}| = 8$. Найти $|\bar{a} + \bar{b}|$ и $|\bar{a} - \bar{b}|$.

Решение. Построим параллелограмм на векторах \bar{a} и \bar{b} (рис.1.22).

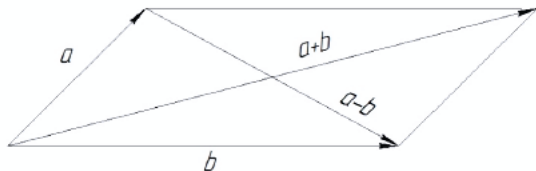


Рис. 1.22.

Воспользуемся свойствами скалярного произведения векторов. Имеем:

$$\begin{aligned} |\bar{a} + \bar{b}|^2 &= (\bar{a} + \bar{b}) \cdot (\bar{a} + \bar{b}) = (\bar{a} \cdot \bar{a}) + (\bar{a} \cdot \bar{b}) + (\bar{b} \cdot \bar{a}) + (\bar{b} \cdot \bar{b}) = \\ &= |\bar{a}|^2 + 2(\bar{a} \cdot \bar{b}) + |\bar{b}|^2 = |\bar{a}|^2 + 2|\bar{a}||\bar{b}| \cos 60^\circ + |\bar{b}|^2 \\ &= 25 + 2 \cdot 5 \cdot 8 \cdot 1/2 + 64 = 129, \end{aligned}$$

откуда $|\bar{a} + \bar{b}| = \sqrt{129}$;

$$\begin{aligned}
 |\bar{a} - \bar{b}|^2 &= (\bar{a} - \bar{b}) \cdot (\bar{a} - \bar{b}) = (\bar{a} \cdot \bar{a}) - (\bar{a} \cdot \bar{b}) + (\bar{b} \cdot \bar{a}) + (\bar{b} \cdot \bar{b}) = \\
 &= |\bar{a}|^2 - 2(\bar{a} \cdot \bar{b}) + |\bar{b}|^2 = |\bar{a}|^2 - 2|\bar{a}||\bar{b}| \cos 60^\circ + |\bar{b}|^2 = \\
 &= 25 - 2 \cdot 5 \cdot 8 \cdot 1/2 + 64 = 49
 \end{aligned}$$

откуда $|\bar{a} - \bar{b}| = 7$.

Пример 3. Даны вершины четырехугольника $ABCD$: $A(1, -2, 2)$, $B(1, 4, 0)$, $C(-4, 1, 1)$, $D(-5, -5, 3)$. Доказать, что его диагонали взаимно перпендикулярны.

Решение. Найдем координаты векторов \overline{AC} и \overline{BD} : $\overline{AC} = (-5, 3, -1)$, $\overline{BD} = (-6, -9, 3)$.

Вычислим скалярное произведение векторов \overline{AC} и \overline{BD} : $(-5)(-6) + 3(-9) + (-1)3 = 0$

Отсюда следует, что $\overline{AC} \perp \overline{BD}$.

1.2.5. Векторное уравнение прямой

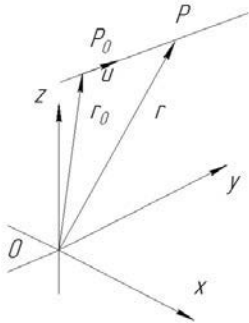


Рис. 1.23

1) Прямая, проведенная через точку P_0 вдоль единичного вектора \bar{u} (рис. 1.23), может быть представлена следующим образом. Обозначим вектор $\overline{OP_0}$ через r_0 , а вектор \overline{OP} , соединяющий точку O с произвольной точкой P , лежащей на рассматриваемой прямой, через r . Тогда $\overline{P_0P} = r - r_0$, причем направление вектора P_0P совпадает с направлением u и, тем самым, $r - r_0 = \lambda u$. Следовательно, уравнение прямой имеет вид:

$$r = r_0 + \lambda u \quad (1.11)$$

Расстояние от P_0 до P определяется соотношением

$$|r - r_0| = |\lambda u| = |\lambda|$$

Различные точки на прямой отличаются друг от друга значением параметра λ . Вектором, задающим направление прямой, может быть и не единичный вектор. Предположим, что таким вектором является некоторый вектор a ($r = r_0 + \mu a$); в этом случае параметр μ пропорционален, а не равен расстоянию от r_0 .

2) Если прямая проходит через две точки P_1 и P_2 (рис. 1.24.), то для произвольной точки P имеем уравнение:

$$r = \overline{OP} = \overline{OP_1} + \overline{P_1P} = \overline{OP_1} + \mu \overline{P_1P_2}$$

Отсюда $r = r_1 + \mu(r_2 - r_1)$, и в конечном итоге

$$r = (1 - \mu)r_1 + \mu r_2 \quad (1.12)$$

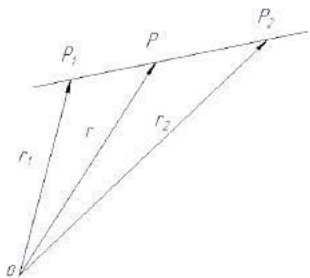
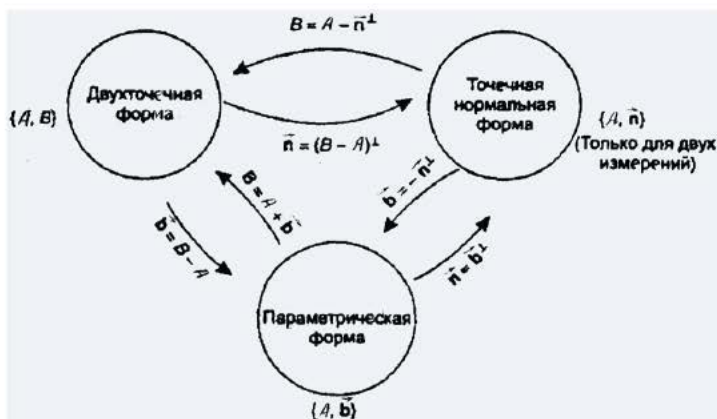


Рис.1.24

Прямая
 Двухточечная форма $\{A, B\}P = A(1-t) + Bt \Rightarrow P(t) = A(1-t) + B(t)$
 Параметрическое представление прямой.

$$L(t) = \begin{cases} x = x_A + b_x t \\ y = y_A + b_y t \end{cases}$$



Пример.

1. Прямая, проходящая через точку $(1, 1, 1)$ в направлении $(1/\sqrt{3}; -1/\sqrt{3}; 1/\sqrt{3})$, определяется соотношением: $r = (1,1,1) + \lambda \left(\frac{1}{\sqrt{3}}; -\frac{1}{\sqrt{3}}; \frac{1}{\sqrt{3}}\right)$
 $x = 1 + \frac{\lambda}{\sqrt{3}}, y = 1 - \frac{\lambda}{\sqrt{3}}, z = 1 + \frac{\lambda}{\sqrt{3}}$ выражениями или $x = 1 + \mu, y = 1 - \mu, z = 1 + \mu$, где $\mu = \frac{\lambda}{\sqrt{3}}$

2. Прямая, проходящая через точки $(1; 0; 2)$ и $(2; 1; 1)$, определяется соотношением: $r = (1 - v)(1,0,2) + v(2,1,1)$

или соотношениями

$$x = 1 + v, \quad y = v, \quad z = 2 - v$$

3. Определим, не будут ли рассмотренные выше прямые пересекаться.

Если эти прямые пересекаются, то должны удовлетворяться одновременно три уравнения:

$$1 + \mu = 1 + v \quad (1)$$

$$1 - \mu = v \quad (2)$$

$$1 + \mu = 2 - v \quad (3).$$

Из уравнений (1) и (3) вытекает, что $1 + v = 2 - v$, следовательно, $v = 1/2$. Подставляя $v = 1/2$ в уравнение (1), мы получаем $\mu = 1/2$. Эти значения удовлетворяют и уравнению (2), следовательно, две прямые пересекаются в точке $(3/2, 1/2, 3/2)$.

1.2.6. Векторное произведение

Векторным произведением векторов \vec{a} и \vec{b} называется вектор, обозначаемый $\vec{c} = [\vec{a}, \vec{b}]$ или $\vec{c} = \vec{a} \times \vec{b}$ и удовлетворяющий следующим трем условиям:

- 1) длина вектора $[\vec{a}, \vec{b}]$ равна площади параллелограмма, построенного на векторах \vec{a} и \vec{b} , т. е. $|\vec{c}| = |\vec{a}||\vec{b}| \sin(\angle \vec{a}, \vec{b})$
- 2) вектор $[\vec{a}, \vec{b}]$ перпендикулярен к плоскости векторов \vec{a} и \vec{b} ;
- 3) упорядоченная тройка $\vec{a}, \vec{b}, [\vec{a}, \vec{b}]$ — правая.

Упорядоченная тройка некопланарных векторов $\vec{a}, \vec{b}, \vec{c}$ называется правой, если наблюдателю, находящемуся внутри телесного угла, образованного этими векторами, кратчайшие повороты от \vec{a} к \vec{b} и от \vec{b} к \vec{c} кажутся происходящими против хода часовой стрелки. В противном случае данная тройка называется левой.

Из определения векторного произведения следует, что если $\vec{a}||\vec{b}$, то $[\vec{a}, \vec{b}] = 0$.

Перечислим основные свойства векторного произведения векторов:

- 1) $[\vec{a}, \vec{b}] = -[\vec{b}, \vec{a}]$
- 2) $[\lambda\vec{a}, \vec{b}] = \lambda[\vec{a}, \vec{b}], \lambda \in R$;
- 3) $[\vec{a} + \vec{b}, \vec{c}] = [\vec{a}, \vec{c}] + [\vec{b}, \vec{c}]$.

Если $\vec{a} = (a_x, a_y, a_z)$, $\vec{b} = (b_x, b_y, b_z)$, то вектор $[\vec{a}, \vec{b}]$ в координатной форме записывают так:

$$[\vec{a}, \vec{b}] = \begin{vmatrix} i & j & k \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = (a_y b_z - b_y a_z)\vec{i} + (b_x a_z - a_x b_z)\vec{j} + (a_x b_y - b_x a_y)\vec{k}$$

Во многих геометрических задачах, а также в задачах по механике приводится находить вектор, перпендикулярный двум заданным векторам. Например,

нам нужно провести перпендикуляр к двум скрещивающимся прямым. Если вектор \bar{v} перпендикулярен $\bar{a}\bar{b}$, то $\bar{v} \cdot \bar{a} = 0$, $\bar{v} \cdot \bar{b} = 0$, и, следовательно,

$$\begin{aligned} a_1 v_1 + a_2 v_2 + a_3 v_3 &= 0 \\ b_1 v_1 + b_2 v_2 + b_3 v_3 &= 0 \end{aligned}$$

Решая эти уравнения для отношений $v_1 : v_2 : v_3$, мы получаем

$$\frac{v_1}{a_2 b_3 - a_3 b_2} = \frac{v_2}{a_3 b_1 - a_1 b_3} = \frac{v_3}{a_1 b_2 - a_2 b_1}$$

Таким образом, вектор $(a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)$ пропорционален \bar{v} и, следовательно, перпендикулярен $\bar{a}\bar{b}$. Этот вектор обозначается через $\bar{a} \times \bar{b}$ и называется *векторным произведением* векторов $\bar{a}\bar{b}$. Итак,

$$\bar{a} \times \bar{b} = (a_2 b_3 - a_3 b_2)\bar{i} + (a_3 b_1 - a_1 b_3)\bar{j} + (a_1 b_2 - a_2 b_1)\bar{k}$$

что можно также записать в виде детерминанта

$$\bar{a} \times \bar{b} = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad (1.13)$$

Найдем модуль вектора $\bar{a} \times \bar{b}$. Из матрицы следует:

$$\begin{aligned} |\bar{a} \times \bar{b}|^2 &= (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2 = \\ &= a_2^2 b_3^2 + a_3^2 b_2^2 + a_3^2 b_1^2 - a_1^2 b_3^2 + a_1^2 b_2^2 - a_2^2 b_1^2 - 2a_2 a_3 b_2 b_3 - \\ &\quad - 2a_3 a_1 b_3 b_1 - 2a_1 a_2 b_1 b_2 = \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - a_1^2 b_1^2 - a_2^2 b_2^2 - a_3^2 b_3^2 - \\ &\quad - 2a_2 a_3 b_2 b_3 - 2a_3 a_1 b_3 b_1 - 2a_1 a_2 b_1 b_2 = |a|^2 |b|^2 - \\ &\quad - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2 = |a|^2 |b|^2 - |a|^2 |b|^2 \cos \theta \end{aligned}$$

Отсюда $|\bar{a} \times \bar{b}|^2 = |a|^2 |b|^2 \cos \theta$. Поэтому $|\bar{a} \times \bar{b}| = |\bar{a}| |\bar{b}| \sin \theta$, если $0 < \theta < 180^\circ$. Таким образом, $\bar{a} \times \bar{b}$ - вектор, перпендикулярный \bar{a} и \bar{b} , длина которого равна $|\bar{a}| |\bar{b}| \sin \theta$. Предположим, что в выбранной нами системе координат $\bar{a} = a\bar{i}$ ($a > 0$) и $\bar{b} = b_1\bar{i} + b_2\bar{j}$. Тогда

$$\bar{a} \times \bar{b} = \begin{vmatrix} i & j & k \\ a_1 & 0 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = (0; 0; ab_2).$$

Откуда следует, что при $b_2 > 0$ направление вектора $\bar{a} \times \bar{b}$ совпадает с направлением положительной полуоси z ; если же $b_2 < 0$ (рис. 1.25 и 1.26), направление вектора $\bar{a} \times \bar{b}$ совпадает с направлением отрицательной полуоси z , причем вектор $\bar{a} \times \bar{b}$ всегда направлен в сторону, противоположную наблюдателю, который видит вращение от \bar{a} к \bar{b} (для $0 < \theta < 180^\circ$) совершающимся по

часовой стрелке. Именно в этом направлении движется винт с правосторонней резьбой, вращающийся по часовой стрелке.

Резюмируем свойства векторного произведения. Пусть θ - наименьший угол, образованный векторами \vec{a} и \vec{b} , $0 < \theta < 180^\circ$; тогда $\vec{a} \times \vec{b} = \overline{ab} \sin \theta \vec{u}$,

где \vec{u} - единичный вектор, перпендикулярный плоскости, в которой лежат векторы \vec{a} и \vec{b} , направление которого совпадает с направлением винта с правосторонней резьбой, вращающегося от \vec{a} к \vec{b} через угол θ (рис. 1.27).

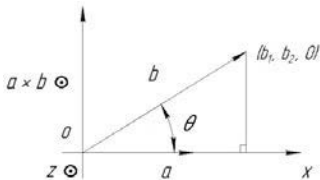


Рис. 1.25

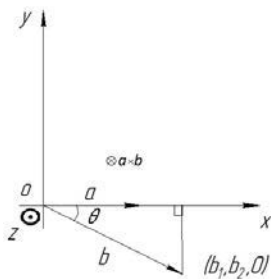


Рис. 1.26

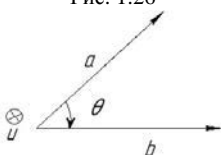


Рис. 1.27

Следует заметить, что если \vec{a} и \vec{b} параллельные векторы, то, поскольку в этом случае $\theta = 0$, получаем $\vec{a} \times \vec{b} = 0$. В частности, $\vec{a} \times \vec{a} = 0$. Но соотношение $\vec{a} \times \vec{b} = 0$ имеет место и в том случае, когда либо \vec{a} , либо \vec{b} равны нулю; следовательно, имея соотношение $\vec{a} \times \vec{b} = 0$, нельзя сделать вывод, что векторы \vec{a} и \vec{b} параллельны. Перечислим остальные алгебраические свойства векторного произведения. Пусть $\vec{a}, \vec{b}, \vec{c}$ - произвольные векторы, а λ - какой-нибудь скаляр; тогда

1. $b \times a = -a \times b$ (поскольку изменение направления вращения приводит к противоположной ориентации вектора);

2. $a \times (b + c) = a \times b + b \times c$
3. $(\lambda \times a) \times b = a \times (\lambda \times b) = \lambda \times (a \times b)$
4. $\bar{i} \times \bar{j} = \bar{k}, \bar{j} \times \bar{k} = \bar{i}, \bar{k} \times \bar{i} = \bar{j}$

Приложения векторного произведения

Пример 1. Вектор площади треугольника (рис. 1.28). Площади треугольника можно придать векторный смысл, и тогда формула площади треугольника будет иметь следующий вид: $S_{ABC} = 0,5AC \cdot Bb = 0,5AC \cdot AB \sin \theta = 0,5|ab|$.

Вектор площади $A = 0,5|ab|$ часто оказывается полезным.

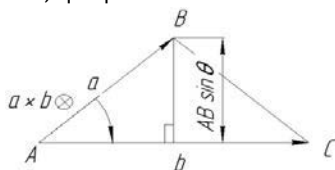


Рис. 1.28.

Пример 2. Общий перпендикуляр к двум скрещивающимся прямым.

Чтобы найти длину общего перпендикуляра Q_1Q_2 к прямым P_1R_1 и P_2R_2 (рис. 1.29),

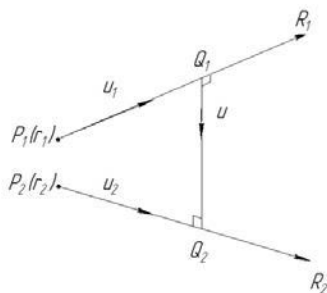


Рис. 1.29.

обозначим направления прямых P_1R_1 и P_2R_2 через единичные векторы \bar{u}_1 и \bar{u}_2 , а радиус-векторы точек P_1 и P_2 - через \bar{r}_1 и \bar{r}_2 соответственно; пусть направляющим вектором перпендикуляра $\overline{Q_1Q_2}$ будет единичный вектор \bar{u} (неизвестный). Тогда:

$$\begin{aligned} \overline{Q_1Q_2} &= \bar{r}_1 + P_1Q_1\bar{u}_1 = \bar{r}_2 + P_2Q_2\bar{u}_2 - Q_1Q_2\bar{u} \\ \bar{r}_1\bar{u} + P_1Q_1\bar{u}_1\bar{u} &= \bar{r}_2\bar{u} + P_2Q_2\bar{u}_2\bar{u} - Q_1Q_2 \end{aligned}$$

Но, поскольку Q_1Q_2 есть перпендикуляр к P_1R_1 и P_2R_2 , произведения $\bar{u}_1 \cdot \bar{u}$ и $\bar{u}_2 \cdot \bar{u}$ равны нулю и, следовательно, $Q_1Q_2 = (\bar{r}_2 - \bar{r}_1)\bar{u}$. Кроме того,

поскольку \bar{u} перпендикулярен \bar{u}_1 и \bar{u}_2 , он пропорционален произведению $\bar{u}_1 \times \bar{u}_2$. Поскольку длина отрезка $Q_1 Q_2$ должна быть величиной положительной, то

$$Q_1 Q_2 = \left| \frac{(\bar{r}_2 - \bar{r}_1) \cdot (\bar{u}_2 \times \bar{u}_1)}{|\bar{u}_1 \times \bar{u}_2|} \right|$$

1.2.7. Векторное уравнение плоскости

В трехмерной графике интенсивно используются различные многоугольники. Поэтому создается впечатление, что плоскости появляются на каждом шагу. Многоугольник (грань объекта) располагается в определенной плоскости и часто приходится отсекалть объекты – относительно плоскостей или находить плоскость, в которой располагается определенная грань.

Подобно прямым, плоскость может задаваться тремя основными формами: трехточечной формой, параметрическим или нормальным представлением.

Вектор \bar{r} , задающий положение произвольной точки $P(\bar{r})$, лежащей в плоскости π , которая проходит через заданную точку $P_0(\bar{r}_0)$ перпендикулярно вектору \bar{u} , можно определить, исходя из того факта, что вектор $\bar{r} - \bar{r}_0$ лежит в плоскости π , и, следовательно, он перпендикулярен вектору \bar{u} (рис. 1.30).

Таким образом,

$$(\bar{r} - \bar{r}_0) \cdot \bar{u} = PP_0 \cos \pi/2 = 0,$$

и уравнение плоскости имеет вид:

$$\bar{r} \cdot \bar{u} = \bar{r}_0 \cdot \bar{u} = p \quad (1.14)$$

где p — расстояние от начала координат до плоскости π , вычисляемое по длине перпендикуляра, опущенного из начала координат на плоскость.

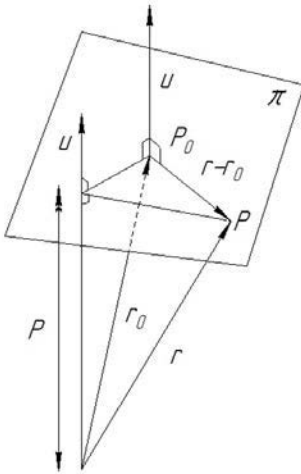


Рис.1.30

Необходимо заметить следующее: поскольку \bar{u} входит в обе части уравнения, то вместо \bar{u} , если это более удобно, можно брать любое его кратное. Например, плоскость, нормаль которой имеет то же направление, что и вектор $(1; 1; 0)$, и которая проходит через точку $(1; 0; 0)$, задается соотношением

$$(x, y, z) \cdot (1, 1, 0) = (1, 0, 0) \cdot (1, 1, 0),$$

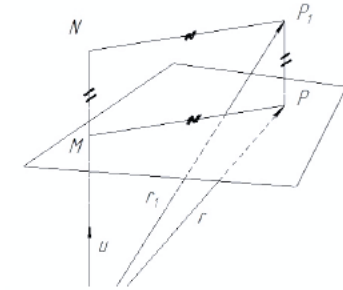


Рис.1.31

или, что то же самое, соотношением $x + y = 1$. (Заметим, что уравнение $\bar{r} \cdot \bar{u} = p$ будет всегда линейным уравнением от x, y, z .) Точка \bar{r} , являющаяся ортогональной проекцией точки $P_1(\bar{r})$ на плоскость $r \cdot u = p$ (рис. 1.31), определяется соотношением:

$$\bar{r}_1 - \bar{r}_0 = NM\bar{u},$$

отсюда $\bar{r}_1 - \bar{r}_0 = (ON - OM)\bar{u} = \bar{r}_1 + (\bar{r} \cdot \bar{u} - \bar{r}_1 \cdot \bar{u})\bar{u}$ и поэтому $\bar{r}_1 = \bar{r}_1 + (p - \bar{r}_1 \cdot \bar{u})\bar{u}$. Этот результат используется при определении параллельной проекции трехмерных объектов

Параметрическое представление плоскости

Параметрическая форма задания плоскости строится на трех составляющих: одной из ее точек (например c), и двух (непараллельных) векторах, \bar{a} и \bar{b} , расположенных в данной плоскости (рис.1.32). Если заданы три (неколлинеарные, не лежащие на одной прямой) точки A, B и C плоскости, то в качестве векторов можно взять:

$$\bar{a} = A - C \text{ и } \bar{b} = B - C.$$

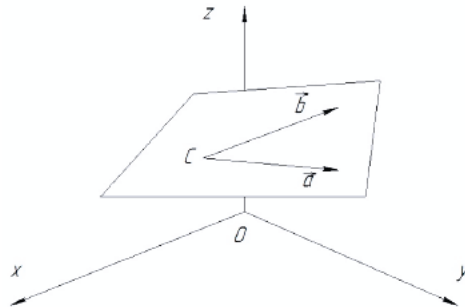


Рис.1.32.

Отметим, что любая точка на данной плоскости может быть представлена в виде следующей векторной суммы: точка c плюс некоторая скалярная величина кратная вектору \bar{a} и плюс еще один скаляр t , кратный вектору \bar{b} .

$$P(s, t) = c + \bar{a}s + \bar{b}t \quad (1.15)$$

Пример: $s = t = 0$. Тогда точка p совпадает с точкой c .

$$s = 2, t = -3. p(2, -3) = c + \vec{a} - 3\vec{b}.$$

Заметим, что для параметрического задания поверхности (в том числе и плоскости) необходимо два параметра, а для задания кривой линии (в том числе и прямой) достаточно одного параметра.

Действительно, зафиксируем один параметр (например, $s = 5$), тогда $P(5, t)$ – функция одной переменной, представляющая собой прямую

$$P(5, t) = (c + 5\vec{a}) + \vec{b}t.$$

Часто удобнее представлять параметрическую форму в виде, когда ее компоненты становятся явными. Этого можно добиться, расписывая компоненты по проекциям:

$$P(s, t) = (c_x + a_x s + b_x t, c_y + a_y s + b_y t, c_z + a_z s + b_z t) \quad (1.16)$$

Можно переписать уравнение (1) в явном виде в терминах заданных точек A, B и C :

$$P(s, t) = c + s(A - C) + t(B - C). \quad \dots\dots\dots (1.17)$$

Равенство (s) может быть перегруппировано в аффинную комбинацию точек:

$$P(s, t) = sA + tB + (1 - s - t)C \quad \dots\dots\dots (1.18)$$

Пример. Рассмотрим плоскость, проходящую через точки $A(5,3,2), B(1,1,4)C(2,7,3)$

Из(3) имеем: $P(s, t) = (2,7,3) + (3, -4, -1)s + (-1, -6,1)t$

(2) $\Rightarrow P(s, t) = (2 + 3s - t)\vec{i} + (7 - 4s - 6t)\vec{j} + (3 - s - t)\vec{k}$

(4) $\Rightarrow P(s, t) = s(5,3,2) + t(1,1,4) + (1 - s - t)(2,7,3).$

Точечная нормальная форма уравнения плоскости.

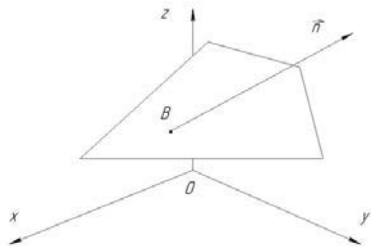


Рис. 1.33

Уравнение плоскости может быть представлено в точечной нормальной форме (откуда легко перейти к классическому уравнению для плоскости)

Плоскость полностью определяется (рис. 1.33):

- одной точкой $B(b_x, b_y, b_z)$ лежащей в ней;
- направлением $\vec{n}(n_x, n_y, n_z)$, перпендикулярным (нормальным) к этой плоскости.

Нормаль к плоскости задает ориентацию этой плоскости в пространстве. Для любой точки $R = (x, y, z)$ плоскости вектор, проведенный из точки R к точке B , задающей плоскость, должен быть перпендикулярен к нормали \vec{n} , откуда

$$\vec{n}(R - B) = 0 \quad (1.19)$$

Это и есть точечное нормальное уравнение плоскости.

Все точки плоскости образуют с точкой B векторы $(R - B)$, составляющие аналогичное скалярное произведение с нормальным вектором \vec{n} . Раскрывая скалярное произведение и используя выражение $\vec{n} = (n_x, n_y, n_z)$ преобразуем нормальную форму в общее уравнение плоскости:

$$\begin{aligned} n_x x + n_y y + n_z z &= D \\ D &= \vec{n} \cdot (B - O). \end{aligned} \quad (1.20)$$

Пример. Из уравнения плоскости $5x - 2y + 8z = 2$ видим, что нормалью к этой плоскости будет вектор $(5, -2, 8)$ или любое скалярное, кратное ему. (Точка на плоскости должна удовлетворять уравнению 1.19).

Уравнение (5) можно представить в несколько ином виде (рис. 1.34).

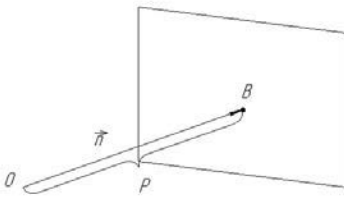


Рис. 1.34

Пусть вектор \vec{n} - перпендикуляр, опущенный из начала координат точки $O(0,0,0)$ на плоскость. Очевидно, что это по-прежнему нормаль к рассматриваемой плоскости. Тогда модуль вектора \vec{n} показывает расстояние от начала координат O до плоскости. Обозначим $|\vec{n}| = P$.

Следовательно, положение любой плоскости определяется ее расстоянием p от начала координат и направлением вектора нормали \vec{n} . Тогда уравнение (1.19) можно записать в виде $\vec{n}\vec{R} - p = 0$, где \vec{R} — радиус вектор текущей точки R плоскости, в координатной форме.

$$x \cos \alpha + y \cos \beta + z \cos \gamma - P = 0,$$

$\cos \alpha = \frac{n_x}{a}, \cos \beta = \frac{n_y}{a}, \cos \gamma = \frac{n_z}{a}, d = \sqrt{n_x^2 + n_y^2 + n_z^2}$ направляющие косинусы единичного вектора $\vec{e}_n(\cos \alpha, \cos \beta, \cos \gamma)$ в направлении нормали $\vec{n}(n_x, n_y, n_z)$. Причем $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$. расстояние ρ от точки $M_o(x_o, y_o, z_o)$ до плоскости равно

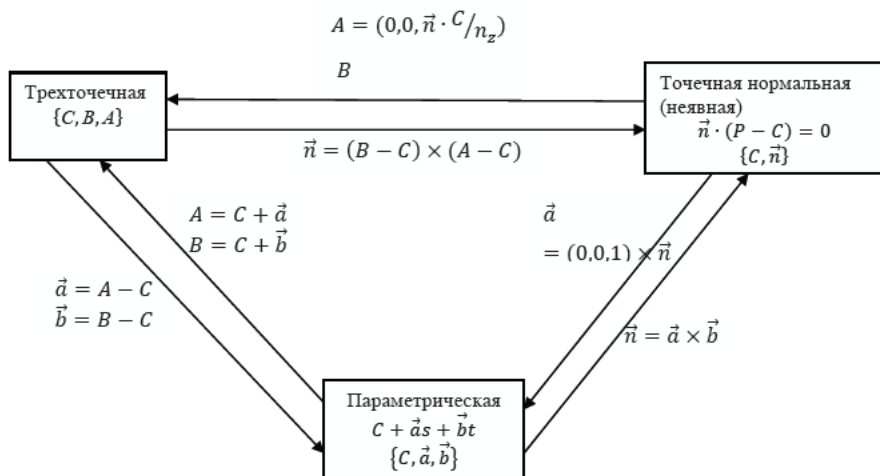
$$\rho = x_o \cos \alpha + y_o \cos \beta + z_o \cos \gamma - p.$$

Знак « - » перед ρ показывает, что точка M_o и начало координат точка O лежат по одну сторону от данной плоскости, знак « + » показывает, что плоскость разделяет точки M_o и O .

Переход от одного представления к другому

Представляется полезным уметь осуществлять переходы между тремя представлениями плоскости и переводить данные, описывающие плоскость, в форму, наиболее подходящую для данной задачи.

Три формы представления плоскости.



Точка C плоскости является общей для всех трех форм. Самым сложным является нахождение двух неколлинеарных (не параллельных) векторов. Рассмотрим его подробнее на конкретном примере.

Задана нормаль \vec{n} к плоскости. Необходимо найти два неколлинеарных вектора \vec{a} и \vec{b} , каждый из которых перпендикулярен \vec{n} . Для этого используем то обстоятельство, что векторное произведение любого вектора на \vec{n} перпендику-

ляно к \vec{n} . Выберем простейший вектор, например $(0,0,1)$, и постоим \vec{a} как векторное произведение этого вектора на \vec{n} :

$$\vec{a} = (0,0,1) \times \vec{n} = (0 \cdot n_z - 1 \cdot n_y)\vec{i} + (1 \cdot n_x - 0 \cdot n_z)\vec{j} + (0 \cdot n_y - 0 \cdot n_x)\vec{k} = -n_y\vec{i} + n_x\vec{j} = (-n_y, n_x, 0)/$$

Используем эту же идею для формирования вектора \vec{b} , нормального к векторам \vec{n} и \vec{a} :

$$\begin{aligned} \vec{b} &= \vec{n} \times \vec{a} = (n_y \cdot 0 - n_z \cdot n_x)\vec{i} + (n_z(-n_y) - 0 \cdot n_x)\vec{j} + \\ &+ (n_x \cdot n_x - n_y(-n_y))\vec{k} = -n_x n_z \vec{i} - n_y n_z \vec{j} + (n_x^2 + n_y^2)\vec{k} = \\ &= (-n_x n_z, -n_y n_z, n_x^2 + n_y^2) \end{aligned}$$

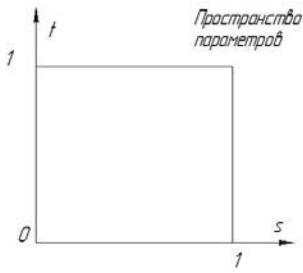
Вектор \vec{b} безусловно неколлинеарен вектору \vec{a} .

Другой случай, требующий пояснений, — это нахождение трех точек плоскости, заданной в нормальной точечной форме. Одна точка, C , уже известна. Оставшиеся две находим при помощи конкретных значений самой точечной нормальной формы, являющейся уравнением $n_x x + n_y y + n_z z = \vec{n}C$. Выберем для простоты $A(0,0, a_z)$ и используем это уравнение для нахождения $a_z = \vec{n} \cdot \frac{C}{n_z}$. Аналогично выберем $B(0, b_y, 0)$ и опять применим это же уравнение для нахождения $b_y = \vec{n} \cdot \frac{C}{n_y}$.

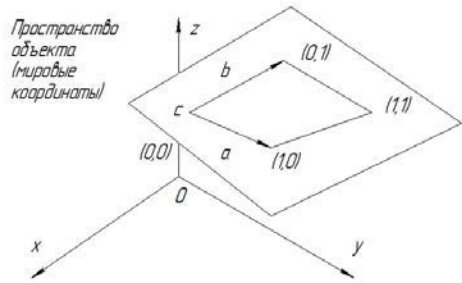
Плоские лоскуты.

В параметрической форме уравнения () значения s и t могут изменяться от $-\infty$ до $+\infty$, поэтому плоскость простирается неограниченно. В машинной графике приходится иметь дело только с «куском» плоскости, например, с расположенным в этой плоскости параллелограммом. Такой кусок называется плоским лоскутом. Позднее будут рассмотрены криволинейные поверхности, сшитые из лоскутов, которые не обязательно будут плоскими. Большая часть практического моделирования объемных тел сводится к собиранию в одно целое лоскутов различной формы для создания поверхности объекта.

Плоский лоскут формируется ограничением диапазона допустимых значений параметров s и t . Как правило требуется, чтобы s и t изменялись в диапазоне от 0 до 1. Лоскут позиционируется и ориентируется в пространстве с помощью соответствующего подбора векторов \vec{a} и \vec{b} и точки C . На рис. 1.35а, показана область допустимых значений s и t в виде квадрата в пространстве параметров, а на рис. 1.35б, изображен лоскут, полученный при таком ограничении в пространстве объекта.



а



б

Рис. 1.35

Каждой точке (s, t) в пространстве параметров соответствует одна трехмерная точка на лоскуте $P(s, t) = C + \vec{a}s + \vec{b}t$. В общем случае такой лоскут является параллелограммом, углы которого соответствуют четырем углам в пространстве параметров:

$$P(0,0) = C,$$

$$P(1,0) = C + \vec{a}$$

$$P(0,1) = C + \vec{b}$$

$$P(1,1) = C + \vec{a} + \vec{b}.$$

Векторы \vec{a} и \vec{b} определяют размер и ориентацию лоскута. Если векторы \vec{a} и \vec{b} взаимно перпендикулярны, то лоскут (и сетка на нем) становится прямоугольным. Если \vec{a} и \vec{b} еще и одинаковой длины, то лоскут становится квадратом. Изменение точки просто сдвигает лоскут без изменения формы и ориентации.

Определение точки пересечения двух отрезков.

В компьютерной графике часто возникает следующая задача:

Дано: два отрезка прямой AB и CD .

Определить, пересекаются ли они, и если пересекаются, то найти точку их пересечения.

На рис.1.36 изображены возможные способы взаимного расположения отрезков AB и CD : (рис.1.36 а,б) – они не пересекаются; (в,г) – имеют одну общую точку; (д) – перекрываются в некоторой области.

Каждый отрезок имеет порождающую прямую – бесконечную линию частью которой он является. Если две порождающие прямые не параллельны, то они пересекаются в некоторой точке. Определим ее.

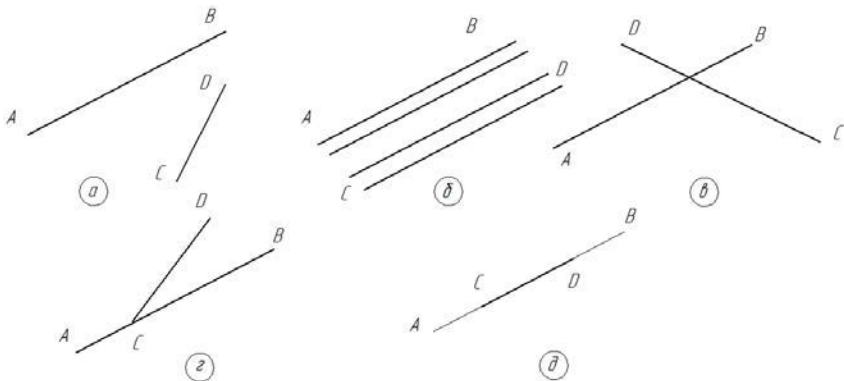


Рис. 1.36

Введем параметрическое представление для каждой из исследуемых отрезков.

$(AB) = B - A \Rightarrow AB(t) = A + \vec{b}t$ где $\vec{b} = B - A$ на отрезке AB $0 \leq t \leq 1$, $(-\infty < +\infty)$

$(CD) = D - C \Rightarrow CD(u) = C + \vec{d}u$, где $\vec{d} = D - C$, $0 \leq u \leq 1$, $-\infty < +\infty$

В случае пересечения порождающих прямых существуют определенные значения t и u , при которых правые части уравнений равны

$$A + \vec{b}t = C + \vec{d}t$$

Для удобства введем вектор $\vec{c} = C - A$. Тогда,

$$\vec{b}t = \vec{c} + \vec{d}u$$

Чтобы решить это уравнение, скалярно умножим обе его части на \vec{d}^\perp , чтобы исключить члены с \vec{d} (вспомним $\vec{d} \cdot \vec{d}^\perp = 0$).

В результате получим:

$$\vec{d}^\perp \cdot \vec{b}t = \vec{d}^\perp \cdot \vec{c}$$

Случай 1. Произведение $\vec{d}^\perp \vec{b}$ отлично от нуля.

$$\vec{d}^\perp \cdot \vec{b} \neq 0 \Rightarrow t \frac{\vec{d}^\perp \vec{c}}{\vec{d}^\perp \vec{b}}$$

$$\vec{b}^\perp \cdot \vec{b} \cdot t = \vec{b}^\perp \vec{c} + \vec{b}^\perp \vec{d}u \Rightarrow \vec{b}^\perp \vec{d} = -\vec{d}^\perp \vec{b} \Rightarrow u = \frac{\vec{b}^\perp \vec{c}}{\vec{d}^\perp \vec{b}}$$

Отрезки AB и CD пересекутся в точке I , если значения t и u лежат в интервале $[0,1]$:

$$I = A + \left(\frac{\vec{d}^\perp \cdot \vec{c}}{\vec{d}^\perp \cdot \vec{b}} \right) \cdot \vec{b} \quad (\text{точка пересечения}).$$

Случай 2 произведение $\vec{d}^\perp \vec{b}$ равно 0.

В этом случае векторы \vec{d} и \vec{b} параллельны (т.к. $\vec{d}^\perp \perp \vec{b}$). Порождающие прямые параллельны, но они могут быть идентичными, тогда отрезки могут перекрываться.

1.2.8. Смешанное произведение

Смешанным произведением векторов $\vec{a}, \vec{b}, \vec{c}$ называется число, обозначаемое $(\vec{a}, \vec{b}, \vec{c})$ и определяемое как скалярное произведение вектора

$$[a, b] \text{ и вектора } \vec{c}: (a, b, c)(\vec{a}, \vec{b}, \vec{c}) = ([a, b] \cdot c)$$

Приведем геометрические свойства смешанного произведения векторов:

1) если V — объем параллелепипеда, построенного на векторах $\vec{a}, \vec{b}, \vec{c}$, то

$(\vec{a}, \vec{b}, \vec{c}) = V$, если тройка $\vec{a}, \vec{b}, \vec{c}$ — правая, и $-V$, если тройка $\vec{a}, \vec{b}, \vec{c}$ — левая.

2) для того чтобы векторы $\vec{a}, \vec{b}, \vec{c}$ были компланарны, необходимо и достаточно выполнение условия $(\vec{a}, \vec{b}, \vec{c}) = 0$.

Перечислим алгебраические свойства смешанного произведения векторов:

- 1) $(\vec{a}, \vec{b}, \vec{c}) = (\vec{b}, \vec{a}, \vec{c}) = (\vec{c}, \vec{a}, \vec{b})$;
- 2) $(\vec{b}, \vec{a}, \vec{c}) = -(\vec{a}, \vec{b}, \vec{c})$; $(\vec{c}, \vec{b}, \vec{a}) = -(\vec{a}, \vec{b}, \vec{c})$; $(\vec{a}, \vec{c}, \vec{b}) = -(\vec{a}, \vec{b}, \vec{c})$;
- 3) $(\alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2, \vec{b}, \vec{c}) = \alpha_1 (\vec{a}_1, \vec{b}, \vec{c}) + \alpha_2 (\vec{a}_2, \vec{b}, \vec{c})$, α_1 и $\alpha_2 \in R$;
- 4) $(i, j, k) = 0$.

Если $\vec{a} = (a_x, a_y, a_z)$, $\vec{b} = (b_x, b_y, b_z)$, $\vec{c} = (c_x, c_y, c_z)$, то

$$(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}$$

Таким образом, циклические перестановки символов не изменяют значения смешанного произведения.

Смешанное произведение, содержащее два равных или параллельных вектора, равно нулю.

Пример 1. Если $\vec{a} = (1; -1; 0)$, $\vec{b} = (2; 0; 1)$, а $\vec{c} = (3; 2; 4)$, то

$$a \cdot (b \times c) = \begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 3 & 2 & 4 \end{vmatrix} = -2 - 2(-4) + 3(-1) = -2 + 8 - 3 = 3$$

Пример 2. Уравнение плоскости, проходящей через концы радиус-векторов \bar{r}_0, \bar{r}_1 , и \bar{r}_2 , можно получить, заметив, что $\bar{r}_1 - \bar{r}_0$ и $\bar{r}_2 - \bar{r}_0$ суть векторы, лежащие в этой плоскости. Тогда вектор $(r_1 - r_0) \times (r_2 - r_0)$ нормален к ней, и можно воспользоваться уравнением (6), заметив, что единичный вектор нормали и пропорционален $(r_1 - r_0) \times (r_2 - r_0)$, так что

$$r((r_1 - r_0) \times (r_2 - r_0)) = r_3 \cdot (r_1 - r_0) \times (r_2 - r_0)$$

Наконец, заметим, что большинство членов правой части равно нулю, так как они содержат вектор r_0 дважды. Следовательно, уравнение плоскости имеет вид

$$r((r_1 - r_0) \times (r_2 - r_0)) = r_0 \cdot (r_1 \times r_2)$$

Примеры:

1. Доказать, что $|[a + b], [a - b]| = 2|[b, a]|$. Выяснить геометрический смысл этого равенства.

Решение. Воспользуемся свойствами векторного произведения:

$$|[a + b], [a - b]| = [a + b, a] - [a + b, b] = [a, a] + [b, a] - [a, b] - [b, b].$$

Так как $[a, a] = |a|^2 \sin 0 = 0$, $[b, b] = 0$, $[a, b] = -[b, a]$, имеем

$$[a + b, a - b] = 2[b, a], \quad |[a + b, a - b]| = 2|[b, a]|.$$

Поскольку $|[b, a]|$ равен площади параллелограмма, построенного на векторах \bar{a} и \bar{b} (рис.1.37), приведенных к общему началу, а $|[a + b, a - b]|$ — площади параллелограмма, построенного на векторах $a + b$, $a - b$, то из доказанного равенства следует, что площадь второго параллелограмма в два раза больше площади первого.

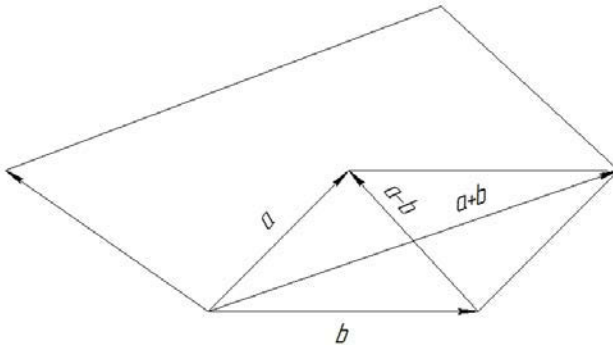


Рис. 1.37

2. Даны вершины треугольника $A(1, -1, 2)$, $B(5, -6, 2)$, $C(1, 3, -1)$. Определить длину высоты, опущенной из B на сторону AC , и площадь треугольника ABC (рис.1.38).

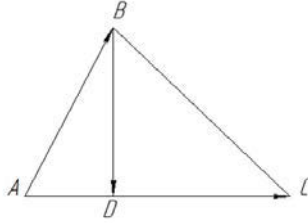


Рис. 1.38

Решение. Требуемую площадь вычислим по формуле:

$$S_{\Delta ABC} = \frac{1}{2} |[\overline{AB}, \overline{AC}]|$$

Найдем координаты векторов \overline{AB} и \overline{AC} и длину вектора \overline{AC} . Имеем:

$$\overline{AB} = (4, -5, 0), \overline{AC} = (0, 4, -3) \quad |\overline{AC}| = \sqrt{16+9} = 5$$

$$\begin{aligned} S_{\Delta ABC} &= \frac{1}{2} |[\overline{AB}, \overline{AC}]| = \frac{1}{2} \operatorname{mod} \begin{vmatrix} j & j & k \\ 4 & -5 & 0 \\ 0 & 4 & -3 \end{vmatrix} = \frac{1}{2} |15i + 12j + 16k| = \\ &= \frac{1}{2} \sqrt{225 + 144 + 256} = \frac{25}{2} = 12,5 \end{aligned}$$

Здесь $\operatorname{mod}|A|$ — модуль определителя матрицы A . Но, так как $S_{\Delta ABC} = |\overline{BD}||\overline{AC}|/2$, то $|\overline{BD}| = 2S/|\overline{AC}| = 25/5 = 5$.

3. Найти вектор x , если известно, что он перпендикулярен к векторам $a_1 = 2, -3, 1$, $a_2 = (1, -2, 3)$ и удовлетворяет условию $(x, i + 2j - 7k) = 10$.

Решение. Так как вектор x перпендикулярен к плоскости векторов a_1 и a_2 , а вектор $[a_1 a_2]$ также перпендикулярен к плоскости этих векторов по определению, то отсюда следует, что $x || [a_1 a_2]$. Имеем: $[a_1 a_2] = \begin{vmatrix} i & j & k \\ 2 & -3 & 1 \\ 1 & -2 & 3 \end{vmatrix} =$

$$= -7i - 5j - k = (-7; -5; -1)$$

Так как $x || [a_1 a_2]$, то координаты этих векторов пропорциональны, т. е. $x = (-7\lambda; -5\lambda; -\lambda)$, $\lambda \in R$. Обозначим $d = (1, 2, -7)$. По условию, $(x, d) = 10$, откуда $-7\lambda - 10\lambda + 7\lambda = 10$; $-10\lambda = 10$, $\lambda = -1$. Тогда $x = (7, 5, 1)$.

4. В пространстве даны четыре точки: $A(1, 1, 1)$, $B(4, 4, 4)$, $C(3, 5, 5)$, $D(2, 4, 7)$. Найти объем тетраэдра $ABCD$.

Решение. Найдем координаты векторов

$$\overline{AB}, \overline{AC}, \overline{AD}: \overline{AB} = (3, 3, 3), \overline{AC} = (2, 4, 4), \overline{AD} = (1, 3, 6).$$

Объем тетраэдра, построенного на векторах $\overline{AB}, \overline{AC}, \overline{AD}$, составляет $1/6$ часть объема параллелепипеда, построенного на тех же векторах:

$$V = \frac{1}{6} |\overline{AB}, \overline{AC}, \overline{AD}| = \frac{1}{6} \operatorname{mod} \begin{vmatrix} 3 & 3 & 3 \\ 2 & 4 & 4 \\ 1 & 3 & 6 \end{vmatrix} = \frac{1}{6} 18 = 3$$

5. Доказать, что четыре точки $A(1, 2, -1), B(0, 1, 5), C(-1, 2, 1), D(2, 1, 3)$ лежат в одной плоскости.

Решение. Достаточно взять три вектора $\overline{AB}, \overline{AC}, \overline{AD}$ (рис. 1.39) и доказать, что они компланарны, т. е. $(\overline{AB}, \overline{AC}, \overline{AD}) = 0$. Так как $\overline{AB} = (-1, -1, 6)$,

$$\begin{aligned} \overline{AC} = (-2, 0, 2), \quad \overline{AD} = (1, -1, 4), \quad \text{то} \quad (\overline{AB}, \overline{AC}, \overline{AD}) &= \begin{vmatrix} -1 & -1 & 6 \\ -2 & 0 & 2 \\ 1 & -1 & 4 \end{vmatrix} = \\ &= 0 - 2 + 12 - 2 - 8 = 0 \end{aligned}$$

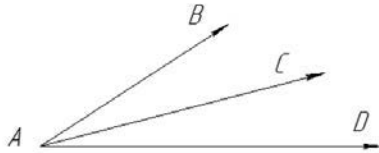


Рис. 1.39

6. Найти длину высоты пирамиды, опущенной на грань BCD , если известно, что ее вершинами являются точки $A(0, 0, 1), B(2, 3, 5), C(6, 2, 3), D(3, 7, 2)$ (рис.1.40).

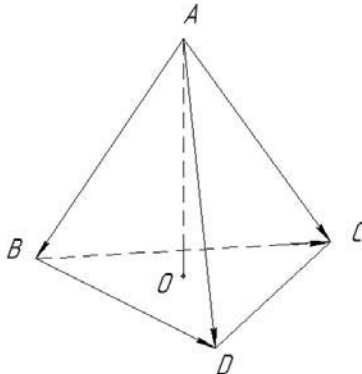


Рис. 1.40

Решение. Так как $V = \frac{1}{3} S_{\Delta BCD} AO$, то $AO = 3V/S$.

$$\text{Имеем: } \frac{1}{6} |(\overline{AB}, \overline{AC}, \overline{AD})| = \frac{1}{6} \operatorname{mod} \begin{vmatrix} 2 & 3 & 4 \\ 6 & 2 & 2 \\ 3 & 7 & 1 \end{vmatrix} = \frac{1}{6} 120 = 20$$

$$S_{\Delta BCD} = \frac{1}{2} |[\overline{BC}, \overline{BD}]| = \frac{1}{2} \text{mod} \begin{vmatrix} i & j & k \\ 4 & -1 & -2 \\ 1 & 4 & -3 \end{vmatrix} = \frac{1}{2} |11i + 10j + 17k| =$$

$$= \frac{1}{2} \sqrt{121 + 100 + 289} = \frac{1}{2} \sqrt{510}.$$

$$\text{Тогда } AO = \frac{3V}{S} = \frac{3 \cdot 20 \cdot 2}{\sqrt{510}} = \frac{4\sqrt{510}}{17}$$

1.2.9 Тройное векторное произведение

Произведение $a \times (b \times c)$ можно вычислить с помощью двух операций нахождения векторного произведения. Однако это можно сделать и проще, воспользовавшись тождеством:

$$a \times (b \times c) = (a \times c)b - (a \times b)c \quad (1.21)$$

Это тождество можно доказать покомпонентно. Докажем, например; что $[a \times (b \times c)]_1 = a_2[b \times c]_3 - a_3[b \times c]_2$

где квадратные скобки с индексами обозначают компоненты соответствующих векторов. Итак,

$$\begin{aligned} [a \times (b \times c)]_1 &= a_2(b_1c_2 - b_2c_1)_3 - a_3(b_3c_1 - b_1c_3)_2 = \\ &= b_1(a_2c_2 + a_3c_3) - c_1(a_2b_2 + a_3b_3) = \\ &= (a_1c_1 + a_2c_2 + a_3c_3)b_1 - (a_1b_1 + a_2b_2 + a_3b_3)c_1 = \\ &= (a \times c)b_1 - (a \times b)c_1. \end{aligned}$$

Другие компоненты получаются перестановкой индексов. Так как перестановки не изменяют значений скалярных произведений, то искомым вектор дается уравнением (1.21).

Пример 1. Пересечение трех плоскостей. Три плоскости $r \times u_1 = p_1$, $r \times u_2 = p_2$ и $r \times u_3 = p_3$ пересекаются в точке $r = \frac{p_1(u_2 \times u_3) + p_2(u_3 \times u_1) + p_3(u_1 \times u_2)}{u_1 \times (u_2 \times u_3)}$ если $u_1 \times u_2 \times u_3 \neq 0$. Это можно проверить, показав, что r удовлетворяет уравнению каждой из этих плоскостей.

Пример 2. Окружность, проходящая через три данные точки (рис.1.41). Пусть радиус-векторы точек P, Q и R равны соответственно $0, a$ и b .

Тогда центр окружности, проходящей через точки P, Q и R , лежит в плоскости $r(a \times b) = 0$. Он также лежит в плоскостях, перпендикулярных отрезкам PQ и PR и делящих эти отрезки пополам; следовательно, $r \times a = a^2/2$ и $r \times b = b^2/2$. Используя результат предыдущего примера, получаем

$$r = \frac{a^2(b \times (a \times b)) + b^2((a \times b) \times b)}{2 \times |a \times b|^2} = \frac{b^2(a^2 - a \times b)a + a^2(b^2 - a \times b)b}{2 \times |a \times b|^2}$$

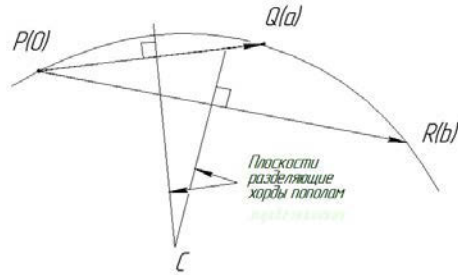


Рис. 1.41

Радиус R задается выражением

$$R = r \times r = \frac{b^2(a^2 - a \times b)(r \times a) + a^2(b^2 - a \times b)(r \times b)}{2 \times |a \times b|^2} =$$

$$= \frac{a^2 b^2 (a^2 - 2 \times a \times b + b^2)}{4 \times |a \times b|^2}$$

Отсюда

$$R = \frac{|a| \times |b| \times |a - b|}{2 \times |a \times b|}$$

Смешанным произведением векторов a, b, c называется число, обозначаемое (a, b, c) и определяемое как скалярное произведение вектора $[a, b]$ и вектора c : $(a, b, c) = ([a, b]c)$

Приведем геометрические свойства смешанного произведения векторов:

1) если V — объем параллелепипеда, построенного на векторах a, b, c , то $(a, b, c) = V$, если тройка a, b, c , — правая, и $-V$, если тройка a, b, c , — левая.

2) для того чтобы векторы a, b, c , были компланарны, необходимо и достаточно выполнение условия $(a, b, c) = 0$.

Перечислим алгебраические свойства смешанного произведения векторов:

1) $(a, b, c) = (b, a, c) = (c, a, b)$;

2) $(b, a, c) = -(a, b, c)$, $(c, b, a) = -(a, b, c)$, $(a, c, b) = -(a, b, c)$;

3) $(\alpha a_1 + \alpha_2 a_2, b, c) = \alpha (a_1, b, c) + \alpha_2 (a_2, b, c)$, α_1 и $\alpha_2 \in R$;

4) $(i, j, k) = 0$.

5) Если $a = (a_x, a_y, a_z)$, $b = (b_x, b_y, b_z)$, $c = (c_x, c_y, c_z)$, то $(a, b, c) =$

$$= \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}$$

РАЗДЕЛ 2. ПРЕОБРАЗОВАНИЯ



ФОМЕНКО А.Т.
РАЗМЫШЛЕНИЯ О МАТЕМАТИКЕ

Кто понимает математику, тот видит в ней не только истину, но и совершенную красоту - красоту холодную и строгую, подобную красоте скульптуры.

Бертран Рассел



Больше, чем когда либо, математика является одновременно культурой в лоне культуры и техникой в сердце техники. Она представляет культурную ценность сама по себе, идеал формальной красоты, заложенной в произведениях искусства - в том, что в них содержится наиболее классического. Этот идеал выражается словами мера, порядок, отношение, пропорция, являющимися математическими терминами. Математика - это школа, в которой обучаются логике на каждом шагу.

В. Серве
(современный бельгийский математик)

Краеугольным камнем компьютерной графики являются аффинные преобразования, составляющие основу большинства графических систем. Однако правильное их применение часто вызывает затруднения.

2.1 Графический конвейер

Преобразования полезны в следующих ситуациях.

- Составляется «сцена» из нескольких объектов, каждый из которых легче всего разрабатывать в собственной системе координат. Сцена составляется посредством помещения нескольких «экземпляров» объекта различного размера в различных местах с использованием нужного преобразования каждого экземпляра.

- Некоторые объекты имеют определенную симметрию. Разработав один «мотив» можно составить всю форму с помощью отражений, поворотов и переносов основного мотива.

- Пользователю может понадобиться рассмотреть объект с различных точек наблюдения и получить изображения с каждой из них. Перемещение и ориентация камеры, как правило, осуществляется с помощью трехмерных аффинных преобразований.

- В компьютерной анимации несколько объектов должны от кадра к кадру двигаться один относительно другого. Эффект может быть достигнут путем сдвига и поворота отдельной локальной системы координат каждого объекта в процессе анимации.

Конечная цель состоит в создании изображений объектов, преобразованных до необходимых размеров, ориентации и положения с тем, чтобы составить желаемую сцену. Для этих целей в целом ряде графических платформ предусмотрен так называемый **«графический конвейер»** - последовательность операций, применяемых ко всем точкам, которые «пропускаются» через такой конвейер.

Все точки P_1, P_2, \dots вначале подвергаются преобразованию, которое называется **«текущим преобразованием»** (*current transformation*) и изменяет их значения, генерируя другую последовательность Q_1, Q_2, \dots . Исходные точки P_1, P_2, \dots описывают заданный геометрический объект, точки Q_1, Q_2, \dots описывают преобразованный вариант того же самого объекта. Эти точки посылаются на следующие этапы преобразования и, в конечном счете, используются для рисования окончательного изображения.

2.2 Преобразование точек и объектов

Преобразования можно рассматривать с двух сторон:

- преобразование объекта,

■ преобразование координат.

При преобразовании объекта координаты каждой его точки изменяются в соответствии с некоторыми законами, а система координат остается неизменной.

При преобразовании координат старая система координат преобразуется в новую, и все точки объекта получают представление в этой системе координат.

Эти два подхода тесно связаны между собой, каждый из них имеет свои преимущества, однако реализуются они по-разному.

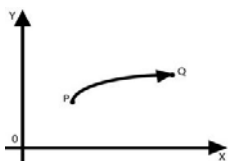


Рис. 2.1

Рассмотрим общую идею преобразования.

Любое преобразование переводит каждую точку P в пространстве (двумерном, трехмерном,) в новую точку Q согласно заданной формуле или алгоритму. Произвольная точка P отображается в другую точку Q . Тогда говорят, что точка Q образ точки P при отображении T (рис. 2.1). Преобразование всего объекта осуществляется преобразованием каждой его точки, причем для каждой точки используется одна и та же функция $T(\)$. За один раз можно отображать целые совокупности точек – например все точки прямой или эллипса.

Более строго: если S - множество точек, то его образ $T(S)$ является множеством всех точек $T(P)$, где P некоторая точка из множества S .

Большинство применяемых в компьютерной графике преобразований непрерывны. В этом случае образ прямой линии есть также связанная линия некоторой формы, не обязательно прямая. Например:

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = T \begin{pmatrix} \cos(P_x)e^{-P_y} \\ \frac{\ln(P_y)}{1 + P_x^2} \\ 1 \end{pmatrix}$$

Это преобразование переводит прямую в кривую линию.

2.3 Аффинные преобразования

При аффинных преобразованиях координаты точки Q являются линейной комбинацией соответствующих координат точки P .

Прямые преобразуются в прямые. Образ прямой линии после аффинного преобразования T есть прямая. Чтобы сохранить прямые линии при выполнении преобразований используется явный координатный фрейм, состоящий из точки V - начала координат, и нескольких взаимно перпендикулярных векторов (в двумерном случае это \vec{i} и \vec{j} , а в трехмерном - $\vec{i}, \vec{j}, \vec{k}$), которые служат осями

этого координатного фрейма. Какой бы фрейм не использовался, точки P и имеют следующее представление (в двумерном случае):

$$P = \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} Q = \begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix}$$

соответственно. Это означает, что точка P находится в положении $P = P_x \vec{i} + P_y \vec{j} + V$, где P_x и P_y - координаты точки P . Преобразование воздействует на это представление точки P и на выходе дает представление точки Q в соответствии с некоторой функцией $T(\bullet)$, т.е.:

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = T \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \text{ или кратко } Q = T(P)$$

Основные преимущества аффинных преобразований:

1. Используются в компьютерной графике;
2. Упрощают масштабирование, поворот и перестановку изображений;
3. Последовательность аффинных преобразований может быть преобразована в суммарное аффинное преобразование;
4. Допускают компактное матричное преобразование.

Форма аффинных преобразований следующая: координаты точки Q (образа) являются линейными комбинациями соответствующих координат точки P .

$$\begin{aligned} Q_x &= m_{11}P_x + m_{12}P_y + m_{13} \\ Q_y &= m_{21}P_x + m_{22}P_y + m_{23}. \end{aligned} \quad (2.1)$$

Матрица линейных преобразований (2.1) в декартовых координатах имеет размерность 2×3 на плоскости или 3×4 в пространстве, то есть она не квадратная. Значит

- невозможно использовать унифицированный механизм работы с матрицами для преобразований точек (нельзя описать перенос);
- декартовы координаты не позволяют применить матричную запись для построения перспективных проекций точек.

Проблема решается введением однородных координат.

2.4 Однородные координаты

Существуют различные способы определения однородных координат. Мы будем исходить из задачи унифицированного представления координат точек в пространстве, включающего бесконечно удаленны **точки**.

Пусть заданы действительных: числа, a и w . Рассмотрим их отношение a/w . Зафиксируем значение a , и будем варьировать значение w . При уменьшении w , значение a/w будет увеличиваться. Заметим, что если w стремится к нулю, то a/w стремится к бесконечности. Таким образом, чтобы включить в рассмотрение понятие бесконечности, для представления значения v используется пара чисел (a,w) , таких, что $v=a/w$. Если $w \neq 0$, значение v в точности равно a/w . В противном случае $v=a/0$, т.е. равно бесконечности.

Представим координаты точки на плоскости $v=(x,y)$ можно через координаты (wx, wy, w) . При $w=1$ эти координаты описывают точку с конечными координатами (x,y) , а при $w=0$ - точку, бесконечно удаленную в направлении (x,y) . Обычным представлением через декартовы координаты (x,y) это сделать невозможно.

Рассмотрим двумерную плоскость, некоторую точку (x,y) на ней и заданную функцию $f(x,y)$. Если заменить x и y на x/w и y/w , то выражение $f(x,y)=0$ заменится на $f(x/w,y/w)=0$. Если $f(x,y)$ - многочлен, то его умножение на w^n (n - степень многочлена) уберет все знаменатели. Например, пусть имеется прямая

$$Ax + By + C = 0$$

Замена x и y на x/w и y/w дает $A(x/w) + B(y/w) + C = 0$. Умножая на w , получаем

$$Ax + By + Cw = 0 \tag{2.2}$$

Другой пример. Пусть задан многочлен 2-го порядка

$$Ax^2 + Bxy + Cy^2 + 2Dx + 2Ey + Fw = 0$$

После замены x и y на x/w и y/w , соответственно, и умножения на w^2 , получаем

$$Ax^2 + Bxy + Cy^2 + 2Dxw + 2Eyw + Fw^2 = 0 \tag{2.3}$$

Если внимательно посмотреть на многочлены (2) и (3) можно заметить, что степени при всех членах равны B случае многочлена 1-го порядка, это степень 1, тогда как для многочлена 2-й степени, все члены (т.е. x^2 , xy , y^2 , xw , yw и w^2) имеют степень 2. Следовательно, для данного многочлена n -го порядка, после введения координаты w все члены будут иметь степень n . Такие многочлены называются однородными, а координаты (x,y,w) называются **однородными координатами** (homogenous coordinates).

Приведенные рассуждения остаются верными и в случае трехмерного пространства. Координаты (x,y,z) заменяются на $(x/w, y/w, z/w)$ и после умножения на w в соответствующей степени n дают однородный многочлен.

Однородные координаты требуют три компоненты для представления точки на плоскости (и четыре компоненты для точки в пространстве).

Какие же однородные координаты соответствуют точке с координатами (x,y) ? Легко видеть, что это будет $(x,y,1)$, т.е. иполагается равной 1. В общем случае, это преобразование не однозначно. Однородные координаты точки (x,y) равны (xw, yw, w) для любого ненулевого w . Аналогично в трехмерном про-

странстве: точке (x, y, z) соответствуют координаты (xw, yw, zw, w) . В то же время, преобразование однородных координат в евклидовы однозначно: точке (x, y, w) соответствует точка $(x/w, y/w)$. Приведем более формальное определение.

Определение Однородными координатами точки $P=(x_1, \dots, x_n)$, $P \in R^n$ называются координаты $P_{\text{одн}} = (wx_1, wx_2, \dots, wx_n, w)$, $P_{\text{одн}} \in R^{n+1}$, причем хотя бы один элемент должен, быть отличен от нуля.

На самом деле, множество векторов $P_{\text{одн}}$ при определенных дополнительных операциях образуют так называемое **проективное пространство**, которое имеет важнейшее значение в компьютерной графике и машинном зрении. Отметим следующее:

преобразование из однородных координат в евклидовы однозначно; преобразование из евклидовых координат в однородные — нет.

За счет введения третьей компоненты в векторы точек $[x \ y]$ и $[x^* \ y^*]$, т. е. представляя их в виде $[x \ y \ 1]$ и $[x^* \ y^* \ 1]$ возможно смещение на плоскости точек, линий, рисунков и контуров из точек или линий. Матрица преобразования

после этого становится матрицей размера 3×2 :
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ m & n \end{bmatrix}.$$

Это необходимо, поскольку число столбцов в матрице, описывающей точку, должно равняться числу строк в матрице преобразования для выполнения операции умножения матриц. Таким образом,

$$[x \ y \ 1] \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ m & n \end{bmatrix} = [x + m \ y + n] = [x^* \ y^*]$$

следовательно, константы m, n вызывают смещение x^* и y^* относительно x и y . Поскольку матрица 3×2 не является квадратной, она не имеет обратной матрицы. Эту трудность можно обойти, дополнив матрицу преобразования до квадратной размера 3×3 .

Например,
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m & n & 1 \end{bmatrix}$$

Заметим, что третья компонента векторов положения точек не изменяется при добавлении третьего столбца к матрице преобразования. Используя эту матрицу в предыдущей формуле, получаем преобразованный вектор $[x^* \ y^* \ 1]$. Добавление третьего элемента к вектору положения и третьего столбца к матрице преобразования позволяет выполнить смещение вектора положения. Третий элемент здесь можно рассматривать как дополнительную координату вектора положения. Итак, вектор положения $[x \ y \ 1]$ при воздействии на него матрицы 3×3 становится вектором положения в общем случае вида $[XYH]$. Представленное преобразование было выполнено так, что $[XYH] = [x^* \ y^* \ 1]$.

Преобразование в трехмерном пространстве, в нашем случае ограничено плоскостью, поскольку $H = 1$.

Если третий столбец $\begin{bmatrix} p \\ q \\ s \end{bmatrix}$ матрицы преобразования T размера 3×3 отличен

от $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, то в результате матричного преобразования получим $[x \ y \ 1] = [XYH]$, где,

$H \neq 1$.

Плоскость, в которой теперь лежит преобразованный вектор положения, находится в трехмерном пространстве. Итак, найденные x^* и y^* получены с помощью пучка лучей, проходящих через начало координат. Результат преобразований показан на рис.2. Из рассмотрения подобных треугольников видно, что $H/X = 1/x^*$ и $H/Y = 1/y^*$. Рассматривая три компоненты, запишем это в виде:

$$[x^* \ y^* \ 1] = \begin{bmatrix} X & Y \\ H & H \end{bmatrix} 1$$

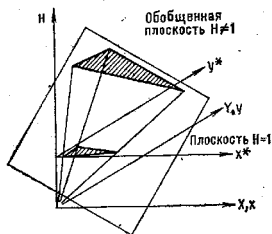


Рис.2.2

Геометрическое представление однородных координат.

Представление двумерного вектора трехмерным или в общем случае n -мерного вектора $(n + 1)$ - мерным называют однородным координатным воспроизведением. При однородном, координатном воспроизведении n -мерного вектора оно выполняется в $(n + 1)$ -мерном пространстве, и конечные результаты в n - мерном пространстве получают с помощью обратного преобразования. Таким образом, двумерный вектор $[x \ y]$ представляется трехкомпонентным вектором $[hx \ yh]$. Разделив компоненты вектора на однородную координату h , получим:

$$x = \frac{hx}{h} \text{ и } y = \frac{yh}{h}$$

Не существует единственного однородного координатного представления точки в двумерном пространстве. Например, однородные координаты $(12, 8, 4)$, $(6, 4, 2)$ и $(3, 2, 1)$ представляют исходную точку $[3 \ 2]$. Для простоты вычисле-

ний выбираем $[x \ y \ 1]$, чтобы представить не преобразованную точку в двумерных однородных координатах.

Преобразование $[x^* \ y^*] = [x \ y] \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ задается выражением в однородных координатах в виде: $[X \ Y \ H] = [x \ y \ 1] \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(Заметим, что левая часть уравнения является результатом преобразования, т. е. здесь записываем матричные уравнения в виде $B = AT$ вместо $AT = B$. Далее будем использовать именно такие соотношения для перевода результатов в компьютерные алгоритмы.) Выполнение указанных выше преобразований показывает, что

$$X = x^*, \ Y = y^*, \ \text{а} \ H = 1.$$

Равенство единице дополнительной координаты означает, что преобразованные однородные координаты равны исходным координатам. В общем случае $H \neq 1$ преобразованные обычные координаты получаются за счет нормализации однородных координат, т. е.

$$x^* = \frac{X}{H} \ \text{и} \ y^* = \frac{Y}{H}$$

Геометрически все преобразования x и y происходят в плоскости $H = 1$ после нормализации преобразованных однородных координат.

Преимущество введения однородных координат проявляется при использовании матрицы преобразований общего вида порядка 3×3 ,

$$\begin{bmatrix} a & b & p \\ c & d & q \\ m & n & s \end{bmatrix}$$

с помощью которой можно выполнять и другие преобразования, такие как смещение, операции изменения масштаба и сдвига, обусловленные матричными элементами a, b, c и d .

Чтобы показать воздействие третьего столбца матрицы преобразований 3×3 , рассмотрим следующую операцию:

$$[XYH] = [x \ y \ 1] \begin{bmatrix} 1 & 0 & p \\ 0 & 1 & q \\ 0 & 0 & 1 \end{bmatrix} = [x \ y \ (px + qy + 1)]$$

здесь $X = x, \ Y = y$, а $H = px + qy + 1$.

Переменная H , которая определяет плоскость, содержащую преобразованные точки, представленное в однородных координатах, теперь образует уравнение плоскости в трехмерном пространстве. Это преобразование показано на рис.2.3, где линия AB , лежащая в плоскости $H=1$, спроектирована на линию CD плоскости $pX + qY - H + 1 = 0$

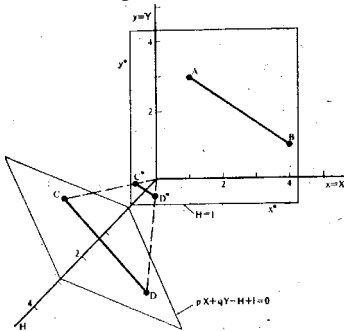


Рис. 2.3

На рис. 2.3. величины $p = q = 1$.

Выполним нормализацию для того, чтобы получить обычные координаты:

$$x^* = \frac{X}{H} = \frac{X}{pX + qY + 1}, y^* = \frac{Y}{H} = \frac{Y}{pX + qY + 1}$$

Полагая $p = q = 1$, для изображенных на рис. 6 точек **A** и **B** с координатами соответственно (1, 3) и (4, 1) получим

$$x^* = \frac{1}{1+3+1} = \frac{1}{5} \text{ и } y^* = \frac{3}{5}$$

После преобразования **A** в **C*** и **B** в **D*** имеем

$$x^* = \frac{4}{1+4+1} = \frac{2}{3} \text{ и } y^* = \frac{1}{6}$$

Однородные координаты для точек **C*** и **D***, показанные на рис.2.3, соответственно равны $(\frac{1}{5}, \frac{3}{5}, 1)$ и $(\frac{2}{3}, \frac{1}{6}, 1)$

Результатом нормализации является перевод трехмерной линии **CD** в ее проекцию **C*D*** на плоскость $H = 1$. Центром проекции является начало координат.

При помощи троек однородных координат и матриц третьего порядка можно описать любое аффинное преобразование плоскости.

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix}. \quad (2.4)$$

Векторы преобразуются так же, как и точки. Очевидно, что когда исходный вектор, преобразуется в другой вектор, его третий компонент всегда равен 0.

$$\begin{pmatrix} W_x \\ W_y \\ 0 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ 0 \end{pmatrix} \quad (2.5)$$

Преобразование неособенное, если $\det M = m_{11}m_{22} - m_{12}m_{21} \neq 0$. Тогда

$$Q = MPM^{-1} = \frac{1}{\det M} \begin{pmatrix} m_{22} & -m_{12} \\ -m_{21} & m_{11} \end{pmatrix} P = M^{-1}Q$$

В дальнейшем, если не оговорено специально, будут рассматриваться неособенные преобразования.

2.5 Элементарные преобразования плоскости

Тожественное преобразование, оставляет точку (геометрический объект) неподвижной

$$T_{тож} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \text{ (I-единичная матрица).}$$

Перемещение (перенос). Часто требуется переместить точку или изображение в другое место (например, на дисплее). Пусть необходимо переместить точку P в положение Q .

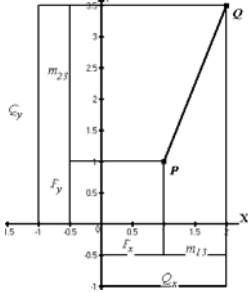


Рис. 2.4.

Из рисунка 4 следует:

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} P_x + m_{13} \\ P_y + m_{23} \\ 1 \end{pmatrix} \Rightarrow Q = P + \bar{d},$$

$$\bar{d} = (m_{13}, m_{23})\text{-вектор перемещения.}$$

$$Q = T \cdot P \tag{2.6}$$

$$T_{\text{пер}} = \begin{pmatrix} 1 & 0 & m_{13} \\ 0 & 1 & m_{23} \\ 0 & 0 & 1 \end{pmatrix} T_{\text{пер}}^{-1} = \begin{pmatrix} 1 & 0 & -m_{13} \\ 0 & 1 & -m_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

Произведение AB двух матриц A и B определено только в том случае, если эти матрицы согласованы, т.е. число столбцов m первой матрицы A равно числу строк n второй матрицы B ($m=n$). Каждый элемент c_{ik} матрицы $C(m \times r)$ произведения матриц $A(m \times n)$ и $B(n \times r)$, $C=AB$ является скалярным произведением i -той строки матрицы A на j -тый столбец матрицы B : $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$. Если точка задана не столбцом, а строкой то выражение (6) примет вид

$$Q = P \cdot T^T \cdot (Q_x, Q_y, 1) = (P_x, P_y, 1) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ m_{13} & m_{23} & 1 \end{pmatrix} \tag{2.7}$$

Масштабирование. Масштабирование относительно начала (координат) отсчета

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} S_x P_x \\ S_y P_y \\ 1 \end{pmatrix} T_M = \begin{pmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{pmatrix} T_M^{-1} = \begin{pmatrix} \frac{1}{S_x} & 0 & 0 \\ 0 & \frac{1}{S_y} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\det|T_M| = S_x S_y$$

Если $S_x = S_y = S$ — равномерное масштабирование.

$|S| > 1$ — увеличение; $|S| < 1$ — уменьшение.

Если $S_x \neq S_y$ — дифференцированное масштабирование

Отражение: частный случай масштабирования,

$S_x = S_y = -1$ $T(P_x, P_y) = (-P_x, -P_y)$ — симметрия относительно начала координат $(\cdot)0$.

$S_x = 1, S_y = -1$ $T(P_x, P_y) = (P_x, -P_y)$ — симметрия относительно оси X .

$S_x = -1, S_y = 1$ $T(P_x, P_y) = (-P_x, P_y)$ — симметрия относительно оси Y .

Поворот. Одной из основных операций в графике является поворот изображения относительно заданной точки на некоторый угол. Из рис.2.5 следует:

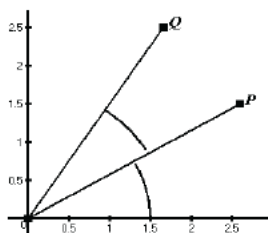


Рис. 2.5

Из рис.2.5 следует:

$$P_x = R \cos \alpha, P_y = R \sin \alpha, Q_x = R \cos(\alpha + \theta), \\ Q_y = R \sin(\alpha + \theta).$$

$$\sin(\alpha + \theta) = \sin \alpha \cos \theta + \cos \alpha \sin \theta, \\ \cos(\alpha + \theta) = \cos \alpha \cos \theta - \sin \alpha \sin \theta. \text{ Д} \\ Q_x = P_x \cos \theta - P_y \sin \theta, Q_y = P_x \sin \theta + \\ + P_y \cos \theta.$$

Из полученных соотношений следует матрица поворотов:

$$T_{\Pi} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad T_{\Pi}^{-1} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\det|T_{\Pi}| = \cos \theta + \sin \theta = 1 \Rightarrow \det|T_{\Pi}^{-1}| = 1$$

Поворот отсчитывается против часовой стрелки.

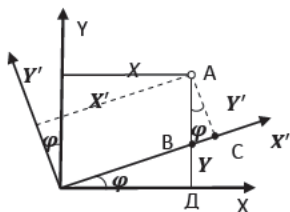


Рис. 2.6.

Повернем систему координат на угол φ против хода часовой стрелки вокруг начала координат точки 0 (рис.2.6).

$$BD = X_A \operatorname{tg} \varphi; OB = \frac{X_A}{\cos \varphi}, AB = Y_A - X_A \operatorname{tg} \varphi,$$

$$X'_A = OB + BCY_A = AB \cos \varphi \\ \begin{cases} X' = X \cos \varphi + Y \sin \varphi \\ Y' = -X \sin \varphi + Y \cos \varphi \end{cases} \Rightarrow \\ \begin{cases} X = X' \cos \varphi - Y' \sin \varphi \\ Y = X' \sin \varphi + Y' \cos \varphi \end{cases}$$

Матрица поворотов (как и матрица любого другого элементарного преобразования) не зависит от того, рассматриваем мы поворот точки или поворот системы координат.

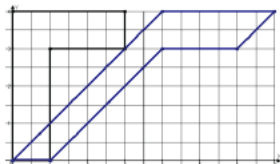


Рис. 2.7.

Сдвиг. Сдвиг в x -направлении (вдоль x) описывается формулами (рис. 2.7)

$$\left. \begin{aligned} Q_x &= P_x + h P_y \\ Q_y &= P_y \end{aligned} \right\} \quad (2.8)$$

h определяет долю координату точки P которая должна быть добавлена к координатам, h может быть как положительным, так и отрицательным.

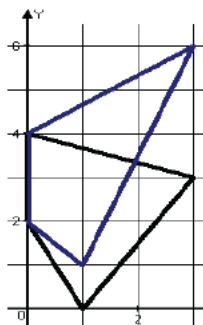


Рис. 2.8.

Матрица этого сдвига имеет вид

$$T_{CX} = \begin{pmatrix} 1 & h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} T_{CX}^{-1} = \begin{pmatrix} 1 & -h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \det|T_{CX}| = 1$$

$$h = 1$$

Аналогично осуществляется сдвиг в y -направлении (рис. 2.8).

$$\begin{aligned} Q_x &= P_x \\ Q_y &= P_y + gP_x \end{aligned} \quad (2.9)$$

$$T_{CY} = \begin{pmatrix} 1 & 0 & 0 \\ g & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} T_{CY}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -g & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \det|T_{CY}| = 1 \quad g = 1$$

Рассмотрим пример: поворот объекта относительно некоторой произвольной точки $P1$. Поскольку нам известно, лишь как поворачивать вокруг начала координат, разобьем исходную задачу на три подзадачи:

- 1) перенос, при котором точка $P1$ перемещается в начало координат;
- 2) поворот;
- 3) перенос, при котором точка из начала координат возвращается в первоначальное положение $P1$.

Последовательность этих преобразований показана на рис.2.9.

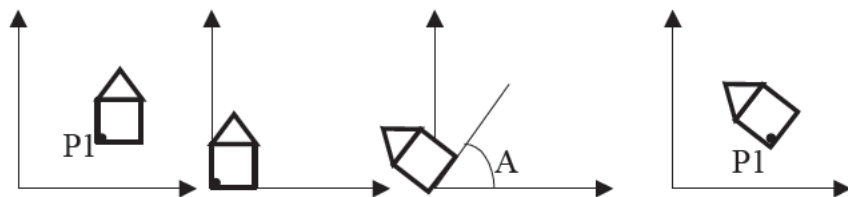


Рис. 2.9

Полезные свойства аффинных преобразований:

■ Последовательное выполнение нескольких аффинных преобразований есть аффинное преобразование.

■ При аффинных преобразованиях сохраняется аффинная комбинация точек

$$W = a_1P_1 + a_2P_2(a_1 + a_2) = 1$$

$T(W) = T(a_1P_1 + a_2P_2) = a_1T(P_1) + a_2T(P_2)$ – определение аффинных преобразований.

$$M(W) = M(a_1P_1 + a_2P_2) = a_1M(P_1) + a_2M(P_2)$$

■ При аффинных преобразованиях сохраняются прямые линии .

$$L(t) = (1 - t)A + tB - \text{прямая } (AB)$$

$$Q(t) = (1 - t)T(A) + tT(B) - \text{прямая } (T(A), T(B));$$

$$P(s, t) = sT(A) + tT(B) + (1 - s - t)T(C)$$

■ Параллельность прямых сохраняется:

$$L = A + \bar{b}t \Rightarrow M(A + \bar{b}t) = MA + (M\bar{b})t$$

\bar{b} — направляющий вектор; $M\bar{b}$ — н.в.

Направление $M\bar{b}$ не зависит от точки $A \Rightarrow$ прямые $A_1 + \bar{b}t$ и $A_2 + \bar{b}t$ имеют одно направление, т. е. параллельны.

Следствие: параллелограммы отображаются в другие параллелограммы; параллелепипеды отображаются в другие параллелепипеды (рис. 2.10).

Преобразование сетки.

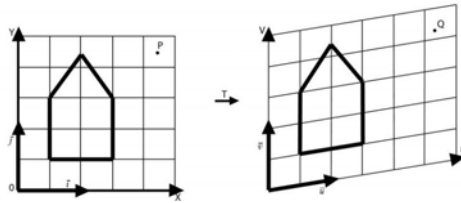


Рис. 2.10

■ Столбцы матрицы показывают преобразованный координатный фрейм.

Рассмотрим столбцы матрицы M аффинных преобразований, поскольку они показывают, как преобразуется координатный фрейм,

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix} = (\bar{m}_1, \bar{m}_2, m_3)$$

Координатный фрейм (\bar{i}, \bar{j}, v)

$$v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \bar{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \bar{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

преобразуется в новый координатный фрейм

$$(\bar{m}_1, \bar{m}_2, m_3):$$

$$\bar{m}_1 = M\bar{i}; \quad \bar{m}_2 = M\bar{j}; \quad m_3 = Mv$$

Оси нового координатного фрейма не обязательно перпендикулярны и не обязательно имеют единичную длину. (Перпендикулярность сохраняется, если

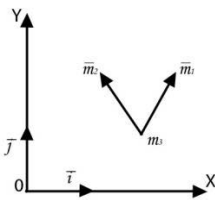


Рис. 2.11

аффинные преобразования включают в себя только повороты и равномерные масштабирования).

■ Относительные пропорции сохраняются.

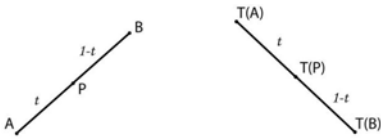


Рис. 2.12

$$P \in (AB); (AB) = A + \bar{b} t$$

где $\bar{b} = B - A$

$$P = A + \bar{b} t_p$$

$$t_p = \frac{P - A}{\bar{b}} t_p = \frac{|P - A|}{|B - A|}$$

$$|P - A| = t_p |B - A|, \quad M(A + \bar{b} t) = MA + (M\bar{b})$$

■ Влияние преобразований на площади и объём фигур.

Перемещения и повороты не влияют на площадь фигур, масштабирование изменяет ее.

$$(2D) \quad \frac{\text{площадь после преобразования}}{\text{площадь до преобразования}} = |\det M|$$

■ Любое аффинное преобразование может быть составлено из элементарных операций. $M(3 \times 3) \Rightarrow$ поворот, масштаб, сдвиг, перемещение.

■ В результате аффинных преобразований порядок алгебраической кривой не увеличивается т.е. остается тем же или уменьшается.

Компоновка двумерных преобразований. Процесс последовательного применения нескольких преобразований с целью формирования единого суммарного преобразования называется компоновкой или композицией (конкатенацией) этих преобразований.

Пусть два преобразования представлены матрицами M_1 и M_2 . тогда точка P вначале преобразуется в точку $M_1 P$, которая затем преобразуется в точку $M_2(M_1 P)$. В силу свойства ассоциативности последнее выражение равно $(M_1 M_2)P$. Тогда имеем $W = MP$, где суммарное преобразование представлено одной матрицей

$$M = M_2 M_1$$

При использовании однородных координат аффинные преобразования компоуются посредством умножения матриц. При этом матрицы должны быть расположены в порядке обратном по отношению к тому, в котором эти матрицы применяются.

При применении последовательности **преобразований точек** суммарное преобразование M получается умножение **слева**.

$$M = M_3 \times M_2 \times M_1.$$

При применении последовательности *преобразований системы координат* суммарное преобразование M получается умножение *справа*

$$M = M_1 \times M_2 \times M_3.$$

2.6. Элементарные трехмерные преобразования

Перемещение. Поворот. Масштабирование. Сдвиг.

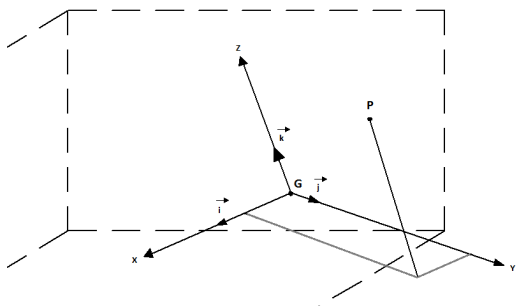


Рис. 2.13

Автоматизированное проектирование геометрического объекта предполагает наличие информации о положении объекта на плоскости или в пространстве и его размерные характеристики. При этом положение геометрического объекта определяется в системе координат, связанной с этим объектом (локальная система координат).

Такая система координат сдвину-

та относительно жестко закрепленной системы координат (глобальная система координат). Локальная система координат получается из глобальной системы координат с использованием таких преобразований, как параллельный перенос или поворот. Найдем зависимость между координатами точки $P(x, y, z)$ в глобальной системе координат и координатами той же точки $P(x', y', z')$ в локальной системе координат.

Рассматривая трехмерные аффинные преобразования будем использовать координатные фреймы, считая, что заданы:

- Начало отсчета – точка G ,
- При взаимно перпендикулярных осях в направлениях $\vec{i}, \vec{j}, \vec{k}$ точка P в таком фрейме задается уравнением:

$$P = G + P_x \vec{i} + P_y \vec{j} + P_z \vec{k}$$

и может быть представлена в виде

$$P = \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix}$$

Пусть $T(\cdot)$ аффинное преобразование, преобразующие точку P в точку Q .

Тогда преобразование $T(\cdot)$ отображается матрицей M размерностью 4×4 :

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Отображение точки Q можно найти, умножив точку P (справа) на матрицу M

$$P = \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix}, \quad \begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = M \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad \text{и} \quad Q = MP$$

Последняя строка матрицы аффинного преобразования состоит из ряда нулей, заканчивающегося одной единицей. Это справедливо только для аффинных преобразований.

Пусть глобальная система координат задана базисом (e_1, e_2, e_3) , отнесенным к точке N . Имеется другая система координат, заданная базисом (s_1, s_2, s_3) , отнесенным к точке K . Радиус-вектор a точки K имеет координаты (a_1, a_2, a_3) .

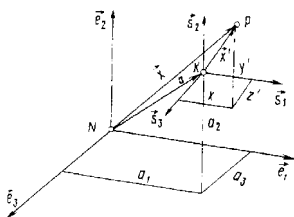


Рис. 2.14

Рассмотрим следующие случаи.

Перемещение. Система координат (K, s_1, s_2, s_3) получается в результате параллельного переноса декартовой прямоугольной системы координат (N, e_1, e_2, e_3) .

Из рис.2.14 очевидна следующая зависимость между векторами: $x = a + x'$

Отсюда получим формулы преобразования координат:

$$x = m_{14} + x';$$

$$y = m_{24} + y'; \quad M = \begin{bmatrix} 1 & 0 & 0 & m_{14} \\ 0 & 1 & 0 & m_{24} \\ 0 & 0 & 1 & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$z = m_{34} + z'.$$

Масштабирование $M = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$

где s_x, s_y, s_z константы, определяющие масштабирование по соответствующим координатам относительно начала отсчета.

Если $S_x = S_y = S_z = S$ – равномерное масштабирование.

$|S| > 1$ -увеличение; $|S| < 1$ -уменьшение.

Если $S_x \neq S_y$ – дифференцированное масштабирование

Отражение: частный случай масштабирования,

$S_x = S_y = S_z = -1$ $\text{P} T(P_x, P_y, P_z) = (-P_x, -P_y, -P_z)$ -симметрия относительно начала координат $(\cdot)\theta$.

$S_x = S_y = 1, S_z = -1$ $\text{P} T(P_x, P_y, P_z) = (P_x, P_y, -P_z)$ - симметрия относительно плоскости XOY ;

$S_x = S_z = 1, S_y = -1$ $\text{P} T(P_x, P_y, P_z) = (P_x, -P_y, P_z)$ - симметрия относительно плоскости XOZ ;

$S_x = -1, S_y = S_z = 1$ $\text{P} T(P_x, P_y, P_z) = (-P_x, P_y, P_z)$ - симметрия относительно плоскости YOZ ;

Сдвиг

Трёхмерные сдвиги более разнообразны, чем их двумерные аналоги.

Матрица элементарного сдвига имеет вид

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ f & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{P} Q = (P_x, fP_x + P_y, P_z)$$

Обобщенный сдвиг состоит из:

- плоскости, проходящей через начало координат и заданной единичной нормалью m ,
- единичного вектора v , лежащего в этой плоскости $\text{P} m \wedge v$,
- угла q (рис. 2.14).

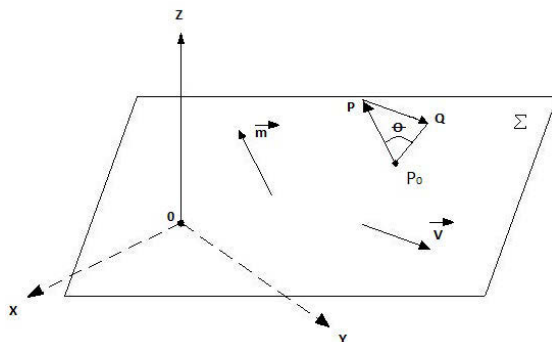


Рис. 2.15

Тогда (рис. 2.15) сдвиг точки P в точку Q осуществляется посредством ее перемещения на нужную величину пропорциональную расстоянию, на которое точка P отстоит от плоскости, а так же тангенсу угла q в направлении v . Матрица преобразования в этом случае имеет вид:

$$M = I + tg(\theta) \begin{pmatrix} m_x v_x & m_x v_y & m_x v_z \\ m_y v_x & m_y v_y & m_y v_z \\ m_z v_x & m_z v_y & m_z v_z \end{pmatrix}$$

где I — единичная матрица 3×3 .

Выразим точку Q через точку P и компоненты сдвига. Расстояние (с учетом знака) от точки P до плоскости определяется выражением $P\vec{m}$, если рассматривать точку P как радиус-вектор, исходящий из начала координат (нормальное уравнение плоскости, проходящей через начало координат имеет вид: $xn_x + yn_y + zn_z = 0$. Если точка P не лежит в плоскости то $|P_x n_x + P_y n_y + P_z n_z| = d$, d - расстояние от точки до плоскости и, следовательно $d = P\vec{n}$, где $P(P_x, P_y, P_z)$, а $\vec{n}(n_x, n_y, n_z)$.

Собирая вместе все компоненты, получим: $Q = P + (P \cdot \vec{m}) \tan \theta \vec{v}$

Точки по другую сторону плоскости сдвигаются в противоположную сторону.

Т.к. $P \cdot \vec{m} = P\vec{m}^T$, то $Q = P(I + \tan \theta \vec{m}^T \vec{v})$. Следовательно, матрица сдвига равна $(I + \tan \theta \vec{m}^T \vec{v})$, где член $\vec{m}^T \vec{v}$ имеет вид:

$$\vec{m}^T \vec{v} = \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} \begin{pmatrix} v_x & v_y & v_z \end{pmatrix} = \begin{pmatrix} m_x v_x & m_x v_y & m_x v_z \\ m_y v_x & m_y v_y & m_y v_z \\ m_z v_x & m_z v_y & m_z v_z \end{pmatrix}$$

Это выражение называется внешним тензорным произведением вектора \vec{m} на вектор \vec{v} .

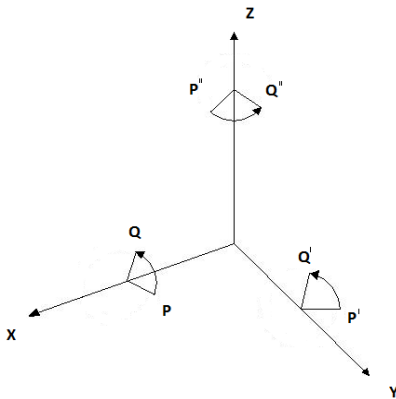


Рис. 2.16

Повороты. Простейшим поворотом является поворот относительно одной из координатных осей. назовем поворот вокруг оси x — « x -вращением», поворот вокруг оси y — « y -поворотом», а вокруг оси z — « z -поворотом».

Рассмотрим по отдельности матрицы, осуществляющие x -, y -, z -вращения.

Положительные углы определяются по правилу «взгляд во внутрь»: положительное значение угла α вызывает враще-

ние против хода часовой стрелки относительно оси, если наблюдатель смотрит во внутрь, по направлению к началу отсчета из точки, находящейся на положительном направлении оси.

Три основных положительных поворота приведены на рис. 2.16 и 2.17.

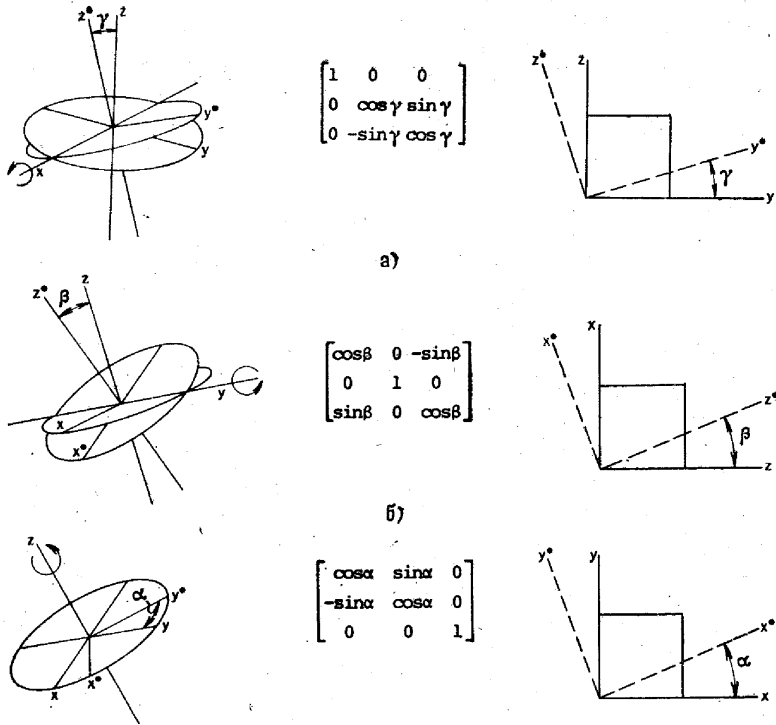


Рис.2.17. Вращение: а — вокруг оси x ; б — вокруг оси y ; в — вокруг оси z

$$R_X(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha_1 & -\sin \alpha_1 & 0 \\ 0 & \sin \alpha_1 & \cos \alpha_1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{- ось } Y \text{ поворачивается в сторону оси } Z.$$

$$R_Y = \begin{pmatrix} \cos \alpha_2 & 0 & \sin \alpha_2 & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \alpha_2 & 0 & \cos \alpha_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{- ось } Z \text{ поворачивается в сторону оси } X.$$

$$R_Z = \begin{pmatrix} \cos\alpha_3 & -\sin\alpha_3 & 0 & 0 \\ \sin\alpha_3 & \cos\alpha_3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \text{ось } X \text{ поворачивается в сторону оси } Z.$$

Композиция поворотов

Одно из самых важных различий между двумерными и трехмерными преобразованиями заключается в способе задания поворотов. В двумерном случае два поворота, например $R(\beta_1)$ и $R(\beta_2)$, при сложении дают $R(\beta_1 + \beta_2)$, при этом порядок, в котором осуществляются эти повороты, значения не имеет. В трехмерном случае, поскольку повороты могут происходить вокруг разных осей, порядок в котором осуществляются два поворота вокруг различных осей, имеет значение: матрицы трехмерных поворотов **некоммутативны** (т.е. порядок умножения нельзя менять, т.к. $AB \neq BA$).

Поворот в трех измерениях обычно строят как композицию трех элементарных поворотов: вначале x -вращение, затем y -вращение и, наконец, z -вращение. Выражение для суммарного вращения имеет вид:

$$M = R_z(\theta)R_y(\beta)R_x(\alpha)$$

Эти углы α, β, θ часто называют углами Эйлера.

Теорема Эйлера 1. гласит, что любой трехмерный поворот может быть получен тремя вращениями вокруг осей x, y, z и, следовательно, может быть записан в виде произведения пяти матриц при соответствующем выборе углов Эйлера. Очевидно, что можно использовать другой порядок вращений, например $R_y(\beta_1)R_z(\theta_1)R_x(\alpha_1)$. В этом случае требуются другие углы Эйлера $\beta_1, \theta_1, \alpha_1$. Всего существует 12 возможных порядков следования трех отдельных вращений, и в каждом из них используются различные величины α, β, θ .

Повороты вокруг произвольной оси.

Теорема Эйлера 2. Любой поворот (или последовательность поворотов) вокруг произвольной точки эквивалентен однократному повороту вокруг некоторой оси, проходящей через эту точку.

Иногда эта теорема формулируется так: если заданы две прямоугольные системы координат с общим началом и произвольными направлениями осей, то всегда существует такая прямая, проходящая через начало координат, что одна система координат может быть совмещена с другой посредством поворота вокруг этой прямой.

Рассмотрим следующий пример

Дано: Ось, заданная вектором \vec{u} , и произвольная точка P которая после поворота на угла вокруг оси \vec{u} должна на перейти в точку Q . Вектор \vec{u} может иметь любое направление (рис. 2.17).

Найти: матрицу, которая описывает этот поворот.

Эта задача может быть решена двумя различными способами: классическим и конструктивным.

Классический способ (рис. 2.18).

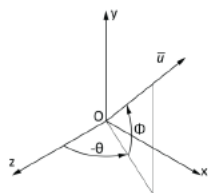


Рис. 2.18

Разобьем искомый поворот на последовательность элементарных поворотов:

1. Выполним два поворота таким образом, чтобы вектор \vec{u} совместился с осью x .
2. Выполним - вращение на угол α .
3. Аннулируем два совмещающих поворота, чтобы вернуть вектор \vec{u} в исходное положение.

Следовательно, нужное преобразование требует

перемножения пяти матриц:

$$R(\alpha) = R_y(-\theta)R_z(\varphi)R_x(\alpha)R_z(\varphi)R_y(\theta.)$$

Конструктивный способ (рис. 2.19).

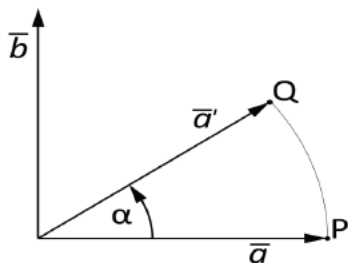
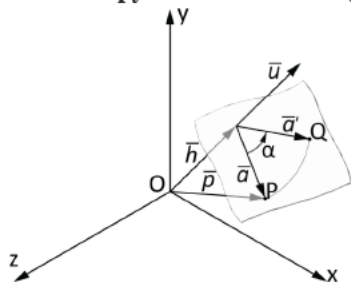


Рис.2.19

Для простоты примем, что вектор \vec{u} орт $|\vec{u}| = 1$. Обозначим через \vec{p} - радиус вектор точки P $\vec{p} = P - 0$, где 0 - начало координат. Спроецируем вектор \vec{p} на направление \vec{u}

Для определения вектора h :

$$pr_{\vec{h}}\vec{p} = (\vec{p} \cdot \vec{u}) \Rightarrow \vec{h} = (\vec{p} \cdot \vec{u})\vec{u} .$$

$$\text{Тогда } \vec{a} = \vec{p} - \vec{h} .$$

Введем вектор \vec{b} в перпендикулярный \vec{a} и лежащий с ним в одной плоскости:

$$\vec{b} = \vec{u} \cdot \vec{a}.$$

$$\text{Т.к. } |\vec{u}| = 1, \text{ то } |\vec{b}| = |\vec{a}| |\vec{u}| \perp \vec{u} \text{ и } \vec{b} \perp \vec{a}' \text{ и } \vec{b} = \vec{u} \cdot (\vec{p} - \vec{h}) = \vec{u} \cdot \vec{p} - \vec{u} \cdot \vec{h}$$

$$\text{Т.к. } \vec{u} \text{ и } \vec{h} \text{ лежат на одной прямой (} \vec{u} \parallel \vec{h} \text{), то } \vec{u} \cdot \vec{h} = 0 \text{ и } \vec{b} = \vec{u} \cdot \vec{p}.$$

Таким методом можно эффективно задавать прямоугольную систему координат в плоскости вращения (или в любой другой плоскости, для которой известен вектор нормали).

$$\text{Повернем вектор } \vec{a} \text{ до } \vec{a}' = \vec{a} \cos \alpha + \vec{b} \sin \alpha.$$

Повернутая точка определяется выражением:

$$Q = h + a \cos \alpha + \vec{b} \sin \alpha.$$

Подставив в последнее уравнение приведенные выше значения, получим:

$$Q = \vec{p} \cos \alpha + (1 - \cos \alpha)(\vec{p} \cdot \vec{u})\vec{u} + \sin \alpha(\vec{u} \times \vec{p})$$

Каждое из трех слагаемых пропорционально вектору \vec{p} . Преобразуем каждый член последнего уравнения таким образом, чтобы вместо вектора получить точку:

$$\vec{p} \cos \alpha = I (\cos \alpha) P$$

где I - единичная матрица 3×3 .

Скалярное произведение двух векторов $\vec{p} \cdot \vec{u}$ может быть представлено в форме произведения точки P на матрицу $(\vec{u}^T P)$, причем

$$(\vec{p} \cdot \vec{u})\vec{u} = \vec{u}^T u P,$$

где $\vec{u}^T u P$ тензорное произведение.

Векторное произведение двух векторов $\vec{u} \times \vec{p}$ может быть записано в форме произведения точки на матрицу

$$\vec{u} \times \vec{p} = \text{Cross}(\vec{u}) P,$$

$$\text{где } \text{Cross}(\vec{u}) = \begin{bmatrix} 0 & u_z & -u_y \\ -u_z & 0 & u_x \\ u_y & -u_x & 0 \end{bmatrix}.$$

Собирая эти члены вместе получим матрицу

$$M = \cos \alpha I + (1 - \cos \alpha) \vec{u}^T \vec{u} + \sin \alpha \text{Cross}(\vec{u}),$$

следовательно, матрица M является суммой трех взвешенных матриц.

Окончательным результатом является матрица

$$R_u(\alpha) = \begin{bmatrix} c + (1-c)u_x^2 & (1-c)u_y u_x - s u_z & (1-c)u_z u_x + s u_y & 0 \\ (1-c)u_x u_y + s u_z & c + (1-c)u_y^2 & (1-c)u_z u_y - s u_x & 0 \\ (1-c)u_x u_z - s u_y & (1-c)u_y u_z + s u_x & c + (1-c)u_z^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

где $c = \cos \alpha$, $s = \sin \alpha$, а (u_x, u_y, u_z) — компоненты единичного вектора \vec{u} .

Основные свойства трехмерных аффинных преобразований

При аффинных преобразованиях

- аффинные комбинации точек сохраняются;
- прямые линии и плоскости сохраняются;
- сохраняется параллельность прямых и плоскостей

■ столбцы матрицы показывают преобразовывающий координатный фрейм. Если столбцами матрицы M являются векторы m_1, m_2, m_3 и точка m_4 то преобразование отображает фрейм из (I, j, k, ϑ) в (m_1, m_2, m_3, m_4) ;

- относительные пропорции сохраняются;
- влияние преобразований на объемы объектов

$$\frac{\text{объем после преобразования}}{\text{объем до преобразования}} = |\det M|;$$

■ Каждое аффинное преобразование состоит из элементарных операций. Трехмерное аффинное преобразование может быть разложено на элементарные преобразования, причем несколькими различными способами.

■ Каждое аффинное преобразование может рассматриваться как последовательность элементарных преобразований (масштабирования S , поворота R , двух сдвигов H_1, H_2), за которыми следует перемещение.

$$M = SRH_1H_2$$

Используемые классы преобразований.

Движения жесткой конструкции.

$$T_{\text{движ}} = \{\text{повороты, отражения, перемещения}\}$$

Интуитивно понятно, что перемещения объекта, его повороты и отражения относительно плоскости не изменяют его форму и размер. Перечисленные преобразования имеют в однородных координатах **ортогональные** матрицы (для которых инверсия равносильна транспонированию). $\widehat{M}^{-1} = \widehat{M}^T$.

Конформные преобразования (преобразования сохраняющие углы)

$$T_{\text{конф}} = \{\text{повороты, отражения, перемещения, равномерные масштабирования}\}$$

Если данное преобразование не изменяет углы то оно не изменяет и скалярные произведения, что важно, например в случае вычисления освещенности.

2.7 Плоские проекции трехмерных объектов

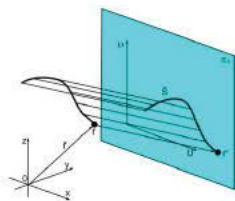


Рис. 2.20

Вычерчивание трехмерных объектов, независимо от того на бумаге ли это происходит или на экране дисплея, осуществляется при помощи двумерных проекций. В плоской проекции каждая точка предмета проецируется определенным образом на плоскость проекции, и ее образ называется точкой проекции. Если **проецирующие прямые**, соединяющие точки предмета с соответствующими **точками проекции**, **параллельны**, то мы имеем **плоскую** параллельную проекцию, как показано на рис. 2.20.

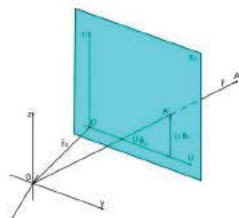
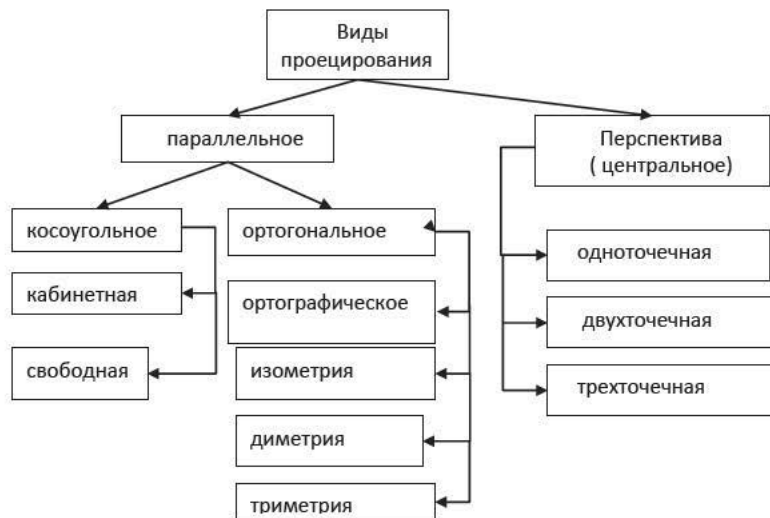


Рис. 2.21

Если же линии проекции сходятся в одной общей точке то, получаемое изображение называется **центральной проекцией**, или **перспективным изображением** (рис. 2.21).

Каждый вид проецирования подразделяется на подвиды в зависимости от взаимного расположения изображаемого объекта, картинной плоскости и способа построения (рис. 2.22).



При конструировании поверхностей и в специальных случаях применяются другие виды проекций, такие как криволинейное проецирование (смотри, например, А. М. Тевлин «Методы нелинейных отображений и их технические приложения», Москва, МАИ. - 1971, 136 с.)

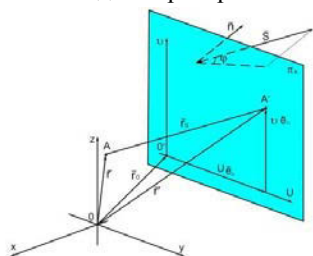
2.7.1. Косоугольное параллельное проецирование



Рис. 2.22. Виды проекций

Примеры косоугольной горизонтальной изометрической проекции

Введем пространственную систему координат $Oxyz$ в которой заданы (рис. 2.23.), (проводя аналогию с решением задач это то что дано, известно):



■ π_k - плоскость проекций (картинная плоскость) в которой строится изображение. Она определяется выбором точки O' с радиус вектором \vec{r}_0 и единичным вектором $\vec{n} \perp \pi_k$:

$$\vec{n}(\cos \alpha_k, \cos \beta_k, \cos \gamma_k | \cos^2 \alpha_k + \cos^2 \beta_k + \cos^2 \gamma_k = 1).$$

Рис. 2.23

■ единичный вектор $\vec{s}(\cos \alpha_s, \cos \beta_s, \cos \gamma_s)$ задающий направление проецирования. Вектор \vec{s} направлен под углом φ к плоскости проекций π_k . При этом между векторами \vec{s} и \vec{n} будет угол $(90 - \varphi)$ и через скалярное произведение он определяется следующим образом:

$$\vec{s} \cdot \vec{n} = |\vec{s}| |\vec{n}| \cos(90 - \varphi) \Rightarrow |\vec{s}| = |\vec{n}| = 1 \Rightarrow \sin \varphi = \vec{s} \cdot \vec{n} \Rightarrow \sin \varphi = \cos \alpha_s \cos \alpha_k + \cos \beta_s \cos \beta_k + \cos \gamma_s \cos \gamma_k. \quad (2.10)$$

■ геометрический объект, изображение которого строится. На рис. 2.23 это точка A , положение которой в пространстве задается радиус вектором \vec{r} .

Для построения изображения необходимо определить положение точки A' - параллельной проекции точки A на плоскость π_k . Точка A' определяется радиус вектором \vec{r}' в пространственной системе координат $OXYZ$ и радиус векто-

ром \vec{r}_k в системе координат $0'u'vw$ связанной с плоскостью π_k так, что оси $0'u$ и $0'v$ лежат в плоскости π_k а $0'w \parallel \vec{s}$. В общем случае это косоугольная система координат, хотя $0'u \perp 0'v$, но $0'w \perp \pi_k$.

Из рис.4. следует:

$$\vec{r}' = \vec{r} + \vec{r}_s \quad (2.11)$$

\vec{r}_s — вектор с началом в точке A , заканчивающийся в A' . Пусть координаты $(\cdot)A'$ в системе $0'_{uvw}$ — $A'(u, v, w)$, тогда т.к. направление проецирования \vec{s} и направление оси $0'w$ - противоположные, то $\vec{r}_s = -w\vec{s}$

$$\vec{r}' = \vec{r}_0 + u\vec{e}_u + v\vec{e}_v \quad (2.12)$$

Подставив \vec{r}' из (2.13) в уравнение (2.11) получим

$$\vec{r} - \vec{r}_0 = u\vec{e}_u + v\vec{e}_v + w\vec{e}_w \quad (2.13)$$

Нам необходимо определить u, v и w - координаты точки A' . Для этого умножим скалярно векторное уравнение (2.13) на вектор $(\vec{e}_u \times \vec{e}_v)$, перпендикулярный и к \vec{e}_u и к \vec{e}_v (а значит слагаемые $\vec{e}_u \cdot (\vec{e}_u \times \vec{e}_v)$ и $\vec{e}_v \cdot (\vec{e}_u \times \vec{e}_v)$ равны нулю)

$$w = \frac{(\vec{r} - \vec{r}_0)(\vec{e}_u \times \vec{e}_v)}{\vec{s}(\vec{e}_u \times \vec{e}_v)} \quad (2.14)$$

Умножив скалярно уравнение (2.13) сначала на вектор $(\vec{e}_u \times \vec{s})$, а затем на $(\vec{s} \times \vec{e}_v)$, учитывая, что циклическая перестановка $(\vec{s} \rightarrow \vec{e}_u \rightarrow \vec{e}_v)$ символов не изменяет значения смешанного произведения получим:

$$\begin{aligned} u &= \frac{(\vec{r} - \vec{r}_0)(\vec{s} \times \vec{e}_v)}{\vec{s}(\vec{e}_u \times \vec{e}_v)} \\ v &= \frac{(\vec{r} - \vec{r}_0)(\vec{e}_u \times \vec{s})}{\vec{s}(\vec{e}_u \times \vec{e}_v)} \end{aligned} \quad (2.15)$$

Для получения проекции заданного геометрического объекта информация относительно w не нужна. Но эта информация необходима для восстановления, в случае надобности, всех координат заданного объекта.

Хотя изображение на плоскости π_k не зависит от того, как введена система координат, а зависит только от взаимного расположения картинной плоскости π_k и направления проецирования вектора \vec{s} , формулы (2.13)-(2.15) существенно зависят от того, каким образом выбрана система координат. Не нарушая общности плоскость проекций π_k можно совместить с одной из координатных плоскостей, например с плоскостью $0xz$. Тогда

$$\pi_k: y = 0; \vec{n}(0,1,0), \quad \vec{r}_0 = 0, \vec{e}_u(1,0,0),$$

при этом ось $0'u$ совпадает с осью $0x, \rightarrow \vec{e}_u(1,0,0)$, а ось $0'v$ с осью $0z, \rightarrow \vec{e}_v(0,0,1)$, и формулы (2.14), (2.15) примут вид:

$$\begin{aligned} x' &= x - y \frac{\cos \alpha}{\cos \beta} \\ z' &= z - y \frac{\cos \gamma}{\cos \beta} \end{aligned} M_{\text{кп}} = \begin{bmatrix} 1 & -\frac{\cos \alpha}{\cos \beta} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{\cos \gamma}{\cos \beta} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.16)$$

Анализ формул (2.16) показывает, что параллельное проецирование можно рассматривать как аффинное преобразование – обобщенный сдвиг, задаваемый соотношениями: $x' = x - k_x y$, $y' = y$, $z' = z - k_z y$,

$$\text{или } P' = P + \vec{k}(k_x, 0, k_z)P, \text{ где } k_x = \frac{\cos \alpha}{\cos \beta}, k_z = \frac{\cos \beta_s}{\cos \beta}.$$

Т.е. проекция P' точки P получается как смещение точки, величина которого пропорциональна ее удалению от картинной плоскости π_K (т.е. пропорциональна координате z), а направление проецирования из (2.10) определяется выражением

$$\sin \varphi = \cos \beta_s.$$

и зависит от направления проецирования, вектора \vec{s} .

Связь между длиной отрезка $d = [(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2]^{1/2}$ прямой AB и длиной проекции $d_K = [(x'_B - x'_A)^2 + (y'_B - y'_A)^2 + (z'_B - z'_A)^2]^{1/2}$ этого отрезка определяется из формулы:

$$d_K = d \left[1 - 2 \frac{\sin \varphi_a}{\sin \varphi} \cos \varphi_{sa} + \left(\frac{\sin \varphi_a}{\sin \varphi} \right)^2 \right]^{1/2}, \quad (2.17)$$

где $\angle \varphi_a$ — угол между прямой a , заданной отрезком AB , и плоскостью проекций π_K ;

$$\sin \varphi_a = \cos \alpha_a \cos \alpha_k + \cos \beta_a \cos \beta_k + \cos \gamma_a \cos \gamma_k;$$

$\vec{a}(\cos \alpha_a, \cos \beta_a, \cos \gamma_a)$ - направляющие косинусы вектора \vec{a} , задающего прямую a ;

$\angle \varphi_{as}$ — угол между направлением проецирования \vec{s} и прямой a ;

$$\cos \varphi_{sa} = \cos \alpha_a \cos \alpha_s + \cos \beta_a \cos \beta_s + \cos \gamma_a \cos \gamma_s.$$

Определим коэффициенты искажения по аксонометрическим осям:

$$e_x = \frac{1}{\sin \varphi} (\sin^2 \varphi - 2 \sin \varphi \cos \alpha_s \cos \alpha_k + \cos^2 \alpha_k)^{1/2};$$

$$e_y = \frac{1}{\sin \varphi} (\sin^2 \varphi - 2 \sin \varphi \cos \beta_s \cos \beta_k + \cos^2 \beta_k)^{1/2};$$

(2.18)

$$e_z = \frac{1}{\sin \varphi} (\sin^2 \varphi - 2 \sin \varphi \cos \gamma_s \cos \gamma_k + \cos^2 \gamma_k)^{1/2},$$

причем

$$u^2 + v^2 + w^2 = 2 + ctg^2 \varphi. \quad (2.19)$$

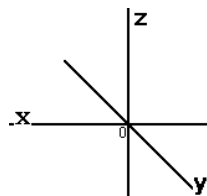


Рис. 2.24.

Косоугольная аксонометрия использовалась при ручном выполнении чертежей. При этом наиболее часто применяются

■ кабинетная проекция (фронтальная диметрия), при которой $e_x=e_y=1, e_z = 0.5$, расположение осей показано на рис. 2.24.

Задание коэффициентов искажения и взаимное расположение аксонометрических осей позволяет однозначно определить направление проецирование т.е.

вектор \vec{s} . Из формулы (9) следует, что $\operatorname{ctg} \varphi = 0.25$, что соответствует $\angle \varphi \approx 76^\circ$ и $\sin \varphi = \cos \beta_s = 0.242$

■ свободная проекция (фронтальная изометрия), при которой $e_x = e_y = e_z = 1$, при том же расположении осей.

В большинстве случаев плоскость проекций π_k расположена перпендикулярно проецирующим прямым, т.е. $\vec{s} = \vec{e}_u \times \vec{e}_v$, в этом случае формулы (2.14), (2.15) примут более простой вид:

$$u = (\vec{r} - \vec{r}_0) \vec{e}_u$$

$$v = (\vec{r} - \vec{r}_0) \vec{e}_v$$

$$w = (\vec{r} - \vec{r}_0) \vec{e}_w$$

Такое проецирование называется ортогональным. Рассмотрим его.

2.2 Ортогональное проецирование

Ортогональное параллельное проецирование подразделяется на два вида, в зависимости от расположения плоскости проекций относительно координатных плоскостей. Если плоскость проекций совмещена с одной из координатных плоскостей, имеем ортографическую проекцию. Процесс проецирования при этом заключается в том, что одна из координат становится равной 0. Так, если π_k совмещена с плоскостью Oxy , то $\pi_k: z = 0$. Проекционные матрицы в однородных координатах имеют вид:

$$M_{\text{орт}}(\pi_1: z = 0) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}, M_{\text{орт}}(\pi_2: y = 0) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix},$$

$$M_{\text{орт}}(\pi_3: x = 0) = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

Если плоскость проекций занимает общее положение, то имеем аксонометрическую проекцию. Аксонометрические проекции в свою очередь делятся на

■ триметрию, π_k наклонена ко всем трем координатным плоскостям под разными углами, а значит и коэффициенты искажения по всем трем осям координат разные;

■ диметрию, π_k равнонаклонена к двум координатным плоскостям, коэффициенты искажения по двум координатным осям равны между собой и не равны третьему;

■ изометрию, π_k равнонаклонена ко всем трем координатным плоскостям и значит коэффициенты искажения по всем трем осям координат равны между собой.

Совместим плоскость проекций π_k общего положения с координатной плоскостью $XOY (z = 0)$. Это осуществляется двумя поворотами:

Поворот на угол β вокруг оси Y , затем на угол α вокруг оси X . Учтем, что поворот вокруг оси Y осуществляется по ходу часовой стрелки (т.е. $\angle\beta$ – отрицательный). Матрица, описывающая эти два преобразования имеет вид

$$M = \begin{vmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \\ = \begin{vmatrix} \cos \beta & \sin \beta \sin \alpha & -\sin \beta \cos \alpha & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ \sin \beta & -\sin \alpha \cos \beta & \cos \alpha \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

Вместе с преобразованием проектирования на плоскость $XOY (z = 0)$ окончательная матрица имеет вид:

$$M' = M M_{\text{орт}}(z = 0) = \begin{vmatrix} \cos \beta & \sin \beta \sin \alpha & 0 & 0 \\ 0 & \cos \alpha & 0 & 0 \\ \sin \beta & -\sin \alpha \cos \beta & 0 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

Применим полученное преобразование к единичным векторам по координатным осям

x, y и z .

$$\begin{aligned} e_x &= [1 \ 0 \ 0 \ 1]M' = [\cos \beta \ \sin \beta \sin \alpha \ 0 \ 0] \\ e_y &= [0 \ 1 \ 0 \ 1]M' = [0 \ \cos \alpha \ 0 \ 0] \\ e_z &= [0 \ 0 \ 1 \ 1]M' = [\sin \beta \ -\sin \alpha \cos \beta \ 0 \ 0] \end{aligned} \quad (2.16)$$

Полученные формулы (10) позволяют определить коэффициенты искажения по всем трем координатным осям

$$\left. \begin{aligned} u &= \sqrt{\cos^2 \beta + \sin^2 \alpha \sin^2 \beta} \\ v &= \cos \alpha \\ w &= \sqrt{\sin^2 \beta + \sin^2 \alpha \cos^2 \beta} \end{aligned} \right\} \quad (2.17)$$

Из теории аксонометрии известно, что

$$u^2 + v^2 + w^2 = 2 \quad (2.18)$$

При выполнении триметрии углы α и β независимы и выбираются из условий построения требуемого изображения.

Диметрия характеризуется одинаковыми коэффициентами искажения по двум осям.

Пусть

$$u = v \text{ } \rho \cos^2 \alpha = (\sin^2 \beta + \sin^2 \alpha \cos^2 \beta) \Rightarrow \sin^2 \beta = \frac{\sin^2 \alpha}{1 - \sin^2 \alpha}, \sin \beta = \text{tg } \alpha \quad (2.19)$$

Выбрав угол α можно вычислить угол β и определить матрицу диметрической проекции.

Для стандартной диметрии, применяемой, как правило, при выполнении изображения вручную $u = v = 2w$ т.е.

$$\cos^2 \alpha = 4(\sin^2 \beta + \sin^2 \alpha \cos^2 \beta) \Rightarrow \sin^2 \beta = \frac{1-5\sin^2 \alpha}{4(1-\sin^2 \alpha)} \quad (2.20)$$

Сравнив правые части выражений (12) и (13) получим:

$$\frac{\sin^2 \alpha}{1 - \sin^2 \alpha} = \frac{1 - 5 \sin^2 \alpha}{4(1 - \sin^2 \alpha)}$$

а т.к. $1 - \sin^2 \alpha \neq 0$ то $4 \sin^2 \alpha = 1 - 5 \sin^2 \alpha \Rightarrow$

$$\sin^2 \alpha = \frac{1}{9} \sin^2 \beta = \frac{1}{8} \Rightarrow \alpha = 19,47122^\circ, \beta = 20,704881.$$

Для изометрии необходимо, чтобы все три коэффициента были равны:

$$\sin^2 \beta + \sin^2 \alpha \cos^2 \beta = \cos^2 \alpha \Rightarrow \sin^2 \beta = \frac{1 - 2\sin^2 \alpha}{1 - \sin^2 \alpha}$$

$$\sin^2 \alpha = 1 - \sin^2 \alpha \Rightarrow \sin^2 \alpha = \frac{1}{3} \sin \alpha = \frac{1}{\sqrt{3}} \Rightarrow \alpha = 35,26439^\circ$$

$$\sin^2 \beta = \frac{1}{2}, \sin \beta = \frac{1}{\sqrt{2}} \Rightarrow \beta = 45^\circ u = v = w$$

$$\sin \alpha = \frac{\sqrt{3}}{3}; \sin \beta = \frac{\sqrt{2}}{2}$$

Центральное проецирование (перспектива)

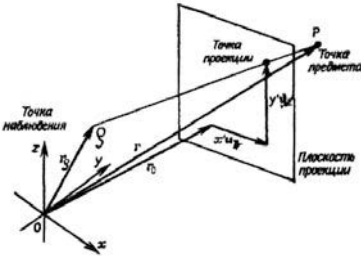


Рис. 2.25.

Рассмотрим метод центрального проецирования, лежащий в основе построения перспективных изображений.

В перспективном изображении точка проекции \vec{r}' , соответствующая заданной точке \vec{r} коллиерна с этой точкой и с точкой \vec{r}_S — точкой наблюдения (центром проецирования). Пусть \vec{r}_0 — начало системы координат, заданной на плоскости проекций (картинной плоскости, рис. 2.25), \vec{e}_u и \vec{e}_v —

единичные векторы осей координат. Тогда $\vec{r}' = \vec{r}_0 + u\vec{e}_u + v\vec{e}_v$. Т. к. точки \vec{r}_S , \vec{r}' и \vec{r} лежат на одной прямой, то $\vec{r}' = w\vec{r} + (1-w)\vec{r}_S \rightarrow \vec{r}' = \vec{r}_0 + u\vec{e}_u + v\vec{e}_v = w\vec{r} + (1-w)\vec{r}_S$ для некоторого значения w . Умножив скалярно последнее выражение на векторы $\vec{e}_v \times (\vec{r} - \vec{r}_S)$, $\vec{e}_u \times (\vec{r} - \vec{r}_S)$, $\vec{e}_u \times \vec{e}_v$, получим

$$x' = \frac{(\vec{r} - \vec{r}_S)[\vec{e}_v \times (\vec{r} - \vec{r}_S)]}{(\vec{r} - \vec{r}_S) \cdot (\vec{e}_u \times \vec{e}_v)}$$

$$y' = \frac{(\vec{r} - \vec{r}_S)[\vec{e}_u \times (\vec{r} - \vec{r}_S)]}{(\vec{r} - \vec{r}_S) \cdot (\vec{e}_v \times \vec{e}_u)}$$

$$z' = \frac{(\vec{r} - \vec{r}_S)[\vec{e}_u \times \vec{e}_v]}{(\vec{r} - \vec{r}_S) \cdot (\vec{e}_u \times \vec{e}_v)}$$

Не нарушая общности, совместим начало координат, точку O , с центром проецирования – точкой $S(O, O, O)$. Систему координат $OXYZ$ расположим так, чтобы все три координатные оси OX, OY и OZ пересекались с картинной плоскостью π_k (рис. 2.26):

$$x \cos \alpha_k + y \cos \beta_k + z \cos \gamma_k - H = 0 \quad (2.21)$$

(здесь $\cos \alpha_k, \cos \beta_k, \cos \gamma_k$ направляющие косинусы вектора \vec{n} , перпендикулярного к плоскости π_k , $OO' = H$ — расстояния от точки O до плоскости π_k). Тогда:

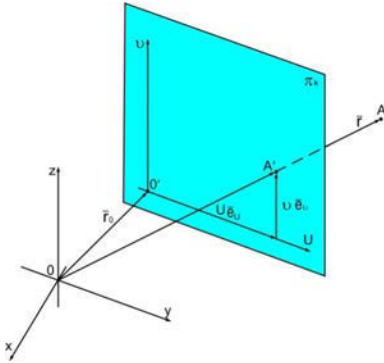


Рис. 2.26

$$M = OX \cap \pi_k \Rightarrow M \left(\frac{H}{\cos \alpha_k}, 0, 0 \right);$$

$$N = OY \cap \pi_k \Rightarrow N \left(0, \frac{H}{\cos \beta_k}, 0 \right);$$

$$P = OZ \cap \pi_k \Rightarrow P \left(0, 0, \frac{H}{\cos \gamma_k} \right).$$

Точки M, N и P называются точками схода следов. Действительно в точке M пересекаются горизонтальный (MN) и фронтальный (PM) следы плоскости π_k , в точке N пересекаются горизонтальный (MN) и профильный (PN) следы, а в точке P фронтальный (PM) и профильный (PN).

Уравнение проецирующего луча, проходящего через центр проецирования точку $S(O, O, O)$ и через текущую точку $A(X_A, Y_A, Z_A)$ имеет вид

$$(SA): \frac{x}{X_A} = \frac{y}{Y_A} = \frac{z}{Z_A}; \quad \text{или} \quad (SA): \frac{x}{\cos \alpha_A} = \frac{y}{\cos \beta_A} = \frac{z}{\cos \gamma_A}; \quad (2.22)$$

где $\vec{e}_A(\cos \alpha_A, \cos \beta_A, \cos \gamma_A)$ - единичный вектор проецирующего луча,

$$\cos \alpha_A = \frac{X_A}{R_A}; \cos \beta_A = \frac{Y_A}{R_A}; \cos \gamma_A = \frac{Z_A}{R_A}; R_A = \sqrt{X_A^2 + Y_A^2 + Z_A^2}.$$

Решая совместно уравнения (2.21) и (2.22) определяем координаты точки $A'(X'_A, Y'_A, Z'_A)$ - центральной проекции точки $A(X_A, Y_A, Z_A)$

$$X'_A = \frac{HX_A}{\rho_A}; Y'_A = \frac{HY_A}{\rho_A}; Z'_A = \frac{HZ_A}{\rho_A}, \quad (2.23)$$

где $\rho_A = X_A \cos \alpha_k + Y_A \cos \beta_k + Z_A \cos \gamma_k$.

Соотношения (2.23) можно представить в эквивалентном виде:

$$X'_A = \frac{H \cos \alpha_A}{\sin \Psi_A}; Y'_A = \frac{H \cos \beta_A}{\sin \Psi_A}; Z'_A = \frac{H \cos \gamma_A}{\sin \Psi_A},$$

где $\sin \Psi_A = \cos \alpha_A \cos \alpha_k + \cos \beta_A \cos \beta_k + \cos \gamma_A \cos \gamma_k, -\angle \Psi_A$ - угол между проецирующим лучом (SA) и картинной плоскостью π_k (на рис. 1 показан угол $v, \angle v = \angle \vec{n}_k, R_A$, тогда $\Psi_A = 90 - v$).

Рассмотрим как выглядит центральная проекция прямой $AB \equiv a$, проходящей через точки $A(x_A, y_A, z_A)$ и $B(x_B, y_B, z_B)$:

$$(AB): \frac{X - X_A}{X_B - X_A} = \frac{Y - Y_A}{Y_B - Y_A} = \frac{Z - Z_A}{Z_B - Z_A}$$

или в параметрическом виде

$$X = X_A + t \cos \alpha_a; Y = Y_A + t \cos \beta_a; Z = Z_A + t \cos \gamma_a;$$

$$\text{где } \cos \alpha_a = \frac{X_B - X_A}{d_a}; \cos \beta_a = \frac{Y_B - Y_A}{d_a}; \cos \gamma_a = \frac{Z_B - Z_A}{d_a}, \quad (2.24)$$

$$d_a = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2 + (Z_B - Z_A)^2}, \vec{n}_a(\cos \alpha_a, \cos \beta_a, \cos \gamma_a) -$$

единичный вектор в направлении прямой $AB(a)$, $\angle \alpha_a = \angle \vec{a}, \vec{X}$; $\angle \beta_a = \angle \vec{a}, \vec{Y}$; $\angle \gamma_a = \angle \vec{a}, \vec{Z}$ - направляющие косинусы орта \vec{n}_a .

Причем $X_B = X_A + d_a \cos \alpha_a$; $Y_B = Y_A + d_a \cos \beta_a$; $Z_B = Z_A + d_a \cos \gamma_a$.

Прямая a спроецируется на плоскость π_k в прямую a_k (т.к. три точки A, B, S не лежат на одной прямой и, следовательно, задают плоскость ΔSBA , а две плоскости π_k и ΔSBA пересекаются по прямой a_k - проекции прямой a).

Длина центральной проекции

$$[A'B'] \equiv d_{a'} = \sqrt{(X'_B - X'_A)^2 + (Y'_B - Y'_A)^2 + (Z'_B - Z'_A)^2}$$

связана с длиной отрезка

$$[AB] \equiv d_a = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2 + (Z_B - Z_A)^2} \text{ соотношением:}$$

$$d_{a'} = H \cdot \frac{[\sin^2 \psi_A + \sin^2 \psi_B - 2 \sin \psi_A \sin \psi_B \cos \varphi]^{\frac{1}{2}}}{\sin \psi_A \sin \psi_B}, \quad (2.25)$$

где $\sin \psi_A = \cos \alpha_A \cos \alpha_K + \cos \beta_A \cos \beta_K + \cos \gamma_A \cos \gamma_K$,

$\angle \psi_A = \angle \vec{a}, \pi_k$ - угол между прямой и плоскостью.

$$\cos \varphi = \cos \alpha_B \cos \alpha_A + \cos \beta_B \cos \beta_A + \cos \gamma_B \cos \gamma_A,$$

$\angle \varphi = \angle SA$ угол между прямой (AB) и проецирующим лучом (SA) ;

$$R_A = \sqrt{X_A^2 + Y_A^2 + Z_A^2} - \text{расстояние от центра проецирования } S \text{ до точки } A.$$

Из уравнения (3) следует: весь проецирующий луч (SA) проецируется на π_k в одну точку

определяющуюся вектором $\vec{R}_A \left(\frac{H \cos \alpha_A}{\sin \psi_A}, \frac{H \cos \beta_A}{\sin \psi_A}, \frac{H \cos \gamma_A}{\sin \psi_A} \right)$. Точка O' задается вектором $\vec{R}_O (H \cos \alpha_K, H \cos \beta_K, H \cos \gamma_K)$.

Тогда можно записать следующее уравнение:

$$\vec{R}_A = \vec{R}_O + \vec{r}_A.$$

Откуда следует

$$\vec{r}_A = \vec{R}_A - \vec{R}_O,$$

$$\text{где } \vec{r}_A \left(H \left(\frac{\cos \alpha_A}{\sin \psi_A} - \cos \alpha_K \right), H \left(\frac{\cos \beta_A}{\sin \psi_A} - \cos \beta_K \right), H \left(\frac{\cos \gamma_A}{\sin \psi_A} - \cos \gamma_K \right) \right).$$

Вектор \vec{r}_A задает положение точки A' - проекции точки A и лежит в плоскости π_k . Чтобы определить положение точки A' на плоскости π_k необходимо

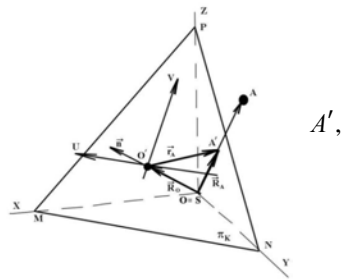


Рис. 2.27.

ввести на картинной плоскости π_K систему координат $O'UV$ следующим образом (рис. 2.27):

■ начало координат, точку $O'(H \cos \alpha_K, H \cos \beta_K, H \cos \gamma_K)$ совместим с основанием перпендикуляра, опущенного из точки O на плоскость π_K ;

■ ось $O'U$ возьмем параллельно прямой (\overline{NM}) : $\frac{X}{\frac{H}{\cos \alpha_K}} = \frac{Y - \frac{H}{\cos \beta_K}}{\frac{H}{\cos \beta_K}}$.

Тогда $\vec{e}_U \left(\frac{\cos \beta_K}{\sin \gamma_K}, -\frac{\cos \alpha_K}{\sin \gamma_K}, 0 \right)$ — единичный вектор прямой (\overline{NM}) , (орт оси $O'U$);

■ ось $O'V$ должна быть перпендикулярна оси $O'U$ и лежать в плоскости проекций π_K . Вектор \vec{V} , задающий направление оси $O'V$, определим из векторного произведения $\vec{V} = \vec{e}_U \times \vec{n}$. Выполнив операцию векторного произведения и произведя нормирование, получим $\vec{e}_V (-\cos \alpha_K \cot \gamma_K, -\cos \beta_K \cot \gamma_K, \sin \gamma_K)$ — орт оси $O'V$.

Вектор \vec{r}_A равен по величине $|\vec{r}_A| = H \cot \Psi_A$, а его направление определяется ортом $\vec{e}_A \left(\frac{\cos \alpha_A}{\cos \Psi_A} - \cos \alpha_K \tan \Psi_A, \frac{\cos \beta_A}{\cos \Psi_A} - \cos \beta_K \tan \Psi_A, \frac{\cos \gamma_A}{\cos \Psi_A} - \cos \gamma_K \tan \Psi_A \right)$ — при этом $\angle \theta_U$ — угол между ортами \vec{e}_U и \vec{e}_A можно определить из скалярного произведения ортов $\cos \theta_U = \vec{e}_U \vec{e}_A$. Откуда $\cos \theta_U = \frac{\cos \alpha_A \cos \beta_K - \cos \beta_A \cos \alpha_K}{\sin \gamma_K \cos \Psi_A}$. Аналогично $\angle \theta_V$ — угол между ортами \vec{e}_V и \vec{e}_A , $\cos \theta_V = \frac{\cos \gamma_A - \cos \gamma_K \sin \Psi_A}{\sin \gamma_K \cos \Psi_A}$. Плоские координаты проекции $A'(U_A, V_A)$ точки $A(x_A, y_A, z_A)$ определяются соотношениями:

$$\left. \begin{aligned} U_A &= |\vec{r}_A| \cos \theta_U; U_A = H \frac{\cos \alpha_A \cos \beta_K - \cos \beta_A \cos \alpha_K}{\sin \gamma_K \sin \Psi_A} \\ V_A &= |\vec{r}_A| \cos \theta_V; V_A = H \frac{\cos \gamma_A - \cos \gamma_K \sin \Psi_A}{\sin \gamma_K \sin \Psi_A} \end{aligned} \right\} \quad (2.26)$$

Используя соотношения (2.22) формулы (2.26) могут быть представлены в другом виде:

$$\left. \begin{aligned} U_A &= \frac{H}{\sin \gamma_K} \cdot \frac{X_A \cos \beta_K - Y_A \cos \alpha_K}{\rho_A} \\ V_A &= \frac{H}{\sin \gamma_K} \cdot \frac{Z_A - \cos \gamma_K \rho_A}{\rho_A} \end{aligned} \right\} \quad (2.27)$$

Во введенной плоской системе координат $O'UV$ точки схода следов имеют следующий вид

$$M \left(\frac{H \cos \beta_K}{\sin \gamma_K}, -H \operatorname{ctg} \gamma_K \right), N \left(-\frac{H \cos \alpha_K}{\sin \gamma_K \cos \beta_K}, -H \operatorname{ctg} \gamma_K \right), P(O, H \operatorname{tg} \gamma_K).$$

Рассмотрим как проецируются прямые. Уравнение проекции прямой имеет вид:

$$\begin{aligned} a: V &= k_a U + b_a, \\ k_a &= \frac{V_B - V_A}{U_B - U_A}, b_a = \frac{U_B V_A - U_A V_B}{U_B - U_A}. \end{aligned} \quad (2.28)$$

Подставив в формулы (2.26) значения координат и (формулы (2.23) с учетом (2.24)) получим

$$\left. \begin{aligned} k_a &= \frac{\sin \Psi_A \cos \gamma_a - \cos \gamma_A \sin \Psi_a}{C_a \sin \Psi_A - \sin \Psi_a C_A}, \\ b_a &= \frac{H}{\sin \gamma_K} \cdot \frac{\cos \gamma_A C_a - \cos \gamma_a C_A - \cos \gamma_K (C_a \sin \Psi_A - C_A \sin \Psi_a)}{C_a \sin \Psi_A - \sin \Psi_a C_A} \end{aligned} \right\} \quad (2.29)$$

где $\sin \Psi_a = \cos \alpha_a \cos \alpha_k + \cos \beta_a \cos \beta_k + \cos \gamma_a \cos \gamma_k$,

$\angle \Psi_a$ - угол между прямой $(AB) \equiv a$ и координатной плоскостью π_K ,

$$C_A = \cos \alpha_A \cos \beta_K - \cos \beta_A \cos \alpha_K$$

$$C_a = \cos \alpha_a \cos \beta_K - \cos \beta_a \cos \alpha_K .$$

Рассмотрим, как в общем случае выглядят центральная проекция параллельных прямых. Пусть прямые a и l параллельны в пространстве:

$$a \parallel l \Rightarrow \cos \alpha_a = \cos \alpha_l; \quad \cos \beta_a = \cos \beta_l, \quad \cos \gamma_a = \cos \gamma_l,$$

причем прямая a проходит через точку A , а прямая l через $(\cdot)L$.

Очевидно, что $A \in a \Rightarrow A_K \in a_K L \in l \Rightarrow L_K \in l_K$.

Предположим, что центральные проекции $a_K l_K$ этих прямых пересекаются в точке $D : D(U_K, V_K) = a_K \cap l_K$.

Тогда

$$\left. \begin{aligned} a_K : V &= K_a U + b_a \\ l_K : V &= K_l U + b_l \end{aligned} \right\} \Rightarrow U_D = \frac{b_a - b_l}{K_l - K_a}; \quad V_D = \frac{K_l b_a - K_a b_l}{K_l - K_a} \quad (2.30)$$

После приведения подобных и несложных преобразований получим:

$$U_D = H \frac{\cos \alpha_a \cos \beta_K - \cos \beta_a \cos \alpha_K}{\sin \gamma_K \sin \Psi_a}; \quad V_D = H \frac{\cos \gamma_a - \cos \gamma_K \sin \Psi_a}{\sin \gamma_K \sin \Psi_a} \quad (2.31)$$

Иследуем полученные формулы (2.31):

■ очевидно, что проекции прямых, параллельных в пространстве, остаются параллельными, если $\sin \Psi_a = 0 \Rightarrow$ (т.е. пересекутся в бесконечности), $\Psi_a = 0$. Но это означает, что прямые $a \parallel l$ параллельны картинной плоскости.

В этом случае соотношение (2.25) имеет вид:

$$d_{a'} = \frac{H d_a}{R_A \sin \Psi_A} \quad (2.32)$$

Т.к. $a \parallel \pi_K$, то $\sin \Psi_A = 0$.

В это соотношение параметры плоскости проекций π_K входят в виде $\frac{H}{\sin \Psi_A}$ — расстояния от проекции A' точки A до центра проецирования S , следовательно, $\frac{d_{a'}}{d_a}$ зависит только от расстояния картинной плоскости от центра проецирования и не зависит от ориентации плоскости π_K .

■ Прямые $a \parallel l$ параллельны координатной оси OX в этом случае $\cos \beta_a = \cos \gamma_a = 0$ и, следовательно $\angle \beta_a = \angle \gamma_a = \frac{\pi}{2}$, а $\cos \alpha_a = 1$, т.е. $\alpha_a = 0$, при этом $\sin \Psi_a = \cos \alpha_K$. В этом случае формулы (10) примут вид

$$U_D = \frac{H \cos \beta_K}{\sin \gamma_K \cos \alpha_K}, \quad V_D = -H \cot \gamma_K \quad (2.33)$$

и точка D совпадает с точкой M — точкой пересечения оси OX с картинной плоскостью π_K .

■ Прямые $a \parallel l$ параллельны координатной оси YO . При этом $\cos \alpha_a = \cos \gamma_a = 0$, $\angle \alpha_a = \angle \gamma_a = \frac{\pi}{2}$, $\cos \beta_a = 1$ т.к. $\angle \gamma_a = 0$, $\sin \Psi_a = \cos \gamma_K$

$$U_D = -\frac{H \cos \alpha_K}{\sin \gamma_K \cos \beta_K}, V_D = -H \cot \gamma_K \quad (2.34)$$

и точка D совпадает с точкой N — точкой пересечения оси YO с плоскостью проекций π_K .

■ Прямые $a \parallel l$ параллельны координатной оси OZ . При этом $\cos \alpha_a = \cos \beta_a = 0$, $\angle \alpha_a = \angle \beta_a = \frac{\pi}{2}$, $\cos \gamma_a = 1$ и $\angle \gamma_a = 0$, $\sin \Psi_a = \cos \gamma_K$

$$U_D = 0, \quad V_D = H \tan \gamma_K. \quad (2.35)$$

и точка D совпадает с точкой P — точкой пересечения оси OZ с картинной плоскостью π_K .

■ Прямые $a \parallel l$ параллельны вектору \vec{n} , т.е. перпендикулярны π_K .

$\cos \alpha_a = \cos \alpha_K$, $\cos \beta_a = \cos \beta_K$, $\cos \gamma_a = \cos \gamma_K$, а следовательно $\sin \gamma_K = 1$ тогда $U_D = (\cos \alpha_a \cos \beta_K - \cos \beta_a \cos \alpha_K) = 0$

$$V_D = H(\cos \gamma_a - \cos \gamma_K) = 0 \quad (2.36)$$

т.е. можно выделить еще одно семейство параллельных прямых, для которых известна точка схода следов (начало координат O' плоской системы координат). Рассмотренный случай расположения плоскости проекций π_K соответствует графической модели трехточечной перспективы. Реально же на картинной плоскости можно выделить проекции четырех совокупностей параллельных прямых и четыре точки, в которых они пересекаются. Это точки: M , в которой пересекаются проекции прямых параллельных оси OX , N — точка пересечения проекций прямых, параллельных оси OY , P — точка пересечения проекций прямых, параллельных оси OZ и точка O' — в которой пересекаются проекции прямых, параллельных вектору \vec{n}_K , задающего положение картинной плоскости.

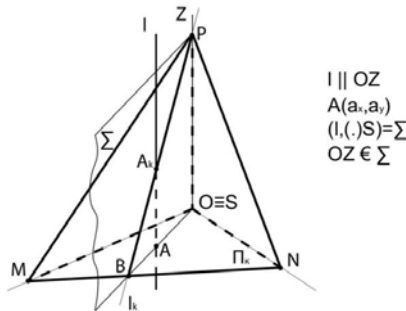


Рис. 2.28. Перспектива прямой $l \parallel OZ$

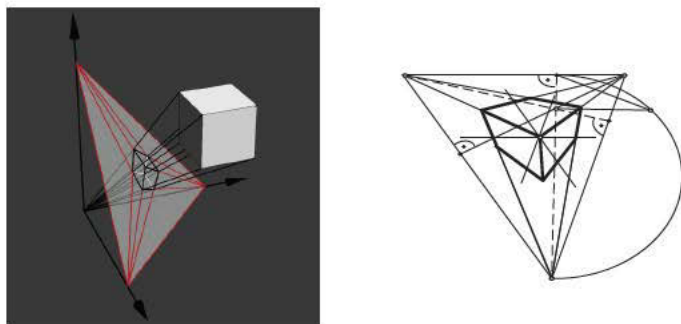


Рис. 2.29. Трех точечная перспектива и ее использование в графике

2.7.3. Перспектива

Двух точечная перспектива

В художественной и архитектурной практике при построении линейных перспектив используется следующий частный случай центрального проецирования: картинная плоскость π_K располагается между изображаемыми предметами и центром проецирования перпендикулярно предметной плоскости π_1 (плоскости XOY). В этом случае рассматриваемая аналитическая модель принимает вид (рис. 2.30):

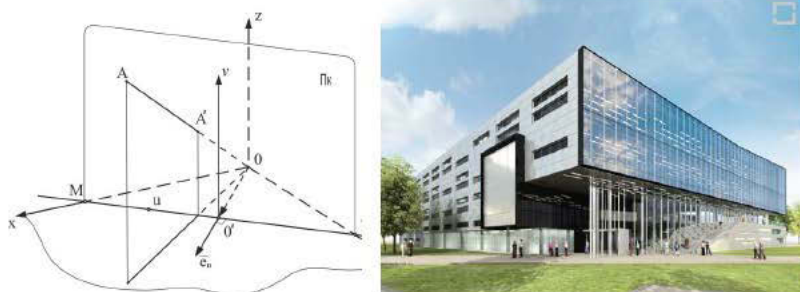


Рис. 2.30. Двухточечная перспектива и ее применение в архитектуре

- Координатная плоскость XOY параллельна предметной плоскости π (рис.31);
- Начало координатной точки O совмещена с центром проецирования точкой $S(0,0,0)$;
- Картинная плоскость $\pi_K: x \cos \alpha_K + y \cos \beta_K - H = 0$ (2.37) (очевидно, что $\alpha_K = 90 - \beta_K$), перпендикулярна координатной плоскости XOY .

Очевидно, что (2.37) получается из (2.21), если положить угол $\gamma = \frac{\pi}{2}$ ($\cos \gamma_K = 0$). Тогда

$$M = OX \cap \pi_K \Rightarrow M \left(\frac{H}{\sin \beta_K}, 0, 0 \right)$$

$$N = OY \cap \pi_K \Rightarrow N \left(0, \frac{H}{\cos \beta_K}, 0 \right).$$

С осью OZ в этом случае плоскость π_K не пересечется.

Уравнение проецирующего луча (2.22) не изменится, а вот выражения (2.23) примут вид:

$$X'_A = \frac{HX_A}{x_A \sin \beta_K + y_A \cos \beta_K}, Y'_A = \frac{HY_A}{x_A \sin \beta_K + y_A \cos \beta_K}, Z'_A = \frac{HZ_A}{x_A \sin \beta_K + y_A \cos \beta_K} \quad (2.38)$$

или

$$X'_A = \frac{H \cos \alpha_A}{\sin \Psi_A}, Y'_A = \frac{H \cos \beta_A}{\sin \Psi_A}, Z'_A = \frac{H \cos \gamma_A}{\sin \Psi_A} \quad (2.39)$$

где $\sin \Psi_A = \cos \alpha_A \sin \beta_K + \cos \beta_A \cos \beta_K$, $\angle \Psi_A$ - угол между проецирующим лучом (SA) и картинной плоскостью π_K .

Векторы \vec{R}_A , \vec{R}_0 , \vec{r}_A (рис) определяются своими компонентами:

$$\vec{R}_A \left(\frac{H \cos \alpha_A}{\sin \Psi_A}, \frac{H \cos \beta_A}{\sin \Psi_A}, \frac{H \cos \gamma_A}{\sin \Psi_A} \right),$$

$$\vec{R}_0 (H \sin \beta_K, H \cos \beta_K, 0),$$

$$\vec{r}_A \left(H \cos \beta_K \cdot \frac{\cos \alpha_A \cos \beta_K - \cos \beta_A \sin \beta_K}{\sin \Psi_A}, -H \sin \beta_K \cdot \frac{\cos \alpha_A \cos \beta_K - \cos \beta_A \sin \beta_K}{\sin \Psi_A}, H \frac{\cos \gamma_A}{\sin \Psi_A} \right).$$

Систему координат на плоскости π_K введем следующим образом:

■ Начало координат, точку $O' (H \sin \beta_K, H \cos \beta_K, 0)$ совместим с основанием перпендикуляра, опущенного из точки O на плоскость π_K ;

■ Ось $O'U$ направим по прямой (\overline{MN}) : $\frac{X}{\sin \beta_K} = \frac{Y - \frac{H}{\cos \beta_K}}{\cos \beta_K}$, тогда

$\vec{e}_U (\cos \beta_K, -\sin \beta_K, 0)$ – орт оси $O'U$;

■ Ось $O'V$ перпендикулярна оси $O'U$ и лежит в плоскости π_K . Следовательно, она параллельна оси OZ .

В этом случае координаты проекции $A' (U_A, V_A)$ точки $A (x_A, y_A, z_A)$ можно получить из

соотношений положив $\cos \beta_K = 0, \sin \beta_K = 1$.

$$\left. \begin{aligned} U_A &= H \frac{\cos \alpha_A \cos \beta_K - \cos \beta_A \sin \beta_K}{\sin \Psi_A} \\ V_A &= H \frac{\cos \gamma_A}{\sin \Psi_A} \end{aligned} \right\} \quad (2.40)$$

Или в координатной форме:

$$U_A = H \frac{x_A \cos \beta_K - y_A \sin \beta_K}{x_A \sin \beta_K + y_A \cos \beta_K}, \quad V_A = H \frac{z_A}{x_A \sin \beta_K + y_A \cos \beta_K} \quad (2.41)$$

Координаты точек схода следов примут вид $M (H \cot \beta_K, 0), N (-H \tan \beta_K, 0)$

Формулы (2.40), (2.41) можно получить и не прибегая к преобразованию координат, а непосредственно из чертежа. Т.к. из формул (2.38)

$$X'_A = \frac{HX_A}{X_A \sin \beta_K + Y_A \cos \beta_K}, \quad Y'_A = \frac{HY_A}{X_A \sin \beta_K + Y_A \cos \beta_K},$$

$$\text{то } O'A' \equiv U_A = \sqrt{X'^2_A + Y'^2_A - H^2} \Rightarrow U_A = H \frac{X_A \cos \beta_K - Y_A \sin \beta_K}{X_A \sin \beta_K + Y_A \cos \beta_K}.$$

Т.к. ось $O'V$ параллельна оси OZ , то $Z'_A = V_A$.

Поскольку нас интересует поведение проекций параллельных прямых, рассмотрим, как изменятся формулы (2.31) после подстановки в них значений (2.40)

$$U_D = H \frac{\cos \alpha_a \cos \beta_K - \cos \beta_a \sin \beta_K}{\sin \Psi_a}; \quad V_D = H \frac{\cos \gamma_a}{\sin \Psi_a}, \quad (2.42)$$

причем $\sin \Psi_a = \cos \alpha_a \sin \beta_K + \cos \beta_a \cos \beta_K$.

Из формул (2.42) следует:

■ Проекция прямых, параллельных в пространстве, будут параллельными, если $\sin \Psi_a = 0$. Следовательно, прямые $a \parallel l$ параллельны картинной плоскости.

■ Пусть прямые $a \parallel l$ параллельны координатной оси OX .

В этом случае $\cos \beta_a = \cos \gamma_a = 0$, т.е. $\angle \beta_a = \angle \gamma_a = \frac{\pi}{2}$, а $\cos \alpha_a = 1$, $\angle \alpha_a = 0$. При этом $\sin \Psi_a = \sin \beta_K$, и $U_D = H \cot \beta_K$, $V_D = 0$ и точка D совпадает с точкой $M (M = OX \cap \pi_K)$.

■ Пусть прямые $a \parallel l$ параллельны координатной оси OY . Тогда $\cos \alpha_a = \cos \gamma_a = 0$, $\angle \alpha_a = \angle \gamma_a = \frac{\pi}{2}$, $\cos \beta_a = 1$, $\angle \beta_a = 0$ и $\sin \Psi_a = \cos \beta_K$. $U_D = -H \tan \beta_K$, $V_D = 0$ и точка D совпадает с точкой $N (N = OY \cap \pi_K)$.

■ Ось OZ параллельна плоскости проекций π_K , а значит и прямые, параллельные оси OZ будут параллельны π_K ($\cos \alpha_a = \cos \beta_a = 0$, $\cos \gamma_a = 1$, $\sin \Psi_a = 0$, $U_D \rightarrow \infty$, $V_D \rightarrow \infty$).

■ Пусть прямые $a \parallel l \parallel \vec{n}_K$ параллельны вектору, перпендикулярному к плоскости проекций π_K . Тогда $\cos \alpha_a = \sin \beta_K$, $\cos \beta_a = \cos \beta_K$, $\cos \gamma_a = 0$. Следовательно, $\sin \Psi_a = 1U_D = 0$, $V_D = 0$.

Точка D совпадает с началом координат точкой $O' (D \equiv O')$. рассмотренный случай расположения картинной плоскости π_K соответствует графическому методу двухточечной перспективы – методу архитектора (рис. 2.31). В действительности в данной ситуации можно выделить проекции трех совокупностей параллельных прямых и указать на π_K три точки, в которых они пересекаются. Это точки: M, N , и O' в которых пересекаются проекции прямых параллельных координатным осям OX, OY . и вектору $\vec{n}_K \perp \pi_K$.

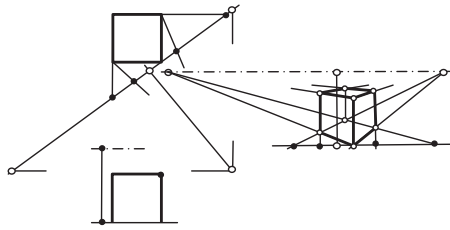


Рис. 2.31. Построение перспективы методом архитектора

Одноточечная перспектива (рис. 2.32). Рассмотрим проецирование трехмерной точки $P(x, y, z)$ в точку $P'(u, v)$ на картинную плоскость $\pi_k: z - N = 0$.

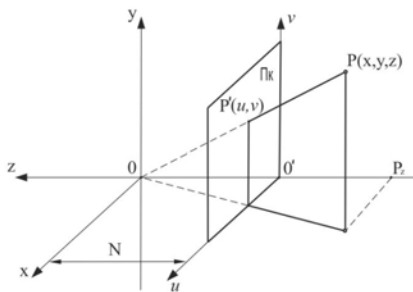


Рис. 2.32. Одноточечная перспектива

Из подобия треугольников $\Delta OP_2P_2 \sim \Delta O'P'_2P'_2$ следует $\frac{u}{P_x} = \frac{N}{-P_z}$.
 ка $P(P_x, P_y, P_z)$ проецируется на картинную плоскость π_k в точку

$$P'(u, v) = \left(N \frac{P_x}{-P_z}, N \frac{P_y}{-P_z} \right)$$

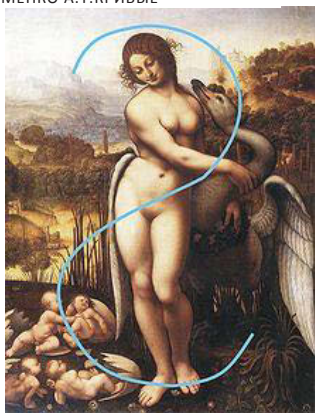
1. Проекция прямых, параллельных в пространстве будут параллельными, если прямые параллельны картинной плоскости.

2. Пусть прямые параллельны координатной оси Ox . Тогда проекции этих прямых пересекаются в точке M . В данном случае имеем одноточечную перспективу.

РАЗДЕЛ 3. КРИВЫЕ ЛИНИИ. ЗАКОНОМЕРНЫЕ КРИВЫЕ



ФОМЕНКО А.Т. КРИВЫЕ

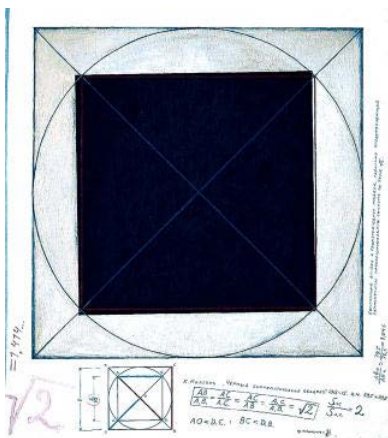


«Леда и лебедь», копия несохранившейся картины Леонардо да Винчи, S-образный изгиб — «Линия красоты» введена британским художником Уильямом Хогартом

Может вызвать удивление обращение к чувствам, когда речь идет о математических доказательствах, которые, казалось бы, связаны только с разумом. Но это означало бы, что мы забываем о чувстве математической красоты, гармонии чисел и форм, геометрической выразительности

А. Пуанкаре

Холод строгих плоскостей,
Чисел нежные кривые....
Велимир Хлебников



ПАНКИН А. ЧЕРНЫЙ КВАДРАТ 1915 Г

Линии могут дать воображению больше уму, потому что линии могут выразить отношения несоизмеримых величин, т.е. отношения которых нельзя узнать, поскольку у них нет общей меры, при помощи, которой можно было бы их сравнить.

Н. Мальбрани (1638-1715) — французский философ

Некоторые аспекты теории кривых линий рассматривались ранее в курсе аналитической геометрии. Графические алгоритмы построения коник (кривых второго порядка) приводились в курсе инженерной графики. Изучению кривых линий уделялось внимание в курсе математического анализа. Достаточно вспомнить, например, что линия служит геометрическим образом функции – важнейшего понятия анализа. Настоящий раздел посвящен вопросам общей теории алгебраических кривых. Рассматриваются различные способы задания и преобразования кривых, их свойства.

Строгое математическое определение кривой базируется на алгебраической геометрии и топологии. Будем мыслить кривую как след движущейся точки или график функции.

Рассмотрены особенности формы различных кривых.

Всякий знает, что такое кривая, пока не выучится математике настолько, что вконец запутается в бесчисленных исключениях.

Ф.Клейн

3.1. Общие сведения о кривых линиях

3.1.1 Способы образования кривых

Исследование особенностей формы кривой и ее свойств средствами дифференциальной геометрии возможно лишь, если кривая представлена в аналитической форме – уравнением. На практике, прежде чем исследовать уравнение кривой необходимо его получить, составить, базируясь на данных определяющих эту кривую и выраженных в условиях поставленной задачи. Эти условия могут быть очень разнообразными. Рассмотрим наиболее важные из них.

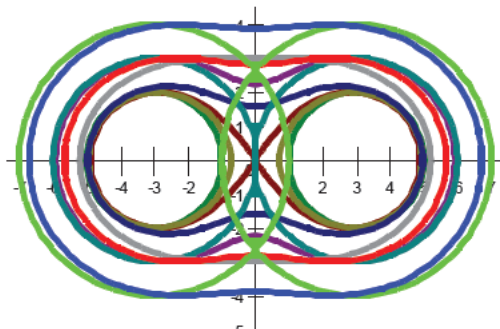


Рис 3.1. Кривые Персея

уравнение тора ($d=4, R=2$), $Y_1=p$ — секущая плоскость с системой координат $ХОУ$ ($p = 0; 1, 8; 2; 3; 4; 5; 5,5; 5,8$) $(x^2+y^2+p^2+d^2-R^2)^2=4d^2(x^2+p^2)$ –кривые Персея

■Кривая определяется как линия пересечения заданной поверхности плоскостью, положение которой известно. Так, древние греки определяли кривые 2-го порядка (коник) как сечения кругового конуса. Кривые Персея (по имени греческого геометра Персея). получаются в результате сечения плоскостью поверхности тора. На рис. 3.1 $(x_1^2+y_1^2+z_1^2+d^2 - R^2)^2=4d^2(x_1^2+y_1^2) -$

■Кривая определяется как геометрическое место точек (ГМТ), обладающих заданными свойствами.

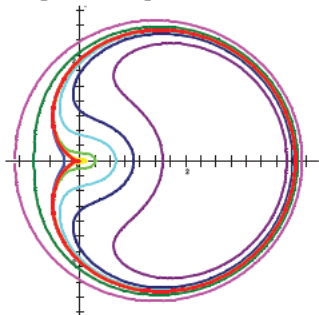


Рис. 1.2. Кривые Декарта

Так Евклид рассматривал коники как ГМТ, сохраняющих постоянное отношение расстояний от данной прямой. Другие примеры: циссоида Диоклеса, конхоида Никомеда, овалы Декарта (ГМТ, расстояние каждой из которых от двух фиксированных точек, фокусов, умноженные на данные числа, имеют постоянную сумму- такое определение дал сам Декарт. Возможно и другое

определение овалы Декарта это ГМТ, расстояние каждой из которых от трех фиксированных точек,

умноженные на некоторые числа, имеют сумму, равную нулю. Уравнение овалов Декарта имеет вид $(x^2+y^2-2rx)^2-l^2(x^2+y^2)-k=0$, r, l, k — константы.. $r=1, l=2, k=0.1; -0.1; 1; 3; -3; 9; -9$; при $k=0$ улитка Паскаля, рис. 3.2. овалы Кассини, строфонда и др.

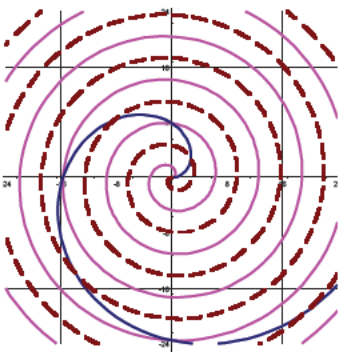


Рис. 3.3. Спираль Архимеда

■Кривая определяется как траектория точки, характер движения которой обусловлен определенным (заданным) образомКинематический способ образования линий был хорош известен греческим математикам Спираль Архимеда ($\rho=r\varphi$) — траектория точки, участвующей одновременно в двух равномерных движениях, одно из которых по прямой, другое по окружности. Циклические кривые являются траекториями точки, жестко связанной с кругом,

который катится без скольжения по окружности

другого радиуса.

■Кривая определяется по способу сопряжения проективно соответствующих элементов (проективная геометрия).

■Кривая определяется заданием ее дифференциальных свойств (цепная линия, изогнутая ось балки).

■Кривая получается в результате того или иного геометрического преобразования: (аффинного, проективного, инверсии, квадратичного, двойственного, касательного...).

■Кривая задается в аналитической форме и представляет собой график.

3.1.2 Классификация кривых

Все множество кривых можно условно разбить на две большие группы.

Закономерные кривые \Leftrightarrow графические кривые.

Кривая закономерная если ее можно задать с помощью уравнения $\Phi(X, Y, Z) = 0$. Кривая графическая, если она задана в виде графика, уравнение которого неизвестно (экспериментальные кривые), или кривая задана в виде таблицы координат ее точек. В этом случае требуется эту дискретно заданную кривую, записать с помощью функции известного класса. Это задача интерполирования, она будет рассмотрена в разделе «Построение обводов».

Естественно положить в основу классификации закономерных кривых природу их уравнений, т.к. характерные особенности формы кривой и ее свойства определяются особенностями и свойствами соответствующего ей уравнения. С учетом природы их уравнений кривые подразделяют на алгебраические и трансцендентные, заданные тригонометрическими функциями, (в прямоугольной системе координат. В полярной системе координат уравнение окружности с центром в полюсе имеет вид $\rho = a(a = const)$ - алгебраическое. Перенесем центр окружности в точку (ρ_1, φ_1) . В этом случае уравнение примет вид: $a^2 = \rho^2 + \rho_1^2 - 2\rho\rho_1 \cos(\varphi - \varphi_1)$, т.е. уравнение стало трансцендентным) и описанные натуральным уравнением. Натуральное уравнение кривой не зависит от наличия системы координат. Оно функционально связывает радиус кривизны и длину ее дуги, т.е. элементы органически связанные с самой природой исследуемой линии.

Следует иметь в виду, что трансцендентные функции могут быть разложены в бесконечный ряд Тейлора, т.е. представляют собой алгебраическую степенную функцию бесконечного порядка.

По расположению в пространстве кривые подразделяются на:

Плоские \Leftrightarrow пространственные.

Кривая плоская, если все ее точки лежат в одной плоскости

Кривая точки которой не лежат в одной плоскости –пространственная

3.2 Способы задания кривой

В дифференциальной геометрии дается следующее определение кривой:

Определение 1. Линией называется одномерный континуум, то есть связное компактное метризуемое пространство, каждая точка которого обладает сколь угодно малой окрестностью с нульмерной границей

Определение 2. Множество μ точек пространства называется элементарной кривой, если это множество является образом интервала при взаимно однозначном отображении δ , которое непрерывно само и обладает непрерывным обратным отображением δ^{-1} .

В теоретических исследованиях наиболее распространенными являются следующие способы задания кривой:

■ $\vec{r} = \vec{r}(t)$ - векторное задание кривой, \vec{r} — радиус вектор точки кривой,

■ $x = x(t), y = y(t), z = z(t)$ — параметрическое задание кривой,

- $y = f(t), z = \varphi(t)$ - явное уравнение кривой,
- $F_1(x, y, z) = 0, F_2(x, y, z) = 0$ - неявное задание кривой,
- $k = k(s), \kappa = \kappa(s)$ -натуральное задание кривой.

Кривая с отличной от 0 кривизной определяется с точностью до положения в пространстве заданием ее кривизны k и кручения κ в функции дуги s кривой – натурального параметра.

Для плоских кривых $\kappa = 0$.

Пример. Натуральное уравнение логарифмической спирали $\rho = a^{\varphi}$ имеет вид $s = kR$, где R - радиус кривизны спирали $k = \frac{1}{\ln a}$.

Неявный способ (в плоском случае) описывает кривую функцией $F(x, y) = 0$, задающей зависимость между координатами x, y . Очевидно, что если кривая задана в явном виде $y = f(x)$ или $x = \varphi(y)$, то легко перейти к заданию в неявной форме. Перенеся все члены в одну сторону знака равенства, т.е. $y - f(x) = 0$ или $x - \varphi(y) = 0$, получим задание кривой в неявном виде. Параметрическая форма напоминает движение точки с координатами, зависящими от времени t .

Рассмотрим эти способы задания кривой.

Неявное задание кривой. Начнем с плоских кривых, т.к. в этом случае именно на кривых линиях можно дать наглядную графическую интерпретацию (рис. 3.4). Точка находится на кривой тогда и только тогда, когда она удовлетворяет уравнению $F(x, y) = 0$.

Например, уравнение прямой линии, проходящей через точки $A(x_A, y_A)$ и $B(x_B, y_B)$, имеет неявную форму:

$$F(x, y) = -(x - x_A)(y_B - y_A) + (y - y_A)(x_B - x_A),$$

а окружность радиуса R с центром в точке $A(x_A, y_A)$ задается в неявном виде с помощью функции:

$$F(x, y) = (x - x_A)^2 + (y - y_A)^2 - R^2.$$

Преимущество записи в неявной форме состоит в том, что для любой точки плоскости легко проверить, лежит ли эта точка на кривой: Для этого достаточно вычислить $F(x, y)$ в исследуемой точке.

Для целого класса кривых имеет смысл понятие «внутренность» или «внешность» кривой. Таковую функцию $F(x, y)$ называют «внутренне-внешней» функцией, т.к. имеет место следующее

$F(x, y) = 0$ для всех (x, y) на кривой,

$F(x, y) > 0$ для всех (x, y) вне кривой (по одну сторону)

$F(x, y) < 0$ для всех (x, y) внутри кривой (по другую сторону).

При неявном способе, чтобы определить точки (x, y, z) на кривой приходится иметь дело с функциями вида

$$x = x; y = f(x), z = g(x)$$

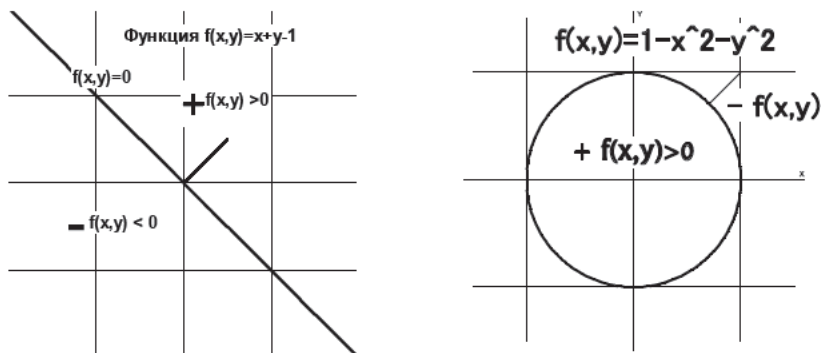


Рис. 3.4. Неявный способ задания прямой и окружности

Это представление, однако, влечет за собой многие осложнения: в некоторой точке кривой тангенс угла наклона может оказаться равным бесконечности. Непростой является и проверка, принадлежит ли точка некоторому сегменту кривой, заданной начальной и конечной точками, особенно если кривая образует петлю. Причем, при вычерчивании кривой с помощью гладкой линии необходимо учитывать наклон кривой. **Параметрическое задание кривой.** В случае параметрического представления трехмерных кривых эти проблемы существенно упрощаются, поскольку при таком представлении легко описываются замкнутые и многозначные функции, а вместо тангенсов углов наклона, которые могут быть бесконечными, используются касательные векторы, всегда имеющие конечную величину. Параметрическое задание кривой. Удобно рассматривать параметрическое задание кривой как траекторию по которой движется точка P .

Определение. Кривая (параметризованная – задана параметрически) в вещественном евклидовом пространстве R^n это отображение

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t))$$

причем предполагается, что каждая функция γ_i имеет производные всех порядков для всех $t \in I$. Переменная t называется параметром (точка $\gamma(t)$ отвечает значению параметра t).

Пример Прямая: $-(x - x_A)(y_B - y_A) + (y - y_A)(x_B - x_A) = 0$

Ее параметрическое представление имеет вид:

$$\begin{cases} x(t) = x_A + (x_B - x_A)t \\ y(t) = y_A + (y_B - y_A)t \end{cases}$$

Кривые естественно возникают как решения дифференциальных уравнений. В этих случаях параметр t имеет специальный смысл (а именно, это время). Всегда можно сделать замену одного параметра другим, не меняя кривой геометрически. Одна и та же кривая может быть задана несколькими различными параметрическими уравнениями.

Пример. Эллипс: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ параметрическое представление эллипса имеет вид:

$$\left. \begin{aligned} x(t) &= a \cos t \\ y(t) &= b \sin t \end{aligned} \right\} 0 \leq t \leq 2\pi$$

Для задания кривой, как плоской, так и пространственной достаточно одного параметра, т.е. кривая – однопараметрическое множество.

Важнейшим применением параметрического метода является описание пути, который проходит объект, перемещаясь.

Основные преимущества параметрического метода:

1. Простое вычисление координат точек по изменяющимся значениям параметра кривой (для кривой, заданной в неявном виде $f(x, y, z,) = 0$ при вычислении текущих координат приходится решать каждый раз нелинейное уравнение).

2. Упрощается расчет информации при воспроизведении кривых на графическом терминале.

3. Простота осуществления операций преобразования объекта: сдвиг, вращение, масштабирование, симметрия, т.к. используется операция преобразования векторов, непосредственно задающих кривую.

Длина дуги.

Определение. Кривая называется спрямляемой в окрестностях точки P , если эта точка имеет элементарную окрестность такую, что все правильно вписанные в нее ломаные равномерно ограничены по длине. Кривая, спрямляемая в окрестности каждой своей точки, называется спрямляемой.

Если гладкая кривая l спрямляема, и если $\vec{r} = \vec{r}(t)$ ее гладкая параметризация и $\tilde{l}(a \leq t \leq b)$ - отрезок кривой l , то длина этого отрезка вычисляется по формуле

$$s(\tilde{l}) = \int_a^b |\vec{r}'(t)| dt.$$

Пусть l - спрямляемая кривая, $\vec{r} = \vec{r}(t)$ - какая-нибудь ее параметризация. Пусть $s(t)$ - длина дуги $t_0 t$ кривой l . Определим функцию $\sigma(t)$ условиями:

$$\begin{aligned} \sigma(t) &= s(t) \text{ если } t_0 < t, \\ \sigma(t) &= -s(t) \text{ если } t_0 > t, \\ \sigma(t) &= 0. \end{aligned}$$

Определение. Функция $\sigma(t)$ строго монотонна. Поэтому σ можно принять в качестве параметра на кривой. Такую параметризацию называют естественной.

$$|\vec{r}'(s)| = 1$$

Приведем формулы для длины дуги регулярной кривой для различных случаев аналитического задания кривой.

$$1. \quad x = x(t), \quad y = y(t), \quad z = z(t)$$

$$s(t_1, t_2) = \int_{t_1}^{t_2} |\vec{r}'(t)| dt = \int_{t_1}^{t_2} \sqrt{x'^2 + y'^2 + z'^2} dt$$

$$2. \quad y = y(x), \quad z = z(x)$$

$$s(x_1, x_2) = \int_{x_1}^{x_2} \sqrt{1 + y'^2 + z'^2} dx$$

Для плоских кривых, лежащих в плоскости xu , в этих формулах надо положить $z' = 0$.

Пример. Окружность радиуса R с центром в точке $(0,0)$ допускает параметризацию

$$\gamma(t) = (R \cos t, R \sin t)$$

и естественную параметризацию

$$\gamma\left(\frac{S}{R}\right) = \left(R \cos\left(\frac{S}{R}\right)\right)$$

Чтобы перейти от параметрического задания к неявному необходимо каким-то образом исключить переменную t . Простых общих правил не существует.

3.3 Основные характеристики кривой

Касательная и нормаль. Кривизна и кручение кривой. Формулы Френе. Кривизна плоской кривой. Особые точки плоских кривых. Эволюта и эвольвента. Огибающая семейства кривых.

3.3.1. Касательная и нормаль.

К основным локальным характеристикам кривой относятся: касательная к кривой в заданной точке, нормаль и кривизна (рис. 3.5).

Определение. Кривую будем называть регулярной (k -раз дифференцируемой) если у каждой точки этой кривой есть окрестность, допускающая регулярную параметризацию, т.е. задание уравнениями в параметрической форме

$$x = f_1(t), \quad y = f_2(t), \quad z = f_3(t),$$

где f_1, f_2, f_3 – регулярные (k -раз дифференцируемые) функции. При $k=1$ такая кривая называется гладкой.

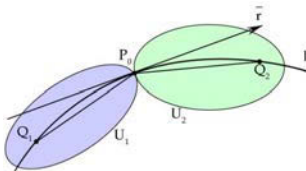


Рис. 3.5

Пусть P_0 - точка кривой l и U некоторая ее окрестность на l такая, что P_0 разбивает ее на две полуокружности U_1 и U_2 . Возьмем на кривой l вторую точку Q_1 . Проведем луч P_0Q_1 с началом P_0 .

Определение. Полукасательной к кривой в точке P_0 , соответствующей полукрестности U_1 будем называть предельное положение лучей P_0Q_1 , если оно существует, когда точка Q стремится к точке P_0 , оставаясь в полукрестности U_1 . Если в точке P_0 существуют обе полукасательные и они являются дополнительными лучами одной прямой, то эта прямая называется касательной прямой в точке P_0 к кривой l .

Пусть задана гладкая кривая в регулярной параметризации

$$\vec{r} = \vec{r}(t) \Rightarrow \begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}$$

Теорема. Гладкая кривая l в каждой своей точке имеет касательную прямую, направляющим вектором которой является вектор $\vec{r}'(t)$.

Пусть точке P_0 соответствует значение параметра t_0 , а точке Q значение параметра $t_0 + \Delta t$. Если $\Delta t > 0$ то $Q \in U_1$, а если $\Delta t < 0$, то $Q \in U_2$. Тогда $\overline{P_0Q} = \vec{r}(t_0 + \Delta t) - \vec{r}(t_0)$. Т.к. кривая гладкая, то имеет место разложение в ряд Тейлора

$$\vec{r}(t_0 + \Delta t) - \vec{r}(t_0) = \vec{r}'(t_0)\Delta t + \vec{\varphi}(\Delta t)$$

причем вектор $\vec{\varphi}(\Delta t)$ имеет бесконечно малую длину и

$$\frac{\vec{\varphi}(\Delta t)}{\Delta t} \rightarrow 0 \quad \text{при } \Delta t \rightarrow 0.$$

Тогда:

$$\frac{\vec{r}(t_0 + \Delta t) - \vec{r}(t_0)}{\Delta t} = \vec{r}'(t_0) \quad (3.1)$$

Что и требовалось доказать.

Следствие. При достаточно малом Δt кривую l можно приближенно заменить на касательную прямую. Касательная прямая является приближением кривой.

Для плоской кривой $\varphi(x, y) = 0, z = 0$ с условием $\varphi_x^2 + \varphi_y^2 \neq 0$ для определения \vec{r}'

$$\varphi_x x'_t + \varphi_y y'_t = 0.$$

В качестве касательного вектора $\vec{r}'(x'_t, y'_t)$ можно взять вектор $(-\varphi_y, \varphi_x)$.

Уравнение касательной прямой имеет вид:

$$\frac{x - x_0}{-\varphi(x_0, y_0)} = \frac{y - y_0}{\varphi(x_0, y_0)}. \quad (3.2)$$

Пусть l - гладкая пространственная кривая.

Определение. Плоскость, проходящая через точку P_0 перпендикулярно касательной к l в этой точке, называется нормальной плоскостью.

Пусть \vec{r} радиус вектор произвольной точки этой плоскости. Так как $\vec{r}'(t_0)$ - нормальный вектор плоскости и вектор $\vec{r} - \vec{r}(t_0)$ лежит в ней, то уравнение нормальной плоскости имеет вид

$$(\vec{r} - \vec{r}(t_0), \vec{r}'(t_0)) = 0 \quad (3.3)$$

Определение. Векторы, ортогональные к касательной прямой, называются нормальными кривой.

Если рассматриваемая кривая плоская, то $P(t) = x(t).y(t),-$

Параметрическое задание кривой – траектория движения точки P. Тогда

$$\vec{V}(t) = \frac{dP(t)}{dt} = \left[\frac{dx(t)}{dt}, \frac{dy(t)}{dt} \right].$$

вектор скорости, описывающий скорость направления функции P(t).

Выражение для касательной к кривой в момент $t = t_0$ в форме $L(u)$ (параметр u означает, что касательная проходит через точку $P(t_0)$ при $u=0$ и движется в направлении $\vec{V}(t_0)$) имеет вид:

$$L(u) = P(t_0) + \vec{V}(t_0)u$$

Нормаль определяется как направление, перпендикулярное к касательной в рассматриваемой точке в тот же момент t_0

$$\vec{n}(t_0) = \overrightarrow{V^{-1}}(t_0) = \left[-\frac{dy(t)}{dt}, \frac{dx}{dt} \right] \Big|_{t=t_0}$$

3.3.2. Соприкасающаяся плоскость к кривой

Пусть P_0 - точка кривой l . Возьмем на ней еще точки Q_1 и Q_2 , расположенные по разные стороны от точки P_0 .

Определение. Предельное положение плоскости, проходящей через эти три точки Q_1, P_0, Q_2 , когда две из них Q_1 и Q_2 стремятся к третьей точке P_0 . Если соприкасающаяся плоскость в точке P_0 существует, то она единственная и проходит через точку P_0 .

Теорема. Пусть $\vec{r} = \vec{r}(t)$ - дважды непрерывно дифференцируемая кривая. Если в точке P_0 векторы $\vec{r}'(t)$ и $\vec{r}''(t)$ не коллинеарны, то в этой точке существует соприкасающаяся плоскость,

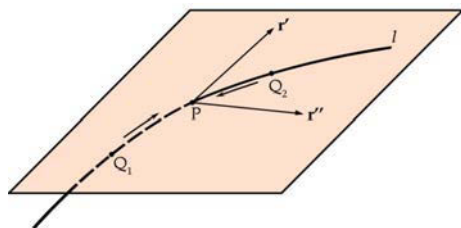


Рис. 3.2. Соприкасающаяся плоскость

и она проходит через векторы $\vec{r}'(t)$ и $\vec{r}''(t)$. Действительно, пусть P_0 соответствует параметру t_0 , а точкам Q_i параметры $t_0 + h_i$. Тогда

$$P_0Q_i = \vec{r}(t_0 + h_i) - \vec{r}(t_0)$$

Разлагая в ряд Тейлора $\vec{r}(t_0)$ в окрестностях P_0 по степеням h_i , получим

$$\vec{r}(t_0 + h_i) = \vec{r}(t_0) + \vec{r}'(t_0)h_i + \vec{r}''(t_0)\frac{h_i^2}{2} + \vec{\phi}(h_i^2),$$

Причем $\frac{\vec{\phi}(h^2)}{h^2} \rightarrow 0$ при $h_i \rightarrow 0$.

Через два вектора $\overline{P_0Q_1}$ и $\overline{P_0Q_2}$ проводим плоскость. Раскладывая в ряд Тейлора эти векторы получим

$$\begin{aligned}\overline{P_0Q_1} &= \vec{r}'(t_0)h_1 + \vec{r}''(t_0)\frac{h_1^2}{2} + \vec{\vartheta}(h_1^2) \\ \overline{P_0Q_2} &= \vec{r}'(t_0)h_2 + \vec{r}''(t_0)\frac{h_2^2}{2} + \vec{\vartheta}(h_2^2)\end{aligned}$$

В плоскости этих векторов лежат и следующие их линейные комбинации:

$$\vec{r}'(t_0) + \vec{r}''(t_0)\frac{h_1^2}{2} + \frac{1}{h_1}\vec{\vartheta}(h_1^2) \quad (3.4)$$

$$\vec{r}''(t_0) + 2\left[\frac{\vec{\vartheta}(h_1^2)}{h_1} - \frac{\vec{\vartheta}(h_2^2)}{h_2}\right] \frac{1}{R_1 - R_2}$$

Векторы (4) имеют предельные положения $\vec{r}'(t_0)$ и $\vec{r}''(t_0)$. Если эти векторы не коллинеарные, то они определяют предельное положение плоскости, проходящей через точки Q_1, P_0, Q_2 , т.е. соприкасающуюся плоскость кривой.

Определение. Точка кривой 1, в которой $\vec{r}'(t_0)$ и $\vec{r}''(t_0)$ коллинеарные, называется точкой распрямления кривой.

Соприкасающаяся плоскость по самому определению не зависит от выбора параметра на кривой. Для плоской кривой соприкасающаяся плоскость совпадает с плоскостью, в которой она лежит.

Запишем уравнение соприкасающейся плоскости.

Пусть \vec{r} - радиус вектор ее произвольной точки. Вектор $\vec{r}'(t_0) \times \vec{r}''(t_0)$ ортогонален к соприкасающейся плоскости. Вектор $\vec{r} - \vec{r}(t_0)$ лежит в ней, поэтому скалярное произведение этих двух векторов равно нулю:

$$\left(\vec{r} - \vec{r}(t_0), \vec{r}'(t_0) \times \vec{r}''(t_0)\right) = 0 \quad (3.5)$$

Это и есть уравнение соприкасающейся плоскости. В координатной форме оно имеет вид:

$$\begin{vmatrix} \tilde{x} - x(t_0) & x'(t_0) & x''(t_0) \\ \tilde{y} - y(t_0) & y'(t_0) & y''(t_0) \\ \tilde{z} - z(t_0) & z'(t_0) & z''(t_0) \end{vmatrix} = 0$$

Соприкасающаяся плоскость позволяет выделить две особые нормали к кривой.

Определение. Нормаль к кривой 1, расположенная в соприкасающейся плоскости называется **главной нормалью**. Она является прямой пересечения нормальной и соприкасающейся плоскостей.

Определение. Нормаль к кривой 1, ортогональная соприкасающейся плоскости называется **бинормалью**

3.3.3. Кривизна и кручение кривой

Дадим определение кривизны кривой. Пусть Р-точка на кривой 1 и Q — точка на 1, близкая к Р.

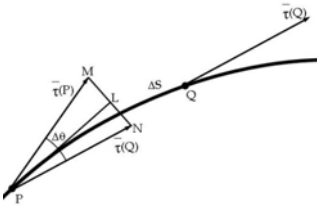


Рис. 3.7. Определение кривизны

Определение. Кривизной k кривой l в точке P называется $k = \lim_{Q \rightarrow P} \frac{\Delta\theta}{\Delta s}$

Теорема. Дважды дифференцируемая кривая $\vec{r} = \vec{r}(s)$ имеет в каждой точке кривизну k , вычисляемую по формуле

$$k = |\vec{r}_{ss}''(s)| \quad (3.6)$$

где $\vec{r}(s)$ - естественная параметризация кривой.

Определение. Единичный вектор $\vec{v} = \frac{\vec{r}_{ss}''}{k}$ называется главной

нормалью. Он определен в точках, где $k \neq 0$.

Доказательство. Пусть точкам P и Q отвечают соответственно значения естественного параметра $-s, s + \Delta s$. $\vec{\tau}(s), \vec{\tau}(s + \Delta s)$ - единичные касательные векторы к кривой в этих точках, а $\Delta\theta$ - угол между $\vec{\tau}(s)$ и $\vec{\tau}(s + \Delta s)$. Переместим вектор $\vec{\tau}(s + \Delta s)$ так, чтобы его начало совпало с точкой P . Концы векторов $\vec{\tau}(s)$ и $\vec{\tau}(s + \Delta s)$ обозначим через M и N соответственно и соединим отрезком прямой MN . Т.к. векторы $\vec{\tau}(s)$ и $\vec{\tau}(s + \Delta s)$ единичные, то ΔPMN равнобедренный. Из точки P опустим на сторону MN высоту PL . Она также будет и биссектрисой и разделит угол $\Delta\theta$ пополам. Из прямоугольного ΔPML находим

$$\frac{\vec{\tau}(s + \Delta s) - \vec{\tau}(s)}{2} = \sin \frac{\Delta\theta}{2}$$

Разделим обе части равенства на $\frac{\Delta\theta}{2}$ и перейдем к пределу при $\Delta s \rightarrow 0$.

Учтем, что $\vec{\tau}(s) = \vec{r}'_s(s)$, т.к. по условию теоремы 1 дважды дифференцируемая кривая, то существует

$$|\vec{r}_{ss}''(s)| = \lim_{\Delta s \rightarrow 0} \frac{\vec{r}'_s(s + \Delta s) - \vec{r}'_s(s)}{\Delta s} = \lim_{\Delta s \rightarrow 0} 2 \frac{\sin \left(\frac{\Delta\theta}{2} \right)}{\Delta\theta/2} \cdot \frac{\Delta\theta}{2\Delta s} = \lim_{\Delta s \rightarrow 0} \frac{\Delta\theta}{\Delta s} = k$$

Теорема доказана.

Определим кручение кривой. Вектор кривизны $\vec{r}_{ss}''(s)$, а значит и главная нормаль \vec{v} ортогональны к касательному вектору $\vec{\tau}$.

Определение. Вектор $\vec{\mu} = \vec{\tau} \times \vec{v}$ называется бинормалью кривой. Он определен в тех точках кривой, в которых определена главная нормаль \vec{v} , т.е. в точках с кривизной $k \neq 0$.

Пусть $\Delta\varphi$ угол между бинормальями $\vec{\mu}(P)$ и $\vec{\mu}(Q)$ в бесконечно близких точках P и Q (т.е. угол между соприкасающимися плоскостями) причем направление отсчета угла положительное от \vec{v} к $\vec{\mu}$, а противоположное — отрицательное.

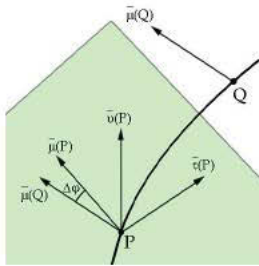


Рис. 3.8. Кривизна кривой

Перенесем вектор $\vec{\mu}(Q)$ параллельно так, чтобы его начало совпадало с точкой P, и спроецируем его на нормальную плоскость. Проекцию обозначим через $\vec{\bar{\mu}}(Q)$. Если при $\Delta S > 0$ бесконечно малый угол в нормальной плоскости от $\vec{\mu}(P)$ к $\vec{\bar{\mu}}(Q)$ положительный, то угол $\Delta\varphi$ берем положительным. Если угол от $\vec{\mu}(P)$ к $\vec{\bar{\mu}}(Q)$ отрицательный, то $\Delta\varphi$ будет отрицательным

Определение. Кривизной κ кривой в точке P называется

$$\kappa = \lim_{Q \rightarrow P} \frac{\Delta\theta}{\Delta s}$$

Теорема. Трехжды дифференцируемая кривая $\vec{r} = \vec{r}(s)$ в каждой точке с отличной от нуля кривизной κ имеет кривизну κ , вычисляемое по формуле

$$\kappa = \frac{1}{\kappa^2} (r'_s, r''_{ss}, r'''_{sss}) \quad (3.7)$$

Для произвольно параметризованной $\vec{r} = \vec{r}(s)$ ее кривизна κ вычисляется по формуле

$$\kappa = \frac{(r'_t, r''_{tt}, r'''_{ttt})}{|r'_t, r''_{tt}|^2}$$

3.3.4. Формулы Френе

Пусть кривая l задана в параметрическом виде, как вектор функции от дуги s . Введем следующие обозначения: единичный касательный вектор, — главная нормаль; — бинормаль (рис. 3.8).

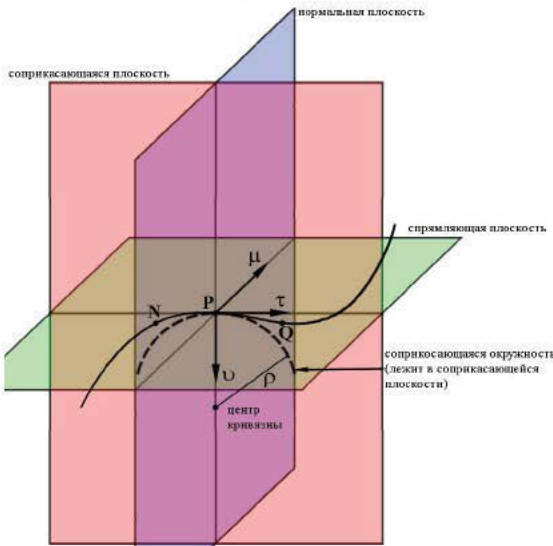


Рис. 3.5. Трехгранник Френе

Эти три вектора изменяются при изменении параметра s и являются вектор-функциями длины дуги s . В каждой точке кривой эти векторы взаимно перпендикулярны, составляют базисный репер пространства. Таким образом три вектора образуют естественный трехгранник. Запишем разложения производных этих векторов по длине дуги s по векторам этого базиса:

$$\begin{aligned}\vec{\tau}'_s &= a_1 \vec{e} + a_2 \vec{v} + a_3 \vec{\mu} \\ \vec{v}'_s &= b_1 \vec{e} + b_2 \vec{v} + b_3 \vec{\mu} \\ \vec{\mu}'_s &= c_1 \vec{e} + c_2 \vec{v} + c_3 \vec{\mu}\end{aligned}$$

найдем коэффициенты разложений. По определению вектора главной нормали имеем:

$$\frac{d^2 \vec{r}}{ds^2} = \frac{d\vec{\tau}}{ds} = k\vec{v} \Rightarrow a_1 = a_3 = 0$$

В векторном виде $\vec{r} = \vec{r}(t, c)$; $|\vec{r}'_t, \vec{r}'_c| = 0$.

Из $\vec{\mu}'_s = -k\vec{v}$ следует, что $c_1 = c_3 = 0$. Найдем производную от вектора \vec{v} , т.к. $\vec{v} = \text{орт}$, то

$$\left(\vec{v}'_s\right)_s = 0 \Rightarrow (\vec{v}, \vec{v}'_s) = b_2$$

Умножим это разложение вектора \vec{v}'_s скалярно на вектор \vec{e} :

$$b_1 = (\vec{v}'_s \cdot \vec{e}) = (\vec{v}, \vec{e})'_s - (\vec{v}, \vec{e}'_s) = -k.$$

Умножим это разложение скалярно на $\vec{\mu}$:

$$b_3 = (\vec{v}'_s \cdot \vec{\mu}) = (\vec{v}, \vec{\mu})'_s - (\vec{v}, \vec{\mu}'_s) = \alpha$$

следовательно $\vec{v}'_s = -k\vec{e} + \alpha\vec{\mu}$.

Полученные разложения векторов $\vec{\tau}'_s, \vec{v}'_s, \vec{\mu}'_s$ называются формулами Френе. Они играют фундаментальную роль в теории пространственных кривых. Выпишем эти формулы:

$$\begin{aligned}\vec{\tau}'_s &= k\vec{v} \\ \vec{v}'_s &= -k\vec{e} + \alpha\vec{\mu} \\ \vec{\mu}'_s &= -\alpha\vec{v}\end{aligned}$$

3.3.5 Кривизна плоской кривой

Плоская кривая – частичный случай пространственной кривой, поэтому для нее применима формула кривизны (3.6) справедлива, но имеет более простой вид.

Пусть кривая лежит в плоскости $z = 0$. Тогда $z' = z'' = 0$, и следовательно

$$k = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}} \quad (3.8)$$

Особенно простой вид эта формула приобретает, если в качестве параметра t взять x . Тогда $x' = 1$; $x'' = 0$. Значит

$$k = \frac{|y''|}{(1 + y'^2)^{3/2}}$$

Для плоской кривой можно определить кривизну k со знаком. Будем ее обозначать той же буквой k . Пусть \vec{e} единичный вектор касательной к гладкой кривой, а \vec{v} единичный вектор, ортогональный к плоскости кривой. Вектор

$\vec{\eta} = [\vec{e}, \vec{\tau}]$ единичный, лежит в плоскости кривой и ортогонален к $\vec{\tau}$. Он может либо совпадать с вектором главной нормали $\vec{\nu}$, либо отличаться от него знаком.

При движении точки P вдоль кривой вектор $\vec{\eta}$ изменяется непрерывно.

Определение. Кривизной k со знаком будем называть коэффициент k в формуле

$$\vec{r}_{SS}'' = \vec{\tau}'_S = k\vec{\eta}$$

где $k = \frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}}$ или $k = -\frac{|x'y'' - y'x''|}{(x'^2 + y'^2)^{3/2}}$ (3.9)

Установим выражение кривизны через производную угла α между касательной и каким-либо фиксированным направлением в плоскости, например положительным направлением оси X. В качестве параметра в формуле (3.9) возьмем длину дуги S. Вектор $\vec{\tau}\{x'_S, y'_S\}$ является единичным касательным вектором. Тогда его компоненты имеют вид

$$x'_S = \cos \alpha, y'_S = \sin \alpha$$

Определим производные:

$$x''_{SS} = -\sin \alpha \frac{d\alpha}{ds}, y''_{SS} = \cos \alpha \frac{d\alpha}{ds}$$

Подставим эти выражения в (7), получим

$$k = \frac{d\alpha}{ds} \quad \text{или} \quad k = -\frac{d\alpha}{ds}.$$

Пример. Определим кривизну окружности радиуса R. Пусть φ — угол между радиус-вектором точки окружности и положительным направлением оси X. Тогда если направление отсчета S на окружности взято против хода часовой стрелки, то $\alpha = \varphi + \frac{\pi}{2}$ кроме того, для окружности $ds = R d\varphi$, следовательно

$$k = \frac{1}{R}.$$

Найдем кривизну плоской кривой l, заданной уравнением $\Phi(x, y) = 0$. Вектор $\{\Phi_x, \Phi_y\}$ — нормаль к этой линии в точке (x, y) . Единичный вектор нормали $\vec{\eta}$ будет равен

$$\vec{\eta} = \left\{ \frac{\Phi_x}{W}, \frac{\Phi_y}{W} \right\}$$

где $W = \sqrt{\Phi_x^2 + \Phi_y^2}$. Единичный вектор $\vec{\tau}$ касательной к l будет равен

$$\vec{\tau} = \left\{ \frac{dx}{ds}, \frac{dy}{ds} \right\} = \pm \left\{ -\frac{\Phi_x}{W}, \frac{\Phi_y}{W} \right\}$$

где $\frac{d}{ds}$ — дифференцирование вдоль кривой l. С помощью формул (3.8) находим выражение для кривизны кривой

$$k = -\langle r_s, \vec{\tau} \rangle$$

Учтем, что $\eta_s = \eta_x \frac{dx}{ds} + \eta_y \frac{dy}{ds}$

$$\text{Получим: } \eta_x = \left\{ \frac{\partial}{\partial x} \cdot \frac{\Phi_x}{W}, \frac{\partial}{\partial x} \cdot \frac{\Phi_Y}{W} \right\}$$

$$\eta_y = \left\{ \frac{\partial}{\partial y} \cdot \frac{\Phi_x}{W}, \frac{\partial}{\partial y} \cdot \frac{\Phi_Y}{W} \right\}$$

Можем записать

$$k = -(\eta, \vec{\tau}) = -(\eta_x, \vec{\tau}) \frac{dx}{ds} - (\eta_y, \vec{\tau}) \frac{dy}{ds} = \left[\Phi_{xx} \left(\frac{dx}{ds} \right)^2 + 2\Phi_{xy} \frac{dx}{ds} \frac{dy}{ds} + \Phi_{yy} \left(\frac{dy}{ds} \right)^2 \right] \frac{1}{W} + \left(\Phi_x \frac{dx}{ds} + \Phi_y \frac{dy}{ds} \right) \left(W_x \frac{dx}{ds} + W_y \frac{dy}{ds} \right) \frac{1}{W^2}.$$

Последнее слагаемое в правой части равно нулю, т.к. $\Phi_x dx + \Phi_y dy = d\Phi = 0$ вдоль l . Подставляя в первое слагаемое выражения компонент вектора $\vec{\tau}$, получим

$$k = \pm \frac{\Phi_{xx}\Phi_y^2 - 2\Phi_{xy}\Phi_x\Phi_y + \Phi_{yy}\Phi_x^2}{(\Phi_x^2 + \Phi_y^2)^{3/2}}$$

Непосредственным вычислением получим правую часть в этой формуле в дивергентном виде

$$k = \pm \left(\frac{\partial}{\partial x} \frac{\Phi_x}{\sqrt{\Phi_x^2 + \Phi_y^2}} + \frac{\partial}{\partial y} \frac{\Phi_y}{\sqrt{\Phi_x^2 + \Phi_y^2}} \right).$$

3.3.6. Соприкасающаяся окружность к плоской кривой

Рассмотрим плоскую кривую l с кривизной $k \neq 0$. В малой окрестности точки P_0 кривую можно приближенно заметить ее касательной прямой в этой точке. Это ее первое приближение. Вторым приближением кривой является некоторая достаточно малая дуга соприкасающейся окружности. Расположим оси координат x, y на плоскости так, чтобы начало координат совпадало с точкой $P_0 \in l$ и ось x касалась кривой l этой точке. Пусть $y=y(x)$ – уравнение данной кривой l , а $y = \bar{y}(x)$ – уравнение некоторой окружности, проходящей через точку P_0 .

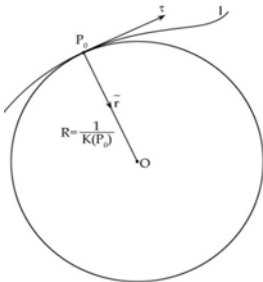


Рис. 3.9. Соприкасающаяся окружность

Определение. Окружность $y = \bar{y}(x)$ называется соприкасающейся, если в точке $x=0$ совпадают между собой значения функций $y(x)$ и $\bar{y}(x)$ и их первых и вторых производных, т.е.

$$y(0) = \bar{y}(0), y'(0) = \bar{y}'(0)$$

Отсюда следует, что разложения $y(x)$ и $\bar{y}(x)$ в ряд Тейлора

$$y(x) = y(0) + y'(0)\Delta x + y''(0) \frac{\Delta x^2}{2} + o(\Delta x^2),$$

$$\bar{y}(x) = y(0) + y'(0)\Delta x + y''(0) \frac{\Delta x^2}{2} + \bar{o}(\Delta x^2),$$

совпадают между собой с точностью до бесконечно малой $\bar{o}(\Delta x^2)$.

Поэтому из всех окружностей, проходящих через точку P_0 , соприкасающаяся окружность наиболее тесно примыкает к кривой.

Так как касательные у данной кривой и у соприкасающейся окружности при $x=0$ совпадают, т.е. соприкасающаяся окружность касается данной кривой, а ее центр O расположен на главной нормали к кривой.

Так как кривизна плоской кривой определяется уравнением

$$k = -\frac{y''(x)}{(1 + y'(x)^2)^{3/2}}$$

и т.к. первые и вторые производные $y(x)$ и $\bar{y}(x)$ совпадают, то кривизна соприкасающейся окружности равна кривизне кривой в точке P_0 . Следовательно, ее радиус $R = \frac{1}{k(P_0)}$, а центр соприкасающейся окружности расположен на главной нормали к кривой на расстоянии $\frac{1}{k(P_0)}$ от точки P_0 .

Определение. Центр соприкасающейся окружности называется центром кривизны кривой к точке P_0 .

3.3.7 Определение непрерывности кривой

Рассмотрим пример.

$$P(t) = (x(t), y(t)) = \begin{cases} (W \cos t, H \sin t) & 0 < t < a \\ (W \cos(3t - 2a), H \sin(3t - 2a)) & a < t < \frac{2(\pi + a)}{3} \end{cases}$$

$$P'(a^-) = (-W \sin(a), H \cos(a))$$

$$P'(a^+) = (-3W \sin(a), 3H \cos(a))$$

Параметрическая непрерывность - C^k -непрерывность.

Определение. Кривая $P(t)$ обладает параметрической непрерывностью k -го порядка в интервале $t \in [a, b]$, если все производные этой кривой, до k -той включительно, существуют и непрерывны во всех точках внутри $\text{int}[a, b]$ ($P(\cdot)$ является кривой гладкой на $[a, b]$).

Геометрическая непрерывность - G_1^K -непрерывность.

G_1^0 — непрерывность — 0-гладкость — $P(t)$ непрерывна по t на $[a, b]$.

G_1^1 — непрерывность на $[a, b]$ означает $P'(C^-) = kP'(C^+)$ для любого C из $[a, b]$; k — некоторая const.

G_1^2 - непрерывность на $[a, b]$ означает $P'(C^-) = kP'(C^+)$ и $P''(C^-) = mP''(C^+)$ для любого C из $[a, b]$, k и m - некоторые const.

3.3.8 Особые точки плоских кривых

Пусть l - регулярная плоская кривая и P - точка на ней.

Определение. Точка P кривой l называется обыкновенной точкой по отношению к данной степени регулярности k , если кривая допускает k раз дифференцируемую параметризацию в окрестности этой точки

$$x = x(t), \quad y = y(t),$$

удовлетворяющую в точке P условию

$$x'^2 + y'^2 \neq 0$$

Если такой параметризации не существует, то P называется особой точкой.

Определение. Точка $M(x_0, y_0)$ плоской кривой $F(x, y) = 0$ называется неособой точкой, если через нее проходит хотя бы одна прямая

$$x = x_0 + at, \quad y = y_0 + bt,$$

такая, что $t=0$ -корень уравнения

$$F(x_0 + at, y_0 + bt) = 0$$

кратности 1, и не больше. В противном случае точка (x_0, y_0) называется особой. Пусть кривая l на плоскости задана уравнением

$$\varphi(x, y) = 0 \text{ и } \varphi \in C^k, k \geq 1.$$

Определение. Точка $M(x_0, y_0)$ плоской кривой $\varphi(x, y) = 0$ называется особой точкой, если в этой точке одновременно выполняются три уравнения

$$\varphi(x_0, y_0) = 0, \quad \varphi_x(x_0, y_0) = 0, \quad \varphi_y(x_0, y_0) = 0.$$

Другими словами точка кривой называется особой, если в ней неопределенно положение касательной.

Особая точка называется - кратной, если в этой точке обращаются в ноль все частные производные до $(K - 1)$ -го порядка включительно и хотя бы одна из производных K -го порядка отлична от нуля.

В двойной точке кривая имеет две касательные, угловые коэффициенты которых являются корнями квадратного уравнения

$$F_{y_0 y_0} K^2 + 2F_{x_0 y_0} K + F_{x_0 x_0} = 0$$

где x_0, y_0 - координаты двойной точки. Двойная точка может быть узловой, изолированной, точкой возврата.

При нахождении особых точек следует придерживаться следующего плана:

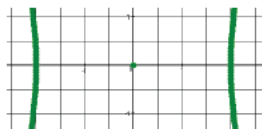
1. Найти координаты особых точек (при этом надо помнить, что они удовлетворяют системе трех, а не двух уравнений $F_x = 0$ и $F_y = 0$).
2. Определить кратность точки.
3. Если точка двойная, определить ее тип; написать уравнения касательных в этой точке и построить кривую (хотя бы в окрестности этой точки).

Полезно знать, что если кривая алгебраическая, то безо всяких вычислений можно сказать, является ли начало координат ее особой точкой (если начало координат принадлежит кривой) и записать уравнения (или уравнение, если точка обыкновенная) касательных в этой точке, для этого достаточно приравнять нулю группу членов низшей степени в уравнении кривой.

Различают несколько основных типов особых точек. Пусть в особой точке $M(x_0, y_0)$ не все производные второго порядка функции φ равны нулю. Введем обозначения:

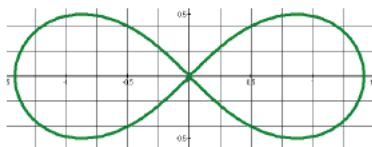
$$A = \varphi_{xx}(x_0, y_0), \quad B = \varphi_{xy}(x_0, y_0), \quad C = \varphi_{yy}(x_0, y_0), \quad \Delta = AC - B^2$$

В зависимости от значения Δ возможны три случая



а) если $\Delta > 0$, то M – изолированная точка; на рис. 3.10 уравнение кривой $y^2 - x^4 + x^2 = 0$ изолированная точка $0(0,0)$,

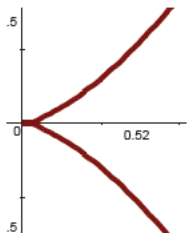
Рис.3.7. Изолированная точка



б) если $\Delta < 0$, то M - двойная точка (узел); кривая пересекает самую себя (является точкой пересечения двух ее ветвей). На рис. 3.11. кривая

Рис.3.8. Точка самопересечения

$$(x^2 + y^2 + 1)^2 - 4x^2 - 1 = 0$$



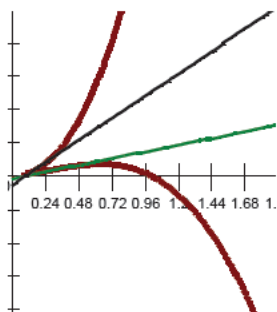
Точка $0(0,0)$ — самокасания.

в) если $\Delta = 0$, то — точка возврата первого рода.

Кривая в окрестности точки O состоит из двух ветвей, которые расположены по разные стороны от полукасательной;

На рис. 3.12. приведена полукубическая парабола $x^3 - y^2 = 0$, в точке $0(0,0)$ имеет две полукасательные

Рис. 3.9. Точка возврата первого рода

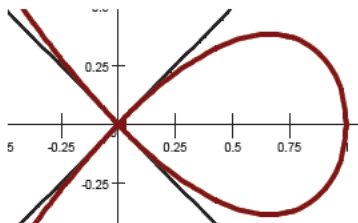


- точка возврата второго рода, в этом случае обе ветви полукасательной расположены по одну сторону от полукасательной;

На рис. 3.12. приведена кривая

Рис 3.10. Точка возврата второго рода

На рис. 3.13 точка возврата второго рода – точка самопересечения.



Уравнение кривой: $x^3 - x^2 + y = 0$, касательные $x+y=0$, $x-y=0$ в точке $0(0,0)$

Рис. 3.13. Точка самопересечения

На рис. 3.14 приведена точка самоприкосновения. (кривая состоит из двух касающихся ветвей)

Уравнение кривой $2x^4 - 3x^2y + y^3 + x^4 = 0$

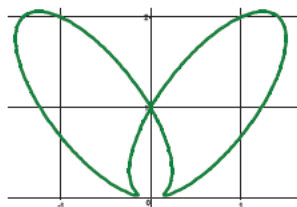


Рис. 3.14. Точка соприкосновения;

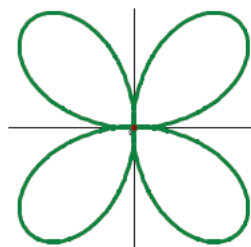
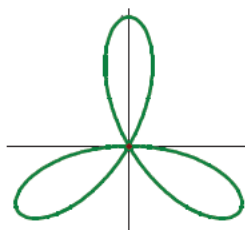


Рис. 3.15. Кратные точки(3-х и 4-х кратные точки)

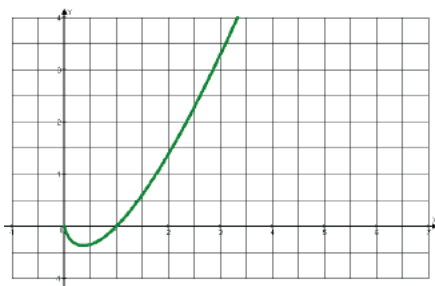
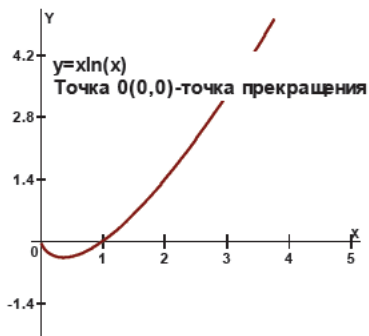
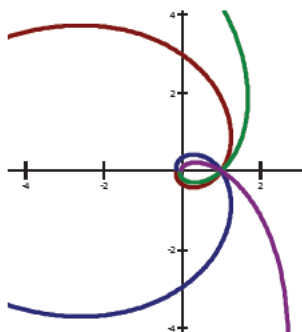


Рис. 3.16. Точка прекращения



Асимптотическая точка - особая точка кривой, вокруг которой кривая закручивается бесконечное число раз, неограниченно к ней приближаясь

Рис. 3.17. Точка излома. Точка встречи

Проведенный анализ справедлив лишь в том случае, когда функция $\varphi(x, y)$ обладает достаточной регулярностью в особой точке кривой. Если функция φ не дифференцируемая, то кривая может иметь особенности, которая не попадает под приведенную классификацию.

Определение. Если в точке M кривой L касательная имеет экстремальное значение (например $=0$), то точка M называется специальной.

К специальным точкам относятся:

- точка перегиба M – ветви k расположены по разные стороны от касательной $t, k = 0$

- вершины кривой – кривизна имеет \max или \min

Например: эллипс A, B – \max ; C, D – \min

- несобственные точки – это действительные точки пересечения кривой с несобственной прямой плоскости. Такие точки задаются асимптотами.

Например гипербола имеет 2 несобственные точки L, M .

3.3.9. Эволюта и эвольвента

Определение. Геометрическое место точек центров кривизны называется эволютой кривой.

Обозначим эволюту \tilde{l} . Дадим ее параметрическое представление. Пусть $\vec{r}(s)$ естественная параметризация кривой \tilde{l} . Тогда радиус-вектор $\vec{\rho}$ центра кривизны можно записать в виде суммы двух векторов: вектора $\vec{r}(s)$ и вектора с началом в точке P_0 и концом в центре кривизны, который коллинеарен главной нормали и имеет длину $1/k$. Следовательно, уравнение эволюты имеет вид

$$\vec{\rho}(s) = \vec{r}(s) + \frac{1}{k}\vec{v}(s)$$

Из формулы $\vec{\tau}_s = k\vec{v}$ для плоской кривой следует $\vec{v}_s = -k\vec{\tau}$. Найдем касательный вектор к эволюте:

$$\vec{\rho}'_s = \vec{\tau} - \frac{1}{k} \vec{v}' + \vec{v} \left(\frac{1}{k}\right)' = \vec{\tau} - \vec{\tau} + \left(\frac{1}{k}\right)' \vec{v} = \left(\frac{1}{k}\right)' \vec{v}.$$

Если $\left(\frac{1}{k}\right)' \neq 0$, то касательный вектор к эволюте \vec{l} кривой направлен по главной нормальной кривой γ .

Определение. Кривая \bar{l} , для которой данная кривая l является эволютой, называется эвольвентой для кривой l .

Найдем параметрическое представление эвольвенты $\vec{r} = \vec{r}(s)$, взяв в качестве параметра s длину дуги кривой l . Т.к. соответствующая точка эвольвенты расположена на касательной к кривой l в точке $\vec{r}(s)$, то

$$\vec{r} = \vec{r}(s) + \lambda(s) \vec{\tau}.$$

Кривая l - эволюта для \bar{l} , ее касательный вектор $\vec{\tau}$ направлен по главной нормали эвольвенты (рис. 3.8). Поэтому вектор $\vec{\tau}$ ортогонален к касательному вектору эвольвенты, т.е. к \vec{r}'_s . Из этого условия найдем $\lambda(s)$. Имеем

$$\vec{r}'_s = \vec{\tau} + \lambda' \vec{\tau} + \lambda k \vec{v},$$

$$0 = (\vec{r}'_s, \vec{\tau}) = 1 + \lambda'$$

Следовательно $\lambda(s) = c - s$, $c = \text{const}$.

Уравнение эвольвенты кривой имеет вид

$$\vec{r}(s) = \vec{r}(s) + (c - s) \vec{\tau}, \quad (3.10)$$

где c - произвольная постоянная. Следовательно, для данной кривой существует бесконечно много эвольвент.

Действительно, из формул (3.10) имеем

$$\vec{r}'_s = \lambda k \vec{v} = (c - s) k \vec{v}.$$

Пусть \bar{s} длина эвольвенты и \bar{k} ее кривизна, т.к. \vec{r}'_s - единичный вектор, то

$$\frac{d\bar{s}}{ds} = (c - s) k$$

причем, $c - s > 0$.

Тогда

$$\begin{aligned} \vec{r}'_s &= \vec{v} \\ \vec{r}''_{ss} &= \vec{v} \frac{ds}{d\bar{s}} = -k \vec{\tau} \frac{1}{(c - s)k} = -\frac{1}{(c - s)} \vec{\tau} \end{aligned}$$

Поэтому кривизна эвольвенты $\bar{k} = \frac{1}{(c - s)}$ и $\vec{\rho} = -\vec{\tau}$ - главная нормаль эвольвенты. Уравнение (3.10) примет вид:

$$\vec{r}(s) = \vec{r}(s) + \frac{1}{k} \vec{v},$$

т.е. кривая $\vec{r}(s)$ действительно является эволютой для $\vec{r}(s)$.

Приведем наглядный способ образования эвольвент. На кривой l возьмем точку Q и отложим от нее дугу кривой длины c . Второй конец кривой обозначим P . Пусть на эту дугу наложена гибкая нерастяжимая нить, закрепленная

одним концом в точке Q . Будем сматывать нить с кривой l , как с шаблона, за конец P , все время натягивая ее. Тогда точка P опишет эвольвенту кривой l . Действительно, на некотором участке QA нить проходит по кривой l . Пусть этот участок имеет длину S . На оставшемся участке AP нить натягивается по прямой, касающейся l . Длина этого участка равна $c - s$. Следовательно, согласно (4.9), точка P опишет эвольвенту кривой l .

Пример

Дано: Эллипс $x^2 + 4y^2 = 4 \Rightarrow (2 \cos t, \sin t)$.

Каждому значению t соответствует единственная точка эллипса. Измерим в каждой точке $p_0 = (2 \cos t, \sin t)$ насколько эллипс «круглый», т.е. насколько точно он аппроксимируется окружностью. Проведем произвольную окружность, проходящую через точку p_0 , центр которой находится в точке (a, b) . Эта окружность задается уравнением

$$g(t) = (2 \cos t - a)^2 + (\sin t - b)^2 - \mu$$

Постоянная μ выбрана, так что $g(t)$ обращалась в нуль в точке p_0 . Уравнение $g(t) = 0$ имеет одно решение $t = t_0$. Но корень $t = t_0$ может оказаться кратным (двойным, тройным т.д.). Это означает, что при $t = t_0$ несколько точек пересечения окружности и эллипса сливаются в одну. Такое явление характеризуется числом производных функции g , обращающихся в нуль при $t = t_0$.

Если $g^{(k)}(t) = 0$, то такое касание называется k -точечным.

Приравняв $g'(t) = 0$ и $g''(t) = 0$, получим:

$$\begin{aligned} g'(t) &= -4(2 \cos t - a) \sin t + 2(\sin t - b) \cos t = 0 \\ &\Rightarrow -3 \sin t \cos t + 2a \sin t - b \cos t = 0 \end{aligned}$$

$$\begin{aligned} g''(t) &= -3 \cos^2 t + 3 \sin^2 t + 2a \cos t + b \sin t = 0 \\ &\Rightarrow -3 \cos 2t \cos t + 2a \cos t + b \sin t = 0 \end{aligned}$$

систему двух уравнений с двумя неизвестными, решая которую, определим:

$$a = \frac{3}{2} \cos^3 t; \quad b = -3 \sin^3 t.$$

Таким образом, точка (a, b) при изменении t , описывает кривую $(\frac{3}{2} \cos^3 t, -3 \sin^3 t)$. Эта кривая (рис 3.18) называется эволютой эллипса и является геометрическим местом точек окружности, имеющих, по меньшей мере, трехточечное касание.

Значительно проще находится эволюта в случае, когда кривая задана параметрическими уравнениями:

$$x = \varphi(t), \quad y = \Psi(t).$$

Параметрические уравнения эволюты имеют вид

$$\xi = x - \dot{y} \frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\dot{y} - \dot{y}\dot{x}}, \quad \eta = y + \dot{x} \frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\dot{y} - \dot{y}\dot{x}}$$

Пример. Эллипс $x = a \cos t; y = b \sin t$ тогда

$$\dot{x} = -a \sin t \quad \dot{y} = b \cos t$$

$$\begin{aligned} \ddot{x} &= -a \cos t & \ddot{y} &= -b \sin t \\ \dot{x}\ddot{y} - \dot{y}\ddot{x} &= ab \sin^2 t + ab \cos^2 t = ab. \\ \xi &= a \cos t - b \cos t \frac{a^2 \sin^2 t + b^2 \cos^2 t}{ab} \\ &= a \cos t (1 - \sin^2 t) - \frac{b^2}{a} \cos^3 t = \frac{a^2 - b^2}{a} \cos^3 t \\ \eta &= \frac{a^2 - b^2}{a} \sin^3 t \end{aligned}$$

исключая параметр t получим

$$\left(\frac{a\xi}{a^2 - b^2}\right)^{2/3} + \left(\frac{b\eta}{a^2 - b^2}\right)^{2/3} = 1$$

частный случай кривых Ламе. (при $a=2, b=1$)

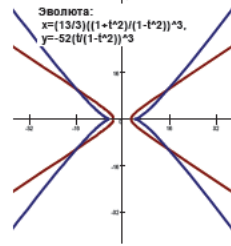
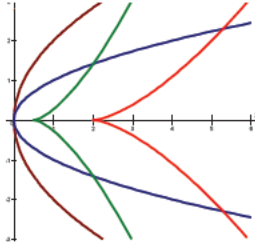
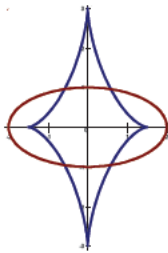


Рис.3.18.Эволюта эллипса Рис.3.19.Эволюта параболы Рис. 3.20. Гиперболаи ее эволюта

Кривая задана натуральным уравнением $R = \varphi(s)$

Пусть R и S радиус кривизны и длины дуги эвольвенты,

R_1 и S_1 радиус кривизны и длины дуги соответствующей эволюты.

$$\left. \begin{aligned} R_1 &= R \frac{dR}{dS} \\ S_1 &= R + c \end{aligned} \right\} \Rightarrow \begin{aligned} R &= S_1 \\ S &= \int \frac{S_1 dS_1}{R} \end{aligned}$$

$c = const.$ Выбором начала отсчета дуги эволюты обеспечиваем $c=0$.

Основные свойства эволют и эвольвент:

- Всякая плоская кривая имеет бесчисленное множество эвольвент.
- Через каждую точку касательной к эволюте проходит одна и только одна эвольвента.
- Касательные эвольвенты являются нормальными эвольвенты.
- Длина дуги эвольвенты равна абсолютному значению разности радиусов кривизны эвольвенты в конце ее дуги.
- Всякая плоская кривая есть геометрическое место центров кривизны своей эвольвенты.

■ Дифференциал дуги эволюты равен дифференциалу радиуса кривизны эвольвенты.

$$dR = dS$$

■ Точке экстремума радиуса кривизны эвольвенты соответствует на эволюте, в общем случае, точка возврата.

■ Если линия L задана уравнением $\vec{r} = \vec{r}(s)$, где s натуральный параметр, то уравнение семейства ее эвольвент имеет вид $\vec{\rho} = \vec{r} + (\alpha - s)\vec{t}$, где \vec{t} единичный вектор касательной L , α — произвольный параметр.

Касательная, нормаль, радиус кривизны, координаты центра кривизны при различных способах задания кривой.

■ $y = f(x)$. Явное задание.

Касательная в точке $A(x_1, y_1)$: $y - y_1 = \frac{dy_1}{dx_1}(x - x_1)$;

Уравнение нормали $A(x_1, y_1)$: $y - y_1 = -\frac{dx_1}{dy_1}(x - x_1)$;

Радиус кривизны: $r = \left[1 + \left(\frac{dy_1}{dx_1} \right)^2 \right]^{3/2} / (d^2y_1/dx_1^2)$;

Координаты центра кривизны: $O(\xi, \eta)$

$$\xi = x_1 - \left\{ \frac{dy_1}{dx_1} \left[1 + \left(\frac{dy_1}{dx_1} \right)^2 \right] \right\} / \frac{d^2y_1}{dx_1^2}$$

■ $f(x, y) = 0$. Не явное задание.

Касательная $\frac{\partial f}{\partial x_1}(x - x_1) + \frac{\partial f}{\partial y_1}(y - y_1) = 0$

Нормаль $\frac{\partial f}{\partial y_1}(x - x_1) + \frac{\partial f}{\partial x_1}(y - y_1) = 0$

Радиус кривизны

$$r = \left[\left(\frac{\partial f}{\partial x_1} \right)^2 + \left(\frac{\partial f}{\partial y_1} \right)^2 \right]^{3/2} / \left[\frac{\partial^2 f}{\partial x_1^2} \left(\frac{\partial f}{\partial y_1} \right)^2 - 2 \frac{\partial^2 f}{\partial x_1 \partial y_1} \cdot \frac{\partial f}{\partial x_1} \cdot \frac{\partial f}{\partial y_1} + \frac{\partial^2 f}{\partial y_1^2} \left(\frac{\partial f}{\partial x_1} \right)^2 \right]$$

эволюта определится уравнениями

$$\xi = x + \rho \frac{\partial f}{\partial x}; \quad \eta = y + \rho \frac{\partial f}{\partial y},$$

где ξ, η координаты точки эволюты, соответствующие точке с координатами x, y заданной кривой и

$$\rho = \frac{\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}{\delta}, \quad \delta = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial f}{\partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial y} & \frac{\partial f}{\partial x} & 0 \end{vmatrix}$$

■ $x = f(t), y = \varphi(t)$. Параметрическое задание.

$$\text{Касательная } (y - y_1) \frac{df_1}{dt} - (x - x_1) \frac{d\varphi_1}{dt} = 0$$

$$\text{Нормаль } (y - y_1) \frac{d\varphi_1}{dt} + (x - x_1) \frac{df_1}{dt} = 0$$

$$\text{Радиус кривизны } r = (df_1^2 + d\varphi_1^2)^{3/2} / (df_1 d^2\varphi_1 - d\varphi_1 d^2 f_1)$$

Координаты центра кривизны:

$$\xi = f_1 - [d\varphi_1(df_1^2 + d\varphi_1^2)] / (df_1 d^2\varphi_1 - d\varphi_1 d^2 f_1)$$

$$\eta = \varphi_1 - [df_1(df_1^2 + d\varphi_1^2)] / (df_1 d^2\varphi_1 - d\varphi_1 d^2 f_1)$$

■ $\rho = \varphi(\theta) \left[\frac{1}{\rho} = u \right] A(\rho_0, \theta_0)$. Задание в полярных координатах.

$$\text{Касательная } u_0 = u \cos(\theta_0 - \theta) + \frac{du}{d\theta} \sin(\theta_0 - \theta)$$

$$\text{Нормаль } u_0 = u \cos(\theta_0 - \theta) - \frac{d\theta}{du} \sin(\theta_0 - \theta)$$

$$\text{Радиус кривизны } r = \left[\rho^2 + \left(\frac{d\rho}{d\theta} \right)^2 \right]^{3/2} / \left[\rho^2 + 2 \left(\frac{d\rho}{d\theta} \right)^2 - \rho \frac{d^2\rho}{d\theta^2} \right]$$

Координаты центра кривизны:

$$\xi = \rho \cos \theta - \left\{ \left[\rho^2 + \left(\frac{d\rho}{d\theta} \right)^2 \right] \left[\rho \cos \theta + \sin \theta \frac{d\rho}{d\theta} \right] \right\} / \left[\rho^2 + 2 \left(\frac{d\rho}{d\theta} \right)^2 - \rho \frac{d^2\rho}{d\theta^2} \right];$$

$$\eta = \rho \sin \theta - \left\{ \left[\rho^2 + \left(\frac{d\rho}{d\theta} \right)^2 \right] \left[\rho \sin \theta - \cos \theta \frac{d\rho}{d\theta} \right] \right\} / \left[\rho^2 + 2 \left(\frac{d\rho}{d\theta} \right)^2 - \rho \frac{d^2\rho}{d\theta^2} \right].$$

3.3.10 Огибающая семейства кривых

Наряду с отдельными кривыми для некоторых задач приходится рассматривать и семейства кривых. При этом, используя тот или иной способ задания кривых, удобно поставить в соответствие каждой кривой семейства одно или несколько чисел – параметров. По количеству параметров семейства кривых делятся на однопараметрические, двухпараметрические и т.д. На пример, каждая прямая из семейства прямых $y=c$ на плоскости x,y , параллельных оси x , определяется одним параметром. Семейство окружностей на плоскости в общем случае зависит от 3 параметров.

Рассмотрим семейство кривых на плоскости, задаваемое уравнением. Например, семейство окружностей постоянного радиуса R и с центром на оси x (рис. 3.21).

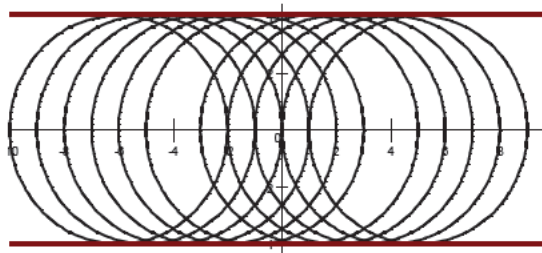


Рис. 3.21. Огибающая семейства окружностей

Уравнение такой окружности будет:

$$(x - \alpha)^2 + y^2 = R^2 \quad (3.11)$$

При фиксированном α получим уравнение некоторой конкретной окружности, а при изменении α получим семейство окружностей.

В общем случае будем считать, что уравнение

$$f(x, y, \alpha) = 0 \quad (3.12)$$

задает однопараметрическое семейство кривых на плоскости. Это означает, что при каждом фиксированном α уравнение (3.12) задает некоторую выбранную кривую l_α из семейства.

Рассматривая семейство кривых на плоскости можно увидеть, что некоторые линии на плоскости являются особыми для семейства. Так, для семейства (3.11) такими особыми линиями являются прямые $y = \pm R$.

Эти линии в каждой своей точке касаются одной из окружностей семейства, иначе говоря, огибают их.

Определение. Будем называть огибающей для семейства линий кривую Γ на плоскости, которая в каждой своей точке касается проходящей через эту точку кривой из семейства.

Найдем уравнение огибающей Γ . Пусть $P(x, y)$ – точка огибающей. Пусть через каждую точку Γ проходит только одна кривая l_α семейства, т.е. каждой точке P можно поставить в соответствие только одно значение параметра α . Тогда координаты x, y можно рассматривать как функции параметра α : $x = x(\alpha)$ и $y = y(\alpha)$.

Т.к. точка P лежит на кривой l_α , то

$$f(x(\alpha), y(\alpha)) = 0. \quad (3.13)$$

Очевидно, что кривая Γ и огибающая Γ в точке P имеет общую касательную. Дифференцируя (3.13) по α получим т.к. кривая задается уравнением (3.12), то нормаль к l_α в точке P имеет вид. Касательный вектор к Γ есть, поэтому

$$f_x(P)x'_\alpha + f_y(P)y'_\alpha = 0$$

Для точки огибающей должно выполняться и уравнение $f_\alpha(P) = 0$. Следовательно, точки огибающей должны удовлетворять следующим двум уравнениям

$$\begin{cases} f(x, y, \alpha) = 0 \\ f'_\alpha(x, y, \alpha) = 0 \end{cases}$$

Рассмотрим пример (рис. 3.22)

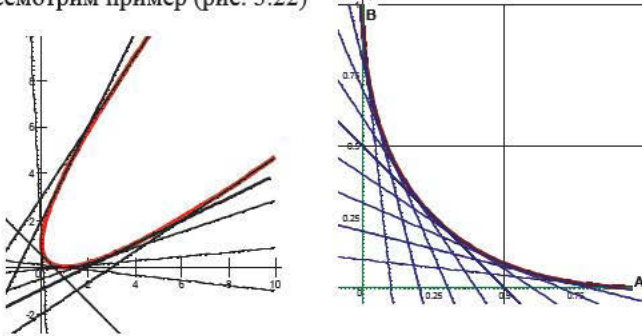


Рис. 3.22. Огибающая семейства прямых. Пропорциональная кривая

$$F(x, y, t) = tx + (1 - t)y + t(t - 1) = 0,$$

тогда

$$F'_t(x, y, t) = x - y + 2t - 1 = 0.$$

Исключив параметр t , получим уравнение огибающей в неявном виде:

$$x^2 - 2xy + y^2 - 2(x + y) + 1 = 0$$

или в параметрическом виде

$$x = (1 - t^2); \quad y = t^2$$

Если кривые семейства $F(x, y, v) = 0$ описываются параметрическими уравнениями:

$$\begin{aligned} x &= x(u, v) \\ y &= y(u, v) \end{aligned} \quad (3.14)$$

где u – параметр, определяющий положение точки на любой заданной кривой семейства, v – параметр, определяющий кривую семейства, то $F(x, y, t) = F(x(u, v), y(u, v), v) = f(u, v)$, откуда $f(u, v) = 0$

тогда:
$$\frac{df}{du} = \frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial u} = 0$$

$$\frac{df}{dv} = \frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial v} = 0$$

для всех точек любой кривой данного семейства. Более того для любой точки огибающей имеет соотношение $\frac{\partial F}{\partial v} = 0$, откуда

$$\frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial u} = 0$$

$$\frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial v} = 0$$

Рассматривая эти соотношения как уравнения для $\frac{\partial F}{\partial x}$ и $\frac{\partial F}{\partial y}$ получим:

$$\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} = 0 \quad (3.15)$$

Если $\frac{\partial F}{\partial x} = 0$ и $\frac{\partial F}{\partial y} = 0$, то касательная к кривой $F(x, y, v) = 0$ не определена, а сама кривая имеет в этом случае заострение.

Чтобы получить параметрическое уравнение огибающей семейства гладких кривых, необходимо, используя (3.15), исключить из уравнения (3.14) либо u , либо v .

Пример: $x = v + u \cos v$; $y = u \sin v$.

$$\frac{\partial x}{\partial u} = \cos v; \quad \frac{\partial y}{\partial u} = \sin v; \quad \frac{\partial x}{\partial v} = 1 - u \sin v; \quad \frac{\partial y}{\partial v} = u \cos v$$

$$\frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} = \cos v (u \cos v) - (1 - u \sin v) \sin v = u - \sin v = 0$$

$$u = \sin v.$$

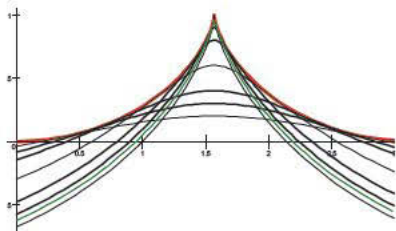


Рис. 3.23. Огибающая семейства кривых

Подставляя это значение в уравнения, определяющее рассматриваемое семейство кривых, получим параметрическое уравнение огибающей (рис. 3.23).

Решение системы уравнений:

$$F(x, y, a) = 0$$

$$F'_a(x, y, a) = 0$$

называется дискриминантой линии семейства. Если F'_x и F'_y в точках дискриминанты не обращаются в нуль, то дискриминанта совпадает с огибающей. Например, уравнение огибающей семейства кривых $(x - a)^3 + (y - a)^3 - 3(x - a)(y - a) = 0$ удовлетворяет прямая $x = y$, которая, однако, не является огибающей. Это дискриминанта, прямая, состоящая из узловых точек кривых семейства (рис. 3.24):

$$\begin{aligned} x &= f_1(x_1, y_1) \\ y &= f_2(x_1, y_1) \end{aligned} \quad (3.16.)$$

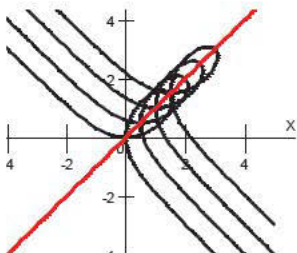


Рис. 3.24. Дискриминанта (не огибающая)



Сама идея, сама возможность выразить прямую или кривую в алгебраических терминах мне показалась не менее прекрасной, чем «Илиада».
Э. Кинэ (1803-1875) – французский писатель и публицист

Теорема 2. Две неприводимые алгебраические кривые, порядка m и n соответственно, пересекаются не более чем в mn точках. Это свойство непосредственно следует из теоремы о числе корней системы двух алгебраических уравнений, одно из которых степени m , а другое – степени n .

Теорема 3. Алгебраическая кривая n -го порядка определяется $\frac{n(n+3)}{2}$ простыми точками.

Действительно, в уравнении кривой n -го порядка имеется $\frac{(n+1)(n+2)}{2}$ коэффициентов. Однако, по крайней мере, один из коэффициентов $A_i \neq 0$ (что следует из определения кривой n -го порядка), а значит, на него можно разделить все члены уравнения (1). Следовательно, число коэффициентов, подлежащих определению, уменьшится на единицу и будет равно

$$\frac{(n+1)(n+2)}{2} - 1 = \frac{n^2 + 3n + 2 - 2}{2} = \frac{n(n+3)}{2}$$

Таким образом, задавая $\frac{n(n+3)}{2}$ точек кривой, получим систему $\frac{n(n+3)}{2}$ уравнений, решая которую можно определить коэффициенты заданной кривой.

Теорема 4. Точка алгебраической кривой кратности i эквивалентна заданию $\frac{i(i+1)}{2}$ простых ее точек.

Алгебраическая кривая определяется уравнением (1), если она не проходит через начало координат. Если кривая проходит через начало координат, то в ее уравнении будет отсутствовать свободный член (подставляя $x=0, y=0$ в (1) получим $M=0$). Если точка $0(0,0)$ - начало координат, является для кривой i -кратной точкой, то инцидентная точке 0 прямая l

$$y = kx \tag{4.2}$$

Будет пересекать кривую еще в $(n - i)$ точках, отличных от 0 . Рассматривая совместно уравнения (1) и (2) получим приводимое уравнение n -й степени вида

$$x^i(A_0x^{n-1} + A_1^{n-1} + \dots + A_{n-i-1}x' + A_{n-1}) = 0$$

Остальные все коэффициенты последних i строк равны 0 . Очевидно, что число этих коэффициентов равно

$$\frac{i(i+1)}{2}$$

Из этой теоремы следует, например, что кривая третьего порядка ($n=3$) определяется или девятью простыми точками или одной двойной ($i=2$) и шестью простыми точками.

Приведем еще ряд теорем без доказательств.

Теорема 5. Каждая кривая n -го порядка проходящая через $\frac{n(n+3)}{2} - 1$ точек, проходит также еще через $\frac{(n-1)(n-2)}{2}$ определенных точек плоскости, положение которых зависит от положения заданных точек.

Теорема 6. Нераспадающаяся плоская алгебраическая кривая n -го порядка может иметь не более $\frac{(n+1)(n+2)}{2}$ двойных точек.

Теорема 7. Точка алгебраической кривой крайности i эквивалентна $\frac{i(i+1)}{2}$ двойным точкам.

4.2. Понятие о тангенциальных координатах

Числа (ξ, η) называются тангенциальными (линейными, плюкеровыми) координатами. Уравнение $F(\xi, \eta) = 0$ определяет семейство прямых которое, в общем случае, имеет огибающую, зависящую от вида функции $F(\xi, \eta)$.

Уравнение

$$\xi x + \eta y + 1 = 0 \quad (4.3)$$

определяет прямую (ξ, η) на плоскости. Координаты ξ, η в этом уравнении фигурируют симметрично с текущими координатами (x, y) . Будем рассматривать ξ, η как координаты прямой, равноправные с координатами x, y точки, и при соответствующих обстоятельствах считать их переменными вместо x, y . При таком истолковании x, y имеют постоянное значение и уравнение (3) выражает условие того, что некоторая переменная прямая проходит через неподвижную точку x, y , т.е. оно является уравнением этой точки в координатах прямой. Не оказывая предпочтения ни одному из обоих образов в способе выражения и не указывая, какая именно пара величин рассматривается как постоянная, а какая – как переменная, получим, что уравнение (3) представляет собой условие соединенного положения (инцидентности) точки и прямой.

Симметрия уравнения (4.3)

$$ix + vy + 1 = 0$$

Относительно точечных и тангенциальных координат влечет за собой соответствие (двойственность) между теоремами, относящимися к взаимному расположению точек и прямых.

Например.

Теорема. Три точки $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ лежат на одной прямой в том и только в том случае, если

$$\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix} = 0.$$

Двойственная ей:

Теорема. Три прямые $(u_1, v_1), (u_2, v_2), (u_3, v_3)$ пересекаются в одной точке в том и только в том случае, если

$$\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{pmatrix} = 0$$

4.3 Класс алгебраической кривой

Алгебраические кривые классифицируются не только по их порядку, но и по их классу и роду (жанру).

Класс алгебраической кривой определяется степенью ее уравнения в тангенциальных координатах.

Геометрически это означает, что класс кривой определяется числом касательных, действительных и мнимых, которые можно провести к этой кривой из произвольной точки плоскости, не лежащей на ней.

Построим тангенциальное уравнение кривой, а, следовательно, определим ее класс. Для этого представим себе, что данная кривая $f(x, y) = 0$ пересечена прямой $\xi x + \eta y + 1 = 0$. Условие того, что две точки ее пересечения с кривой совпадают между собой, записанное в виде равенства, связывающего ξ и η и будет искомым тангенциальным уравнением кривой.

Пример 1. Найти тангенциальное уравнение окружности $x^2 + y^2 = 1$. Пересечем окружность прямой $\xi x + \eta y + 1 = 0$. Исключим y из уравнения прямой и окружности.

$$y = -\left(\frac{1}{\eta} + \frac{\xi}{\eta}x\right) \Rightarrow x^2 + \frac{1}{\eta^2} + \frac{\xi^2}{\eta^2}x^2 + \frac{2\xi x}{\eta^2} = 1 \Rightarrow \\ \Rightarrow x^2(\xi^2 + \eta^2) + 2\xi x + (1 - \eta^2) = 0 \Rightarrow x_{1,2} = \frac{-2 \pm \sqrt{4\xi^2 - 4(\xi^2 + \eta^2)(1 - \eta^2)}}{2(\xi^2 + \eta^2)}$$

Условием касания прямой и окружности будет совпадение корней этого квадратного уравнения.

$$4\xi^2 - 4(\xi^2 + \eta^2)(1 - \eta^2) = 4\xi^2\eta^2 - 4\eta^2 + 4\eta^4 = -4\eta^2(1 - \xi^2 - \eta^2) \\ \eta \neq 0; \quad 1 - \xi^2 - \eta^2 = 0 - \text{тангенциальное уравнение окружности.}$$

Исключая из уравнения прямой и окружности из-за симметрии уравнения прямой, получим то же самое выражение.

Если необходимо перейти от тангенциального уравнения кривой $f(x, y) = 0$ к ее обычному уравнению, то необходимо к этому уравнению присоединить уравнение $\xi x + \eta y + 1 = 0$ пучка всех касательных к кривой, проходящих через точку $M(x, y)$. Равенство, определяющееся условием совпадения двух касательных в одну и будет искомым уравнением кривой в прямоугольной системе.

Пример 2. В тангенциальных координатах задана кривая $\xi + \eta + \xi\eta = 0$. Найти ее уравнение в прямоугольных координатах.

Рассмотрим уравнение пучка прямых $\xi x + \eta y + 1 = 0$, проходящих через точку $M(x, y)$. Найдем те прямые этого пучка, которые касаются кривой. Исключим из рассматриваемых уравнений.

$$\eta = -\left(\frac{1}{y} + \xi \frac{x}{y}\right) \Rightarrow \xi \left(1 - \frac{1}{y} - \xi \frac{x}{y}\right) - \left(\frac{1}{y} + \xi \frac{x}{y}\right) = 0 \\ y \neq 0, \quad \xi^2 x + (1 + x - y) + 1 = 0$$

$$\xi_{1,2} = \frac{-(1+x-y) \pm \sqrt{(1+x-y)^2 - 4x}}{2x}$$

Для того, чтобы две прямые, определяемые двумя значениями ξ в полученном равенстве совпали в одну, необходимо, чтобы эти значения ξ были равны между собой. Это будет достигнуто, если $(1+x-y)^2 - 4x = 0$. Полученное выражение и представляет собой искомое обычное уравнение кривой.

Порядок и класс линии в общем случае не совпадают, за исключением кривых второго порядка, которые одновременно являются кривыми 2 класса.

4.5 Формулы Пюккера

Вводя тангенциальные координаты прямых ξ, η Пюккер рассматривает уравнение

$$F(\xi, \eta) = 0$$

определяющее однопараметрическое семейство прямых, как уравнение кривой, огибающей это семейство. Это уравнение называется тангенциальным, т.к. кривая определяется в этом случае не своими точками, а касательными. Пюккер показал, что на комплексной проективной плоскости порядок кривой n , ее класс k , число ее двойных точек d (узловых и изолированных), число ее точек возврата $-r$, число ее двойных касательных (т.е. прямых, касающихся кривой в двух точках) $-t$ и число ее точек перегиба w при отсутствии других особенностей связаны формулами:

$$k = n(n-1) - 2d - 3r, \quad n = k(k-1) - 2t - 3w, \\ w = 3n(n-2) - 6d - 8r, \quad r = 3k(k-2)6t - 8w,$$

которые теперь носят его имя – формулы Пюккера. Они были приведены им впервые в его «Системе алгебраической геометрии на плоскости» в 1834 году.

Юлиус Пюккер (16.08.1801 – 22.05.1868г.) – выдающийся немецкий математик и физик, создатель «новой аналитической геометрии» (Ф. Клейн).

4.5 Род алгебраической кривой

Род, или жанр алгебраической кривой определяется числом p – разностью между наибольшим числом двойных точек, которые может иметь кривая этого порядка, и их фактическим числом u данной кривой.

$$p = d - d_1$$

где d – теоретическое число двойных точек для кривой этого порядка, d_1 – число двойных точек у рассматриваемой кривой.

С ранее приведенными числовыми характеристиками кривой число p связано соотношениями:

$$p = \frac{(n-1)(n-2)}{2} - d - r,$$

$$p = \frac{(k-1)(k-2)}{2} - t - w.$$

Если рассматриваемая кривая имеет наибольшее число двойных точек, возможных для кривых ее порядка, то это будет кривая нулевого рода (рациональная или уникурсальная). Такие кривые обладают весьма важным свойством:

Теорема. Координаты точки, двигающейся по кривой нулевого рода, могут быть выражены рациональными функциями некоторого параметра.

Доказательство. Кривая n -го порядка и нулевого рода должна иметь $d = \frac{(n-1)(n-2)}{2}$ двойных точек. Через эти точки и через $(n-3)$ простых точек проведем кривые $n-2$ порядка, образующие пучок, т.к. задано $\frac{(n-1)(n-2)}{2} + (n-3) = \frac{(n-1)(n-2)}{2} + \frac{2(n-2)}{2} - 1 = \frac{(n+1)(n-2)}{2} - 1$ точек, а для определения кривой $(n-2)$ порядка необходимо $\frac{(n-1)}{2(n+1)}$ точек (см. теорему 3). Точек задано на 1 меньше, чем необходимо для определения кривой $(n-2)$ порядка, следовательно, через эти точки можно провести не одну кривую, а пучок. Пусть $f_1(x, y) + \lambda f_2(x, y) = 0$ - уравнение того пучка. Чтобы найти точки пересечения данной кривой с кривыми пучка, исключим из уравнения одну из координат. В результате получим для определения второй координаты уравнение степени $n(n-2)$. В этом уравнении будут известны корни, являющиеся координатами точек, через которые были проведены кривые пучка. Таких точек будет $2 \frac{(n-1)(n-2)}{2} + (n-3) = n(n-2) - 1$, т.е. число известных корней на единицу меньше его степени. Подставив эти корни в уравнение, получим для определения искомой координаты уравнение первой степени, координаты которого будут рациональными функциями параметра λ , и, следовательно, искомая координата выразится через параметр λ рационально.

В силу этого свойства кривые нулевого рода называются рациональными кривыми.

$p = 0$ для кривых 2-го порядка.

$p = 1 - d_n$ для кривых 3-го порядка.

Анализ формул Плюккера позволяет получить ряд интересных выводов, имеющих существенное значение при конструировании технических кривых – динамических обводов.

Одномерные обводы – оси всевозможных трубопроводов, предназначенных для перемещения с определенной скоростью жидких и сыпучих материалов с минимальным сопротивлением. Кроме того одномерные и двумерные обводы перемещающиеся (обтекаемые) в сплошной среде и испытывающие при данных условиях минимальное сопротивление называют динамическими. Такие

обводы применяются при конструировании всевозможных аэро- и гидродинамических профилей, осей трубопроводов, скоростных автомагистралей и т.п.

Движение материальной точки по плоской кривой описывается дифференциальными уравнениями Лагранжа первого рода, которые в естественной форме имеют вид

$$m \frac{dV}{dt} = F_t, \quad \frac{mV^2}{\rho} = F_n + R_n$$

где m - масса точки, ρ - радиус кривизны в данной точке кривой, F_t, F_n - проекции активной силы на касательную и нормаль, R_n - проекция силы сопротивления на нормаль, V - скорость движения точки, t - время. Из приведенных уравнений следует, что проекции сил, скорости и ускорения изменяется монотонно, если угловой коэффициент касательной и радиус круга кривизна траектории изменяются монотонно. Следовательно, из всего множества кривых наиболее динамичными будут те, у которых законы изменения первых и вторых производных будут наиболее плавным, с минимальным количеством экстремальных точек и точек перегиба.

Анализ формул Плюккера показывает, что рациональные кривые ($p = 0$) n -го порядка имеют наименьшее количество точек, перегиба, по сравнению с любыми другими кривыми. Действительно, если $p = 0$ то $d + r = \frac{(n-1)(n-2)}{2} \Rightarrow$ число точек перегиба = \min .

Если $d + r = 0$, то жанр будет максимальным

$$p = \frac{(n-1)(n-2)}{2}$$

А количество точек перегиба $\omega = 3n(n-2) \Rightarrow \max$.

4.6 Циклические точки и циркулярные кривые

Уравнение окружности в однородных координатах имеет

$$x_1^2 + x_2^2 + 2Ax_1x_2 + 2Bx_2x_3 + Cx_3 = 0$$

Точки пересечения этой окружности с несобственной прямой $x_3 = 0$ определяются

системой $\frac{x_2}{x_1} = i, x_3 = 0$ и $\frac{x_2}{x_1} = -i, x_3 = 0$.

Полагая $x_1 = 1$, находим эти точки $j_1(1, i, 0)j_2(1, -i, 0)$.

Т.к. координаты точек J_1, J_2 не зависят от коэффициентов А,В,С то, следовательно, всякая окружность проходит через точки J_1 и J_2 , которые являются несобственными и мнимыми точками этой окружности и называются круговыми или циклическими точками плоскости. Алгебраическими кривые, проходящие через циклические точки, называются циркулярными кривыми.

Фокусы, диаметры, центр

С понятием циклических точек плоскости связано понятие фокусов алгебраических кривых, которое обобщает понятие фокусов кривых 2-го порядка. Кривые 2-го порядка обладают тем свойством, что касательные, проведенные к ним из циклических точек плоскости, пересекаются в их фокусах. Следовательно, фокусы кривой 2-го порядка можно определить как точки пересечения касательных к этой кривой, проведенных из циклических точек плоскости. Это определение, без каких-либо изменений, распространяется на кривые n -го порядка.

Если класс кривой равен k , то число фокусов этой кривой равно k^2 , соответственно числу точек пересечения k касательных, проведенных из циклической точки J_1 , и k касательных, проведенных из точки J_2 . Из этих k^2 фокусов действительными будут только k , ибо в действительной точке F будут пересекаться лишь касательные, уравнение которых соответственно $U + V_i = 0$ и $U - V_i = 0$, где U и V - линейные функции x и y . Корни системы $U=0, V=0$ будут координатами действительных фокусов.

Диаметром кривой второго порядка называется прямая, являющаяся геометрическим местом среди параллельных хорд. Обобщим это понятие для алгебраической кривой n -го порядка.

Центром средних расстояний некоторой системы точек есть точка, координаты которой равны средним арифметическим соответствующих координат точек системы. Если кривую n -го порядка пересечь системой параллельных прямых и на каждой из них определить центр средних расстояний точек ее пересечения с данной кривой, то геометрическим местом этих точек прямая, направление которой определяется направлением параллельных хорд. Эту прямую называют диаметром кривой n -го порядка.

Если все диаметры данной кривой пересекаются в называют ньютоновым центром этой кривой. Если кривая обладает центральной симметрией, то ньютонов центр совпадает с центром ее симметрии одной и той же точке, то ее.

Теорема Шаля. Если к кривой провести систему параллельных касательных, то центр средних расстояний точек касания есть постоянная точка, положение которой не зависит от направления касательных. Эту точку называют абсолютным центром алгебраической кривой.

Полюсы и поляры

На кривые n -го порядка обобщается и понятие поляры точки относительно кривой 2-го порядка. При этом заданной точке плоскости ставится в соответствие $(n - 1)$ поляр относительно кривой n -го порядка. Первая из этих поляр является кривой порядка $(n - 1)$, вторая, являющаяся полярной заданной точки относительно первой поляры, имеет порядок $(n - 2)$ и т.д., наконец $(n - 1)$ -я поляра является прямой линией.

Т.к. координаты точек J_1, J_2 не зависят от коэффициентов A, B, C то, следовательно, всякая окружность проходит через точки J_1 и J_2 , которые являются несобственными и мнимыми точками этой окружности и называются круговыми или циклическими точками плоскости. Алгебраическими кривые, проходящие через циклические точки, называются циркулярными кривыми.

Фокусы, диаметры, центр

С понятием циклических точек плоскости связано понятие фокусов алгебраических кривых, которое обобщает понятие фокусов кривых 2-го порядка. Кривые 2-го порядка обладают тем свойством, что касательные, проведенные к ним из циклических точек плоскости, пересекаются в их фокусах. Следовательно, фокусы кривой 2-го порядка можно определить как точки пересечения касательных к этой кривой, проведенных из циклических точек плоскости. Это определение, без каких-либо изменений, распространяется на кривые n -го порядка.

Если класс кривой равен k , то число фокусов этой кривой равно k^2 , соответственно числу точек пересечения k касательных, проведенных из циклической точки J_1 , и k касательных, проведенных из точки J_2 . Из этих k^2 фокусов действительными будут только k , ибо в действительной точке F будут пересекаться лишь касательные, уравнение которых соответственно $U + V_i = 0$ и $U - V_i = 0$, где U и V - линейные функции x и y . Корни системы $U=0, V=0$ будут координатами действительных фокусов.

Диаметром кривой второго порядка называется прямая, являющаяся геометрическим местом среди параллельных хорд. Обобщим это понятие для алгебраической кривой n -го порядка.

Центром средних расстояний некоторой системы точек есть точка, координаты которой равны средним арифметическим соответствующих координат точек системы. Если кривую n -го порядка пересечь системой параллельных прямых и на каждой из них определить центр средних расстояний точек ее пересечения с данной кривой, то геометрическим местом этих точек прямая, направление которой определяется направлением параллельных хорд. Эту прямую называют диаметром кривой n -го порядка.

Если все диаметры данной кривой пересекаются в называют ньютоновым центром этой кривой. Если кривая обладает центральной симметрией, то ньютонов центр совпадает с центром ее симметрии одной и той же точке, то ее.

Теорема Шаля. Если к кривой провести систему параллельных касательных, то центр средних расстояний точек касания есть постоянная точка, положение которой не зависит от направления касательных. Эту точку называют абсолютным центром алгебраической кривой.

Полюсы и поляры

На кривые n -го порядка обобщается и понятие поляры точки относительно кривой 2-го порядка. При этом заданной точке плоскости ставится в соответствие $(n - 1)$ поляр относительно кривой n -го порядка. Первая из этих поляр является кривой порядка $(n - 1)$, вторая, являющаяся полярной заданной точки относительно первой поляры, имеет порядок $(n - 2)$ и т.д., наконец $(n - 1)$ -я поляр является прямой линией.



Ты видишь, конечно, что все это с первого взгляда очень похоже на пустой вздор, и, однако же, выполненные сообразно с этим алгебраические действия приводят к изображениям поистине удивительным.

*М. Штифель
(1487-1567)-немецкий математик*

Кривые можно построить путем преобразования других линий. Линии принятые для преобразования называют исходными или базовыми. Они могут быть как кривыми, так и прямыми линиями.

Форма и свойства новой кривой определяются двумя факторами:

- формой и свойствами базовой кривой;
- свойствами самого преобразования, т.е. природой функции f_1 и f_2 .

Поэтому одной из основных задач теории кривых является детальное исследование каждого конкретного преобразования, соответствующего определенному виду функции f_1 и f_2 .

Здесь важно установить:

- конструктивное, геометрическое содержание рассматриваемого преобразования;
- новое выражение и новые свойства преобразуемой кривой;
- инвариант преобразования, т.е. такие характеристики и свойства кривых, которые сохраняются при этом преобразовании. Эти свойства являются глубинными, внутренними, отражающими родство таких кривых, которые как по своей форме, так и по своим индивидуальным качествам могут быть весьма и весьма различными.

Следовательно, геометрические преобразования открывают не только широкие возможности для образования новых кривых, но и позволяют обнаружить новые свойства уже известных кривых.

Рассмотрим в общих черта наиболее часто встречающиеся точечные преобразования.

5.1 Точечные преобразования

Аналитически точечное преобразование осуществляется тем, что в уравнении $f(x, y) = 0$ базовой кривой производится замена переменных.

В простейшем случае точечного преобразования, каждой точке заданной кривой, ставится в соответствие по определенному закону точка плоскости. Геометрическое место таких точек и представляет собой новую кривую, соответствующую заданной. Очевидно, что в результате преобразования исходной форма кривой искажается в той или иной степени. Полученное уравнение выражает в этой же системе координат новую кривую, соответствующую исходной. Следует обратить внимание на то, что система координат неподвижна, а плоскость деформируется.

5.2. Применение аффинных преобразований к плоским кривым

Аффинные преобразования уже подробно рассматривались выше. Здесь мы остановимся на их применении к плоским кривым. На плоскости эти преобразования осуществляются формулами

$$\begin{aligned}x &= a_{11}x_1 + a_{12}y_1 + a_{13} \\ y &= a_{21}x_1 + a_{22}y_1 + a_{23}\end{aligned}\tag{5.1}$$

где x и y — координаты точки преобразующей кривой,

x_1 и y_1 — координаты соответствующей ей точки новой кривой.

Если $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$, то преобразование будет обратным и взаимно однозначным.

Общее аффинное преобразование, определяемое формулами (5.1) сводится к движению плоскости (плоскопараллельному перемещению и вращению) и к преобразованию ее по формулам:

$$\begin{aligned} x &= \lambda x_1 \\ y &= \mu y_1 \end{aligned} \quad (5.2)$$

т.е. к сжатию или растяжению плоскости в направлении осей (если λ или отрицательно, то еще дополнительно происходит зеркальное отражение деформируемой плоскости относительно оси y – или оси x). При $\lambda = \mu$ — получим преобразование подобия.

Нас интересует следующее свойства аффинного преобразования:

1. Бесконечно удаленные точки кривой переходят в бесконечно удаленные точки.

При этом ветви параболические (не имеющие асимптот) перейдут в параболические, а гиперболические – в гиперболические.

Порядок и класс алгебраической кривой – инварианты аффинного преобразования.

В частности, прямая переходит в прямую.

Простое отношение трех точек, лежащих на одной прямой сохраняется.

Параллельные прямые переходят в параллельные. (Но угол, под которым пересекаются прямые изменяется).

Если U площадь базовой фигур, а U_1 площадь аффинно-преобразованной, то справедливо равенство

$$U = \Delta U_1$$

где $\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

Перечисленные свойства преобразований определяют так называемые аффинные свойства кривых, т.е. свойства, присущие всем аффинно родственным кривым.

5.3. Проективные преобразования

5.3.1. Трилинейные (проективные) координаты

В практике применений как проективных преобразований, определенные преимущества по сравнению с декартовой системой координат имеет проективная (трилинейная) система координат. Проективными координатами точки $M(x,y)$ на плоскости называют три числа пропорциональные линейным функциям ее декартовых координат x,y , т.е. по определению:

$$\begin{cases} \rho \xi_1 = a_1 x + b_1 y + c_1 \\ \rho \xi_2 = a_2 x + b_2 y + c_2 \\ \rho \xi_3 = a_3 x + b_3 y + c_3 \end{cases} \quad (5.3)$$

где a_i, b_i, c_i — определенные числа, удовлетворяющие условию

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0,$$

а ρ - произвольный, не равный нулю множитель.

Очевидно, что одновременное увеличение (или уменьшение) всех трех координат в одно и то же число раз не изменяет положение этой точки на плоскости. Это дает определенные преимущества при их применении.

Разрешая формулы (5.3) относительно x и y , получим:

$$x = \frac{\begin{vmatrix} \xi_1 & b_1 & c_1 \\ \xi_2 & b_2 & c_2 \\ \xi_3 & b_3 & c_3 \end{vmatrix}}{\Delta}, \quad y = \frac{\begin{vmatrix} a_1 & \xi_1 & c_1 \\ a_2 & \xi_2 & c_2 \\ a_3 & \xi_3 & c_3 \end{vmatrix}}{\Delta}, \quad \Delta = \begin{vmatrix} a_1 & b_1 & \xi_1 \\ a_2 & b_2 & \xi_2 \\ a_3 & b_3 & \xi_3 \end{vmatrix} \quad (5.4)$$

По этим формулам происходит переход от проективных координат к декартовым.

Геометрический смысл проективных координат заключается в том, что они пропорциональны расстояниям от точки M до трех неподвижных прямых, определяемых уравнениями:

$$\begin{aligned} a_1x + b_1y + c_1 &= 0 \\ a_2x + b_2y + c_2 &= 0, \\ a_3x + b_3y + c_3 &= 0 \end{aligned}$$

которые при условии $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0$ образуют $\Delta O_1 O_2 O_3$ координатный треугольник проективной системы.

Из (5.3) следует, что уравнения $\xi_1 = 0, \xi_2 = 0, \xi_3 = 0$ являются уравнениями сторон координатного треугольника, вершины которого находятся в точках $O_1(1,0,0), O_2(0,1,0), O_3(0,0,1)$.

Частным случаем проективной системы координат является однородная система координат, рассмотренная ранее. Она получается, когда одна из сторон координатного треугольника является бесконечно удаленной. Если при этом две другие стороны пересекаются под прямым углом, то однородные координаты точки (x_1, x_2, x_3) связаны с проективными следующими соотношениями:

$$x = \frac{x_1}{x_3}, y = \frac{x_2}{x_3},$$

Откуда следует, что для перехода от однородной системы координат к декартовой достаточно x_1 заменить на x, x_2 на y , а x_3 положить равным 1. Однородные координаты связаны с проективными следующими соотношениями:

$$\begin{aligned} \rho \xi_1 &= a_1 x_1 + b_1 x_2 + c_1 x_3 \\ \rho \xi_2 &= a_2 x_1 + b_2 x_2 + c_2 x_3 \\ \rho \xi_3 &= a_3 x_1 + b_3 x_2 + c_3 x_3 \end{aligned}$$

$$x_1 = \begin{vmatrix} \xi_1 & b_1 & c_1 \\ \xi_2 & b_2 & c_2 \\ \xi_3 & b_3 & c_3 \end{vmatrix}, \quad x_2 = \begin{vmatrix} a_1 & \xi_1 & c_1 \\ a_2 & \xi_2 & c_2 \\ a_3 & \xi_3 & c_3 \end{vmatrix}, \quad x_3 = \begin{vmatrix} a_1 & b_1 & \xi_1 \\ a_2 & b_2 & \xi_2 \\ a_3 & b_3 & \xi_3 \end{vmatrix}$$

Основным преимуществом проективной системы координат является то, что в ней бесконечно удаленные точки плоскости имеют конечные координаты. Это позволяет исследовать поведение кривой на бесконечности так же, как это делается при анализе формы кривой вблизи ее обычной точки. Однако при решении вопросов метрического характера преимущества остаются за декартовой системой координат.

Рассмотрим примеры уравнений линий в проективных координатах.

1. Получим уравнение прямой $ax + by + c = 0$ в проективных координатах. Для этого заменим x и y их выражениями (2):

$$a \begin{vmatrix} \xi_1 & b_1 & c_1 \\ \xi_2 & b_2 & c_2 \\ \xi_3 & b_3 & c_3 \end{vmatrix} + b \begin{vmatrix} a_1 & \xi_1 & c_1 \\ a_2 & \xi_2 & c_2 \\ a_3 & \xi_3 & c_3 \end{vmatrix} + c \begin{vmatrix} a_1 & b_1 & \xi_1 \\ a_2 & b_2 & \xi_2 \\ a_3 & b_3 & \xi_3 \end{vmatrix} = 0$$

После соответствующих преобразований получим однородное относительно проективных координат уравнение

$$u_1 \xi_1 + u_2 \xi_2 + u_3 \xi_3 = 0$$

$$u_1 = \begin{vmatrix} a & b & c \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, u_2 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a & b & c \\ a_3 & b_3 & c_3 \end{vmatrix}, u_3 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix}.$$

2. Общее уравнение кривой второго порядка в проективных координатах будет однородным уравнением вида:

$$a_{11}\xi_1^2 + 2a_{12}\xi_1\xi_2 + a_{22}\xi_2^2 + 2a_{13}\xi_1\xi_3 + 2a_{23}\xi_2\xi_3 + a_{33}\xi_3^2 = 0$$

Пусть рассматриваемая коника проходит через три заданные точки. Примем эти точки за вершины координатного треугольника. Тогда, чтобы кривая прошла через вершину $0_1(1,0,0)$ положим в общем уравнении $\xi_2 = 0$ и $\xi_3 = 0$. В результате получим $a_{11} = 0$. Подобным образом учитывая, что кривая проходит через точки 0_1 и 0_2 получим, $0_2(0,1,0) \Rightarrow a_{22}, 0_3(0,0,1) \Rightarrow a_{33}$. Следовательно, искомое уравнение запишем в виде:

$$a_{12}\xi_1\xi_2 + a_{13}\xi_1\xi_3 + a_{23}\xi_2\xi_3 = 0$$

Таким образом, переменные ξ_1, ξ_2, ξ_3 в уравнении любой исследуемой кривой можно рассмотреть как сокращенные обозначения линейных функций. Такой взгляд положен в основу так называемого «способа сокращенных обозначений», предложенного Плюккером. Сущность его заключается в том, что выражения первого или высших порядков, входящие в декартово уравнение линии, обозначают одной буквой и рассматривают и в качестве своеобразных координат, чтобы затем непосредственным рассмотрением уравнения, без дополнительных выкладок, усматривать свойства кривых – «читать уравнения», по выражению Плюккера.

Пример. Всякое уравнение кривой третьего порядка может быть представлено в виде

$$pqr - s^2 = 0$$

где p, q, r, s - линейные функции декартовых координат. Из этого уравнения следует, что в общем случае кривая имеет три точки перегиба, причем эти точки лежат на одной прямой. Действительно, пересекая кривую прямой $p = 0$ получим три точки пересечения, так как порядок кривой равен 3. Но эти точки будут принадлежать и прямой $s = 0$. С другой стороны, прямые $p = 0$ и $s = 0$ пересекаются только в одной точке. Следовательно, все указанные точки совпадают в одну, которая и является точкой перегиба кривой. Пересекая кривую прямыми $q = 0$ и $r = 0$ получим еще две точки перегиба. Причем все полученные точки лежат на одной прямой $s = 0$.

5.3.2. Проективные преобразования.

В декартовой системе координат проективные преобразования определяются формулами

$$x = \frac{a_{11}x_1 + a_{12}y_1 + a_{13}}{a_{31}x_1 + a_{32}y_1 + a_{33}}; \quad y = \frac{a_{21}x_1 + a_{22}y_1 + a_{23}}{a_{31}x_1 + a_{32}y_1 + a_{33}} \quad (5.5)$$

где x, y – координаты преобразуемой кривой, а соответствующей ей точки. При этом

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

При этом условии преобразование будет обратимым и взаимно однозначным. В проективных координатах рассматриваемое преобразование определяется формулами:

$$\left. \begin{aligned} \rho \xi_1 &= a_1 \xi'_1 + b_1 \xi'_2 + c_1 \xi'_3 \\ \rho \xi_2 &= a_2 \xi'_1 + b_2 \xi'_2 + c_2 \xi'_3 \\ \rho \xi_3 &= a_3 \xi'_1 + b_3 \xi'_2 + c_3 \xi'_3 \end{aligned} \right\} \quad (5.6)$$

где ξ_1, ξ_2, ξ_3 проективные координаты точки преобразуемой кривой, а ξ'_1, ξ'_2, ξ'_3 координаты соответствующей ей точки по отношению к тому же координатному треугольнику.

Основные свойства проективного преобразования:

Из (5.5) следует, что точкам прямой $a_{31}x_1 + a_{32}y_1 + a_{33} = 0$ соответствуют бесконечно удаленные точки. (При аффинном преобразовании образами бесконечно удаленных точек могли быть только бесконечно удаленные точки).

В силу линейности формул (5.6) прямая при проективном преобразовании перейдет в прямую, линия n -го порядка – в линию того же порядка, линия K -го класса – в линию того же класса.

1. Порядок и класс кривой – инварианты проективного преобразования.

2. Двойное отношение четырех точек и четырех точек и четырех лучей сохраняется.

Следствие. Гармоническим четверкам элементов при проективном преобразовании соответствует так же гармонические четверки элементов. Сложное (двойное) отношение четырех точек A, B, C и D прямой – величина $\frac{AC}{BC} : \frac{AD}{BD}$ (все отрезки направленные).

Гармоническое деление: точки C и D делят отрезок AB гармонически, если $\frac{AC}{BC} : \frac{AD}{BD}$ (отрезки направленные) т.е. точки C и D делят отрезок в одном и том же отношении, но одна внешним, а другая внутренним образом.

5.3.2. Частные случаи проективного преобразования.

Определим точки плоскости, переходящие при преобразовании в себя (остающиеся на месте) и каково их число. Очевидно, что проективные координаты таких точек должны быть связаны зависимостью или $\xi_1 : \xi_2 : \xi_3 = \xi'_1 : \xi'_2 : \xi'_3$.

$$\left. \begin{aligned} \xi_1 &= k\xi_1' \\ \xi_2 &= k\xi_2' \\ \xi_3 &= k\xi_3' \end{aligned} \right\} \quad (5.7)$$

Подставив эти значения ξ_1, ξ_2, ξ_3 в формулы (5.6) и положив $\rho = 1$, получим три уравнения

$$\left. \begin{aligned} (a_1 - k)\xi_1' + b_1\xi_2' + c_1\xi_3' &= 0 \\ a_2\xi_1' + (b_2 - k)\xi_2' + c_2\xi_3' &= 0 \\ a_3\xi_1' + b_3\xi_2' + (c_3 - k)\xi_3' &= 0 \end{aligned} \right\} \quad (5.8)$$

Для совместности этих уравнений необходимо, чтобы

$$\begin{vmatrix} a_1 - k & b_1 & c_1 \\ a_2 & b_2 - k & c_2 \\ a_3 & b_3 & c_3 - k \end{vmatrix} = 0 \quad (5.9)$$

Получим кубическое уравнение относительно k . Определив k и подставив полученное значение в (5.8), получим систему совместных уравнений из решения которых и определяются координаты неподвижных точек. Различные типы решений уравнения (5.9) определяют различные частные случаи проективного преобразования.

Рассмотрим наиболее интересные из этих случаев.

1. Все три корня k_1, k_2, k_3 действительны и различны.

В этом случае уравнения (5.8) дадут три двойные точки P_1, P_2, P_3 и соответственно три двойные прямые, соединяющие эти точки и образующие инвариантный $\Delta P_1 P_2 P_3$ проективного преобразования. Примем его за координатный треугольник. При этом формулы преобразования (5.6) примут вид:

$$\left. \begin{aligned} \xi_1 &= k_1\xi_1' \\ \xi_2 &= k_2\xi_2' \\ \xi_3 &= k_3\xi_3' \end{aligned} \right\} \quad (5.10)$$

2. Один из корней действительный, два других – мнимые.

В этом случае инвариантный треугольник имеет одну действительную вершину и одну действительную сторону противоположную этой вершине.

3. Два действительных корня $k_1=k_2$ равны между собой, при этом все три уравнения (5.8) сводятся к одному. В этом случае задача определения двойных точек становится неопределенной: двойной точкой может быть любая точка, находящаяся на прямой $P_1 P_2$, определяемой уравнениями (5.8), которая будет двойной прямой. При этом формулы преобразования (5.10) принимают вид:

$$\left. \begin{aligned} \xi_1 &= k_1\xi_1' \\ \xi_2 &= k_1\xi_2' \\ \xi_3 &= k_3\xi_3' \end{aligned} \right\} \quad (5.11)$$

Эти формулы определяют частный, но чрезвычайно важный вид проективного преобразования – гомологию. Все прямые, соединяющие соответствующие в гомологии точки плоскости, проходят через одну точку – центр гомологии P_3 ; соответствующие в гомологии прямые будут пересекаться на одной и

той прямой P_1P_2 , - оси гомологии. Это вытекает из формул (5.11). Из соотношения $\frac{\xi_3}{\xi_2} = \frac{\xi'_3}{\xi'_2}$ следует: если какая либо точка $M(\xi_1, \xi_2, \xi_3)$ лежит на прямой $a\xi_1 + b\xi_2 = 0$, проходящей через P_3 , то и точка $M_1(\xi'_1, \xi'_2, \xi'_3)$ будет лежать на этой прямой. Точно так же, если точка $M(\xi_1, \xi_2, \xi_3)$ пробегает какую-либо прямую $a\xi_1 + b\xi_2 + c\xi_3 = 0$, то соответствующая ей точка M_1 будет перемещаться по прямой $k_1(a\xi'_1 + b\xi'_2) + ck_3\xi'_3 = 0$, причем обе эти прямые проходят через точку пересечения прямой $\xi_3 = 0$ с прямой $a\xi_1 + b\xi_2 = 0$.

4. Частным случаем гомологии является инволюционное преобразование, которое получается, если в выражениях (5.11) положить $k_1 = 1, k_3 = -1$. При этом формулы преобразования (5.11) принимают вид:

$$\xi_1 = \xi'_1; \xi_2 = \xi'_2; \xi_3 = \xi'_3.$$

Если при этом прямая $\xi_3 = 0$ бесконечно удалена, то проективные координаты обратятся в однородные, переходя от которых к декартовым x и y , получим формулы преобразования в виде

$$x = -x_1; y = -y_1$$

т.е. центральную симметрию.

По сравнению с аффинными свойствами проективные являются более глубокими, общими свойствами кривых. Так, эллипс, гипербола и парабола с проективной точки зрения эквивалента.

5.4 Преобразование инверсии

Рассмотрим преобразования, которые имеют с аффинными и проективными преобразованиями то общее, что преобразующие функции их являются так же рациональными, но это уже будут функции 2-й степени. Частным случаем таких преобразований является инверсия.

Определение. Инверсией (от латыни inversion – превращение, преобразование) кривой линии относительно окружности радиуса R называют преобразование, при котором произведение радиус – векторов соответствующих точек данной (базовой) кривой и точек строящейся кривой постоянно и равно R^2 .

Число R^2 – степень (коэффициент) инверсии. Центр окружности радиуса R называется центром (полусом) инверсии.

$$OA \cdot OA_1 = R^2 \tag{5.12}$$

Пусть $(\cdot)A(x, y)$, а координаты $(\cdot)A_1(x_1, y_1)$ тогда формула (5.1) примет вид:

$$\sqrt{x^2 + y^2} \cdot \sqrt{x_1^2 + y_1^2} = R^2 \tag{5.13}$$

Учтем, что $(\cdot)A$ и $(\cdot)A_1$ инциденты одной и той же прямой OA_1 . Тогда $\frac{x}{y} = \frac{x_1}{y_1}$. Подставив последнее соотношение в (5.4) получим формулы преобразования в виде

$$x_1 = \frac{xR^2}{x^2 + y^2}, y_1 = \frac{yR^2}{x^2 + y^2} \tag{5.14}$$

Если окружность, по отношению к которой инвертируется плоскость, отнесена к полярной системе координат с полюсом в центре окружности, то формула соответствия радиус-векторов точек $M(\rho, \theta)$ и $M_1(\rho_1, \theta)$ запишется в виде

$$\rho\rho_1 = R^2.$$

Инверсией любой окружности радиуса R_1 , с концентрической окружностью радиуса R относительно данного полюса является так же концентрическая окружность радиуса $R_2 = \frac{R^2}{R_1}$.

Если $R_1 = nR$, то $R_2 = \frac{R^2}{nR} = \frac{R}{n}$, где n – целое число.

Основные свойства инверсии.

Инверсия есть взаимно однозначное преобразование всех точек плоскости, за исключением одной – полюса 0 инверсии, который преобразуется в несобственную точку.

Инверсия есть преобразование, двукратное применение которого дает тождественное преобразование, т.е. инверсия есть инволюционное преобразование.

Прямая, проходящая через полюс, преобразуется в ту же прямую.

Окружность, не проходящая через полюс инверсии преобразуется в окружность, так же не проходящую через полюс.

Окружность, проходящая через полюс инверсии, преобразуется в прямую, не проходящую через полюс.

Пример.

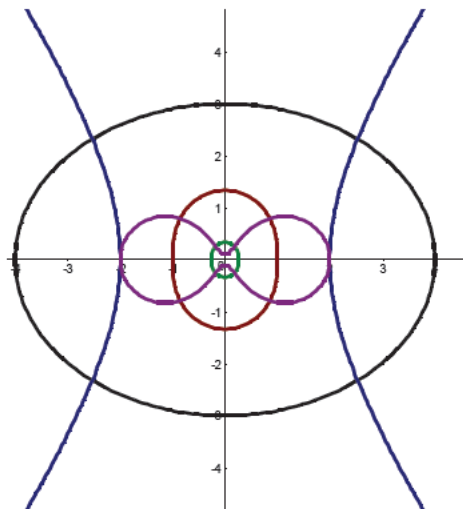


Рис. 5.1. Инверсия эллипса и гиперболы

На рис. 5.1 приведены инверсии эллипса $\left(\frac{x}{4}\right)^2 + \left(\frac{y}{3}\right)^2 = 1$ ($R = 1,2$) и инверсия гиперболы $\left(\frac{x}{2}\right)^2 - \left(\frac{y}{3}\right)^2 = 1$; ($R=2$).

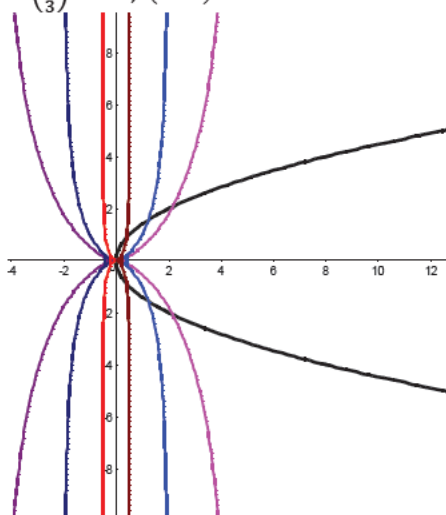


Рис. 5.2. Инверсия параболы

На рис 5.2 приведены инверсии параболы $y^2=2x$ ($r = 1, 2, 3$), относительно окружности $x^2+y^2=R^2$ и относительно мнимой окружности $x^2+y^2=-R^2$.

Основная окружность при инверсии может быть мнимой, тогда как преобразование и в том же случае остается действительным. Так, если в качестве основной взять мнимую окружность $x^2 + y^2 = -R^2$, то формулы преобразования примут вид:

$$x = -\frac{x_1 R^2}{x_1^2 + y_1^2}, \quad y = -\frac{y_1 R^2}{x_1^2 + y_1^2}$$

Гиперболическая инверсия с центром в начале координат O , со степенью k , осью которой является ось x , записывается формулами:

$$\begin{aligned} x_1 &= x \\ y_1 &= k/y \end{aligned}$$

Гиперболическая инверсия сводится к симметрии относительно пары параллельных прямых или же симметрии относительно равнобочной гиперболы или параболы.

На рис. 5.3 приведены гиперболические инверсии эллипса $\left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 = 1$ ($k = 2,3$) и гиперболы $\left(\frac{x}{2}\right)^2 - \left(\frac{y}{3}\right)^2 = 1$; ($k = 2, 3$).

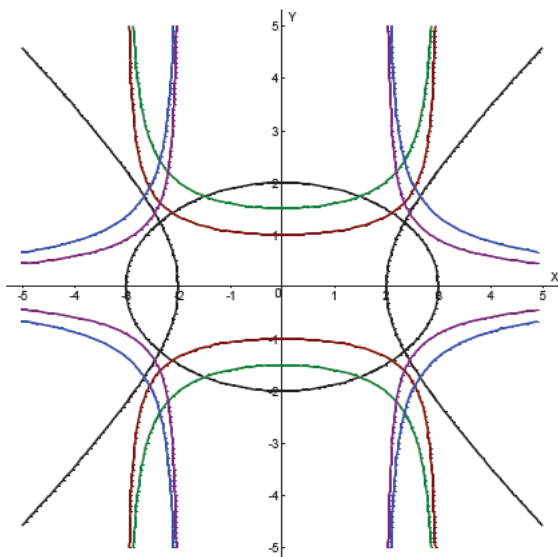


Рис. 5.3. Гиперболическая инверсия

Двойственные преобразования. Полярное преобразование.

В основу двойственных преобразований положена идея соответствия точки и прямой, реализуемая п्लюккеровыми координатами – коэффициенты u и v в уравнении прямой

$$ux+vy+1=0 \quad (5.15)$$

Если заданные u и v как функции x и y – координат точки, то тем самым каждой точке (x,y) ставится в соответствие определенная прямая (5.15). Это соответствие реализуется если координаты u и v прямой определены как дробно-линейные функции x и y .

Пусть на плоскости задана кривая

$$f(x,y) = 0 \quad (5.16)$$

Пусть точка принадлежит этой кривой. Если точка M перемещается по кривой (3), то соответствующая ей прямая так же перемещается, определяя при этом новую кривую – огибающую семейства тех прямых, с которыми она совпадает при своем перемещении. При этом производная прямая в каждом своем положении является касательной к новой кривой, уравнение которой в тангенциальных координатах

$$f(u,v) = 0 \quad (5.17)$$

При двойственном преобразовании не только точке соответствует прямая, но и прямой соответствует точка.

Это заключение выражает принцип двойственности, широко используемый в проективной геометрии. Им определяется и название рассмотренного преобразования.

Рассмотрим один частный случай.
Пусть коэффициенты таковы, что

$$a_{12} = a_{21}, a_{13} = a_{31}, a_{23} = a_{32}.$$

Тогда преобразование становится инволюционным, т.е. если точке M соответствует прямая l , то при том же преобразовании прямой l будет соответствовать та же самая точка M , а новая кривая обратится в первоначально заданную. Этот частный случай называется полярным преобразованием. Соотношение между преобразуемой кривой и новой кривой заключается в том, что каждая из них представляет собой геометрическое место полюсов, полярными которыми относительно некоторой кривой 2-го порядка служат касательные к другой кривой.

При полярном преобразовании соотношения между особенностями форм соответствующих кривых будут взаимно двойственны.

Если в качестве преобразуемой линии взять кривую n -го порядка, то уравнение (5.17) новой кривой окажется уравнением n -ой степени относительно u и v , следовательно: кривая n -го порядка переходит при двойственном преобразовании в кривую n -го класса.

При движении точки по первой кривой соответствующая ей прямая будет перемещаться по плоскости, касаясь второй кривой. Это позволяет определить особенности формы новой кривой по особенностям формы заданной кривой.

Точке возврата заданной кривой соответствует на новой кривой точка перегиба; двойной точке – двойная касательная (прямая, касающаяся линии в двух точках); изолированной точке, в которой пересекаются мнимые ветви кривой – изолированная касательная.

Если k точек заданной кривой лежат на одной прямой, то k соответствующих им касательных ко второй кривой пройдет через одну и ту же точку.

5.5 Подерное преобразование

Определение. Подерой кривой m относительно заданной точки C плоскости – центра подерного преобразования, называется новая кривая m^* , являющаяся геометрическим местом оснований перпендикуляров, опущенных из этой точки на касательные к заданной кривой. Когда точка A пробегает линию m , соответствующая ей точка A^* пробегает линию m^* – подеру кривой.

1. Пусть m – некоторая заданная кривая на плоскости m :

$$f(x, y) = 0 \quad (5.18)$$

и пусть точка $C(a, b)$ – центр подерного преобразования. Проведем из точки $A(\xi, \eta)$ кривой m касательную τ :

$$(\xi - x) \frac{\partial f}{\partial y} + (\eta - y) \frac{\partial f}{\partial x} = 0 \quad (5.19)$$

Уравнение прямой, проходящей через точку $C(a, b)$ и перпендикулярной к касательной τ , имеет вид

$$n: (\xi - a) \frac{\partial f}{\partial y} - (\eta - b) \frac{\partial f}{\partial x} = 0 \quad (5.20)$$

Исключая из уравнений x и y получим уравнение подерной кривой (рис. 5.4).

Пример 1 Дано: $0(0,0)$ – полюс поперного преобразования;

Окружность l радиуса r с центром в точке $0_1(a,0)$.

Определить уравнение кривой, которая получится в результате поперного преобразования этой окружности.

$$l \equiv f(x, y): (x - a)^2 + y^2 - r^2 = 0. \quad (5.21)$$

$$\frac{\partial f}{\partial x} = 2(a - x); \quad \frac{\partial f}{\partial y} = 2y.$$

Уравнение касательной τ к окружности l в точке $A(\xi, \eta)$:

$$(\xi - x)(x - a) + (\eta - y)y = 0, \quad (5.22)$$

Уравнение нормали к τ , проходящей через полюс $0(0,0)$:

$$\xi y - \eta(x - a) = 0. \quad (5.23)$$

Следовательно

$$y = \frac{r}{\xi}(x - a).$$

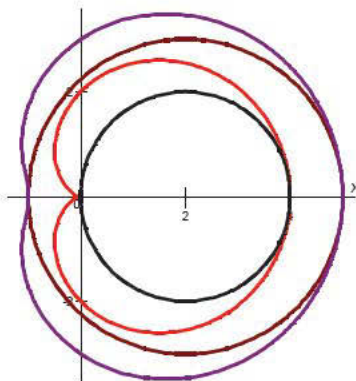


Рис. 5.4. Подеры окружности

Определим координаты точки пересечения касательной с нормалью: подставив полученный результат в (5.21) найдем

$$x = f_1(\xi, \eta) \Rightarrow (\xi - x)(x - a) + \left[\eta - \frac{\eta}{\xi}(x - a) \right] \frac{\eta}{\xi}(x - a) = 0$$

$$(x - a) \left[\xi - x + \frac{\eta^2}{\xi} - \frac{\eta^2}{\xi^2} x + \frac{\eta^2}{\xi^2} a \right] = 0 \Rightarrow x = \frac{\xi(\eta^2 + \xi^2) + \eta^2 a}{\xi^2 + \eta^2},$$

подставив полученное выражение для x в (6) получим

$$y = \frac{\eta(\eta^2 + \xi^2) - \eta\xi a}{\xi^2 + \eta^2}$$

Подставив полученные формулы для x и y в уравнение окружности (5.21) получим уравнение подеры окружности:

$$\left[\frac{\xi(\eta^2 + \xi^2) + \eta^2 a}{\xi^2 + \eta^2} - a \right]^2 + \eta \left[\frac{(\eta^2 + \xi^2) - \xi a}{\xi^2 + \eta^2} \right]^2 - r^2 = 0$$

$$[\xi(\eta^2 + \xi^2) - \xi^2 a]^2 + [(\eta^2 + \xi^2) - \xi a]^2 - r^2(\eta^2 + \xi^2)^2 = 0$$

$$[(\eta^2 + \xi^2) - \xi a]^2(\eta^2 + \xi^2) - r^2(\eta^2 + \xi^2)^2 = 0$$

$$(\eta^2 + \xi^2 - \xi a)^2 - r^2(\eta^2 + \xi^2) = 0$$

т.к. точка $A(\xi, \eta)$ пробегает всю окружность, т.е. координаты ξ и η изменяются, то полученное уравнение и есть уравнение подеры окружности.

Рассмотрим другой подход к определению уравнения подеры, базирующийся на тангенциальных координатах.

Пусть точка, относительно которой строится подера совпадает с началом координат. Пусть заданная кривая выражается в плюкеровых координатах уравнением $f(u, v) = 0$, т.е. является огибающей прямых

$$ux + vy + 1 = 0, \quad (5.22)$$

а, следовательно, каждая из прямых (5.22) есть касательная к этой кривой. Перпендикуляр к такой касательной, проходящий через начало координат и точку $M(x, y)$ плоскости выразится уравнением

$$-vx + uy + 1 = 0 \quad (5.23)$$

Из равенств (5.22) и (5.23) находим:

$$u = -\frac{x}{x^2 + y^2}; \quad v = -\frac{y}{x^2 + y^2} \quad (5.24)$$

Эти выражения определяют соотношение между плюкеровыми координатами кривой l и точечными координатами ее подеры.

Соответствие, устанавливаемое формами (5.24) между прямыми и точками плоскости определяют подерное преобразование. Оно складывается из полярного преобразования относительно окружности единичного радиуса с центром в начале координат

$$(u = -\xi, \quad v = -r)$$

и инверсии относительно той же окружности заданной кривой

$$\left(\xi = \frac{x}{x^2 + y^2}; \quad \eta = \frac{y}{x^2 + y^2} \right).$$

Пример 2. Используем для решения задачу, рассмотренную в примере 1.

Получим тангенциальное уравнение окружности

$$(x - a)^2 + y^2 - r^2 = 0 \Rightarrow x^2 - 2ax + a^2 + y^2 - r^2 = 0$$

Для этого пересечем ее прямой

$$ux + vy + 1 = 0$$

Исключим из уравнений прямой и окружности

$$y = -\frac{1}{v}(1 + ux) \quad x = -\frac{1}{u}(1 + vy)$$

$$x^2(u^2 + v^2) - 2x(av^2 - u) + v^2(a^2 - r^2) + 1 = 0$$

$$y^2(u^2 + v^2) + 2yv(1 + au) + u^2(a^2 - r^2) + 2au + 1 = 0$$

Условием касания прямой и окружности будет совпадение корней полученного квадратного уравнения (детерминант должен быть равным нулю).

После приведения подобных

$$D = v^2[r^2(u^2 + v^2) - (ua + 1)^2] = 0 \quad D = u^2[r^2(u^2 + v^2) - (ua + 1)^2] = 0$$

Полученные равенства будут удовлетворяться, если

$$(u^2 + v^2)r^2 - (ua + 1)^2 = 0$$

$$\frac{(x^2 + y^2)r^2}{(x^2 + y^2)^2} - \left(1 - \frac{ax}{x^2 + y^2}\right)^2 = 0 \Rightarrow \frac{(x^2 + y^2 - ax)^2 - r^2(x^2 + y^2)}{(x^2 + y^2)^2} = 0$$

Это и есть тангенциальное уравнение рассматриваемой окружности.

Чтобы получить искомое уравнение подеры окружности используем формулы подерного преобразования

$$u = -\frac{x}{x^2 + y^2}; \quad v = -\frac{y}{x^2 + y^2}$$

Подерой окружности $(x - a)^2 + y^2 - r^2 = 0$ будет улитка Паскаля: $(x^2 + y^2 - ax)^2 - r^2(x^2 + y^2) = 0$.

Подера эллипса, гиперболы и параболы показаны на рис. 5.5; 5.6; 5.7.

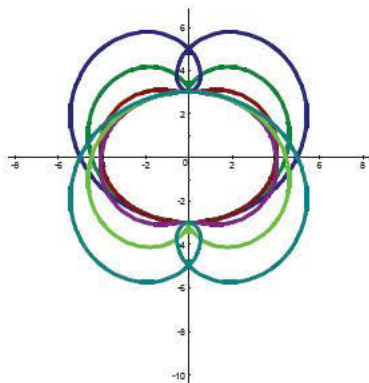
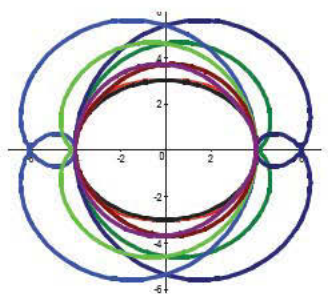


Рис. 5.5. Подера эллипса

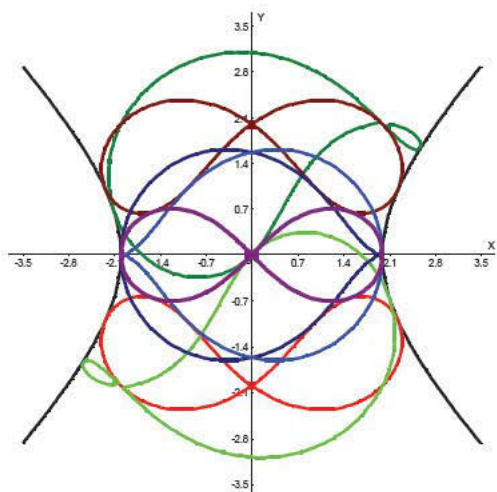


Рис. 5.5. Гипербола и ее подеры

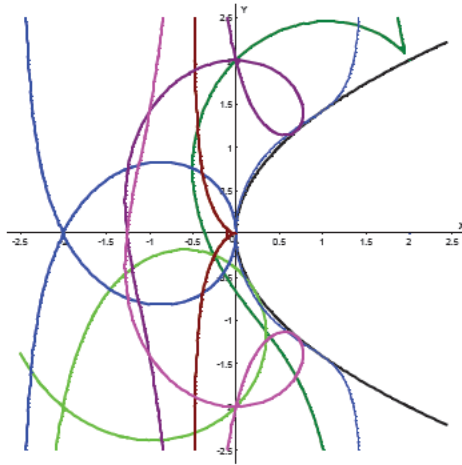


Рис. 5.5. Парабола и ее подеры

Параллельные (конгруэнтные) кривые.

Пусть α - угол составляемый с осью абсцис касательной к некоторой кривой l , а $p = \varphi(\alpha)$ - расстояние этой касательной от начала координат. Тогда касательная выразится уравнением:

$$\begin{aligned} x \cos(90 - \alpha) + y \sin(90 + \alpha) - p &= 0 \\ -x \sin \alpha + y \cos \alpha - \varphi(\alpha) &= 0 \end{aligned} \quad (5.25)$$

При различных значениях α , меняющихся от $-\infty$ до $+\infty$, уравнение (5.25) выразит множество касательных к заданной кривой. Следовательно уравнение (5.25) определяет кривую как огибающую ее касательных.

Если продифференцировать уравнение (5.25) по параметру α то получится уравнение

$$-x \cos \alpha + y \sin \alpha - \varphi'(\alpha) = 0 \quad (5.26)$$

которому соответствует прямая, перпендикулярная к прямой (5.25).

Прямая (5.26), проходящая через точку касания и будучи перпендикулярной к касательной является нормалью l .

При меняющихся значениях α уравнение (5.26) выразит семейство нормалей к заданной кривой и одновременно семейство касательных к новой кривой, которая является огибающей прямой (5.26). Следовательно, кривая, определяемая как огибающая семейства прямых

$$\begin{aligned} -x \cos \alpha + y \sin \alpha - \varphi'(\alpha) &= 0, \\ -x \sin \alpha + y \cos \alpha - \varphi(\alpha) &= 0, \end{aligned}$$

будет двойной кривой, заданной как огибающая семейства прямых

а эта последняя в свою очередь является эвольвентой первой кривой.

Чтобы от уравнения (5.26) перейти к обычному уравнению эволюты, достаточно продифференцировать уравнение (5.26) по параметру α и, присоединяя к полученному уравнению уравнение (5.26), исключить параметр α .

Рассмотрим уравнение

$$y \cos \alpha - x \sin \alpha - [\varphi(\alpha) + c] = 0 \quad (5.27)$$

в котором $c = \text{const}$.

По аналогии с уравнением (5.25) это уравнение при меняющихся значениях α выразит огибающую прямых (5.27), которые можно получить, фиксируя α . Назовем эту огибающую кривую l_1 . Между кривой l и второй кривой l_1 существуют определенные соотношения: в точках, соответствующих одному и тому же значению α , прямые (5.25) и (5.27), т.е. касательные к кривым l и l_1 параллельны между собой. Назовем такие точки соответствующими точками кривых l и l_1 и обозначим их через N и M . Кроме того, продифференцировав (5.27) по параметру α получим уравнение нормали к кривой l_1 , которое не отличается от уравнения (5.26). Следовательно, нормаль к кривой l в ее точке N совпадает с нормалью к кривой l_1 имеют в соответствующей точке M . Таким образом кривая l и кривая l_1 имеют в соответствующих точках параллельные касательные и общую нормаль. Такие кривые называются параллельными.

Действительно, расстояние между кривыми l и l_1 , считаемое по общей нормали, остается постоянным, равным c . Эти кривые называют так же эквидистантными.

Кривую, параллельную заданной, можно построить, откладывая на нормалях заданной кривой отрезки одной и той же длины. Параллельная кривая может быть определена также как траектория центра круга, которой катится вдоль заданной кривой или как огибающая окружностей неизменного радиуса, центры которых лежат на данной кривой.

Пример. Построить кривую параллельную эллипсу

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (5.28)$$

Будем рассматривать ее как огибающую окружностей

$$(x - \alpha)^2 + (y - \beta)^2 - c^2 = 0 \quad (5.29)$$

центры которых лежат на заданном эллипсе. Следовательно, параметры α и β удовлетворяют соотношению

$$\frac{\alpha^2}{a^2} + \frac{\beta^2}{b^2} - 1 = 0$$

используем следующий прием: умножим обе части уравнения (5.30) и сложим полученное с уравнением (5.29):

$$(x - \alpha)^2 + (y - \beta)^2 - c^2 + \lambda \left[\frac{\alpha^2}{a^2} + \frac{\beta^2}{b^2} - 1 \right] = 0.$$

Продифференцировав полученное выражение по α и β получим:

$$x - \alpha - \frac{\alpha\lambda}{a^2} = 0; \quad y - \beta - \frac{\beta\lambda}{b^2} = 0 \quad (5.31)$$

Определим из уравнений (5.31) α и β .

$$\alpha = \frac{x a^2}{a^2 + \lambda}; \quad \beta = \frac{y b^2}{b^2 + \lambda}.$$

Подставим полученные значения в (5.28) и (5.29)

$$\left. \begin{aligned} \frac{x^2 \lambda^2}{(a^2 + \lambda)^2} + \frac{y^2 \lambda^2}{(a^2 + \lambda)^2} &= c^2 \\ \frac{x^2 a^2}{(a^2 + \lambda)^2} + \frac{b^2 \lambda^2}{(b^2 + \lambda)^2} &= 1 \end{aligned} \right\}$$

Если из полученных двух уравнений исключить параметры в результате будет получено уравнение параллельной к эллипсу кривой – алгебраической кривой 8-го порядка. Если в качестве параметра взять λ , то параметрическое уравнение будет иметь вид:

$$x = \left(1 + \frac{a^2}{\lambda}\right) \sqrt{\frac{\lambda^2 - b^2 c^2}{a^2 - b^2}},$$

$$y = \left(1 + \frac{b^2}{\lambda}\right) \sqrt{\frac{a^2 c^2 - \lambda^2}{a^2 - b^2}}$$

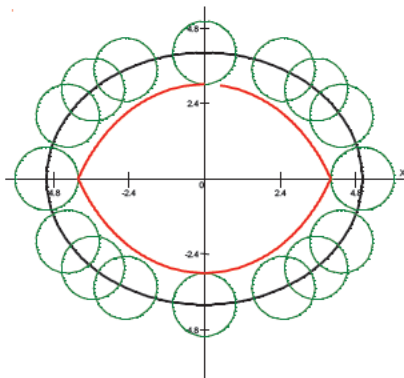


Рис. 5.8. Эллипс и его параллель

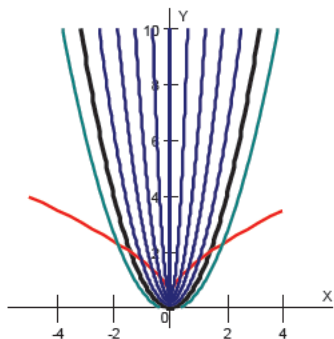


Рис. 5.9. Парабола и ее параллели

5.6 Катакаустика

Каустика – огибающая лучей, отраженных или преломленных данной линией.

Пусть $l(t)$ плоская кривая единичной скорости. Рассмотрим огибающую нормалей этой кривой. Такая огибающая представляет определенный физический интерес. Представим себе световые лучи, исходящие одновременно из всех точек $l(t)$ и распространяющихся вдоль нормалей. Тогда называется исходным волновым фронтом, а кривая, на которой свет фокусируется – каустикой («каустика» означает «жгучая»).

Огибающая нормалей – это в точности геометрическое место центров кривизны кривой (эволюта).

Катакаустика – огибающая прямых, образующая в точках заданной кривой тот же угол с нормалью, как ряд прямых, выходящих из данной точки, или ряд параллельных прямых. Она является огибающей лучей, полученных путем отражения пучка лучей выходящих из данной точки или параллельных.

Диакатаустика – огибающая лучей преломленных.

Катакаустикой данной кривой называется огибающая лучей, отраженных этой кривой.

Рассмотрим катакаустики кривых для частного случая, когда падающие лучи параллельны между собой.

Пусть $f(\xi, \eta)$ (5.32)

- уравнение отражающей кривой, катакаустикой которой нужно построить.

Пусть падающие лучи параллельны оси ординат, тогда уравнение любого из этих лучей имеет вид:

$$x - \xi = 0.$$

Уравнение нормали к кривой в точке $M(\xi, \eta)$ запишется в виде:

$$x - \xi - (y - \eta) \frac{d\eta}{d\xi} = 0.$$

Если θ угол, составленный нормалью с осью абсцисс, а γ - угол, образованный отраженным лучом с той же осью, то [т.к. угол падения равен углу то $\gamma - 180 - 2\theta$]

Откуда

$$\tan \gamma = \tan(2\theta - 90^\circ) = -\frac{1 - \tan^2 \theta}{2 \tan \theta}$$

Т.к. $\tan \theta = -\frac{d\xi}{d\eta}$, то $\tan \gamma = -\frac{1 - \left(\frac{d\xi}{d\eta}\right)^2}{2 \frac{d\xi}{d\eta}}$, следовательно, отраженный луч выра-

зится уравнением:

$$(x - \xi) \left[1 - \left(\frac{d\eta}{d\xi} \right)^2 \right] + 2(y - \eta) \frac{d\eta}{d\xi} = 0 \quad (5.33)$$

т.к. катакаустика является огибающей лучей, то для ее определения необходимо продифференцировать уравнение (5.33) по ξ , а затем из полученного уравнения и уравнения (5.33) выразить x и y координаты точки P катакаустики, соответствующей точке M заданной кривой.

Дифференцируя (5.33) по ξ получим

$$-(x - \xi) \frac{d\eta}{d\xi} + (y - \eta) = \frac{1 + \left(\frac{d\eta}{d\xi}\right)^2}{2 \frac{d^2\eta}{d\xi^2}} \quad (5.34)$$

Из уравнений (5.33) и (5.34) имеем

$$x = \xi - \frac{\frac{d\eta}{d\xi}}{\frac{d^2\eta}{d\xi^2}}; \quad y = \eta + \frac{1 + \left(\frac{d\eta}{d\xi}\right)^2}{2 \frac{d^2\eta}{d\xi^2}} \quad (5.35)$$

Уравнения (5.35) выражают координаты x и y точки P катакаустики в функциях ξ и η - координатах точки M заданной кривой. Присоединяя к этим

уравнениям уравнение (5.32) отражающей кривой и исключая параметры ξ и η , получим уравнение катакаустики.

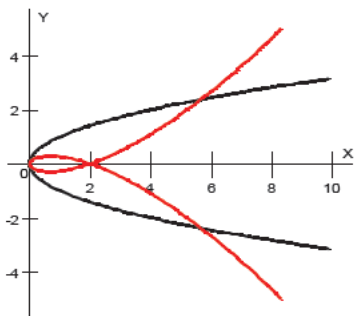


Рис. 5. 10. Парабола и ее катакаустика

Пример. Отражающая кривая – парабола

$$\eta^2 = 2p\xi$$

а падающие лучи параллельны оси ординат. Используя (5.35) получим

$$x = 3\xi; \quad y = \frac{\eta(3p - 2\xi)}{2p}$$

Присоединяя к этим уравнениям уравнение параболы и используя ξ и η получим уравнение катакаустики параболы: $54py^2 = x(2x - 9p)^2$.

5.7. Конхондное преобразование

Определение. Конхондальным называется преобразование кривой линии при котором радиус-векторы ее точек, исходящие из хаданного полюса,увеличиваются или уменьшаются на одну и ту же величину d .

Если базовая кривая задана в полярной системе координат и ее полярное уравнение имеет вид: $r = f(\varphi)$, то уравнение ее конхонды будет иметь вид: $r = f(\varphi) \pm d$. Обычно под конхондой понимают конхонду прямой, так называемую конхонду Никомеда — по имени древнегреческого геометра Никомеда (ок. 250—150 гг. до н. э.), который использовал ее для решения задач о трисекции угла и удвоении куба. Если расстояние от полюса O полярной системы до данной прямой равно a , то уравнение конхонды Никомеда имеет вид:

$$\rho = \frac{a}{\sin\varphi} \pm d$$

или в декартовых координатах с началом в полюсе:

$$(x-a)^2(x^2+y^2) - d^2x^2 = 0$$

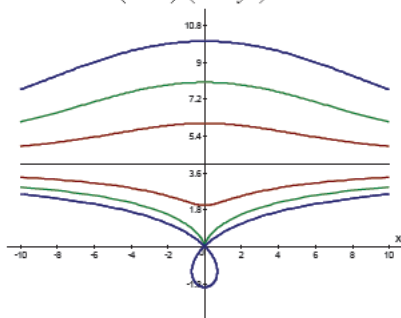


Рис. 5.11. Конхонда Никомеда

Конхонда Никомеда (рис. 5.11) - это алгебраическая кривая 4-го порядка, состоящая из двух ветвей. В зависимости от соотношения a и d ветви конхонды

Никомеда будут иметь разную форму; при этом полюс O может быть узловой точкой ($d > a$), точкой заострения ($d = a$) и изолированной точкой ($d < a$). Обе ветви конхоиды Никомеда асимптотически приближаются в обоих направлениях к прямой.

Конхоида окружности относительно полюса, лежащего на окружности, называют улитками Паскаля (рис. 5.12).

На рис. 5.13 приведены изображения конхоид эллипса относительно полюса, совпадающего с началом координат.

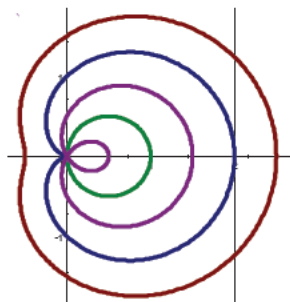


Рис. 5.12 Конхоиды окружности.

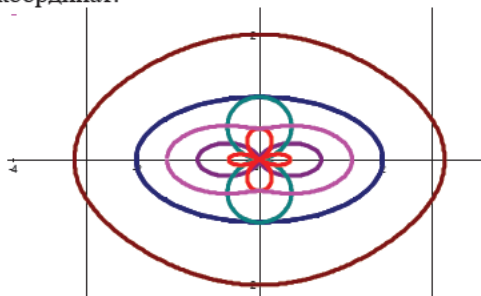


Рис. 5.13. Конхоиды эллипса

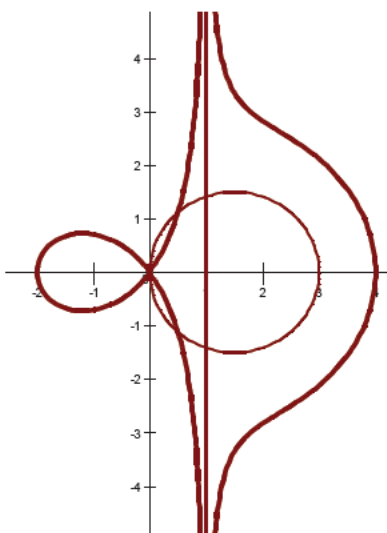


Рис. 5.14. Конхоида Слюза

Открытие конических сечений приписывается Менехму (4 век до н.э.). Их теорию обстоятельно развил Аполлоний Пергский (2 век до н.э.) рассматривая плоские сечения произвольных конусов с круговым основанием.

*Б. Л. ван дер Варден
«Пробуждающаяся наука»*

Кривые второго порядка (коники) определяются алгебраическим уравнением 2-го порядка в декартовой системе координат:

$$F(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0 \quad (6.1)$$

где $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$.

Разделив обе части уравнения (3) на коэффициент какого либо из его членов (или на свободный член) при условии, что $a_{ij} \neq 0$, получим пять отношений коэффициентов, которые и будут параметрами уравнения (6.1) и одновременно параметрами коники. Следовательно, кривая второго порядка может быть задана:

- пятью линейно независимыми точками (никакие три из них не лежат на одной прямой);
- пятью касательными к ней (при этом никакие три касательные не должны пересекаться в одной точке);
- любой комбинацией из касательных прямых и точек, удовлетворяющих указанным выше условиям.

6.1. Классификация кривых 2-го порядка

Подвергнем выражение (6.1) аффинным преобразованиям.

Перенесем начало координат в произвольную точку $M(X_0, Y_0)$, подставив

$$x = X - X_0, y = Y - Y_0 \quad (6.2)$$

После приведения подобных формула (6.3) примет вид

$$F(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2x(a_{11}x_0 + a_{12}y_0 + a_{13}) + 2y(a_{22}x_0 + a_{12}y_0 + a_{13}) + (a_{11}x_0^2 + 2a_{12}x_0y_0 + a_{22}y_0^2 + a_{31}x_0 + a_{32}y_0 + a_{33}) = 0. \quad (6.3)$$

Из полученного выражения видно, что коэффициенты при членах второй степени не изменились.

Определение. Точка $M(x_0, y_0)$ называется центром линии (6.1), если

$$\begin{aligned} a_{11}x_0 + a_{12}y_0 + a_{13} &= 0, \\ a_{21}x_0 + a_{22}y_0 + a_{23} &= 0 \end{aligned} \quad (6.4)$$

т.е. если при переносе начала координат в эту точку в уравнении (1) пропадают члены первой степени.

Соотношения (6.4) определяют две прямые линии

$$a_{11}x + a_{12}y + a_{13} = 0 \quad \text{и} \quad a_{21}x + a_{22}y + a_{23} = 0$$

Если эти прямые пересекаются, то имеется только один центр. Определим его координаты (x_0, y_0) :

$$x_0 = (a_{12}a_{23} - a_{13}a_{22}) / (a_{11}a_{22} - a_{12}^2),$$

$$y_0 = (a_{12}a_{23} - a_{11}a_{13}) / (a_{11}a_{22} - a_{12}^2). \quad (6.5)$$

Если прямые (6.4) параллельны, то центр отсутствует. Наконец, прямые могут совпадать. В этом случае центры заполняют целую прямую. Таким образом, для линии (6.1) возможны следующие три случая:

- Существует единственное решение $(a_{11}a_{22} - a_{12}^2 \neq 0)$.
- Не существует ни одного центра. $(a_{11}a_{22} - a_{12}^2 = 0)$.
- Имеется целая прямая центров.

$$(a_{11}a_{22} - a_{12}^2 \neq 0, a_{12}a_{23} - a_{13}a_{22} = 0 \text{ или } a_{12}a_{23} - a_{11}a_{13} = 0)$$

Если существует единственный центр, то линия (1) называется центральной линией второго порядка, в противном случае - нецентральной. Любой центр линии (1) является ее центром симметрии.

Перенесем начало координат в точку O. Тогда выражение (1) примет вид

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + a_{33} = 0 \quad (6.6)$$

Введем в рассмотрение определители:

$$I = a_{11} + a_{22} \quad D = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad (6.7)$$

где $a_{ik} = a_{ki}$ ($i, k = 1, 2, 3$).

D и Δ называются малым и большим определителями уравнения (6.1).

Теорема 1. Линия второго порядка тогда и только тогда является центральной, когда $D \neq 0$.

Следствие. Центральные линии второго порядка:

- Эллипс: $D = 1$ и $\Delta = 1$. Имеет один центр $(0, 0)$.
- Гипербола: $D = -1$ и $\Delta = 1$. Имеет один центр $(0, 0)$.
- Пара действительных пересекающихся прямых: $D = -1$ и $\Delta = 0$. Имеет один центр $(0, 0)$.

(Мнимый эллипс и пара мнимых пересекающихся прямых. Имеет чисто теоретический интерес.)

■ Нецентральной линией без центра является только парабола.

Рассмотрим взаимное расположение кривой (6.1) и произвольной прямой

$$x = x_0 + tl, \quad y = y_0 + tm. \quad (6.8)$$

Общие точки линии (6.1) и прямой (6.8) определяются из уравнения

$$(a_{11}l^2 + a_{12}lm + a_{22}m^2)t^2 + 2[a_{11}lx_0 + a_{12}(ly_0 + mx_0) + a_{22}my_0 + a_{13}l + a_{23}m]t + F(x_0, y_0) = 0 \quad (6.9)$$

Определение. При $a_{11}l^2 + 2a_{12}lm + a_{22}m^2 \neq 0$ прямая (6.8) имеет неасимптотическое направление [по отношению к линии (6.1)], а при $a_{11}l^2 + 2a_{12}lm + a_{22}m^2 = 0$ эта прямая имеет асимптотическое направление.

Для прямой неасимптотического направления выражение (6.9) является уравнением второй степени. Оно может иметь: - два корня, в этом случае прямая (6.8) пересекает кривую (6.1) в двух различных точках, один (двойной) корень, в этом случае прямая (6.8) имеет с линией (6.1) одну общую точку (касается кривой), либо не имеет ни одного корня - прямая (6.8) и линия (6.1) не пересекаются

Определение. Прямая асимптотического направления, не имеющая общих точек с линией (6.1), называется ее асимптотой.

Определение. Линия (6.1) называется линией

- Эллиптического типа, если $D > 0$,
- Гиперболического типа, если $D < 0$,
- Параболического типа, если $D = 0$.

Гиперболической ветвью я называю ту, которая бесконечно приближается к какой-нибудь из асимптот, а параболической ту, которая лишена асимптоты. Эти ветви лучше всего распознаются с помощью касательных

В самом деле, если точка касания уходит в бесконечность, касательная гиперболической ветви совпадает с асимптотой, а касательная ветви параболической сама удаляется в бесконечность, исчезает и не находится нигде.

Определение. Диаметр кривой второго порядка называется прямой, проходящая через середины параллельных хорд.

Диаметр, сопряженный хордам, образующим угол α с положительным направлением оси Ox , определяется уравнением:

$$(a_{11}x + a_{12}y + a_{13}) \cos \alpha + (a_{21}x + a_{22}y + a_{23}) \sin \alpha = 0. \quad (6.10)$$

Для центральных кривых ($D \neq 0$) все диаметры пересекаются в одной точке-центре. Для нецентральных кривых ($D = 0$) все диаметры параллельны или совпадают.

Определим из уравнения (6.10) угол α

$$\operatorname{tg} \alpha = a_{12} / (a_{22} - a_{11}) \quad (6.11)$$

Полученные результаты приведены в таблице.

			Невырожденные конические сечения $\Delta \neq 0$	Вырожденные (распадающиеся) Конические сечения $\Delta = 0$
Центральные конические сечения $D \neq 0$	$D > 0$	$\frac{\Delta}{I} < 0$	Действительный эллипс (окружность, если $I^2 - 4D, a_{11} = a_{22} = 0$)	
		$\frac{\Delta}{I} > 0$	Минимальный эллипс (ни одной действительной точки)	
		$\frac{\Delta}{I} = 0$		Действительная точка пересечения двух мнимых прямых (эллипс, выродившийся в точку)
	$D < 0$		Гипербола	Пара действительных пересекающихся прямых (выродившаяся гипербола)
Конические сечения без центра или с неопределенным центром $D = 0$		$A' > 0$	Парабола	Пара мнимых параллельных прямых (ни одной действительной точки)
		$A' < 0$		Пара действительных параллельных прямых
		$A' = 0$		Одна действительная прямая (пара совпавших прямых)

$$\text{где } A' = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

Кривые 2-го порядка делятся на центральные (эллипс, гипербола) – имеющие центр симметрии, и параболические кривые – центр симметрии не существует.

Для центральных кривых ($D \neq 0$) все диаметры кривой пересекаются в одной точке (центре). При $D = 0$ все диаметры параллельны или совпадают.

Каждый из двух сопряженных диаметров центральной кривой 2-го порядка делит пополам хорды, параллельные другому диаметру.

Диаметр, перпендикулярный к сопряженным ему хордам, называется главной осью кривой 2-го порядка и является осью симметрии кривой. Кривая имеет либо 2 взаимно перпендикулярные оси (гиперболы, эллипс), либо каждый диаметр является главной осью (окружность). При $D = 0$ (парабола) кривая имеет одну главную ось. Точки пересечения кривой с главными осями называются вершинами этой кривой.

Направление главных осей и сопряженных им хорд называется главным направлением кривой 2-го порядка. Угол между положительным направлением оси Ox и каждым из двух главных направлений кривой определяется формулой (6.11). Окружность имеет неопределенные главные направления.

6.2. Приведение уравнения кривой 2-го порядка к стандартному (каноническому) виду

Если ввести новую систему координат, совершив поворот осей на угол α удовлетворяющей уравнению (11) и перенос начала по формулам (5), то уравнение (1) любой невырожденной кривой 2-го порядка может быть приведено к следующему стандартному виду:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a^2 = -\frac{1}{\lambda_2} \frac{A}{D} = -\frac{A}{\lambda_1 \lambda_2^2}; b^2 = -\frac{A}{\lambda_2 \lambda_1^2} - \text{эллипс} \quad (\text{рис. 6.1})$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, a^2 = -\frac{1}{\lambda_1} \frac{A}{D} = -\frac{A}{\lambda_2 \lambda_1^2}; b^2 = -\frac{A}{\lambda_1 \lambda_2^2} - \text{гипербола} \quad (\text{рис. 6.2})$$

$$y^2 = 2pxp = \frac{1}{i} \sqrt{-\frac{A}{I}} = \frac{1}{\lambda_1^2} \sqrt{-\frac{A}{\lambda_1}} > 0 \quad \lambda_2 = 0 - \text{парабола} \quad (\text{рис. 6.3})$$

где $\lambda_1 \geq \lambda_2$ – корни характеристического уравнения.

$$F_0(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$$

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0 \text{ или } \lambda^2 - I\lambda + D = 0.$$

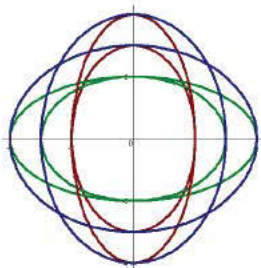


Рис. 6.1. Эллипсы

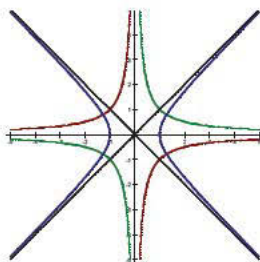


Рис. 6.2. Гиперболы

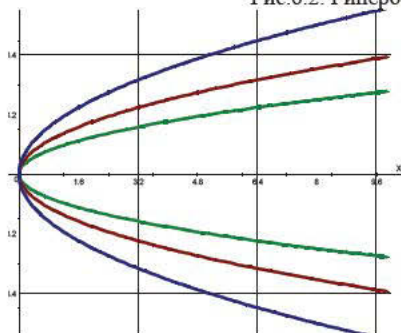


Рис. 6.3. Параболы

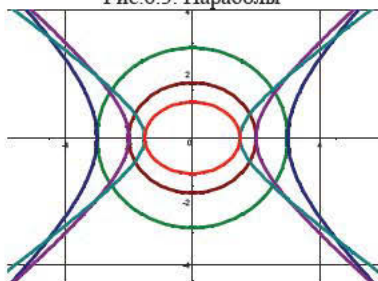


Рис. 6.4. Софокусные эллипсы и гиперболы

6.3. Геометрическое определение невырожденной кривой 2-го порядка

Преобразованием начала координат любое уравнение (6.1) может быть приведено к виду:

$$y^2 = 2px + (1 - e)x^2 \quad (p > 0) \quad (6.12)$$

В этом случае кривая проходит через начало координат новой системы, ось Ox - ось симметрии кривой (рис. 6.5).

На рис. 6.5. приведены коники с общей вершиной:

p - масштабный коэффициент,

e - эксцентриситет,

$e=0$ - окружность, $0 \leq e < 1$ - эллипс ($e=0.5, 0.7$)

$e=1$ - парабола, $e > 1$ - гипербола ($e=2, 3$)

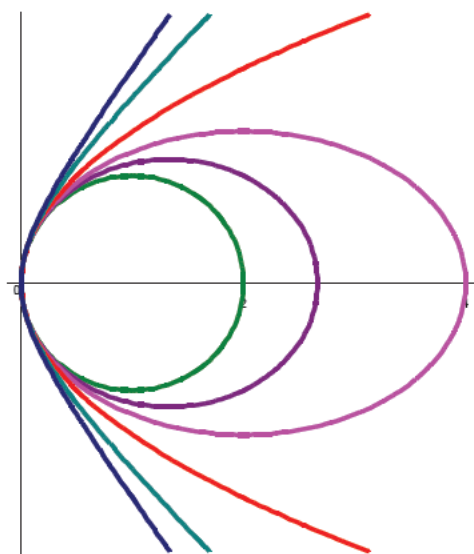


Рис. 6.5. Коники с общей вершиной

Невырожденная кривая 2-го порядка является геометрическим местом точек, отношение расстояний которого $e \geq 0$ (эксцентриситет) от данной точки (фокуса) и от данной прямой (директрисы) постоянно. При $e < 1$ - эллипс ($e = 0$ окружность), $e > 1$ - гипербола; $e = 1$ - парабола.

Уравнение директрисы кривой:

$$x = -\frac{p}{e(1+e)} \quad (6.13)$$

координаты фокуса:

$$x = \frac{p}{(1+e)}, \quad y = 0 \quad (6.14)$$

директриса перпендикулярна к оси симметрии, проходящей через фокус и вершину кривой – фокальную ось.

Если кривая центральная (эллипс или гипербола), то прямая

$$x = \frac{p}{(1-e^2)} = a$$

- ось симметрии кривой, а, следовательно, кривая имеет два фокуса и две директрисы.

Уравнение касательной к кривой (3) в ее точке (x_1, y_1) :

$$(a_{11}x_1 + a_{12}y_1 + a_{13})x + (a_{21}x_1 + a_{22}y_1 + a_{23})y + (a_{31}x_1 + a_{32}y_1 + a_{33}) = 0 \quad (6.15)$$

Уравнение нормали в той же точке (x_1, y_1)

$$\frac{x-x_1}{a_{11}x_1+a_{12}y_1+a_{13}} = \frac{y-y_1}{a_{21}x_1+a_{22}y_1+a_{23}} \quad (6.16)$$

Уравнение (6.7) определяет прямую, называемую полярной точки (x_1, y_1) относительно кривой второго порядка (6.1), независимо от того, лежит ли точка (x_1, y_1) на кривой или нет. Точка (x_1, y_1) называется полюсом прямой (6.7). Поляра точки кривой есть касательная к кривой в этой точке.

$$\rho = \frac{p}{1 + e \cos \varphi} \quad (6.17)$$

- уравнение кривых 2-го порядка в полярных координатах.

- параметр, e - эксцентриситет, полюс находится в фокусе, полярная ось направлена от фокуса к ближайшей вершине.

6.4. Параметризация кривых второго порядка

Параметрическая форма для кривой генерирует различные точки этой кривой в зависимости от значения параметра. Параметрическая форма напоминает движение материальной точки во времени, которое можно преобразовать в движение пера, рисующего кривую. Траектория движения точки вдоль кривой фиксируется двумя функциями $x(\cdot), y(\cdot)$. Можно рассматривать пару $x(t), y(t)$ как текущие координаты точки в момент времени t .

Пусть прямая проходит через точки $A(x_A, y_A)$ и $B(x_B, y_B)$ тогда

$$x(t) = x_A + (x_B - x_A)t$$

$$y(t) = y_A + (y_B - y_A)t$$

При $t = 0$ $x(0) = x_A, y(0) = y_A$; $t = 1$ $x(1) = x_B, y(1) = y_B$.

При изменении $0 \leq t \leq 1$ текущая точка P пробежит все точки прямой от точки A до B .

Кривая второго порядка, приведенная к каноническому виду, может быть параметризована с помощью тригонометрических или гиперболических функций:

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 \Rightarrow x(t) = a \cos t; y(t) = b \sin t \quad 0 \leq t \leq 2\pi - \text{эллипс}$$

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1 \Rightarrow x(t) = a \operatorname{ch}(t), y(t) = b \operatorname{sh}(t) - \text{гипербола.}$$

Положив: $\operatorname{sint} = 2u/(1+u^2)$, $\operatorname{cost} = (1-u^2)/(1+u^2)$ получим для эллипса рациональную параметризацию:

$$x = 2at/(1+t^2), y = b(1-t^2)/(1+t^2).$$

Положив: $\operatorname{th}(t/2) = u$ получим рациональную параметризацию для гиперболы:

$$x = a(1+t^2)/(1-t^2), y = 2bt/(1-t^2).$$

Парабола имеет параметризацию:

$$y^2 - 4ax = 0 \Rightarrow x(t) = at^2, y(t) = 2at - \text{парабола}$$

6.5. Суперконики

Особой разновидностью эллипса является суперэллипсы – семейство эллипсоподобных форм, которые во многих ситуациях, в особенности в дизайне, дают замечательный эффект. Неявная форма для суперэллипса имеет вид:

$$\left(\frac{x}{a}\right)^n + \left(\frac{y}{b}\right)^n = 1$$

где n - параметр, называемый выпуклостью (следует брать n - четным числом) (рис. 6.6) При $n = 2$ суперэллипс становится обычным эллипс. Параметрическое представление суперэллипса

$$x(t) = a \cos t \left| (\cos t)^{2/n-1} \right|;$$

$$y(t) = b \sin t \left| (\sin t)^{2/n-1} \right|,$$

для $0 \leq t \leq 2\pi$. Показатели степени при $\sin t$ и $\cos t$ на самом деле равны $\frac{2}{n}$, а данная специфическая форма записи использована для того, чтобы избежать возведения отрицательного числа в дробную степень. При $n > 1$ выпуклость направлена наружу, а при $n < 1$ - вовнутрь.

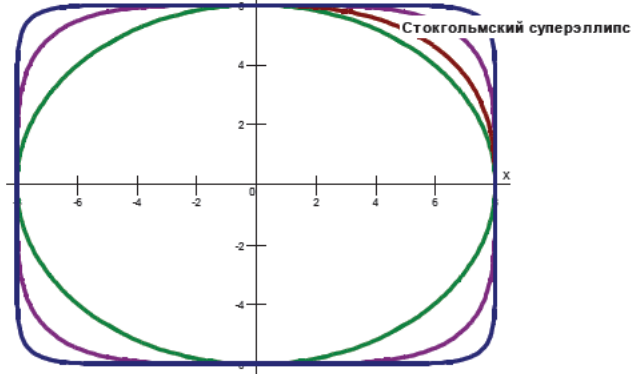


Рис. 6.6. Суперэллипсы

Аналогичным образом может быть определена супергипербола (рис. 6.7)

$$\left(\frac{x}{a}\right)^n - \left(\frac{y}{b}\right)^n = 1$$

$$x(t) = a \sec t \left| (\sec t)^{2/n-1} \right|;$$

$$y(t) = b \tan t \left| (\tan t)^{2/n-1} \right|.$$

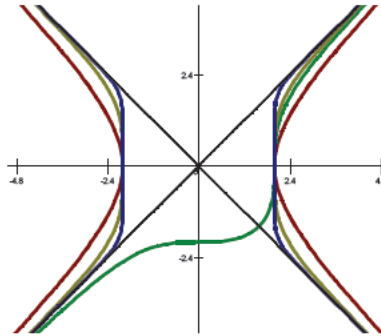


Рис. 6.7. Супергиперболы

Историческая справка. Суперэллипсы впервые были исследованы в 1818 г. французским инженером, математиком, физиком, членом Парижской и член-корреспондентом Петербургской Академий Наук Габриелем Ламе (22.07.1795-01.03.1870) в работе «Исследование различных методов, применяе-

мых для решения геометрических задач». Позже, в 1959 г. выдающийся изобретатель Пит Хейн (известный как создатель куба Some и игры Hex) решал задачу проектирования кольцевой траекторной развязки в Стокгольме. Окружность должна была вписаться в прямоугольник (с отношением сторон $\frac{a}{b} = \frac{6}{5}$), который определялся другими дорогами. Кроме того было необходимо обеспечить плавный транспортный поток, а также быть приятной для глаза. Эллипс получался слишком заостренным на концах транспортной магистрали, поэтому Хейн использовал более плоскую кривую с более прямыми сторонами – суперэллипс с $n = 2,5$, при котором выпуклость была наиболее подходящей (рис. 6.5) Стокгольм мгновенно принял идею суперэллипса для своего нового центра. Эти кривые были «удивительно подходящими, ни слишком округлыми, ни слишком угловатыми, удачное сочетание красоты эллипсов и прямоугольников». (М. Гарднер «Мозаики Пенрауза для дверных шифров»).

...Мне кажется совершенно ясным, что если -как это и делают – почитать геометрическим то, что вполне определено и точно, а механическим то, что не таково, и если рассматривать геометрию как науку, которая учит вообще познанию мер всех тел.то из нее также мало следует исключать самые сложные, как и самые простые линии, если только можно представить себе, что эти линии описаны непрерывным движением или же несколькими такими последовательными движениями, из которых последующие вполне определяются им предшествующими,- ибо этим путем всегда можно точно узнать их меру.

*Рене Декарт
«Геометрия» (1637)*

Отдельные кривые третьего и высших порядков были известны еще древним грекам, хотя и не выражавших их с помощью уравнений в прямолинейных координатах. Первые

исследования кривых методами новой аналитической геометрии были проведены Декартом и его ближайшими последователями. Ньютон построил подробно развитую теорию кривых третьего порядка в мемуаре «Перечисление линий третьего порядка», опубликованном в 1704 году. Именно классификация кривых 3-го порядка, предложенная Ньютоном (были предложены и другие классификации, например, Плюккером (1835г.),Адамовым А.А. (1918) и другие), является и сегодня наиболее удобной и наиболее часто используемой.

7.1.Классификация кривых третьего порядка.

При помощи элементарных преобразований Ньютон привел общее уравнение кривой третьего порядка

$$Ax^3+3Bx^2y+3Cxy^2+Dy^3+3Ex^2+6Fxy+3Gy^2+3Hx+3Qy+L=0 \quad (7.1)$$

к одной из четырех канонических форм:

A: $xy^2+ey=ax^3+bx^2+cx+d$

B: $xy= ax^3+bx^2+cx+d$

C: $y^2= ax^3+bx^2+cx+d$

D: $y= ax^3+bx^2+cx+d$

По коэффициентам канонического уравнения он составил вспомогательные уравнения 4-й и 3-й степени:

$$ax^4+bx^3+cx^2+dx+1/4e^2=0$$

$$ax^3+bx^2+cx+d=0$$

которые можно назвать характеристическими. В зависимости от различных соотношений между корнями характеристического уравнения Ньютон делит все

кривые на 7 классов, 14 родов, 72 типа. Рассмотрим предложенную классификацию более подробно.

Класс 1. Гиперболическая гипербола.

Гиперболическая гипербола :

$$xy^2 + ey = ax^3 + bx^2 + cx + d \quad (a > 0) \quad (7.2)$$

имеет три обыкновенные асимптоты (т. е. она имеет с кривой в ее бесконечно удаленной точке касание 1-го порядка):

$$x=0, \quad y = \pm \sqrt{a} \left(x + \frac{b}{2a} \right);$$

которые пересекаются в точках $(0, \frac{b}{2\sqrt{a}})$, $(0, -\frac{b}{2\sqrt{a}})$, $(-\frac{b}{2a}, 0)$, которые являются вершинами асимптотического треугольника.

Род 1. Адиаметральная гиперболическая гипербола ($e \neq 0$).

Соотношение (7.2) можно рассматривать как квадратное уравнение относительно переменной y , тогда

$$y = \frac{-e \pm 2 \sqrt{ax^4 + bx^3 + cx^2 + dx + \frac{1}{4}e^2}}{2x} = \frac{-e \pm \sqrt{a(x-x_1)(x-x_2)(x-x_3)(x-x_4)}}{2x}. \quad (7.3)$$

Приравняв нулю подкоренное выражение, получим характеристическое уравнение для кривых первого рода:

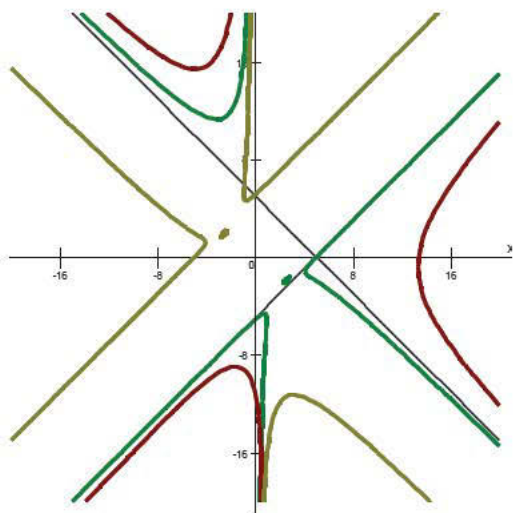
$$ax^4 + bx^3 + cx^2 + dx + \frac{1}{4}e^2 = a(x-x_1)(x-x_2)(x-x_3)(x-x_4) = 0 \quad (7.4)$$

где x_1, x_2, x_3, x_4 — корни уравнения (7.4). Переменная y принимает действительные или мнимые значения в зависимости от знака подкоренного выражения в (7.3). При этом возможны девять случаев. Рассмотрим их.

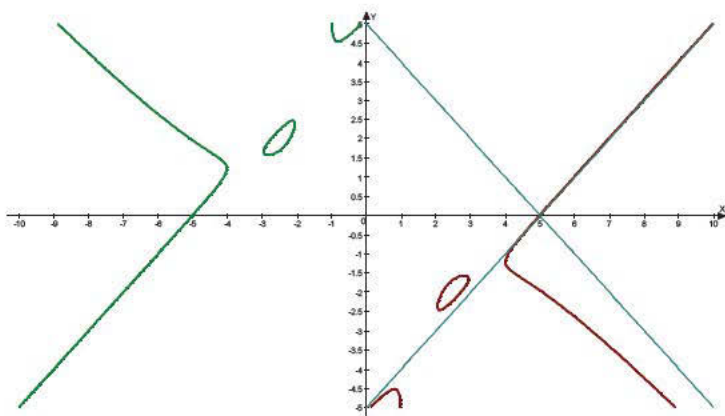
1. Все корни характеристического уравнения - действительные, различные и все одного знака. Пусть $x_1 < x_2 < x_3 < x_4$. Подкоренное выражение (7.3) неотрицательное при $x \leq x_1$; $x_2 \leq x \leq x_3$; $x_4 \leq x$. В промежутке $[x_2, x_3]$ y принимает только конечные значения, кривая имеет замкнутую форму - овал (рис.7.1 б), находящийся внутри асимптотического треугольника. Кроме овала кривая содержит три бесконечные гиперболические ветви в промежутках $[x_4, +\infty)$, $(-\infty, x_1)$, $(-\infty, 0)$. Кривая пересекает асимптоты в точках

$$\left(0, \frac{d}{e} \right), \left(\frac{4ad \pm 2be\sqrt{a}}{b^2 - 4ac \pm 4ae\sqrt{a}}, \frac{8a^2 d + b^2 - 4abc}{2a(b^2 - 4ac \pm 4ae\sqrt{a})} \right)$$

и пересекает ось абсцисс (рис. 7.1а).



а



а

Рис. 7.1. 6

2. Все корни характеристического уравнения - действительные, различные, два из них положительные, два - отрицательные. Кривая состоит из трех бесконечных ветвей, одна из которых змеевидная, по определению Ньютона, две другие - гиперболические (рис. 7.2). Между собой ветви не пересекаются, но все пересекают ось абсцисс. Змеевидная ветвь пересекает все три асимптоты на конечном расстоянии.

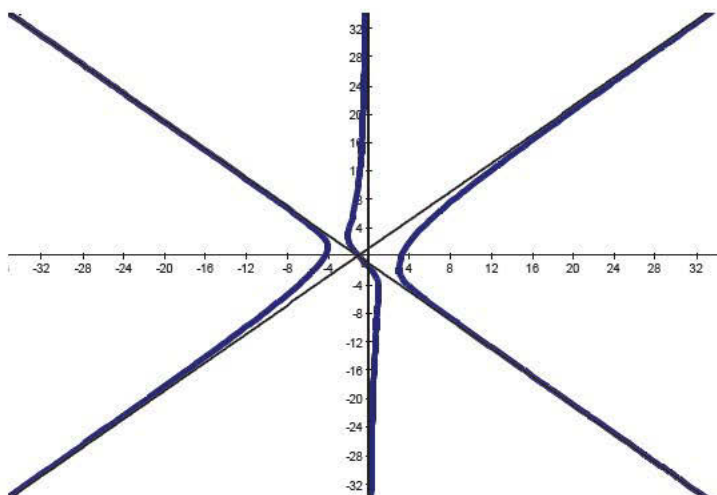


Рис. 7.2.

3. Все корни характеристического уравнения - действительные, два из них равны между собой, причем кратный корень больше или меньше обоих разных корней и отличается от них знаком. На рис. 7.3. изображена кривая, заданная уравнением:

$$xy^2 + 2\sqrt{6}y = x^3 - 3x^2 - 3x + 7$$

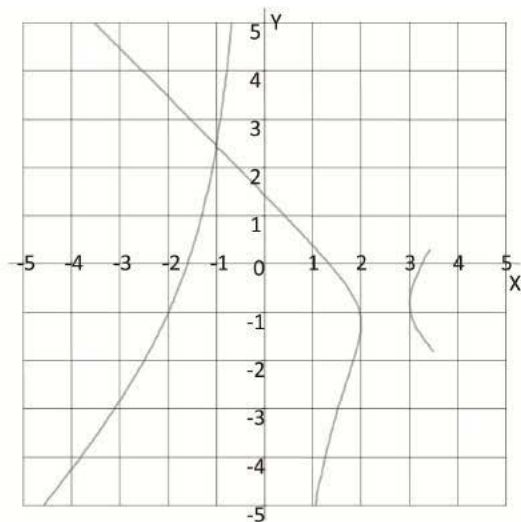


Рис. 7.3.

Кривая состоит из трех бесконечных ветвей; две из них пересекаются в узловой точке (абсцисса которой равна кратному корню характеристического уравнения).

4. Знаки всех корней характеристического уравнения одинаковы (рис. 7.4)

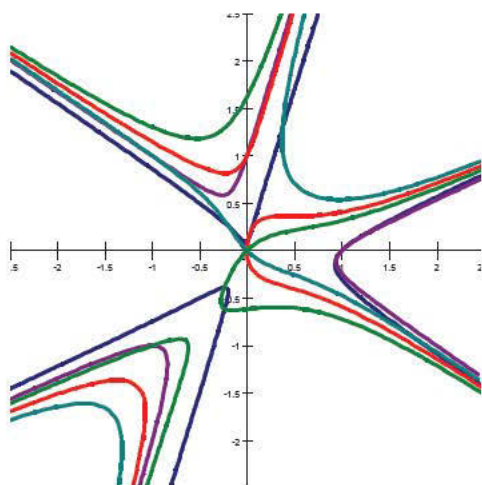


Рис. 7.1. Кривые третьего порядка.

синяя-

$$x^3 - x^2y - 3xy^2 + y^3 - x^2 = 0$$

фиолетовая-

$$x^3 - x^2y - 3xy^2 + y^3 - x^2 - y^2 = 0$$

красная-

$$x^3 - x^2y - 3xy^2 + y^3 - x^2 - y^2 + x = 0$$

голубая

$$x^3 - x^2y - 3xy^2 + y^3 - x^2 - y^2 + x + y = 0$$

зеленая-

$$x^3 - x^2y - 3xy^2 + y^3 - x^2 - y^2 + x - y = 0$$

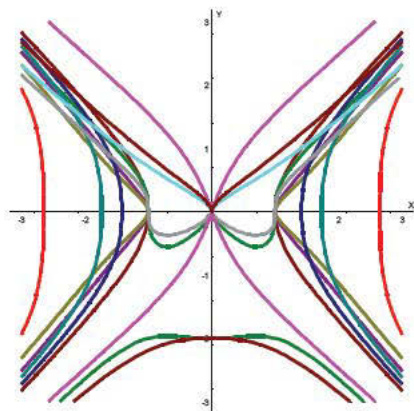


Рис. 7.4. Кривые четвертого порядка

Важнейшие кривые

Полукубическая парабола (рис.1)

$$a^2x^3 - y^2 = 0, a > 0 \quad a = 1, 2, 4 \quad x = t^2, y = at^3$$

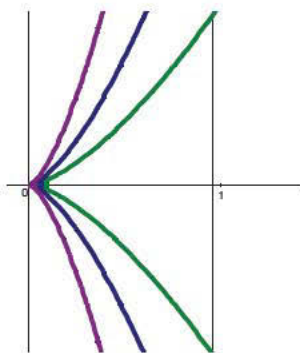


Рис. 7.6.

Локон Аньези

$$(x^2 + a^2)y - a^3 = 0 \quad a > 0 \quad a = 1, 2, 4$$

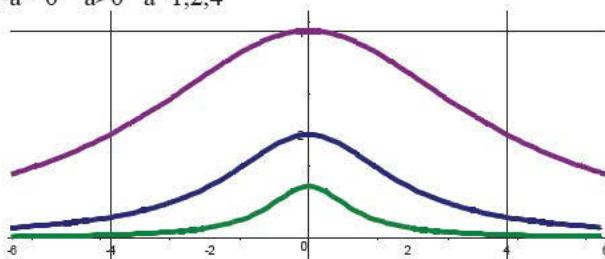


Рис. 7.7

Декартов лист

$$x^3 + y^3 - 3axy = 0 \quad a > 0$$

$$x = 3at / (1 + t^3), y = 3at^2 / (1 + t^3)$$

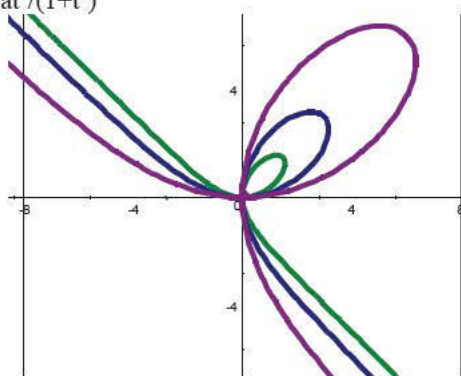


Рис. 7.8

Циссоида

$$x^3 + (x-a)y^2 = 0 \quad a > 0$$

$$x = at^2/(1+t^2), \quad y = at^3/(1+t^2)$$

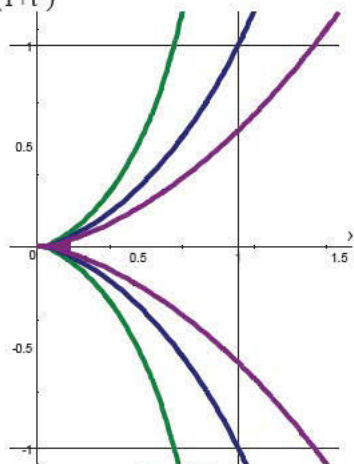


Рис. 7.9.

Строфоида

$$(x+a)x^2 + (x-a)y^2 = 0 \quad a > 0$$

$$x = a(t^2-1)/(t^2+1) \quad y = at(t^2-1)/(t^2+1)$$

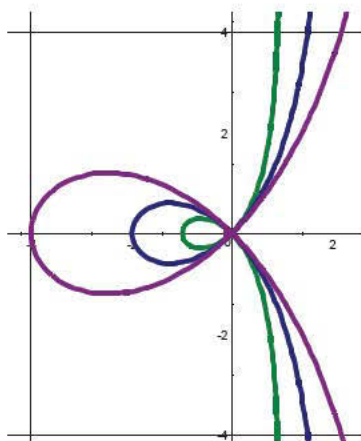


Рис. 7.10.

Офнурида (змеиный хвост) (рис. 7.11)

$$x(x^2+y^2)-y(ay-bx)=0$$

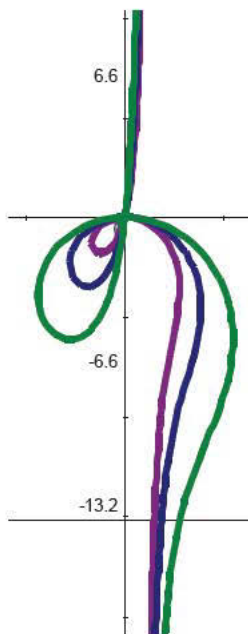


Рис. 7.11

Трисектриса Маклорена (рис. 7.12)

$$x(x^2+y^2)-a(y^2-3x^2)=0$$

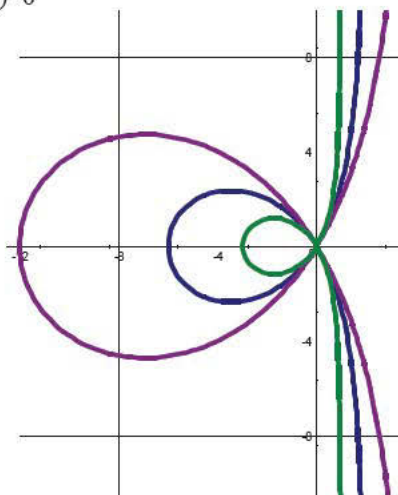


Рис. 7.12

Кубика Чирнгаузена
 $2(2a+x)^3 - 27a(x^2+y^2) = 0$

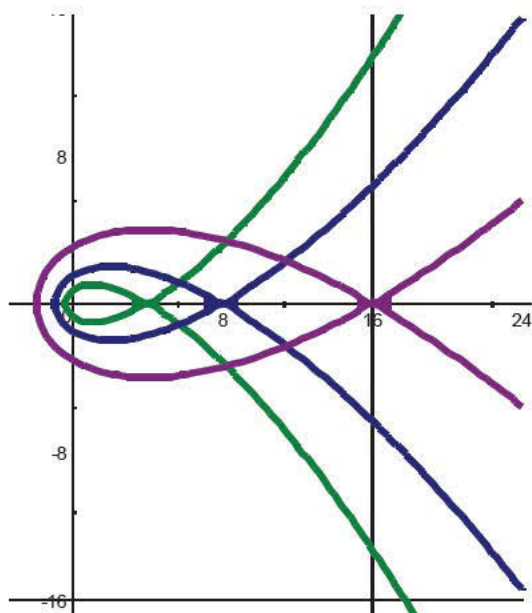


Рис. 7.13.

Овалы Кассини

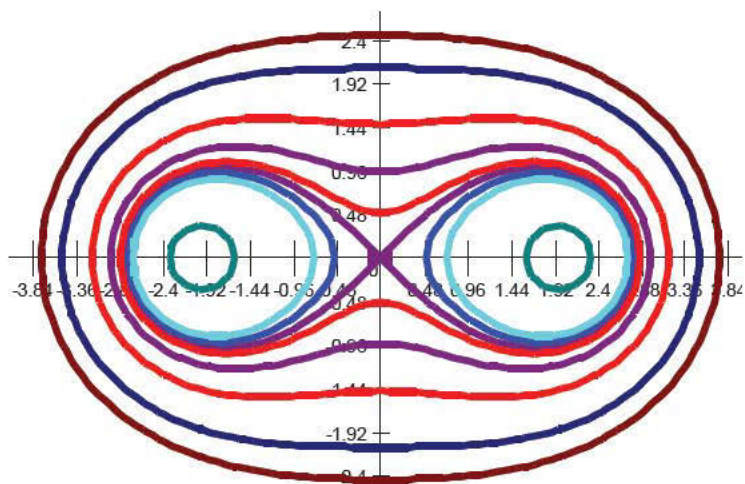


Рис. 7.14.

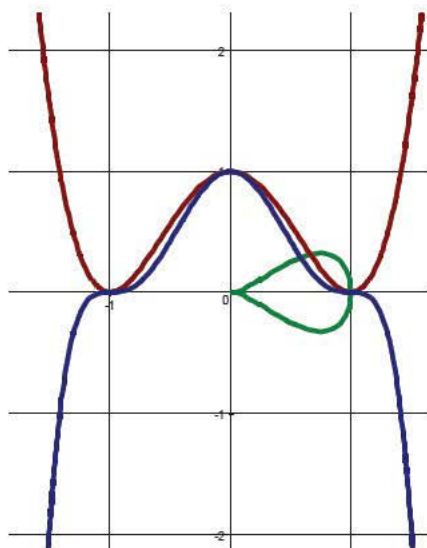
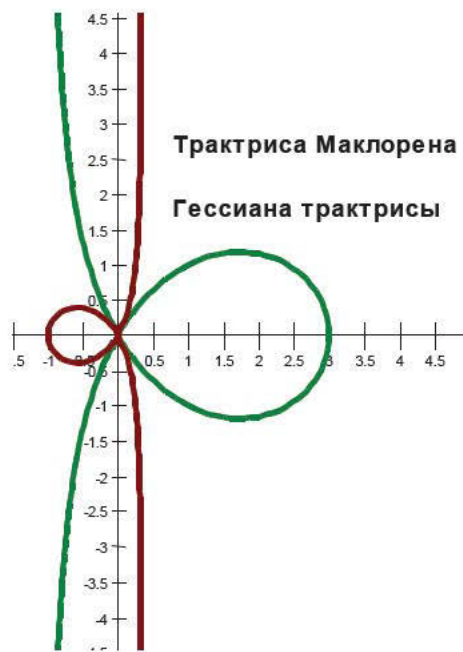


Рис. 7.15. Кривые 4-го порядка

зеленая $x^4 - x^3 + y^2 = 0$
 Красная $y + (1 - x^2)^2 = 0$
 Синяя $y - (1 - x^2)^2 = 0$

РАЗДЕЛ 8. КОНСТРУИРОВАНИЕ КРИВЫХ (ОБВОДЫ И СПЛАЙНЫ)



Свет, который несет нам математика, можно сравнить со светом звезд - прекрасным и ясным, но холодным и чуждым всему человеческому. Впрочем, ирония состоит в том, что человеческий ум как будто тем лучше знает, умеет обращаться с вещами, чем дальше они удалены от самого центра его существования.

*Герман Вейль
Пол века математики. — М.: «Знание», 1969, - с. 6*

геометрия – это настоящая естественная наука, только более простая, а следовательно, и более совершенная, чем какая – либо другая. Мы не должны думать, что, допуская употребление математического анализа, она становится из – за этого чисто логической наукой, не зависящей от наблюдения.

О. Конт (1798-1857)



Безье, Пьер Этьенн (*Pierre Étienne Bézier*, 1.09.1910 — 25.11.1999) — французский инженер и патентообладатель (но не первооткрыватель) кривых Безье и поверхностей Безье, которые в настоящее время широко используются в компьютерных системах проектирования и графических редакторах.



Якоб Исаак Шёнберг (*Isaac Jacob Schoenberg*; 21.04.1903 — 21.02.1990 г) — румынский и американский математик, известный, прежде всего, открытием [сплайнов](#). (Статья о сплайнах опубликована в 1946 г.)

В течение 1943—1945 он находился в отпуске, выполняя для военных работу как математик на Абердинском полигоне. В это время он приступил к работе, которая его прославила, теории сплайнов.

В конструировании и в дизайне часто встречается задача: через ряд точек, заданных своими координатами провести кривую, отвечающую определенным геометрическим условиям. Чтобы записать в терминах математики введем необходимые понятия и дадим основные определения.

8.1. Основные понятия и определения

Определение. Обводом (сплайном) точек $A^1, A^2, A^3, \dots, A^n$ называется кривая, составленная из последовательности дуг $A^1A^2, A^2A^3, \dots, A^{n-1}A^n$ различных кривых, состыкованных между собой определенным образом.

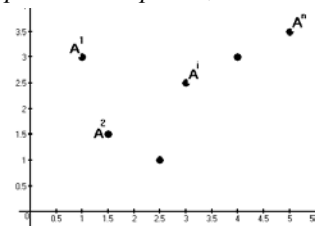


Рис. 8.1 Заданные точки

Точки стыка A^2, A^3, \dots, A^{n-1} дуг обвода называются узлами (изломами) обвода. Гладкость обвода определяются порядком гладкости в его узлах. Если в точках стыка составляющие обвода имеют общие касательные, то обвод называется обводом первого порядка гладкости. Обвод A^1A^n

будет n -го порядка гладкости, если координаты всех его точек удовлетворяют уравнению $y = f(x)$, где абсцисса x принимает значения $a \leq x \leq b$ и на этом отрезке функция $f(x)$ однозначна, непрерывна и имеет непрерывные производные n -го порядка.

Если $n = 0$ и первая производная терпит разрыв, то непрерывная сама по себе дуга A^1A^2 имеет хотя бы одну точку излома, и представляет собою ломаную линию.

Если $n = 1$, то A^1A^2 - дуга первого порядка гладкости, и, следовательно, в каждой ее точке касательная единственная, а угол φ ее наклона к оси x определяется из выражения $\varphi = \text{arctg}(Y')$.

Если $n = 2$, то A^1A^2 - дуга второго порядка гладкости, т.е. сохраняется непрерывность первой и второй производной, что обеспечивает единственность радиуса кривизны кривой в каждой ее точке, а величина радиуса кривизны определяется из выражения

$$\rho = \frac{(1+y'^2)^{3/2}}{|y''|}.$$

Если $n = 3$, то A^1A^2 - дуга третьего порядка гладкости, т.е. сохраняется непрерывность производных от первой до третьей, что обеспечивает единственность скорости V изменения радиусов кривизны кривой в каждой ее точке, а величина V , определяется из выражения

$$V = (1 + y'^2)^{1/2} \frac{3y'y''^2 - y'^2y''' - y''^3}{y'^2}.$$

Порядок гладкости зависит от требований, предъявляемых к качеству разрабатываемых кривых. Каждая дуга обвода однозначно задается конечным числом геометрических элементов: точек, касательных, кругов кривизны и т.д. которые являются ее определителем.

Определение. *Определителем кривой называется совокупность независимых условий, однозначно определяющих эту кривую. Число условий, определяющих данную кривую, называется ее параметрическим числом.*

Плоская алгебраическая кривая n -го порядка имеет параметрическое число, равное $\frac{n(n+3)}{2}$ из которых параметры положения (исключая прямую и окружность) будут 3. Прямая не имеет параметров формы, окружность имеет один параметр формы (радиус) и два параметра положения (координаты центра).

Существуют следующие способы задания дуг обвод

- Табличный – кривая задается координатами множества своих точек.
- Графический – кривая вычерчена на некотором носителе (бумаге, экране).
- Аналитический – кривая задана уравнением.

Определение. *Обвод, заданный координатами своих точек называется дискретным.*

Замена дискретного обвода математическими уравнениями всегда осуществляются с каким-то приближением.

Определение. *Геометрический образ, заменяющий с определенной степенью точности исходный геометрический образ, называется аппроксимирующим, а процесс его нахождения аппроксимацией. Аппроксимирующая кривая приближается поочередно к каждой узловой точке, однако не проходит через все эти точки.*

Определение. *Если аппроксимирующий обвод проходит через все узловые точки дискретного обвода, то он называется интерполирующим.*

Две основные задачи на построение обводов

■ Вычерченную кривую с замеренными координатами некоторых ее точек описать с помощью функции известного класса;

■ Заменить данную функцию $f(x)$ функцией другого класса $\varphi(x)$.

Это задачи интерполирования.

Определение. *Интерполяционной формулой называется формула, которая сопоставляет значения функции $f(x)$, заданной на некотором множестве аргументов x^i , со значениями функции $\varphi(x^i)$. Точки, соответствующие значениям $x^0, x^1, x^2, \dots, x^k$ называются узлами интерполирования.*

Основная задача интерполирования: по координатам узловых точек x_i, x_j некоторой кривой: определить коэффициенты интерполирующей функции.

8.2. Построения обводов методом интерполяции

Интерполяция кривыми второго порядка. Интерполирование обводов дугами окружности. Интерполирование обводов дугами окружности с весами. Понятие о графическом дискриминанте. «Степенные» и «специальные» функции. Тригонометрические функции.

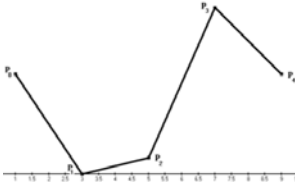


Рис.2.1. Интерполирование прямыми

Интерполяция кривыми второго порядка. В качестве интерполирующей функции могут быть использованы: прямая, дуга окружности, кривая второго порядка, кривые более высоких порядков. Рассмотрим ряд примеров. **Интерполирование**

прямыми. Интерполирующей функцией является прямая, проходящая две точки (x_i, y_i) и (x_{i+1}, y_{i+1}) . Составим уравнение прямой, проходящей через данные точки:

$$\frac{x - x_i}{x_{i+1} - x_i} = \frac{y - y_i}{y_{i+1} - y_i}$$

Уравнение прямой в общем виде описывается соотношением:

$$Ax + By + C = 0, \text{ где } A = y_{i+1} - y_i; B = x_i - x_{i+1}; C = x_{i+1} \cdot y_i - x_i \cdot y_{i+1}.$$

На рис. 8.2 приведен пример: кривая задана таблицей массива точек: $P_0(1,4), P_1(3,0), P_2(5,1), P_3(7,6), P_4(9,4)$. Интерполяция выполнена прямыми.

Интерполирование обводов дугами окружности. Интерполирующей функцией является окружность, проходящая через три точки $(x_i, y_i), (x_{i+1}, y_{i+1}), (x_{i+2}, y_{i+2})$. Запишем уравнение окружности в общем виде

$$(x - x_0)^2 + (y - y_0)^2 - R^2 = 0$$

Здесь (x_0, y_0) - координаты центра окружности, а R – ее радиус.

Окружность должна проходить через заданные точки. Выполняя это условие, получим относительно искомых величин $(x_0, y_0), R$ нелинейные (квадратные) уравнения. Это не совсем удобно. Преобразуем уравнение (2а)

$$\begin{aligned} x^2 - 2xx_0 + x_0^2 + y^2 - 2yy_0 + y_0^2 - R^2 &= 0 \\ a(x^2 + y^2) + bx + cy + 1 &= 0, \\ x_0^2 + y_0^2 - R^2 &\neq 0 \end{aligned}$$

где $a = \frac{1}{x_0^2 + y_0^2 - R^2}$; $b = -2ax_0$; $c = -2ay_0$ - новые неизвестные, подлежащие определению. Подставив координаты данных точек в уравнение окружности, получим систему трех линейных уравнений с тремя неизвестными:

$$\begin{cases} a(x_i^2 + y_i^2) + bx_i + cy_i + 1 = 0 \\ a(x_{i+1}^2 + y_{i+1}^2) + bx_{i+1} + cy_{i+1} + 1 = 0 \\ a(x_{i+2}^2 + y_{i+2}^2) + bx_{i+2} + cy_{i+2} + 1 = 0 \end{cases}$$

Решив полученную систему, определим коэффициенты a, b, c используя которые вычислим координаты центра окружности $O(x_0, y_0)$ и величину R - радиуса окружности

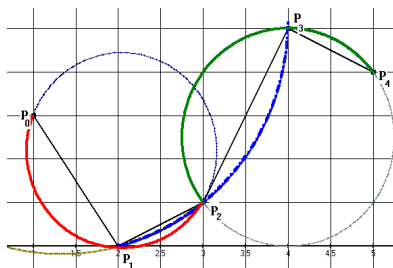


Рис. 8.2. Интерполирование дугами окружностей

$$x_0 = -\frac{b}{2a}; y_0 = -\frac{c}{2a};$$

$$R = \sqrt{\frac{b^2 + c^2 - 4a}{4a^2}}$$

Пример. На рис.2 приведен пример интерполирования окружностями.

Кривая задана таблицей массива точек:

$$P_0(1,1.5), P_1(2,0), P_2(3,0.5), P_3(4,2.5), P_4(5,2)$$

$$P_0P_1P_2: (y - 1.1)^2 + (x - 2.04)^2 - 1.12^2 = 0$$

$$P_1P_2P_3: (y - 2.58)^2 + (x - 1.33)^2 - 2.58^2 = 0$$

$$P_2P_3P_4: (y - 1.25)^2 + (x - 4)^2 - 1.25^2 = 0$$

Из рисунка видно, что на участках P_1P_2 и P_2P_3 проходят две дуги. Как поступить в этом случае рассмотрим в следующем ниже.

Интерполирование обводов дугами окружности с весами.

Интерполяционная кривая в этом случае строится на каждом участке между i -ой и $(i + 1)$ -й узловыми точками дискретного обвода. На участке $(i, i + 1)$ можно построить:

■ окружность радиуса R_i по координатам точек $i - 1, i, i + 1$ и определить координаты центра x_{0i}, y_{0i} . Уравнение этой окружности:

$$y_i(x) = y_{0i} \pm \sqrt{R_i^2 - (x - x_{0i})^2}$$

■ окружность радиуса R_{i+1} по координатам точек $i, i + 1, i + 2$ и определить координаты центра x_{0i+1}, y_{0i+1} . Уравнение этой окружности:

$$y_2(x) = y_{0i+1} \pm \sqrt{R_{i+1}^2 - (x - x_{0i+1})^2}$$

На участке $(i, i + 1)$ имеем две дуги окружности с радиусами R_i, R_{i+1} , причем в точке i больше влияние окружности радиуса R_i , а в точке $(i + 1)$ - R_{i+1} . Отобразим это влияние весовыми функциями $P_i(x) = Ax + B$ и $P_{i+1}(x) = Cx + D$, линейными по x . На отрезке (x_i, x_{i+1}) функция $P_i(x)$ убывает от 1 до 0, а $P_{i+1}(x)$ возрастает от 0 до 1, причем в сумме $P_i(x) + P_{i+1}(x) = 1$. Таким образом, уравнение интерполирующей кривой на участке $(i, i + 1)$ геометрически представляет собой сумму двух окружностей, взятых с их весами.

Условия $\left\{ \begin{array}{l} P_i(x_i) = 1 \quad P_i(x_{i+1}) = 0 \\ P_{i+1}(x_i) = 0 \quad P_{i+1}(x_{i+1}) = 1 \end{array} \right\}$ дают для коэффициентов A, B, C и D следующие значения:

$$A = -\frac{1}{x_{i+1} - x_i}; \quad B = \frac{x_i}{x_{i+1} - x_i}; \quad C = \frac{1}{x_{i+1} - x_i}; \quad D = -\frac{x_i}{x_{i+1} - x_i}$$

$$\text{Тогда: } P_i(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i}; \quad P_{i+1}(x) = \frac{x - x_i}{x_{i+1} - x_i}.$$

Окончательно, уравнение участка $[i, i + 1]$ примет вид:

$$Y = P_i(x)y_1(x) + P_{i+1}(x)y_2(x) = (Ax + B)y_1(x) + (Cx + D)y_2(x)$$

С переходом к следующему участку вновь вычисляется все коэффициенты A, B, C, D весовых функций и получаем уравнение этого участка

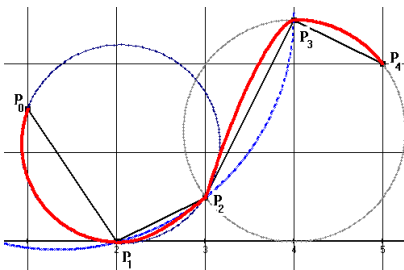


Рис. 8.4. Интерполирование дугами окружностей с весами

и точки. В результате получается непрерывная кривая, проходящая через все заданные точки. Направление касательной также непрерывно и в узловых точках совпадает с направлением касательной к окружности, для которой данная точка является средней из трех. Непрерывная кривизна на данном участке близится к кривизне реальной. При переходе к нелинейным видам порядок гладкости повышается.

Пример. На рис. 8.4 кривая задана таблицей массива точек: $P_0(1,1.5)$, $P_1(2,0)$, $P_2(3,0.5)$, $P_3(4,2.5)$, $P_4(5,2)$ интерполирована дугами окружностей с весами:

$$BC: y = \frac{5-x}{3[4-(x-3)^2]^{0.5}} + \frac{x-2}{3\{-2.9+[(4.61)^2-(x-1.4)^2]^{0.5}\}};$$

$$CD: y = (6-x)\{-2.9+[(4.61)^2-(x-1.4)^2]^{0.5}\} + (x-5)\{-0.8-[(2.15)^2-(x-7)^2]^{0.5}\}.$$

Понятие о графическом дискриминанте. Кривые второго порядка получили наибольшее распространение при задании обводов, поскольку эти кривые детально изучены, для них разработан хороший графоаналитический аппарат конструирования сложных обводов.

Кривые второго порядка (коники) определяются алгебраическим уравнением 2-го порядка в декартовой системе координат:

$$F(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0 \quad (8.1)$$

где $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$.

Кривая второго порядка может быть задана:

■ Пятью линейно независимыми точками (ни какие 3 из ни не лежат на одной прямой)

Пусть заданы 5-точек кривой $M_i(x_i, y_i), i = 1, \dots, 5$;

Соединим любые две заданные точки отрезком прямой. Разделим отрезок пополам (или в любом другом отношении). Перенесем в полученную точку D начало координат - $O(0,0)$. Тогда кривая не проходит через начало координат и, следовательно, коэффициент a_{33} в уравнении (8.1) отличен от 0, т.е. $a_{33} \neq 0$. Разделив все координаты на (не меняя обозначений), получим:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + 1 = 0 \quad (8.2)$$

Тогда для определения коэффициентов уравнения необходимо решить систему уравнений:

$$a_{11}x_i^2 + 2a_{12}x_iy_i + a_{22}y_i^2 + 2a_{13}x_i + 2a_{23}y_i + 1 = 0 \quad i = 1, \dots, 5$$

■ Пятью касательными к ней (ни какие 3 касательных не должны пересекаться в одной точке);

Как правило, используется при чисто графических построениях, так как касательная задается в конкретной точке кривой.

■ Любой комбинацией из касательных прямых и точек, удовлетворяющих указанным выше условиям.

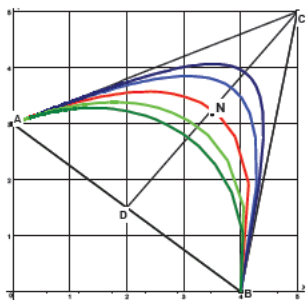


Рис. 8.5. Графический дискриминант

Задать касательную к конике в некоторой точке это значит задать координаты точки и угловой коэффициент прямой k (либо угол α наклона прямой к оси $OX \Rightarrow \tan \alpha$). Уравнение касательной к кривой второго порядка в точке $M(x_M, y_M)$ имеет вид,

$$(a_{11}x_M + a_{12}y_M + a_{13})(x - x_M) + (a_{12}x_M + a_{22}y_M + a_{23})(y - y_M) = 0 \Rightarrow \\ \Rightarrow y = -\frac{(a_{11}x_M + a_{12}y_M + a_{13})}{(a_{12}x_M + a_{22}y_M + a_{23})}x + y_M + \left(y_M + \frac{(a_{11}x_M + a_{12}y_M + a_{13})}{(a_{12}x_M + a_{22}y_M + a_{23})}x_M\right) \quad (8.3)$$

тогда

$$k_M = -\frac{a_{11}x_M + a_{12}y_M + a_{13}}{a_{12}x_M + a_{22}y_M + a_{23}} \quad (8.4)$$

k_M угловой коэффициент касательной в точке M .

Задание 2 точек и касательных к конике в них дают нам 8.4 уравнения. Чтобы задать конкретную кривую 2 порядка (эллипс, гиперболу или параболу) для задания еще одной точки воспользуемся инженерным дискриминантом.

В точке C пересекаются касательные t_A и t_B , касающиеся коники в точках A и B соответственно (рис. 8.5). Проведем хорду AB и медиану CD . Задав на медиане точку N полностью определим кривую 2-го порядка. Отношение отрезков ND и CD (рис. 8.5) называется дискриминантом коники

$$f = \frac{ND}{CD}$$

Переведем качественное рассмотрение, дискриминанта на язык формул.

Дано: Точки $A(x_A, y_A), B(x_B, y_B)$ и касательные t_A, t_B в этих точках и точка $N(x_N, y_N)$. По заданным условиям провести определенную (в зависимости от значения f) кривую второго порядка.

Соединим точки A и B прямой и разделим отрезок пополам точкой D : $AD = DB = d$. Совместим начало координат точку $O(0,0)$ с точкой D . Направим ось OX по прямой от точки D к B . Получим

$$A(-d, 0), t_A, B(d, 0), t_B$$

Определим координаты $(\cdot)C$. Для этого запишем уравнения прямых AC и BC :

$$AC: y = k_A x + b_A (\cdot)A \in AC \Rightarrow 0 = -k_A d + b_A \Rightarrow b_A = k_A d \Rightarrow k_A x + k_A d \\ BC: y = k_B x + b_B (\cdot)B \in BC \Rightarrow 0 = k_B d + b_B \Rightarrow b_B = -k_B d \Rightarrow k_B x - k_B d \\ (\cdot)C = AC \cap BC \Rightarrow \begin{cases} y_C = k_A x_C + k_A d \\ y_C = k_B x_C - k_B d \end{cases} \Rightarrow x_C = \frac{k_B + k_A}{k_B - k_A} d, y_C = \frac{2k_B k_A d}{k_B - k_A}.$$

Из определения инженерного дискриминанта следует

$$x_N = f x_C; y_N = f y_C \Rightarrow y_N = \frac{k_B + k_A}{k_B - k_A} f d, y_N = \frac{2k_B k_A f d}{k_B - k_A} \quad (8.5)$$

Искомая кривая l должна проходить через точки A и B :

Подставив значения координат точек A и B в уравнение (8.2) получим:

$$\left. \begin{aligned} (\cdot)A(-d, 0) \in l &\Rightarrow a_{11}d^2 + 2a_{13}(-d) + 1 = 0 \\ (\cdot)B(d, 0) \in l &\Rightarrow a_{11}d^2 + 2a_{13}d + 1 = 0 \end{aligned} \right\}$$

Полученные уравнения позволяют определить коэффициенты

$$a_{11} = -\frac{1}{d^2}; a_{13} = 0 \quad (8.6)$$

В точках A и B заданы касательные, т.е. известны угловые коэффициенты k_A и k_B прямых AC и BC . Подставив значения координат точек A и B в уравнение (8.3) и учитывая (8.4) и (8.6) получим:

$$\left. \begin{aligned} t_A: \left(-\frac{1}{\alpha}\right)(x+d) + (-a_{12}d + a_{23})y = 0 &\Rightarrow k_A d a_{23} - k_A d^2 a_{12} - 1 = 0 \\ t_B: \frac{1}{\alpha}(x-d) + (a_{12}d + a_{23})y = 0 &\Rightarrow k_B d a_{23} + k_B d^2 a_{12} + 1 = 0 \end{aligned} \right\}$$

Решая полученную систему уравнений определим коэффициенты $a_{12}a_{23}$:

$$a_{12} = -\frac{k_B+k_A}{2k_Bk_A d^2}; a_{23} = -\frac{k_B-k_A}{2k_Bk_A d} \quad (8.7)$$

Для определения остающегося неизвестным коэффициента a_{22} учтем, что коника должна проходить через точку N , управляющую выбором, какую именно кривую (эллипс, параболу или гиперболу) требуется провести. Подставив значения (8.5, 8.6 и 8.7) в формулу (8.2) и разрешив результат относительно a_{22} , после не сложных преобразований получим

$$a_{22} = (2f-1) \left(\frac{k_B-k_A}{2k_Bk_A f d}\right)^2 - \left(\frac{k_B+k_A}{2k_Bk_A d}\right)^2 \quad (8.8)$$

Из теории кривых 2-го порядка известно (смотри раздел 3), что тип кривой зависит от величины и знака малого определителя:

$$D = a_{11}a_{22} - a_{12}^2. \quad (8.9)$$

Подставив значения (8.6-8.7) в (8.9) после приведения подобных получим:

$$D = -(2f-1) \left(\frac{k_B-k_A}{2k_Bk_A f d^2}\right)^2 \text{ т.к. } \left(\frac{k_B-k_A}{2k_Bk_A f d^2}\right)^2 > 0$$

всегда положительно, то:

$$D = 0 \Rightarrow f = 0,5 - \text{имеем параболу}$$

$$D > 0 \Rightarrow f < 0,5 - \text{эллипс}$$

$$D < 0 \Rightarrow f > 0,5 - \text{гипербола.}$$

При параллельных касательных (что возможно при $k_A = k_B$ и только для случаев задания эллипса или гиперболы) этой формой задания пользоваться нельзя.

Если кривая задана в параметрическом виде, то ее уравнение в системе координат D и V (рис. 8.6) будет иметь вид:

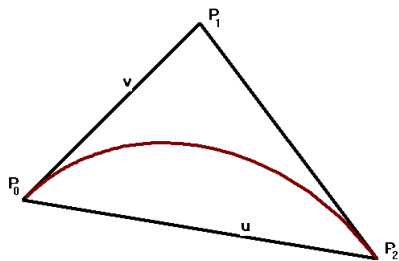


Рис. 8.6. Кривая в параметрическом виде

$$u = \frac{1}{\gamma t^2 + t + 1}, v = \frac{t}{\gamma t^2 + t + 1},$$

где $t = \frac{v}{u}$ - параметр, γ - коэффициент кривой.

При изменении параметра t от 0 до $+\infty$ получим все множество точек внутри ΔABC . Исключив параметр t из уравнения, получим уравнение кривой в неявном виде:

$$u^2 + \gamma v^2 + uv - u = 0$$

Из этого уравнения можно определить коэффициент кривой:

$$\gamma = \frac{u(1-u-v)}{v^2}.$$

Параметр t и коэффициент γ связаны соотношением:

$$t_{1,2} = \frac{-1 \pm \sqrt{1-4\gamma}}{2\gamma}; \gamma = \left(\frac{1-f}{2f}\right)^2; f = \frac{1}{1+2\sqrt{\gamma}}.$$

Если кривая плоская, то в плоскости XOY уравнение принимает вид:

$$x = \frac{\gamma t^2 x_A + x_B t + x_C}{\gamma t^2 + t + 1}, y = \frac{\gamma t^2 y_A + y_B t + y_C}{\gamma t^2 + t + 1}.$$

Если задан соприкасающийся ΔABC и точка $N(x_N, y_N)$, то можно определить коэффициент кривой:

$$\gamma = \frac{\begin{vmatrix} x_B - x_C & x_N - x_C \\ y_B - y_C & y_N - y_C \end{vmatrix} \begin{vmatrix} x_A - x_N & x_A - x_B \\ y_A - y_N & y_A - y_B \end{vmatrix}}{\begin{vmatrix} x_A - x_C & x_A - x_N \\ y_A - y_C & y_A - y_N \end{vmatrix}}.$$

$$x = x_A + u(x_C - x_A) + v(x_B - x_A) = \frac{\gamma t x_A + x_B + x_C}{\gamma t^2 + t + 1}$$

$$y = y_A + u(y_C - y_A) + v(y_B - y_A)$$

$$z = z_A + u(z_C - z_A) + v(z_B - z_A)$$

«Степенные» и «специальные» функции.

Для решения задачи, в которой функции $y = f(x)$ заменяются функцией другого класса, часто применяют в инженерной практике «степенные» функции вида:

$$y = [4\bar{x}(1-\bar{x})]^m. \quad (8.10)$$

Эти функции аэродинамичны и используются для разработки различных агрегатов самолета. Решение задачи осуществляется в интервале от 0 до l и проводится в следующем порядке

$$\left. \begin{aligned} \bar{x}_i &= \frac{x_i}{2l} \quad (0 \leq x_i \leq l) \\ \bar{y}_i &= \frac{y_i}{y_{max}} \quad (0 \leq y_i \leq l) \end{aligned} \right\} \quad (8.11)$$

где, y_{max} и l - исходные параметры.

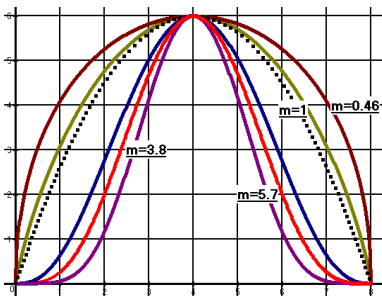


Рис. 8.7. Интерполирующая степенная функция

Логарифмируя уравнение (8.10), с учетом (8.11), получим

$$m = \frac{\lg \bar{y}}{\lg [4\bar{x}(1-\bar{x})]} \quad (8.12)$$

m - параметр полноты кривой.

Интерполирующая функция в интервале $0 \div 1$ примет вид:

$$y = y_{max} \left[\frac{2x}{l} \left(1 - \frac{x}{2l} \right) \right]^m \quad (8.13)$$

При $m < 1$ уравнение дает кривую с вертикальной касательной в начальной точке, при $m = 1$ – параболу с наклонной касательной в точке $(0,0)$; при $m > 1$ – кривую с точкой перегиба и горизонтальными касательными в начальной и конечной точках. Касательная в любой точке кривой имеет вид:

$$\bar{y}' = m [4\bar{x}(1-\bar{x})]^{m-1} (4-8\bar{x}).$$

Или
$$\bar{y}' = \frac{4m(1-2\bar{m})}{[4\bar{x}(1-\bar{x})]^{1-m}}. \quad (8.14)$$

Пример. Составить интерполяционную формулу для кривой, определенной точкой $(2, 2)$, $l = 4, y_{max} = 6$, (рис. 8.7).

$$x_1 = \frac{1}{4}, \quad \bar{y}_1 = \frac{1}{3}, \quad m = \frac{\lg \bar{y}_1}{\lg x + \lg \bar{x}_1 + \lg(1-\bar{x}_1)} = \frac{-\lg 3}{\lg 3 - \lg 4} \approx 3.8$$

Интерполяционная формула имеет вид: $y = 6 \left[\frac{1}{2} x \left(1 - \frac{x}{8} \right) \right]^{3.8}$.

Для задания выпуклых и вогнутых кривых, приближающихся к осям координат, применяют «специальные» функции вида:

$$\bar{y} = \sqrt{\frac{1-\bar{x}^2}{1-m\bar{x}^2}} \quad (8.15)$$

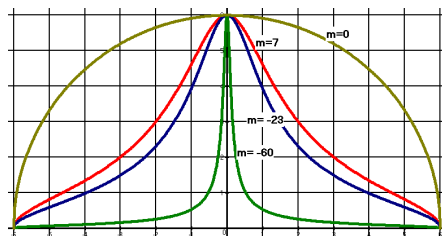


Рис. 8.8. Интерполирующая степенная функция

Если задаться интервалом от 0 до 1 и значением y , соответствующем $x = 0$, то в системе координат (x, y) получим: $\bar{x} = \frac{x}{l}, \quad \bar{y} = \frac{y}{y_{max}}$.

В зависимости от значений параметра m получаем кривые различной формы.

При $t = 0$ функция (8.15) описывает окружность;

$m > (-3)$ выпуклая замкнутая кривая

$m \rightarrow (-\infty)$ кривая неограниченно приближается к осям координат.

Параметр m кривой определяется по формуле:

$$m = \frac{\bar{x}^2 + \bar{y}^2 - 1}{\bar{x}^2 \bar{y}^2}.$$

Рассмотренные выше функции широко применяются в решении задач построения очерков и сечений агрегатов различных аппаратов.

Пример. Составить интерполяционную формулу для кривой, определяемой точкой (2,2), $l=4$; $y_{max} = 6$, (рис. 8.8).

$$\bar{x} = \frac{1}{2}; \bar{y} = \frac{1}{3}; m = \left(\frac{1}{4} + \frac{1}{9 \pm 1}\right) : \frac{1}{4} \cdot \frac{1}{9} = -\frac{36 - 13}{36} \cdot \frac{36}{1} = -23.$$

Интерполяционная формула имеет вид:

$$y = y_{max} \sqrt{\frac{l^2 - x^2}{l^2 + 23x^2}}.$$

Функция вида $\bar{y} = tg \varphi \left(\bar{x} - \frac{1}{m} \bar{x}^m\right)$ обеспечивает заданное значение касательной $tg \varphi$ в начальной точке обвода. По исходным данным $tg \varphi$, l, y_{max} , определяются $\bar{y} = \frac{y}{l}$, $\bar{x} = \frac{x}{l}$. Обозначим $\lambda = \frac{l}{2y_{max}}$, тогда $y_{max} = \frac{l}{2\lambda}$. Откуда $\bar{y} = \frac{y_{max}}{l} = \frac{l}{2\lambda l} = \frac{1}{2\lambda}$. Параметр полноты кривой определяется по формуле

$$m = \frac{1}{1 - \frac{1}{y} tg \varphi} \Rightarrow y = tg \varphi \left[\frac{x}{l} - \frac{1}{m} \left(\frac{x}{l}\right)^m\right]. \quad (8.16)$$

Тригонометрические функции. При задании некоторых частных обводов и поверхностей применяют тригонометрические функции различных видов.

1. Тригонометрические полиномы общего вида

$$y = a_0 + \sum_{i=1}^n (a_i \cos i\Psi + b_i \sin i\Psi) \quad (8.17)$$

В частном случае функция может состоять из одних косинусов

$$y = a_0 + \sum_{i=1}^n a_i \cos i\Psi \quad (8.18)$$

2. Трансцендентные уравнения в параметрическом виде $x = \frac{1}{a} \cos \alpha$;

$$y = y_{max} (\sin \alpha)^2 + b \quad (8.19)$$

где a, b - коэффициенты уравнения, η - показатель степени, определяется по точке $M(x_M, y_M)$, принимающей ободу:

$$\eta = \frac{lg \bar{y}_M}{lg \sin \alpha}, \text{ где } \bar{y}_M = \frac{y_M}{y_{max}}.$$

8.3. Интерполяция полиномиальными кривыми

Полиномиальные кривые. Полиномы принадлежат к фундаментальным математическим объектам, широко используемым в компьютерной графике. Они позволяют эффективно организовать вычислительный процесс.

Рассмотрим взаимодействие между неявной и параметрической формами простых полиномов.

Определение. Полиномом L -ой степени от t называется функция вида

$$a_0 + a_1 t + a_2 t^2 + \dots + a_L t^L,$$

где константы a_0, a_1, \dots, a_L - коэффициенты полинома, каждый из которых связан с одной из степеней t . Степенью полинома называется наибольшая степень L , в которую возводится t , т.е. $a_L \neq 0$. Порядком полинома называется число его коэффициентов - $(L + 1)$. порядок полинома всегда на единицу больше его степени.

Линейный полином $y = ax + b, (a \neq 0)$ определяет прямую, проходящую через точки $A\left(-\frac{b}{a}, 0\right)$ и $B(0, b)$. При $b = 0$ точки A и B совпадают и прямая проходит через начало координат $O(0,0)$.

Квадратичный полином $y = ax^2 + bx + c, (a \neq 0)$ определяет параболу с осью симметрии, параллельной оси y , и вершиной $c\left(-\frac{b}{2a}, \frac{4ac-b^2}{4a}\right)$. Кривая пересекает ось y в точке $B(0, c)$.

При $\Delta = 4ac - b^2 < 0$ парабола пересекает ось x в точках $A_1\left(\frac{-b+\sqrt{\Delta}}{2a}, 0\right)$ и $A_2\left(\frac{-b-\sqrt{\Delta}}{2a}, 0\right)$.

При $\Delta = 0$ кривая касается оси x в точке $\left(-\frac{b}{2a}, 0\right)$ (касание второго порядка).

При $\Delta > 0$ точек пересечения с осью x нет.

$$\text{Полином 3 степени } y = ax^3 + bx^2 + cx + d \quad (8.20)$$

Задаёт различные кривые, вид которых определяется значением коэффициентов a, b, c, d ; $\Delta = 3ac - b^2$ и дискриминантом $D = b^2c^2 - 4ac^3 - 4bd - 27a^2b^2 + 18abcd$.

Если $a > 0$, то $y \rightarrow -\infty$ при $x \rightarrow -\infty$ и $y \rightarrow +\infty$ при $x \rightarrow +\infty$

Если $a < 0$, то $y \rightarrow +\infty$ при $x \rightarrow -\infty$ и $y \rightarrow -\infty$ при $x \rightarrow +\infty$.

При $\Delta > 0, a < 0$ кривая имеет точку перегиба E . (рис. 1а)

$\Delta = 0$ кривая (кубическая парабола) имеет точку перегиба E (касательная в точке E параллельна оси x) (рис. 8.9 б).

$\Delta < 0, a > 0$ кривая имеет 1 максимум в точке $Ax_{max} = \left(-b - \frac{\sqrt{-\Delta}}{3a}\right)$,

один минимум в точке $Bx_{min} = \left(-b + \frac{\sqrt{-\Delta}}{3a}\right)$ и точку перегиба E (рис. 1в)

Целые полиномиальные кривые.

При $\Delta > 0$ кривая пересекает ось в трех точках.

При $\Delta = 0$ у кривой две или одна точка пересечения с осью X , причем ровно в одной точке пересечения имеет место касание (в первом случае имеет место касание второго, а во втором третьего порядка).

При $\Delta < 0$ имеет место одна (простая) точка пересечения с осью X .

Если $n \geq 0$ целое число, то кривые не имеют особых точек, без асимптот, имеют не более n точек пересечения с осью x , не более $(n - 1)$ экстремумов и не более $(n - 2)$ точек перегиба. В случае нескольких экстремумов максимумы и минимумы чередуются.

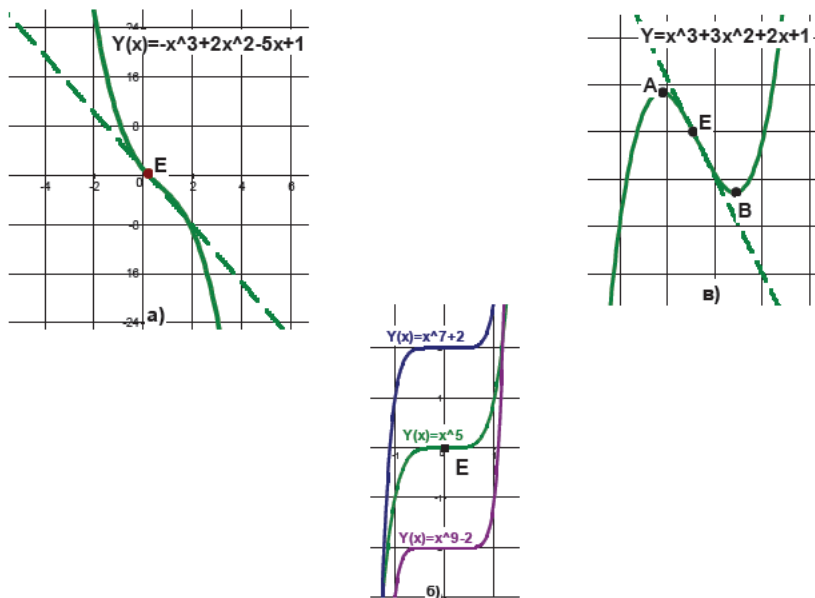


Рис. 8.9. Полиномы 3 степени

■ n – нечетное. Существует, по крайней мере, одно пересечение с осью X и при $n \geq 3$ по меньшей мере одна точка перегиба. Число экстремумов при $n \geq 3$ всегда четно, а число точек перегиба нечетно.

■ Если $a > 0$ то при $x \rightarrow +\infty \Rightarrow y \rightarrow +\infty$, а при $x \rightarrow -\infty \Rightarrow y \rightarrow -\infty$
 $a < 0$ то при $x \rightarrow +\infty \Rightarrow y \rightarrow -\infty$, а при $x \rightarrow -\infty \Rightarrow y \rightarrow +\infty$.

■ n - четное. При $n \geq 2$ существует по меньшей мере один экстремум. Число экстремумов всегда нечетно, а число точек перегиба всегда четно.

При $a > 0$ то и при $x \rightarrow +\infty$ и при $x \rightarrow -\infty \Rightarrow y \rightarrow +\infty$

$a < 0$ при $x \rightarrow +\infty$ и при $x \rightarrow -\infty \Rightarrow y \rightarrow -\infty$.

Интерполяционная формула Лагранжа

Задача состоит в том, чтобы точно провести кривую через заданные точки.

Пусть задан массив точек: $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), x_0 < x_1 < x_2 \dots < x_n$.

Решение будем искать в виде полинома:

$$y = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 .$$

Для определения неизвестных коэффициентов $a_n, a_{n-1} \dots a_1, a_0$ воспользуемся условием прохождения кривой через каждую заданную точку и получим систему из $n + 1$ уравнений с $n + 1$ неизвестными $a_n, a_{n-1} \dots a_1, a_0$:

$$\begin{aligned} y_0 &= a_n x_0^n + a_{n-1} x_0^{n-1} + \dots + a_1 x_0 + a_0 \\ y_1 &= a_n x_1^n + a_{n-1} x_1^{n-1} + \dots + a_1 x_1 + a_0 \\ &\dots \\ y_n &= a_n x_n^n + a_{n-1} x_n^{n-1} + \dots + a_1 x_n + a_0 \end{aligned}$$

Полученная система уравнений используется в численных методах математики, но имеет целый ряд недостатков. В геометрическом моделировании применяются более наглядные методы. Рассмотрим их.

Лагранж¹ предложил интерполяционную формулу:

$$L(x) = y_0 L_n^0(x) + y_1 L_n^1(x) + y_2 L_n^2(x) + \dots + y_n L_n^n(x), \quad (8.21)$$

$$\text{где } L_n^k(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)} \quad (8.22)$$

$L_n^k(x)$ - фундаментальные полиномы интерполяции Лагранжа.

Для интерполяции по $(n + 1)$ заданным точкам (x_i, y_i) $i = 0, 1, \dots, n$ таким, что $x_0 < x_1 < \dots < x_n$ нужен полином степени n с $(n + 1)$ коэффициентами.

Покажем, как вычисляется полином, проходящий точно через три заданные точки (x_0, y_0) , (x_1, y_1) и (x_2, y_2) где $x_0 < x_1 < x_2$. Рассмотрим выражение

$$L_3^0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}$$

Если $x = x_0$, то числитель равен знаменателю и $L_0(x_0)=1$.

Если же $x = x_1$ или $x = x_2$ то числитель равен нулю и $L_0(x_1) = L_0(x_2)=0$.

Аналогично

$$L_3^1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \begin{cases} 1, & \text{если } x = x_1 \\ 0, & \text{если } x = x_0 \text{ или } x = x_2 \end{cases}$$

¹Жозеф Луи Лагранж (Joseph Louis Lagrange, 25 января 1736, Турин — 10 апреля 1813, Париж) — французский математик, астроном и механик. Особенно прославился исключительным мастерством в области обобщения и синтеза накопленного научного материала.

$$L_3^2(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \begin{cases} 1, & \text{если } x = x_2 \\ 0, & \text{если } x = x_0 \text{ или } x = x_1 \end{cases}$$

Используя полученные выражения, построим функцию, содержащую заданные значения y :

$$L(x) = L_3^0(x)y_0 + L_3^1(x)y_1 + L_3^2(x)y_2.$$

При $x = x_0$ ненулевое значение имеет только первый член правой части этого выражения и $P_L(x_0) = y_0$. При $x = x_1$ ненулевое значение имеет второй член и $P_L(x_1) = y_1$, а при $x = x_2$, $P_L(x_2) = y_2$.

Следовательно, значения функции $P_L(x)$ совпадают со значениями функции $y(x)$ во всех узлах. Три заданные значения функции позволяют определить только три коэффициента квадратичного интерполяционного полинома.

Пример. Записать формулу Лагранжа для случая $n = 4$ при заданных точках $0(1,2)$, $1(2,5)$, $2(3,6)$

$$L(x) = y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

Построим фундаментальные полиномы интерполяции Лагранжа.

$$L_2^0(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = 0.5x^2 - 2.5x + 3$$

$$L_2^1(x) = \frac{(x-1)(x-3)}{(2-1)(2-3)} = -x^2 + x - 3$$

$$L_2^2(x) = \frac{(x-1)(x-2)}{(3-1)(3-2)} = 0.5x^2 - 1.5x + 1$$

И получим формулу Лагранжа:

$$L(x) = -x^2 + 6x - 3$$

Между опорными точками эта функция может вести себя самым произвольным образом (особенно с ростом n) Более того, как всякий многочлен $y(x)$ неограниченно работает по абсолютной величине при стремлении x к $+\infty$ или $-\infty$ бесконечности.

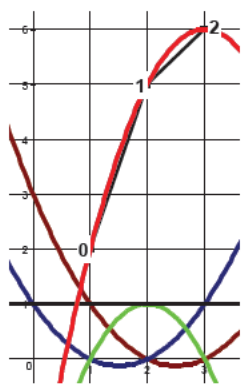


Рис. 8.10. Полиномы Лагранжа

Полиномиальная интерполяция обладает одним неприятным свойством – малые изменения исходных данных могут привести к непредвиденно большим изменениям функции. Поэтому, используя полиномы Лагранжа в практических расчетах, где данные известны не точно, никогда нельзя заранее сказать, получится ли пристойная поверхность или бурное море.

(А.А. Мигдал)

Пример. Записать формулу Лагранжа для случая $n = 4$ при заданных точках $0(1,1.5)$, $1(2,0)$, $2(3,0.5)$, $3(4,8)$, $4(5, 1)$ (рис. 8.11)

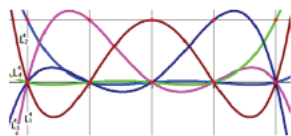


Рис. 8.11. Полиномы Лагранжа

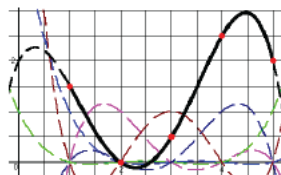


Рис. 8.12. Интерполяционная кривая Лагранжа

Интерполяционная формула Эрмита. Пусть в $(n + 1)$ точках $x_i, i = 0, 1, 2, \dots, n$ нам даны не только значения функции y_i но и ее производные y'_i . Единственным полиномом наименьшей степени, удовлетворяющим этим значениям во всех точках x_i , определяется по формуле

$$P_H(x) = \sum_{i=0}^n H_i(x)y_i + \sum_{i=0}^n H_i^*(x)y'_i, \quad (8.23)$$

где $H_i(x) = [1 - 2L_i(x)]L_i^2(x)$, $H_i^* = (x - x_i)L_i^2(x)$, а $L_i(x)$ определена по формуле, как фундаментальный полином Лагранжа. Поскольку $L_i(x)$ - полином степени n , то $P_H(x)$ - полином степени $2n + 1$, и для определения всех коэффициентов требуется $2n + 2$ исходных данных, т.е. известные значения y, y' в $(n + 1)$ точках x_i .

Метод Ньютона (разделенные разности).

Интерполирующий полином, впервые предложенный Ньютоном, имеет следующий вид:

$$P(x) = \alpha_0 + (x - x_0)\alpha_1 + (x - x_0)(x - x_1)\alpha_2 + \dots + (x - x_0)(x - x_1) \dots (x - x_{n-1})\alpha_n \quad (3.5)$$

$P(x)$ - полином степени n . Необходимо определить все α_n . Потребуем, чтобы в каждой заданной точке имело место соотношение $P(x_i) = y_i$. Полагая $x = x_0, x_1, \dots, x_n$ соответственно, получим систему из $(n + 1)$ уравнений:

$$\begin{aligned} y_0 &= \alpha_0 \\ y_1 &= \alpha_0 + (x_1 - x_0)\alpha_1 \\ y_2 &= \alpha_0 + (x_2 - x_0)\alpha_1 + (x_2 - x_0)(x_2 - x_1)\alpha_2 \\ &\dots \end{aligned}$$

$$y_n = \alpha_0 + (x_n - x_0)\alpha_1 + (x_n - x_0)(x_n - x_1)\alpha_2 + \dots + (x_n - x_0)(x_n - x_1) + (x_n - x_{n-1})\alpha_n.$$

Эти уравнения однозначно разрешимы. Первое уравнение дает $\alpha_0 = y_0$. Подставляя полученный результат во второе уравнение, получим

$$\alpha_1 = \frac{y_1 - y_0}{x_1 - x_0}, \text{ т.к. } x_1 - x_0 \neq 0.$$

Из третьего уравнения находим $\alpha_2 = y[x_0, x_1, x_2]$ и т.д. Заданная дополнительная точка (x_{n+1}, y_{n+1}) увеличивает систему еще на одно уравнение, содержащее α_{n+1} . Остальные уравнения остаются неизменными.

Значения $\alpha_0, \alpha_1, \dots, \alpha_n$ не зависят от числа заданных точек:

α_0 зависит только от y_0 ;

α_1 зависит от y_0 и y_1

α_2 - от y_0, y_1, y_2 и т.д.,

т.е. $\alpha_i = y[x_0, x_1, \dots, x_i]$

Такие выражения связаны соотношением:

$$y[x_s, \dots, x_i] = \frac{y[x_{s+1}, \dots, x_t] - y[x_s, \dots, x_{t-1}]}{x_t - x_s} \quad (8.25)$$

и называются разделенными **разностями**. В этих обозначениях формула Ньютона принимает вид:

$$P(x) = y[x_0] + (x - x_0)y[x_0, x_1] + (x - x_0)(x - x_1)y[x_0, x_1, x_2] + \dots + (x - x_0)(x - x_1)(x - x_2) \dots (x - x_{n-1})y[x_0, x_1, x_2, \dots, x_n] \quad (8.26)$$

Разделенные разности можно определить другим (хотя и эквивалентным) способом:

$$\Delta_i^1(x_{i-1}, x_i) = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$$

Разделенные разности первого порядка являются аналогом первой производной в i -й точке. Аналогом второй производной будут вторые разделенные разности (разделенные разности от разделенных разностей):

$$\Delta_i^2(x_{i+1}, x_i, x_{i-1}) = \frac{\Delta_{i+1}^1(x_i, x_{i+1}) - \Delta_i^1(x_{i-1}, x_i)}{x_{i+1} - x_{i-1}}$$

Разделенные разности r -го порядка

$$\Delta^r(x_0, x_1, x_2, \dots, x_r) = \frac{\Delta_{r-1}(x_1, x_2, \dots, x_r) - \Delta_{r-1}(x_0, x_1, \dots, x_{r-1})}{x_2 - x_0}$$

$r = 2, 3, \dots (n - 1)$.

В этих обозначениях интерполяционная формула Ньютона имеет вид:

$$y(x) = y_0 + (x - x_0)\Delta^1(x_0, x_1) + \dots + \prod_{k=1}^{n-1} (x - x_k)\Delta^n(x_0, x_1, \dots, x_n).$$

$$y(x) = y_0 + (x - x_0)\{\Delta^1(x_0, x_1) + (x - x_1)[\Delta^2(x_0, x_1, x_2) + \dots]\}.$$

Пусть задан дискретный обвод координатами своих точек

$$x_i, y_i; i = 1, 2, 3, 4$$

Таблица 8.1.

№	x_i	y_i	Δ_i^1	Δ_i^2	Δ_i^3
1	1	1	$\frac{y_2 - y_1}{x_2 - x_1} = \frac{7 - 1}{6 - 1} = 1,2^*$	$\frac{0,5 - 1,2}{10 - 1} = -0,0778^*$ $\frac{0,25 - 0,5}{18 - 6} = -0,0208$	$\frac{-0,0208 + 0,0778}{18 - 1} = 0,00365^*$
2	6	7	$\frac{y_3 - y_2}{x_3 - x_2} = \frac{9 - 7}{10 - 6} = 0,5$		
3	10	9	$\frac{y_4 - y_3}{x_4 - x_3} = \frac{11 - 9}{18 - 10} = 0,25$		
4	18	11			
Или в обобщенном виде					
Значения x	Значения y	Значения первых разностей	Значение вторых разностей	Значение третьих разностей	
x_0	$y[x_0]$	$y[x_0, x_1]$	$y[x_0, x_1, x_2]$	$y[x_0, x_1, x_2, x_3]$	
x_1	$y[x_1]$	$y[x_1, x_2]$			
x_2	$y[x_2]$	$y[x_2, x_3]$	$y[x_1, x_2, x_3]$		
x_3	$y[x_3]$				

Звездочкой обозначают числа, использующиеся как опорные значения при вычислении интерполирующегося многочлена

$$y(x) = y_1 + (x - x_1)\{\Delta_1^1 + (x - x_2)[\Delta_1^2 + (x - x_3)\Delta_1^3]\};$$

Вычислим для данной дискретной функции значения y при $x = 10$ и $x = 5$

$$y(x) = 1 + (x - 1)\{1,2 + (x - 6)[-0,0778 + (x - 10)0,00365]\}$$

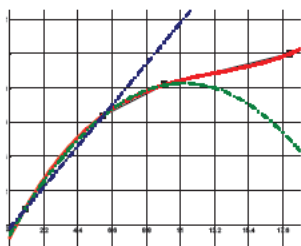


Рис. 8.13. Формула Ньютона

$$\begin{aligned}
 y(10) &= 1 + 9\{1,2 + 4[-0,0778 + 0 \cdot \\
 &\quad 0,00365]\} = 1 + 7,92 = 9 \\
 y(5) &= 1 + 4 \times \\
 &\quad \times \{1,2 + (-1)[-0,0778 + (-5)0,00365]\} \\
 &= 6,18
 \end{aligned}$$

Итерационно-интерполяционный метод Эйткена. При практической реализации задач интерполирования вместо интерполяционных

формул Лагранжа, Ньютона и т.п. удобнее применять интерполяционную схему Эйткена и Невилля, в которой искомые величины вычисляются с помощью итераций.

Запишем интерполяционные многочлены возрастающих степеней. Многочлен 2-й степени:

$$y_{01} = \frac{1}{x_1 - x_0} \begin{vmatrix} x - x_0 & y_0 \\ x - x_1 & y_1 \end{vmatrix}.$$

Если даны три значения аргумента x_0, x_1, x_2 в которых известны y_0, y_1, y_2 , то предварительно строят два линейных многочлена:

$$y_{01} = \frac{1}{x_1 - x_0} \begin{vmatrix} x - x_0 & y_0 \\ x - x_1 & y_1 \end{vmatrix}, y_{02} = \frac{1}{x_2 - x_0} \begin{vmatrix} x - x_0 & y_0 \\ x - x_2 & y_2 \end{vmatrix}.$$

Сделаем еще одну линейную интерполяцию между значениями этих выражений:

$$y_{012} = \frac{1}{x_2 - x_0} \begin{vmatrix} x - x_0 & y_{01} \\ x - x_2 & y_{02} \end{vmatrix}.$$

Т.к. y_{01} и y_{02} многочлены первой степени, то y_{012} многочлен второй степени относительно x . Общая формула интерполяционного многочлена по схеме Эйткена имеет вид:

$$y_{012\dots n} = \frac{1}{x_n - x_0} \begin{vmatrix} x - x_0 & y_{012\dots n-1} \\ x - x_n & y_{012\dots n} \end{vmatrix}. \quad (8.27)$$

Эта формула совпадает с формулой Лагранжа, но более удобна в вычислительном отношении. Первая итерация дает величину, соответствующую линии интерполяции между точками (x_0, y_0) и (x_1, y_1) . Последовательные итерации дают результаты соответствующие квадратичной интерполяции по трем точкам, кубической интерполяции по четырем точкам и т.д. Гладкая сходимость результатов последовательных приближений показывает, что окончательная погрешность мала, но для каждого нового значения x эту процедуру необходимо проводить от начала до конца.

Докажем в заключение анализа метода разделенных разностей важную теорему, имеющую отношение к теории B -сплайнов.

Теорема. Для полинома $P(x)$ степени n все n -е разделенные разности равны, а все $(n + 1)$ -е разделенные разности, следовательно, равны нулю.

Доказательство. Положим $P(x) = \sum_{i=0}^n a_i x^i$. Первая разделенная разность полинома $P(x)$ между x_0 и произвольной точкой x дается формулой

$$P[x_0, x] = \frac{P[x] - P[x_0]}{x - x_0} = \frac{1}{x - x_0} \sum_{i=0}^n a_i (x^i - x_0^i) = \sum_{i=0}^n a_i (x^{i-1} + x_0 x^{i-2} + \dots + x_0^{i-1}) \quad (8.28)$$

Это полином по x степени $(n - 1)$. Подобным образом вторая разделенная разность $P[x_0, x_1, x]$ будет полиномом степени $(n - 2)$ от x . Продолжая это рассуждение, можно показать, что n -я разделенная разность $P[x_0, x_1, x_2, \dots, x_{n-1}, x_n]$ является полиномом нулевой степени от x и поэтому имеет постоянное значение для всех x . Отсюда непосредственно вытекает, что любая $(n + 1)$ -я разность полинома $P(x)$ равна 0.

8.4. Интерполяционный сплайн

Сплайн-функции получили широкое распространение при задании кривых и поверхностей, т.к. обладают следующими преимуществами: универсальностью, однородностью, алгоритмичностью. Они широко применяются во всех областях науки и техники. Теория сплайнов, история которой начинается с работы Шёнберга [Schoenberg I.J.] 1946 г., в настоящее время продолжает бурно развиваться. Это обусловлено:

1. Устойчивостью сплайнов относительно локальных возмущений, т.е. поведение сплайна в окрестностях точки не сказывается на поведении сплайна в целом, как, например, это имеет место при полиномиальной интерполяции;

2. Хорошая сходимость сплайн-интерполяции в отличие от многочленной. В частности, для функций с нерегулярными свойствами гладкости целесообразно использовать сплайн.

При точном проведении гладких кривых линий, проходящих через заданные узловые точки, конструкторы применяют в качестве лекал длинные тонкие рейки (spline), изготовленные из дерева или металла. Эти рейки, или сплайны закрепляются в узловых точках с помощью специальных грузил (криц). Изменением положения точек, в которых устанавливается грузик для фиксации рейки, а также изменением положения сплайна при достаточном количестве грузиков, можно добиться того, чтобы сплайн точно проходил через заданные узловые точки. Если рассматривать рейку (сплайн) как тонкую балку, то имеет место закон Эйлера.

$$M(x) = \frac{EI}{R(x)}$$

где $M(x)$ - изгибающий момент;

E - модуль Юнга;

I - геометрический момент инерции;

$R(x)$ - радиус кривизны кривой, совпадающей с осью балки.

Из дифференциальной геометрии известно, что

$$R(x) = \frac{[1 + (y')^2]^{3/2}}{y''}$$

При малых изгибах $(y')^2 \ll 1$, и $R(x) = \frac{1}{y''(x)}$,

где $y(x)$ - деформированная ось балки.

$$y''(x) = \frac{1}{EI} M(x).$$

Изгибающий момент $M(x)$ между узловыми точками изменяется линейно, поскольку точки закрепления рейки можно рассматривать как простые опоры. В этом случае математический кубический сплайн непрерывен и имеет непрерывную первую и вторую производные. Будем искать некоторую функцию $S(x)$, непрерывную на отрезке $[a, b]$ вместе со своими первой и второй производной, совпадающую с кубическим полиномом на каждом участке между узловыми точками и удовлетворяющую условиям $S(x_i) = y_i (i = 1, 2, \dots, n)$. Такую функцию и называют кубическим сплайном.

По аналогии обозначим через M_i «моменты» $S''(x) (i = 1, 2, \dots, n)$. У кубического полинома вторая производная является линейной функцией т.е. на каждом участке $[x_i, x_{i+1}]$ можно записать

$$S''(x) = M_i \frac{x_{i+1} - x}{h_i} + M_{i+1} \frac{x - x_i}{h_i},$$

где $h_i = x_{i+1} - x_i$.

Проинтегрировав дважды обе части равенства и вычислив константы интегрирования (из условия $S(x_i) = y_i$), получим

$$S'(x) = -M_i \frac{(x_{i+1} - x)^2}{2h_i} + M_{i+1} \frac{(x - x_i)^2}{2h_i} + C_1,$$

$$S(x) = M_i \frac{(x_{i+1} - x)^3}{6h_i} + M_{i+1} \frac{(x - x_i)^3}{6h_i} + C_1 x + C_2.$$

Константы интегрирования определим из граничных условий

$$S(x_i) = y_i \text{ и } S(x_{i+1}) = y_{i+1}$$

$$S(x) = M_i \frac{(x_{i+1} - x)^3}{6h_i} + M_{i+1} \frac{(x - x_i)^3}{6h_i} + \left(y_i - \frac{M_i h_i^2}{6} \right) \frac{x_{i+1} - x}{h_i} + \left(y_{i+1} - \frac{M_{i+1} h_i^2}{6} \right) \frac{x - x_i}{h_i}.$$

Это уравнение третьей степени относительно переменной x - кубическая парабола вида:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

Определение. Математическим сплайном называется приближенное представление деформированной оси рейки кусками кубических парабол с определенными разрывами производных, допускаемых в узловых точках. Кубический сплайн – это полином третьей степени между i -й и $(i + 1)$ -й точками:

$$S^i(x) = a_0^i + a_1^i(x - x_i) + a_2^i(x - x_i)^2 + a_3^i(x - x_i)^3$$

или

$$S^i(x) = \sum_{k=0}^3 a_k^i (x - x_i)^k \quad (8.29)$$

Математический сплайн простейшего вида непрерывен и имеет непрерывные первую и вторую производные во всех точках интервала $[a, b]$. Они определяются следующим образом:

первая производная

$$S'^i(x) = a_1^{(i)} + 2a_2^{(i)}(x - x_i) + 3a_3^{(i)}(x - x_i)^2 \quad (8.30)$$

вторая производная

$$S''^i(x) = 2a_2^{(i)} + 6a_3^{(i)}(x - x_i)^2 \quad (8.31)$$

Третья производная претерпевает в узловых точках разрыв с конечным скачком

$$S'''^i(x) = 6a_3^{(i)} = const.$$

Это соответствует физическому состоянию упругой равнопрочной постоянного сечения рейке, имеющей непрерывную кривизну. Скорость изменения разрыва в точках крепления.

Оптимальные свойства сплайн-интерполяции обоснованы теоремой Холлидея, которая утверждает, что среди всех интерполирующих функций, имеющих в интервале $[a, b]$ непрерывную вторую производную u сплайнов минимальная кривизна. Для сплайнов доказаны также теоремы существования, единственности и наилучшего приближения.

Рассмотрим математическую постановку задачи.

Дано. Упорядоченные координаты точек дискретного обвода x_i, y_i .

Определить: коэффициенты интерполирующего сплайна.

Решение задачи сводится к нахождению коэффициентов полиномов третьей степени для каждого отдельно взятого участка обвода. Для того, чтобы обвод был гладким, необходимо, чтобы в каждой узловой точке слева и справа выполнялись следующие условия:

$$S(x_i) = y_i; \quad (8.32)$$

$$S'(x_i - 0) = S'(x_i + 0) = y'_i \quad (8.33)$$

$$S''(x_i - 0) = S''(x_i + 0) = y''_i \quad (8.34)$$

Предположим, что известны производные в начальной и конечной точках обвода:

$$S'(x_0) = y'_0, S'(x_n) = y'_n; \quad (8.35)$$

$$S''(x_0) = y''_0, S''(x_n) = y''_n. \quad (8.36)$$

Случай бесконечных производных в граничных точках исключается.

Определим коэффициенты a_0, a_1, a_2, a_3 на каждом i -м участке между узловыми точками. По определению сплайна $S(x)$ во внутренних точках $x_i (i = 1, 2, 3, \dots, n - 1)$ из выражений (8.29-8.31) и выполнению условий (8.35), (8.36) напомним:

$$\begin{aligned} a_0^i &= a_0^{i-1} + a_1^{i-1}h_{i-1} + a_2^{i-1}h_{i-1}^2 + a_3^{i-1}h_{i-1}^3 \\ a_1^i &= a_1^{i-1} + 2a_2^{i-1}h_{i-1} + 3a_3^{i-1}h_{i-1}^2 \\ a_2^i &= a_2^{i-1} + 3a_3^{i-1}h_{i-1} \end{aligned} \quad (8.37)$$

где $h_i = x_{i+1} - x_i$

Определим коэффициенты a_0, a_1, a_2, a_3 для каждого участка: соотношения (8.32) для всех $(i = 1, 2, 3, \dots, n)$ дают:

$$a_0^i = y_i \quad (8.38)$$

Из третьего соотношения (8.37) следует

$$a_3^{i-1} = \frac{a_2^i - a_2^{i-1}}{3h_{i-1}}, i = 1, 2, 3, \dots, n \quad (8.39)$$

Из первого соотношения (8.37) и формул (8.38) и (8.39) находим

$$y_i = y_{i-1} + a_1^{i-1}h_{i-1} + a_2^{i-1}h_{i-1}^2 + \frac{a_2^i - a_2^{i-1}}{3h_{i-1}}h_{i-1}^3;$$

Упрощаем:

$$\left. \begin{aligned} y_i - y_{i-1} &= a_1^{i-1}h_{i-1} + a_2^{i-1}h_{i-1}^2 + \frac{a_2^{i-1}h_{i-1}^2}{3} + \frac{a_2^i h_{i-1}^2}{3} \\ a_1^{i-1} &= \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{h_{i-1}}{3} (2a_2^{i-1} + a_3^i) \end{aligned} \right\} \quad (8.40)$$

где $\frac{y_i - y_{i-1}}{h_{i-1}} = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} = \Delta(x_{i-1}, x_i)$ - первые разделенные разности.

Из второго соотношения (8.37) имеем

$$a_1^i = a_1^{i-1} + 2a_2^{i-1} + 3h_{i-1}^2 \left(\frac{a_2^i - a_2^{i-1}}{3h_{i-1}} \right)$$

Вместо значения a_1^{i-1} подставим выражение (12):

$$a_1^{i-1} = \Delta(x_{i-1}, x_i) - \frac{h_{i-1}}{3} (2a_2^{i-1} + a_2^i) + 2a_2^{i-1}h_{i-1} + a_2^i h_{i-1} - a_2^{i-1}h_{i-1}.$$

Преобразуем:

$$\begin{aligned} a_1^i &= \Delta(x_{i-1}, x_i) - \frac{2a_2^{i-1}}{3} h_{i-1} + 2a_2^{i-1}h_{i-1} + a_2^i h_{i-1} - a_2^{i-1}h_{i-1}, \\ a_1^i &= \Delta(x_{i-1}, x_i) - \frac{h_{i-1}}{3} (a_2^{i-1} + 2a_2^i) \end{aligned} \quad (8.41)$$

Определяем коэффициент a_2^i следующим образом: запишем формулу (8.40) для a_1^i :

$$a_1^i = \Delta(x_i, x_{i+1}) - \frac{h_i}{3} (2a_2^i + a_2^{i+1})$$

и, приравняв это выражение к формуле (13), получим:

$$\begin{aligned} \Delta(x_i, x_{i+1}) - \frac{h_i}{3} (2a_2^i + a_2^{i+1}) &= \Delta(x_{i-1}, x_i) + \frac{h_{i-1}}{3} (a_2^{i-1} + 2a_2^i), \\ \Delta(x_{i-1}, x_i) - \Delta(x_i, x_{i+1}) &= \frac{h_{i-1}}{3} a_2^{i-1} + \frac{h_{i-1}}{3} 2a_2^i + \frac{h_i}{3} 2a_2^i + \frac{h_i}{3} a_2^{i+1}, \\ 3(\Delta(x_{i-1}, x_i) - \Delta(x_i, x_{i+1})) &= 2a_2^i(h_i + h_{i-1}) + a_2^{i+1}h_i + a_2^{i-1}h_{i-1} \end{aligned}$$

Разделим на $h_i + h_{i-1}$

$$3\Delta^2(x_{i-1}, x_i, x_{i+1}) = a_2^{i-1} \frac{h_{i-1}}{h_i + h_{i-1}} + 2a_2^i + a_2^{i+1} \frac{h_i}{h_i + h_{i+1}}$$

где $\Delta^2(x_{i-1}, x_i, x_{i+1}) = \frac{\Delta(x_{i-1}, x_i) - \Delta(x_i, x_{i+1})}{h_i + h_{i+1}}$ - вторые разделенные разности.

Обозначим

$$\alpha_i = \frac{h_{i-1}}{h_i + h_{i-1}}; \beta_i = \frac{h_i}{h_i + h_{i+1}} \Rightarrow \alpha_i + \beta_i = 1$$

$$\text{Получим } \alpha_i a_2^{i-1} + 2a_2^i + \beta_i a_2^{i+1} = 3\Delta^2(x_{i-1}, x_i, x_{i+1}) \quad (8.42)$$

В последнее уравнение (14), при $i = n - 1$ входит дополнительная величина

$$a_2^n = a_2^{n-1} + 3a_3^{n-1}h_{n-1}$$

Добавим к уравнению (8.42) концевые условия (8.35), (8.36). Опустив промежуточные выкладки, приведем окончательные формулы дополнительной системы:

$$\left. \begin{aligned} 2a_2^0 + \beta_0 a_2^1 + \alpha_0 a_2^n &= (2 + \alpha_0 + \beta_0) \Delta^2(x_{-1}, x_0, x_1) \\ \beta_n a_2^0 + \alpha_n a_2^{n-1} + 2a_2^n &= (2 + \alpha_n + \beta_n) \Delta^2(x_{n-1}, x_n, x_{n+1}) \end{aligned} \right\} \quad (8.43)$$

где $\alpha_n = 0; \beta_n = 0$.

Решив систему (8.42) с концевыми условиями (8.43), получим коэффициенты a_2^i . Запишем эту систему в матричном виде $Q_2 \cdot a_2 = d$, где d - вектор правых частей уравнений (8.42) и (8.43). В развернутом виде:

$$\begin{bmatrix} 2 & \beta_0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 & \alpha_0 \\ \alpha_1 & 2 & \beta_1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \\ 0 & \alpha_2 & 2 & \beta_2 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_3 & 2 & \beta_3 & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 2 & \beta_{n-3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \alpha_{n-2} & 2 & \beta_{n-2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \alpha_{n-1} & 2 & \beta_{n-1} \\ \beta_n & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & \alpha_n & 2 \end{bmatrix} \times \begin{bmatrix} a_2^0 \\ a_2^1 \\ a_2^2 \\ a_2^3 \\ \cdot \\ a_2^{n-3} \\ a_2^{n-2} \\ a_2^{n-1} \\ a_2^n \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \\ \cdot \\ d_{n-3} \\ d_{n-2} \\ d_{n-1} \\ d_n \end{bmatrix}$$

Для граничных условий (8.35) и (8.36) матрица является трех диагональной, т.е. $\alpha_0 = 0$; $\beta_0 = 0$. Такую систему уравнений удобно решать методом прогонки.

По найденным коэффициентам a_2^i определяем из уравнений (8.40) или (8.41) коэффициенты a_1^i :

$$a_1^i = \Delta(x_i, x_{i+1}) - \frac{h_i}{3} (2a_2^i + a_2^{i+1})$$

Из (11) определяем a_3^i

$$a_3^i = \frac{a_2^{i+1} - a_2^i}{3h_i}$$

В ходе расчета коэффициенты a_0^i и a_2^i хранятся, а a_1^i и a_3^i вычисляются вместе с координатами необходимых промежуточных точек. $a_0^i = y_i$, $a_2^i = 0,5y_i''$ т.е. a_2^i - определяет радиус кривизны в точках обвода.

8.5. Интерполирование параметрическими кубическими кривыми. Параметрические кривые

8.5.1. Полиномиальные параметрические кривые степени 1.

$$x = at + b$$

$$y = ct + d$$

Линейная параметрическая форма для $x(t)$ и $y(t)$ описывает прямую линию, соответствующая ей неявная форма линейна по x и по y . Действительно, из выражения для x определим t : $t = (x - b)/a$, подставив его в выражение для y получим:

$$y = c \frac{x - b}{a} + d \Rightarrow y = \frac{c}{a}x + d - \frac{cb}{a} \Rightarrow \frac{c}{a}x - y + \left(d - \frac{cb}{a}\right) = 0$$

Явное и неявное уравнение прямой линии.

■ Полиномиальные параметрические кривые степени 2

$$x = at^2 + bt + c$$

$$y = dt^2 + et + f$$

Выясним, какие кривые описываются в этом случае. Для этого перейдем от параметрической к неявной форме уравнений. Выделим в выражении для $x(t)$ полный квадрат

$$x = a \left(t^2 + 2 \frac{b}{2a} t + \frac{b^2}{4a^2} \right) + c - \frac{b^2}{4a} = a \left(t + \frac{b}{2a} \right)^2 + \left(c - \frac{b^2}{4a} \right)$$

Произведем подстановку как в выражении для $x(t)$ так в $y(t)$

$$u = t + \frac{b}{2a}, \quad t = u - \frac{b}{2a}, \quad l = c - \frac{b^2}{4a}$$

получим:

$$x = au^2 + l,$$

$$y = du^2 + mu + n,$$

где $m = e - \frac{db}{a}$, $n = f + \frac{db^2}{4a^2} - \frac{bl}{2a}$.

Чтобы перейти к неявной форме выразим параметр и через переменную x :

$$x = au^2 + l \Rightarrow u^2 = \frac{x-l}{a} \Rightarrow u = \sqrt{\frac{x-l}{a}}$$

и подставим полученное выражение в уравнение для y :

$$y = \frac{d}{a}(x-l) + m \sqrt{\frac{x-l}{a}} + n$$

Избавимся от иррациональности:

$$\left[y - \frac{d}{a}(x-l) - n \right]^2 = m^2 \frac{x-l}{a} \Rightarrow$$

$$y^2 - 2 \frac{d}{a}xy + \frac{d^2}{a^2}x^2 + x \left(\frac{2d}{a} - \frac{m}{a} \right) - 2ny + \left(n^2 + \frac{ml}{a} + 2 \frac{dl}{a} \right) = 0$$

Очевидно, что получили уравнение 2-го порядка

$$A = a_{11}a_{22} - a_{12}^2, \text{ но } a_{11} = \frac{d^2}{a^2}, a_{22} = 1, a_{12} = \frac{d}{a} \Rightarrow A = 0$$

При использовании полиномов степени 2 при любом наборе констант a, b, c, d, e, f описываемая кривая является параболой.

■ Параметрические кубические кривые.

Параметрической кубической кривой является кривая, в которой (x, y, z) -многочлены третьего порядка (т.е. кубические) относительно некоторого параметра t . Т.к. рассматриваются конечные отрезки кривой, то без потери

общности можем ограничить диапазон изменения параметра t и считать $0 \leq t \leq 1$. Следовательно,

$$\begin{aligned}x(t) &= a_x t^3 + b_x t^2 + c_x t + d_x; \\y(t) &= a_y t^3 + b_y t^2 + c_y t + d_y; \quad 0 \leq t \leq 1 \\z(t) &= a_z t^3 + b_z t^2 + c_z t + d_z.\end{aligned}\tag{8.44}$$

Производные функций $x(t), y(t), z(t)$ по параметру t имеют один и тот же вид. Например,

$$\frac{dx}{dt} = 3a_x t^2 + 2b_x t + c_x$$

Три производные определяют касательный вектор. Тангенсы углов наклона задаются как отношения компонент этого вектора:

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}; \quad \frac{dz}{dz} = \frac{dz/dt}{dz/dt}; \quad \text{и т.д.}$$

При поиске способов определения a_x, b_x, c_x, d_x будем рассматривать только производные от $x(t)$. Для $y(t)$ и $z(t)$ производные и окончательные формулы аналогичны формулам для $x(t)$ и приводиться не будут.

Параметрический кубический многочлен с 4 коэффициентами является параметрической кривой наиболее низкой степени, которая при соответствующем выборе коэффициентов может удовлетворить четырем условиям (положению которого из концов сегмента и касательным векторам к ним). При использовании параметрического представления более высокого уровня появляется волнистость и возникают осцилляции. Кубический многочлен является параметрической функцией наиболее низкой степени, с помощью которой можно представить неплоскую кривую, необходимую для описания пространственных кривых.

Из многих способов описания параметрических кубических кривых рассмотрим три:

■ Метод Эрмита, в котором задаются положения конечных точек кривой и касательные векторы в них;

■ Метод Безье – задаются положения конечных точек кривой, а для неявного задания касательных в этих точках используются две другие точки, обычно лежащие не на кривой;

■ Метод В-сплайнов – конечные точки не лежат на кривой, в результате чего как первая, так и вторая производные оказываются непрерывными на концах сегмента.

Каждая из этих трех форм описания кривой имеет свои достоинства и недостатки.

Форма Эрмита (Фергюсона).

Дано: точки P_1 и P_4 и касательные векторы R_1 и R_4 .

Определить: коэффициенты a_x, b_x, c_x, d_x из выражения (1), удовлетворяющие условиям

$$x(0) = P_{1x}; \quad x(1) = P_{4x}; \quad x'(0) = R_{1x}; \quad x''(1) = R_{4x} \quad (8.45)$$

Индекс x использован для ссылки на x -компоненты точек и касательных векторов.

$$x(t) = [t^3 t^2 t 1] \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}_x, \Rightarrow x(t) = [t^3 t^2 t 1] C_x \Rightarrow x(t) = T C_x, \quad (8.46)$$

где T - вектор-строка степени t , C_x - вектор-столбец коэффициентов $x(t)$.

Запишем условия (8.45), используя уравнения (8.46):

$$\begin{aligned} x(0) = P_{1x} &= [0 \quad 0 \quad 0 \quad 1] C_x, \\ x(1) = P_{4x} &= [1 \quad 1 \quad 1 \quad 1] C_x. \end{aligned} \quad (8.47)$$

Чтобы записать выражения для ограничений на касательные векторы, продифференцируем выражение (8.45) по t и получим

$$x'(t) = [3t^2 \quad 2t \quad 1 \quad 0] C_x$$

тогда

$$\begin{aligned} x'(0) = R_{1x} &= [0 \quad 0 \quad 1 \quad 0] C_x, \\ x'(1) = R_{4x} &= [3 \quad 2 \quad 1 \quad 0] C_x. \end{aligned} \quad (8.48)$$

Выражения (8.47) и (8.48) можно объединить в одно матричное уравнение

$$\begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix}_x = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 \end{bmatrix} C_x$$

Обращая матрицу 4×4 получим искомое выражение для C_x

$$C_x = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix} = M_h G_{h_x} \quad (8.49)$$

Здесь M_h -эрмитова матрица,

G_{h_x} -геометрический вектор Эрмита.

Подставив этот результат в выражение (3) получим

$$\begin{aligned} x(t) &= T M_h G_{h_x}, \\ y(t) &= T M_h G_{h_y}, \\ z(t) &= T M_h G_{h_z}. \end{aligned} \quad (8.50)$$

Уравнения для $x(t), y(t), z(t)$ можно записать в виде

$$P(t) = T M_h G_n$$

Если заданы P_1, P_4, R_1, R_4 , можно определить $x(t), y(t), z(t)$ для $0 \leq t \leq 1$ и найти все точки на сегменте кубической кривой от P_1 до P_4 , у которого касательный вектор в начальной точке R_1 , а в конечной R_4 .

Найдем произведение TM_h :

$$TM_h = [(2t^3 - 3t^2 + 1)(-2t^3 + 3t^2)(t^3 - 2t^2 + t)(t^3 - t^2)].$$

Умножая это выражение, справа на G_{n_x} , получим

$$x(t) = TM_h G_{n_x} = P_{1x}(2t^3 - 3t^2 + 1) + P_{4x}(-2t^3 + 3t^2) + R_{1x}(t^3 - 2t^2 + t) + R_{4x}(t^3 - t^2)$$

4 функции переменной t в произведении TM_h называют **функциями сопряжения**, т.к. с помощью первых двух функций сопрягаются точки P_1 и P_4 , а посредством двух других – векторы R_1 и R_4 , в результате чего получается «сглаженное» объединение $x(t)$.

Алгоритм де Кастельо

Задана последовательность точек P_0, P_1, \dots, P_n .

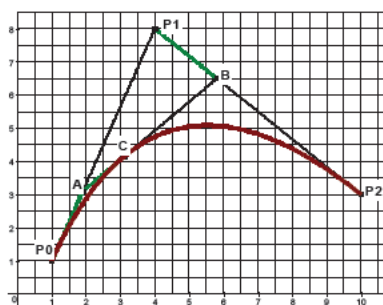


Рис. 8.15. Трех точечная кривая

Необходимо определить положение точки $P(t)$ для каждого значения t от 0 до 1. Кривая должна проходить через заданные точки P_0 и P_n . Начнем с трех точек P_0, P_1, P_2 выберем некоторое значение параметра t в интервале от 0 до 1 (например, $t = 0,3$). Определим положение точки A , находящейся в t -той части прямой P_0P_1 . Аналогично определим положение точки B на прямой P_1P_2 (при том же значении t). Тогда

$$A(t) = (1 - t)P_0 + P_1$$

$$B(t) = (1 - t)P_1 + P_2t \quad (8.51)$$

Повторим для полученных точек A и B процесс линейной интерполяции (при том же самом t) и найдем точку $P(t)$, лежащую на той части между точками A и B :

$$P(t) = (1 - t)A + tB$$

Выполнив этот процесс для каждого t в диапазоне от 0 до 1 сформируем кривую $P(t)$. Определим параметрическую форму этой кривой, подставив в выражение (8.51) формулы (8.52)

$$P(t) = (1 - t)^2P_0 + 2t(1 - t)P_1 + t^2P_2 \quad (8.52)$$

Параметрическая форма $P(t)$ квадратичная относительно t . Следовательно, полученная кривая будет параболой, причем она остается параболой и при изменении t от $-\infty$ до $+\infty$. При $t = 0$ она проходит через точку P_0 , а при $t = 1$ через точку P_2 .

Таким образом, имеем корректно определенный процесс для генерации плавной параболической кривой на базе трех заданных точек.

Кривые Безье. Безье (1970) перегруппировал члены параметрического кубического многочлена Эрмита-Фергюсона таким образом, что физический смысл векторных коэффициентов стал более ясным.

Это особенно важно при конструировании кривой, а не ее подгонки.

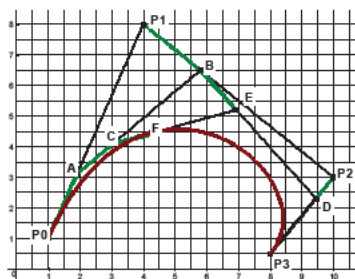


Рис. 8.15. Четырехточечная кривая

Применим алгоритм де Кастельо для построения кривых Безье на базе четырех контрольных точек.

$$\begin{aligned} A(t) &= (1-t)P_0 + P_1t \\ B(t) &= (1-t)P_1 + P_2t \\ C(t) &= (1-t)P_2 + P_3t \\ D(t) &= (1-t)A + Bt \\ E(t) &= (1-t)B + Ct \end{aligned}$$

Для точки F получим:

$$F(t) = (1-t)D + Et$$

Подставим в последнее выражение формулы для точек A, B, C, D, E

$$\begin{aligned} F(t) &= (1-t)D + Et = \\ &= (1-t)[(1-t)A + Bt] + t[(1-t)B + Ct] = (1-t)^2A + 2t(1-t)B + \\ &+ t^2C = (1-t)^3P_0 + 3t(1-t)^2P_1 + 3t^2(1-t)P_2 + t^3P_3. \end{aligned}$$

Четырехточечная кривая Безье имеет следующую параметрическую форму:

$$P(t) = (1-t)^3P_0 + 3t(1-t)^2P_1 + 3t^2(1-t)P_2 + t^3P_3,$$

которая является кубическим полиномом относительно t . Члены, входящие в данный полином, носят название полиномов Бернштейна.

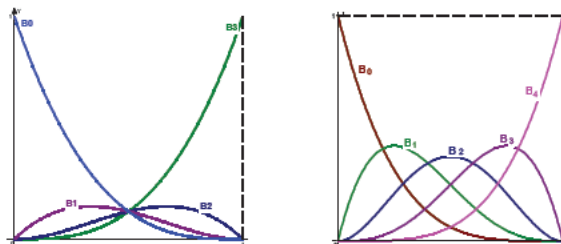


Рис. 8.16. Полиномы Бернштейна

Четыре кубических полинома Бернштейна (рис. 8.16) имеют вид:

$$\begin{aligned} B_0^3(t) &= (1-t)^3 \\ B_1^3(t) &= 3t(1-t)^2 \\ B_2^3(t) &= 3t^2(1-t) \\ B_3^3(t) &= t^3 \end{aligned}$$

Здесь верхний индекс B_s^k указывает на степень полинома. Кубические полиномы Бернштейна можно получить из разложения следующего выражения:

$$\begin{aligned} [(1-t) + t]^3 &= (1-t)^3 + 3t(1-t)^2 + 3t^2(1-t) + t^3 \\ \sum_{i=0}^3 B_i^3(t) &= 1. \end{aligned} \quad (8.53)$$

Точка $P(t)$ является аффинной комбинацией точек и, следовательно, эти взвешенные точки можно складывать. На рис.3а приведены полиномы Бернштейна третьего порядка, а на рис.3б четвертого порядка.

Обобщение алгоритма де Кастельо на произвольное число точек.

Обобщим данный алгоритм на случай $(L+1)$ контрольной точки P_0, P_1, \dots, P_L . для каждого значения t строится последовательность «поколений», для формирования каждого из которых используется твининг соседних точек, созданных в предыдущем поколении.

$$\begin{aligned} P_i^k(t) &= (1-t)P_i^{k-1}(t) + tP_{i+1}^{k-1}(t) \\ P_i^L(t) &= (1-t)P_i^{L-1}(t) + tP_{i+1}^{L-1}(t) \end{aligned}$$

где $i = 0, 1, \dots, L$.

Верхний индекс k в выражении $P_i^k(t)$ обозначает номер поколения. Процесс начинается с $P_i^0(t) = P_i$ и заканчивается готовой кривой Безье $P(t) = P_i^L(t)$. Кривую Безье можно записать в терминах полиномов Бернштейна следующим образом:

$$P(t) = \sum_{i=0}^L P_i B_i^L(t)$$

где i -тый полином Бернштейна степени L имеет вид:

$$B_i^L(t) = C_L^k (1-t)^{L-k} t^k,$$

а $C_L^k = \frac{L!}{k!(L-k)!}$ для $L \geq k$ (при $L > k$ этот коэффициент равен 0),

$[(1-t) + t]^L$ -бином Ньютона.

$$\sum_{i=0}^L B_i^L(t) = 1.$$

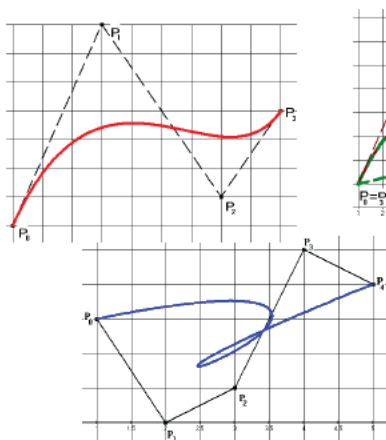


Рис. 8.17. Кривые Безье

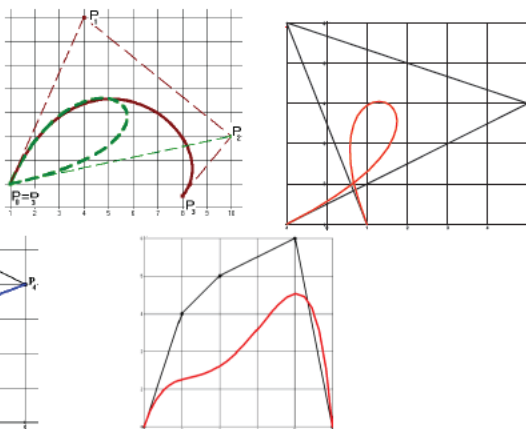


Рис. 8.18. Пяти точечная кривая Безье

Пусть заданы следующие 5 точек $P_0(1,1.5)$, $P_1(2,0)$, $P_2(3,0.5)$, $P_3(4,2.5)$, $P_4(5,2)$. Построим на их базе кривую Безье:

$$P(t) = P_0(t)(1-t)^4 + 4P_1(t)t(1-t)^3 + 6P_2(t)t^2(1-t)^2 + 4P_3(t)t^3(1-t) + P_4(t)t^4$$

$$P(t) = P_0(t)(1-t)^2 + 2P_1(t)t(1-t) + P_2(t)t^2$$

$$P(t) = P_2(t)(1-t)^2 + 2P_3(t)t(1-t) + P_4(t)t^2$$

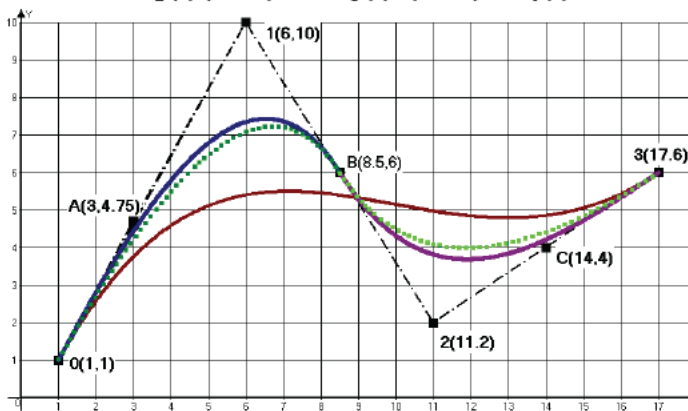


Рис. 8.19. Составные двух и трех точечные кривые

Кривая Безье 3-го порядка: $P(t) = t^3P_0 + 3(1-t)t^2P_1 + 3t(1-t)^2P_2 + (1-t)^3P_3$.

Замена составной кривой:

- из двух кривых Безье второго порядка (введены две вспомогательные точки В и С)

- из двух кривых Безье третьего порядка (введены три вспомогательные точки А, В и С)

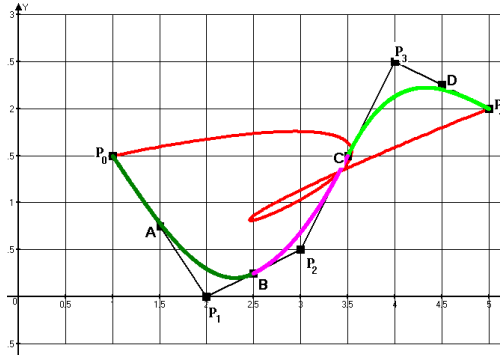


Рис. 8.20. Кривая Безье 4-го порядка. Замена составной кривой: 2-3-го и 1-2-го порядка

Чтобы составная кривая Безье была геометрически непрерывной, необходимо чтобы каждые три точки вместе стыковки сегментов лежали на одной прямой. Например, пусть имеется пять базовых точек P_0, P_1, P_2, P_3, P_4 . Для построения геометрически-непрерывной составной кривой дополним этот набор вспомогательными точками P_5, P_6 , взятыми на серединах отрезков P_1P_2 и P_2P_3 (рис. 8.19). Составную кривую построим из трех сегментов элементарных квадратичных кривых Безье для троек вершин P_0, P_1, P_5 ; P_5, P_2, P_6 и P_6, P_3, P_4 .

В форме Безье используются четыре точки. Касательные векторы в конечных точках задаются отрезками P_1P_2 и P_3P_4 . В частности, касательные векторы R_1 и R_4 , введенные выше, определялись таким образом, чтобы соответствовать четырем точкам Безье P_1, P_2, P_3, P_4 :

$$R_1 = 3(P_2 - P_1) = P'_1(0), \quad R_4 = 3(P_4 - P_3) = P'_3(1).$$

Поэтому, соотношение между геометрической матрицей G_h и геометрической матрицей Безье G_b записывается следующим образом:

$$G_h = \begin{bmatrix} P_1 \\ P_4 \\ R_1 \\ R_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -3 & 3 & 0 & 0 \\ 0 & 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = M_{hb} G_b$$

Подставляя полученное соотношение в формулу (7), получим

$$x(t) = T M_h G_{hx} = T M_h M_{hb} G_{bx}.$$

Обозначим произведение $M_h M_{hb}$ через M_b получим выражение

$$x(t) = T M_b G_{bx},$$

которое имеет теперь форму Безье, где

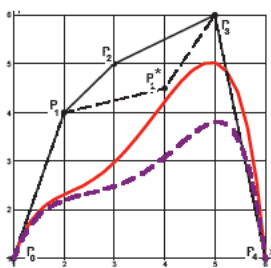
$$M_b = M_n M_{nb} = \begin{bmatrix} -1 & 3 & -3 & 10 \\ 3 & -6 & 1 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Форма Безье чаще используется в графических системах из-за следующих двух своих свойств:

■ Геометрическая матрица (в случае четырех точек) привлекательна в интерактивном режиме, т.е. перемещая точки с помощью курсора, можно легко привести кривую к желаемой форме. (В случае эрмитовой формы касательные векторы должны задаваться в явном виде. В режиме диалога их определять труднее. Однако если касательный вектор известен, то привести кривую в соответствие с ним легче при использовании эрмитовой формы).

■ Четыре управляющие точки определяют выпуклый многоугольник (выпуклую оболочку), внутри которого находится кривая Безье.

Выпуклая оболочка точек на плоскости есть область, задаваемая резиновой нитью, натянутой вокруг всех точек. Она полезна при отсечении по окну. Вместо того, чтобы сразу проводить отсечение кривой, стоит проверить ее выпуклую оболочку, и только если она пересекает окно следует проверять саму кривую. Чтобы лучше понять свойство выпуклой оболочки формы Безье, возьмем произведение $TM_b G_b$:



$$(1-t)^3 P_1 + 3t(1-t)^2 P_2 + 3t^2(1-t) P_3 + t^3 P_4$$

Исследуя коэффициенты многочлена при четырех точках, находим, что каждый из них изменяется в диапазоне 0-1, а их сумма равна 1 для всех $0 \leq t \leq 1$. Это выражение описывает взвешенное среднее значение для 4 управляющих точек.

Основной недостаток кривых Безье заключается в том, что они не предоставляют достаточного

локального контроля – изменение любой контрольной точки приводит к изменению всей кривой. Проблема решается с помощью необходимого набора стыковочных функций $R_0(t), R_1(t) \dots R_5(t)$, которые должны быть найдены, но которые имеют поддержку только на части интервала $[0,1]$.

Определение. Интервал, в котором функция отлична от 0 называют ее поддержкой.

Выясним, какими свойствами должен обладать набор стыковочных функций.

$$V(t) = \sum_{k=0}^L P_k R_k(t) \quad t \in [a, b], \quad R_0(t), \dots, R_L(t) - \text{стыковочные функции.}$$

■ Быть простыми в вычислении и численно устойчивыми; (полиномы достаточно низкой степени);

■ Составлять в сумме единицу для любого t из $[a, b]$: $(\sum_{k=0}^L R_k(t) = 1, \text{ т.к. } V(t) \text{ является аффинной суммой точек. При каждом значении } t \text{ кривая } V(t) \text{ является взвешенной суммой точек } (R_k(t) - \text{ вес } k\text{-той точки}), \text{ а это имеет смысл, если } V(t) \text{ является аффинной } \sum(\cdot) \text{ для любого } t.)$

■ Иметь поддержку лишь на небольшом участке интервала $[a, b]$ – для обеспечения локального контроля;

■ Интерполировать определенные контрольные точки, выбираемые дизайнером;

■ Быть достаточно гладкими, чтобы образовать нужную форму.

8.6. Кусочно-полиномиальные кривые.



Рис. 8.22. Заданная кривая кубический полином, форма которого показана на рис.1 и который удовлетворял бы определенным требованиям. Для этого зададим функцию

$$R(t) = at^3 + bt^2 + ct + d$$

И попробуем подобрать коэффициенты так, чтобы функция $R(t)$ и ее первая производная обращались в нуль при $t=0$ и при $t=1$. Это требование накладывает на коэффициенты следующие четыре условия:

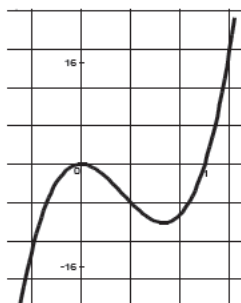


Рис. 8.23. Полином 4 степени

$$\left. \begin{aligned} R(0) &= d = 0 \\ R(1) &= a + b + c + d = 0 \\ R'(0) &= c = 0 \\ R'(1) &= 3a + 2b + c = 0 \end{aligned} \right\}$$

Из этих условий следует, что $a = b = c = d = 0$. Следовательно, подобной формы не существует.

У кубической формы недостаточно гибкости, чтобы «изогнуться» требуемым образом.

Рассмотрим полином четвертой степени (рис. 8.23):

$$R(t) = at^4 + bt^3 + ct^2 + dt + e$$

$$R'(t) = 4at^3 + 3bt^2 + 2ct + d$$

При тех же граничных условиях:

$$R(0) = e = 0$$

$$R(1) = a + b + c + d = 0$$

$$R'(0) = d = 0$$

$$R'(1) = 4a + 3b + 2c + d = 0$$

Получим следующую систему уравнений:

$$e = d = 0$$

$$\begin{cases} a + b + c = 0 \\ 4a + 3b + 2c = 0 \end{cases} \left| \begin{array}{l} a = -(b + c) \\ -4b - 4c + 3b + 2b = 0 \Rightarrow -b - 2c = 0 \end{array} \right. \\ \begin{cases} b = -2c \\ a = -(-2c + c) = c; \end{cases} \begin{cases} a = c \\ b = -2c \end{cases}$$

Пусть $c = 1$; $\Rightarrow a = 1$;

$$R(t) = t^4 + 2t^3 + t^2 \Rightarrow R(t) = t^2(t - 1)^2$$

Рассмотрим параметрическую форму

$$P(t) = \sum_{L=0}^L P_L R_L(t)$$

На базе $(L + 1)$ контрольных точек и $(L + 1)$ стыковочных функций $R_0(t), \dots, R_L(t)$. Пусть $P(t)$ - аффинная сумма точек. В качестве стыковочных функций используем кусочные полиномы, которые определены в более общей последовательности узлов, именуемой узловым вектором

$$\vec{T} = (t_0, t_1, t_2 \dots),$$

Списки узловых значений t_0, t_1, t_2, \dots , для которых предполагается, что они не убывают, т.е. $t_i \leq t_{i+1}$. Некоторые из узлов могут иметь одинаковые значения, однако и в этом случае им присваиваются отдельные имена.

Пусть каждая стыковочная функция $R_k(t)$ - кусочный полином, равный 0 до момента времени t , отличен от 0 (на протяжении нескольких диапазонов узлов) на протяжении поддержки, а затем снова возвращается к 0 значению. Потребуем, чтобы каждая $R_k(t)$ была сплайн-функцией, то есть обладала достаточной степенью гладкости при всех поддерживающих ее значениях t . В этом случае вся кривая

$$P(t) = P_0 R_0(t) + P_1 R_1(t) + \dots$$

это сумма кусочных полиномов с весами, определяемыми контрольными точками. В соседнем диапазоне эта кривая определяется другой суммой полиномов, но все ее сегменты соединяются между собой, сохраняя непрерывность кривой. Такая кривая называется сплайн-кривой.

Отметим! Сплайн-функция – это обычный кусочный полином с определенной степенью гладкости.

Сплайн-кривая – это аффинные сопряжение точек, в котором используется кусочно-полиномиальные функции сопряжения. Сплайн-кривая непрерывна в своих узлах, однако производные в узлах могут терпеть разрыв.

Для достижения большей гибкости составим вместе несколько полиномов низкого порядка. Такие кривые определяются различными полиномами на различных интервалах изменения t и называются кусочными полиномами.

Определение. Сплайн-функция L -той степени, это кусочно-полиномиальная функция степени L , обладающая в каждом узле $(L - 1)$ гладкостью.

Рассмотрим параметрическую форму

$$P(t) = \sum_{l=0}^L P_l R_l(t)$$

На базе $(L + 1)$ контрольных точек и $(L + 1)$ стыковочных функций $R_0(t), \dots, R_L(t)$. Пусть $P(t)$ - аффинная сумма точек. В качестве стыковочных функций используем кусочные полиномы, которые определены в более общей последовательности узлов, именуемой узловым вектором

$$\vec{T} = (t_0, t_1, t_2 \dots),$$

Списки узловых значений t_0, t_1, t_2, \dots , для которых предполагается, что они не убывают, т.е. $t_i \leq t_{i+1}$. Некоторые из узлов могут иметь одинаковые значения, однако и в этом случае им присваиваются отдельные имена.

Пусть каждая стыковочная функция $R_k(t)$ - кусочный полином, равный 0 до момента времени t , отличен от 0 (на протяжении нескольких диапазонов узлов) на протяжении поддержки, а затем снова возвращается к 0 значению. Потребуем, чтобы каждая $R_k(t)$ была сплайн-функцией, т.н. обладала достаточной степенью гладкости при всех поддерживающих ее значениях t . В этом случае вся кривая

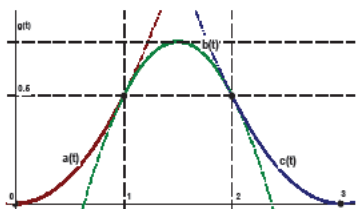
$$P(t) = P_0 R_0(t) + P_1 R_1(t) + \dots$$

Это сумма кусочных полиномов с весами, определяемыми контрольными точками. В соседнем диапазоне эта кривая определяется другой суммой полиномов, но все ее сегменты соединяются между собой, сохраняя непрерывность кривой. Такая кривая называется сплайн-кривой.

Отметим! Сплайн-функция – это обычный кусочный полином с определенной степенью гладкости. Сплайн-кривая – это аффинные сопряжение точек, в котором используется кусочно-полиномиальные функции сопряжения. Сплайн-кривая непрерывна в своих узлах, однако производные в узлах могут терпеть разрыв.

Для достижения большей гибкости составим вместе несколько полиномов низкого порядка. Такие кривые определяются различными полиномами на различных интервалах изменения t и называются кусочными полиномами.

Рассмотрим следующий пример (рис.3).Кривая $l(t)$ состоит из трех полиномиальных сегментов:



$$a(t) = \frac{1}{2}t^2 \text{ определена в диапазоне } [0,1]$$

$$b(t) = \frac{3}{4} - \left(t - \frac{3}{2}\right)^2 \text{ - в диапазоне } [1,2]$$

$$c(t) = \frac{1}{2}(3-t)^2 \text{ - в диапазоне } [2,3].$$

Рис. 8.24. Кусочно-полиномиальные функции сопряжения

Поддержка всей кривой $l(t)$ равна $[0,3]$. Точки, в которых встречается пара отдельных сегментов, называются стыками (на рисунке точки A и B), а значения t , при которых происходит стыковка, называются узлами $(0,1,2,3)$.

Функция $l(t)$ и ее производная непрерывны внутри каждого диапазона и на стыках. Действительно:

$$\begin{aligned} a(t) &= \frac{1}{2}t^2 & a' &= t & a'' &= 1 \\ b(t) &= \frac{3}{4} - \left(t - \frac{3}{2}\right)^2 & b' &= -2t + 3 & b'' &= -2 \\ c(t) &= \frac{1}{2}(3-t)^2 & c' &= -3 + t & c'' &= 1 \end{aligned}$$

$$\begin{aligned} a(1) &= b(1) = \frac{1}{2} \\ b(2) &= c(2) = \frac{1}{2} \\ a'(1) &= b'(1) = 1 \\ b'(2) &= c'(2) = -1 \end{aligned}$$

Следовательно, ни сама функция, ни наклон кривой при движении от одного полиномиального участка к другому не изменяется скачкообразно. Вторая производная в узлах совершает скачок.

$$\begin{aligned} a''(1) &= 1 & b''(2) &= -2 \\ b''(1) &= -2 & c''(2) &= -1 \end{aligned}$$

Форма $l(t)$ - пример сплайн-функции - кусочно полиномиальной функции, обладающей «достаточной» гладкости.

Определение. Сплайн-функция L -той степени, это кусочно-полиномиальная функция степени L , обладающая в каждом узле $(L - 1)$ гладкостью.

Требования к стыковочным функциям

Пусть для генерирования кривой используется алгоритм, осуществляющий сопряжение контрольных точек по формуле

$$V(t) = \sum_{k=0}^L P_k R_k(t) \quad t \in [a, b],$$

где $R_0(t), \dots, R_L(t)$ - стыковочные функции, которые должны обладать следующими желательными свойствами:

- быть простыми в вычислении и численно устойчивыми, т.е. чтобы вычислялись просто и были как можно меньше чувствительными к ошибкам округления. Следовательно, стыковочные функции должны быть полиномами достаточно малой степени. Использование тригонометрических функций не отвечает этому требованию.

- составлять в сумме единицу для любого t из $[a, b]$, т.к. только в этом случае при каждом значении t кривая $V(t)$ является взвешенной суммой точек, т.е. является аффинной суммой точек для любого t из промежутка $[a, b]$

$$\sum_{k=0}^L R_k(t) = 1$$

- иметь поддержку лишь на небольшом участке интервала $[a, b]$ -для обеспечения локального контроля: т.е. поддержка каждой стыковочной функции была сконцентрирована на небольшом участке интервала $[a, b]$. (Интервал, в котором функция отлична от нуля, часто называют ее поддержкой)

- Интерполировать определенные контрольные точки, выбираемые дизайнером, т.е. пользователь может пожелать, чтобы кривая $V(t)$ проходила через некоторые из контрольных точек, а к остальным она должна только приближаться (притягиваться). Механизм изменения стыковочных функций с целью обеспечения интерполяции определенных точек рассмотрим ниже

- Быть достаточно гладкими, чтобы образовать желаемую форму; гладкость $V(t)$ зависит от гладкости стыковочных функций. Важно, чтобы они плавно начинались и заканчивались (т.е. начинались и заканчивались с нулевым значением производной).

***V*-сплайны.** Базисом сплайн-кривых, задаваемых уравнением

$$P(t) = \sum_{k=0}^L P_k R_k(t) \tag{8.54}$$

называется семейство стыковочных функций, которые можно использовать для генерирования всех возможных сплайн-кривых, определяемых данным узловым вектором.

Базис, для которого стыковочные функции имеют наименьшую поддержку и, в силу этого, обеспечивают наибольший локальный контроль, образуют

B-сплайны (буква *B* взята из слова «базис»). Каждая *B*-сплайн-функция базируется на полиномах определенного порядка *m*. Если *m* = 3, то полиномы будут 3 порядка, и, следовательно, степени 2, и т.д.

Уравнение для задания *B*-сплайны кривой имеет вид

$$P(t) = \sum_{k=0}^L P_k N_{k,m}(t) \quad (8.55)$$

Заданы: ■ узловой вектор $T = (t_0, t_1, t_2 \dots)$;

■ $(L + 1)$ контрольная точка P_k ;

■ *m* -порядок *B*-сплайн функций.

Формула, описывающая *B*-сплайн функцию $N_{k,m}(t)$, имеет следующий вид:

$$N_{k,m}(t) = \left(\frac{t - t_k}{t_{k+(m-1)} - t_k} \right) N_{k,m-1}(t) + \left(\frac{t_{k+m} - t}{t_{k+m} - t_{k+1}} \right) N_{k+1,m-1}(t)$$

Это определение рекурсивно и описывает построение функции *m*-го порядка из двух -сплайн функций (*m* - 1)-го порядка. Для инициации этой формулы необходимо определить функцию первого порядка:

$$N_{k,1}(t) = \begin{cases} 1 & t_k < t \leq t_{k+1} \\ 0 & \text{в противном случае} \end{cases} \quad (8.56)$$

$$\sum_{k=0}^L N_{k,1} = 1.$$

Линейные *B*-сплайны. Определим форму первой (*k* = 0) *B*-сплайн функции порядка *m* = 2 $N_{0,2}(t)$, если все ее узлы – равноотстоящие,

т.е. $T = (0,1,2, \dots)$.

При этих значениях параметров уравнение (3) имеет вид:

$$N_{0,2}(t) = \frac{t}{1} N_{0,1}(t) + \frac{2-t}{1} N_{1,1}(t) \quad (8.57)$$

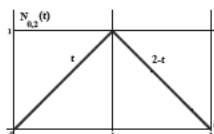
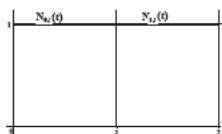


Рис. 8.25. Построение-сплайн функции порядка *m* = 2 $N_{0,2}(t)$

$N_{0,2}(t)$ - аффинная комбинация $N_{0,1}(t)$ и $N_{1,1}(t)$ следовательно:

$$N_{0,2}(t) = \begin{cases} t & 0 \leq t \leq 1 \\ 2 - t & 1 \leq t \leq 2 \\ 0 & \text{при остальных } t \end{cases}$$

При равноотстоящих узлах любой линейный сплайн будет смещенной версией нулевого сплайна, т.е. $N_{i,2}(t) = N_{0,2}(t - 1)$.

$$\text{Так } N_{1,2}(t) = (t - 1)N_{1,1}(t) + \frac{3-t}{1}N_{2,1}(t).$$

Линейные B -сплайны используются в процессе построения B -сплайнов более высоких порядков.

Квадратичные B -сплайны.

Построим квадратичные ($m = 3$) B -сплайн функции $N_{k,3}(t)$ на базе тех же равноотстоящих узлов, что и для линейных B -сплайнов. Достаточно построить только функцию $N_{0,3}(t)$, т.к. остальные функции могут быть получены простым смещением. Из уравнения (3) получим

$$\begin{aligned} N_{0,3}(t) &= \frac{t}{2}N_{0,2}(t) + \frac{3-t}{2}N_{1,2}(t) = \\ &= \frac{t}{2} [tN_{0,1}(t) + (2-t)N_{1,1}(t)] + \frac{3-t}{2} [(t-1)N_{1,1}(t) + (3-t)N_{2,1}(t)] \\ &= \frac{t^2}{2}N_{0,1}(t) + \left[-t^2 + 3t - \frac{9}{4} - \frac{6}{4} + \frac{9}{4} \right] N_{1,1}(t) + \frac{(3-t)^2}{2}N_{2,1}(t) \\ &= \frac{t^2}{2}N_{0,1}(t) + \left[\frac{3}{4} - \left(t - \frac{3}{2} \right)^2 \right] N_{1,1}(t) + \frac{(3-t)^2}{2}N_{2,1}(t) \end{aligned}$$

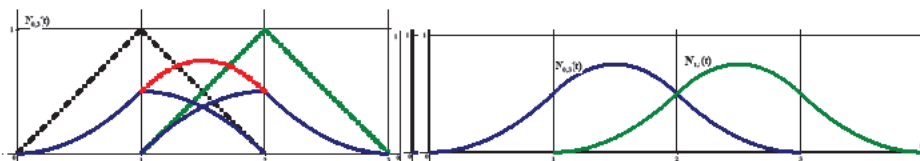


Рис. 8.26. Построение B -сплайн функции порядка $m = 3$ $N_{0,3}(t); N_{1,3}(t)$

Для равноотстоящих узлов, если $t_k = k$, то $N_{k,m}(t) = N_{0,m}(t - k)$

$$N_{0,3}(t) = \begin{cases} \frac{t^2}{2} & 0 \leq t \leq 1 \\ \frac{3}{4} - \left(t - \frac{3}{2} \right)^2 & 1 \leq t \leq 2 \\ \frac{1}{2}(3-t)^2 & 2 \leq t \leq 3 \\ 0 & \text{при остальных } t \end{cases}$$

$N_{0,3}(t)$ зависит от четырех узлов 0,1,2,3, а его поддержкой является промежуток $[0,3]$. Произведение подъема и спуска на треугольные импульсы представляют собой две параболы, пересекающиеся под углом

Кубические B-сплайны. Построим кубические ($m = 4$) B-сплайн функции $N_{k,4}(t)$ на базе тех же равноотстоящих узлов, что и для предыдущих линейных и квадратичных B-сплайнов. Достаточно построить только функцию $N_{0,4}(t)$, т.к. остальные функции могут быть получены простым смещением.

$N_{0,4}(t)$ - для равноотстоящих узлов.

$$N_{0,4}(t) = \begin{cases} u(1-t) & 0 \leq t \leq 1 \\ v(2-t) & 1 \leq t \leq 2 \\ v(t-2) & 2 \leq t \leq 3 \\ u(t-3) & 3 \leq t \leq 4 \\ 0 & \text{при остальных } t \end{cases} \Rightarrow \begin{cases} u(t) = \frac{1}{6}(1-t)^3 \\ v(t) = \frac{1}{6}(3t^3 - 6t^2 + 4) \end{cases}$$

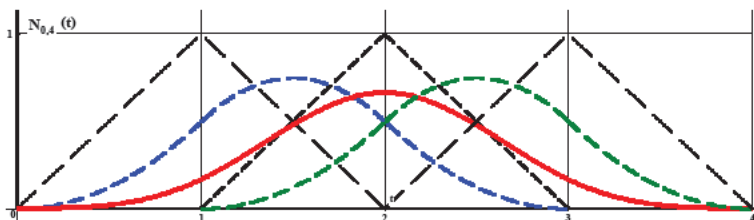


Рис. 8.27. Построение B-сплайн функции порядка $m = 4$ $N_{0,4}(t)$

Стандартный узловой вектор

Для B-сплайна любого порядка при равноотстоящих узлах имеет место равенство: если $t_k = k \Rightarrow N_{k,m} = N_{0,m}(t - k)$

Изменяя расстояние между узлами, пользователь получает большую возможность контролировать форму конструируемой кривой. Рассмотрим, что происходит в случае, когда два узла оказываются очень близко друг к другу.

Рассмотрим пример: пусть узловой вектор имеет вид

$T = (0, 1, 2, 3, 3 + \varepsilon, 4 + \varepsilon, \dots)$ где ($\varepsilon > 0$) малое положительное число.

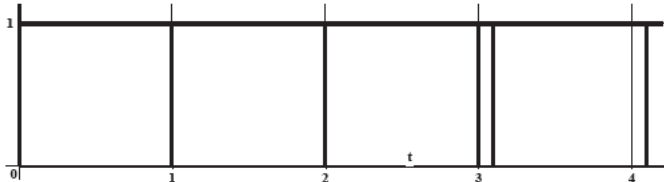


Рис. 8.28. Новый узловой вектор

Очевидно, что соответствующий отрезок кусочного полинома, лежащего в интервале $[3, 3 + \varepsilon]$, "вжимается" в очень узкий диапазон. При этом стыковочные функции не будут собственными смещенными версиями.

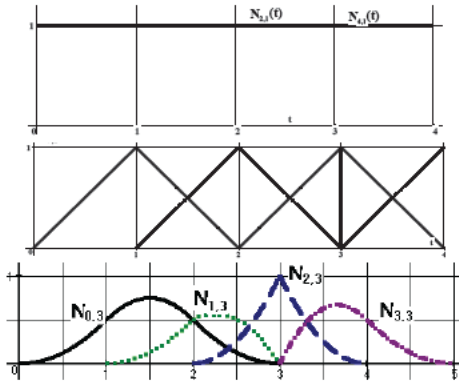


Рис. 8.29

Положив $\varepsilon = 0$, получим следующую ситуацию:

- диапазон $[3, 3 + \varepsilon]$ исчезает;
- при $t = 3$ получим кратный узел. В рассматриваемом примере он имеет кратность 2;
- узловой вектор принимает вид $T=(0,1,2,3,3,4,5)$, т.е. и $t_3=3$ и $t_4=3$;
- две линейные формы при $t = 3$ имеют разрыв, т.к.

$$N_{1,3}(t) = (t - 2)N_{1,2}(t) + \frac{3-t}{1}N_{2,1}(t).$$

- квадратичные формы – разрывную производную.

В общем случае в кратном узле i -гладкая кривая переходит в $(i - 1)$ -гладкую кривую.

Для квадратичного сплайна в рассматриваемом случае кривая будет интерполировать контрольную точку P_2 , т.к. $N_{2,3}(3) = 1$, а все остальные стыковочные функции равны 0.

При приближении t к кратному узлу возникает сильное притяжение к управляющей контрольной точке. При конструировании кривых в качестве стандартного принят узловой вектор специального вида при котором кривая интерполирует первую и последнюю контрольные точки, но усиливает возможность предсказать поведение конструирующей кривой.

Определение. *Стандартный узловой вектор B-сплайна порядка m начинается и заканчивается узлом кратности m , а между остальными узлами задается едичное расстояние.*

Пример. Пусть имеется 8 контрольных точек. Необходимо использовать кубические ($m = 4$) B-сплайны. В этом случае стандартный узловой вектор имеет вид:

$$T = (0, 0, 0, 0, 1, 2, 3, 4, 5, 5, 5, 5)$$

Стандартный узловой вектор для $(L + 1)$ контрольной точки В-сплайнов - го порядка описывается так:

1. Всего в нем содержится $L + m + 1$ узлов: t_0, t_1, \dots, t_{L+m}
2. Все первые m узлов t_0, \dots, t_{m-1} равны 0 (начинаются при $t = 0$)
3. Узлы t_m, \dots, t_L возрастают с единичным шагом от 1 до $L - m + 1$
4. Все последние m узлов t_{L+1}, \dots, t_{L+m} равны $L - m + 2$

Ограничение: порядок не может превосходить число контрольных точек

8.8. Рациональные сплайны и NURBS - кривые

Рациональные параметрические формы. Рассмотрим вкратце класс функций, называемых рациональными дробями. Именно эти функции являются основой NURBS и неоднородного рационального В-сплайна, которые будут рассмотрены позднее.

Пусть параметризация для каждой из функций $x(t)$ и $y(t)$ определена как отношение двух полиномов.

Рассмотрим некоторые частные случаи.

■ **Линейный случай:** $x = \frac{a+bt}{c+dt}, y = \frac{e+ft}{g+ht}$.

Из $x = \frac{a+bt}{c+dt}$ определим $t = \frac{a-cx}{-b+dx}$ и подставим в выражение для y получим, принимая что $xd - b \neq 0$, после простейших преобразований

$$y[(ah - bg) + (gd - hc)x] = (af - eb) + (ed - fc)x: \text{или}$$

$$Axy + By + Cx + D = 0,$$

$$A = gd - hc; B = (ah - bg); C = fc - ed; D = eb - af.$$

Получим уравнение, дающее прямую при $A = 0$, кривую второго порядка - гиперболу, если $A \neq 0$.

(детерминант $\Delta = a_{11}a_{22} - a_{12}^2$ коники $\Delta = -A^2 < 0$)

■ **Случай** квадратичных полиномов, особо выделяя следующего параметрическую форму:

$$P(t) = \frac{P_0(1-t)^2 + 2\omega P_1(1-t)t + P_2t^2}{(1-t^2) + 2\omega t(1-t) + t^2}, \quad (8.58)$$

где P_0, P_1, P_2 - три произвольные точки плоскости, ω - весовой коэффициент.

Эти контрольные точки и вес определяют форму кривой.

Уравнение (2) представляет собой сокращенную запись следующего выражения

$$(x(t), y(t)) = \left\{ \frac{x_0(1-t)^2 + 2\omega x_1(1-t)t + x_2t^2}{(1-t^2) + 2\omega t(1-t) + t^2}, \frac{y_0(1-t)^2 + 2\omega y_1(1-t)t + y_2t^2}{(1-t^2) + 2\omega t(1-t) + t^2} \right\},$$

где $P_0(x_0, y_0), P_1(x_1, y_1), P_2(x_2, y_2)$.

Полиномы в знаменателе одинаковы для $x(t)$, $y(t)$, не зависят от контрольных точек, а зависят только от «весового» параметра ω .

При $t = 0$ справа получим (x_0, y_0) , следовательно кривая проходит через точку P , интерполирует ее. При $t = 1$ кривая проходит через точку P_2 . Для всех $0 \leq t \leq 1$ функция $P(t)$ зависит от всех трех точек.

Кривая - одно из конических сечений, тип которого зависит от значения ω :

$\omega < 1$ - кривая является эллипсом

$\omega = 1$ - параболой

$\omega > 1$ – гиперболой.

■ Случай кубических полиномов, особо выделяя следующего

$$P(t) = \frac{\omega_0 P_0 (1-t)^3 + 3\omega_1 P_1 (1-t)^2 t + 3\omega_2 P_2 (1-t) t^2 + \omega_3 P_3 t^3}{\omega_0 (1-t)^3 + 3\omega_1 (1-t)^2 t + 3\omega_2 (1-t) t^2 + \omega_3 t^3} \quad (8.59)$$

Коэффициенты $\omega_i (i = 0, 1, 2, 3)$ называются весом (значимостью) точки P_i .

Кривые, характеристические точки, которых обладают весами, называются рациональными.

Рациональная сплайн-кривая формируется приведенным ранее методом сопряжения контрольных точек

$$P(t) = \sum_{k=0}^L P_k R_k(t).$$

Пользователь задается набором весов $\{\omega_0, \omega_1, \dots, \omega_k\}$, после чего формирует следующие стыковочные функции:

$$R_k(t) = \frac{\omega_k N_{k,m}(t)}{\sum_{k=0}^L \omega_k N_{k,m}(t)} \quad P(t) = \frac{\sum_{k=0}^L \omega_k N_{k,m}(t)}{\sum_{k=0}^L \omega_k N_{k,m}(t)}$$

где $(\omega_0, \omega_1 \dots \omega_l)$ -параметры формы.

Веса, параметры формы, $\omega_k > 0$ обычно задаются не отрицательными чтобы не опасаться обращения в 0 знаменателя. Поскольку узловой вектор, используемый при определении B -сплайн функций, как правило, является неравномерным (узлы в нем – не равноотстоящие), то такое семейство кривых получило название неравномерных рациональных B –сплайнов, т.е *NURBS*.

Если все веса одинаковы, то знаменатель уравнения превращается в константу, вследствие чего эта форма сводится к уже знакомому B -сплайну,

$$P(t) = \sum_{k=0}^L P_k N_{k,m}(t).$$

Действительно, пусть $\omega_0 = \omega_1 = \dots = \omega_k = \omega$, тогда

$$\sum_{k=0}^L \omega_k N_{k,m}(t) = \omega \sum_{k=0}^L N_{k,m}(t), \text{ а т.к. } \sum_{k=0}^L N_{k,m}(t) = 1, \text{ то}$$

$$\sum_{k=0}^L \omega_k N_{k,m}(t) = \omega = \text{const.}$$

и отличается от него только при неравных весов.

Преимущества неравномерных рациональных B -сплайнов $NURBS$:

- при правильном выборе контрольных точек и весов функция $P(t)$ является в точности коническим сечением
- инвариантны относительно проективных преобразований, обеспечивающих перспективную проекцию сцен.

8.9. Аппроксимация

Метод наименьших квадратов.

При замерах координат точек неизбежно происходят ошибки измерений: грубые, систематические, случайные. В процессе обработки данных грубые и систематические ошибки устраняются. Случайные же ошибки влияют на все производимые расчеты. При этом интерполирующая функция повторяет случайные ошибки исходных данных. Поэтому интерполяция применяется только тогда, когда этими ошибками можно пренебречь. В противном случае возникает задача аппроксимации, когда требуется найти кривую, которая наиболее точно соответствует исходной кривой, заданной дискретной информацией. Для решения этой задачи надо установить критерий приближения. Математически доказано, что наилучшим критерием аппроксимации является критерий, в котором оценивается минимум суммы квадратов отклонений аппроксимирующих и исходных значений функцией (невязок):

$$S = \sum_{i=1}^n (y_i - y_{\text{appr}})^2 = \min. \quad (8.60)$$

Постановка задачи. По координатам узловых точек x_i, y_i отвода определить коэффициенты a_0, a_1, \dots, a_m аппроксимирующей функции.

Поставленная задача решается по следующей схеме:

1. Составляется функция S , зависящая от коэффициентов a_0, a_1, \dots, a_n и удовлетворяющая условию минимума суммы квадратов невязок;
2. Записываются условия наличия минимума функции S и определяются коэффициенты аппроксимирующей функции.

Аппроксимация прямой линией.

Дано: k точек (x_i, y_i)

Требуется: задать прямую, которая по критерию наименьших квадратов наилучшим образом «проходит» через эти точки.

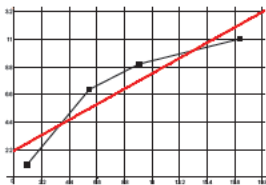


Рис. 8.30. Аппроксимация прямой

Запишем уравнение прямой

$$y = a_0 + a_1 x.$$

Составим функцию S

$$S = \sum_{i=1}^k (y_i - y_{\text{апр.}})^2 = \sum_{i=1}^k (y_i - a_0 - a_1 x_i)^2 = \min. \quad (8.61)$$

В данном случае условия наличия \min соответствуют обращению в нуль частных производных функции S по переменным a_0 и a_1 .

$$\left. \begin{aligned} \frac{\partial S}{\partial a_0} &= -2 \sum_{i=1}^k (y_i - a_0 - a_1 x_i) = 0 \\ \frac{\partial S}{\partial a_1} &= -2 \sum_{i=1}^k (y_i - a_0 - a_1 x_i) x_i = 0 \end{aligned} \right\}$$

или
$$\begin{aligned} k a_0 + a_1 \sum_{i=1}^k x_i &= \sum_{i=1}^k y_i. & \rightarrow & \quad k a_0 + a_1 \sum_{i=1}^k x_i = \sum_{i=1}^k y_i. \\ a_0 \sum_{i=1}^k x_i + a_1 \sum_{i=1}^k x_i^2 &= \sum_{i=1}^k x_i y_i & \rightarrow & \quad a_0 \sum_{i=1}^k x_i + a_1 \sum_{i=1}^k x_i^2 = \sum_{i=1}^k x_i y_i \end{aligned}$$

Аппроксимация с помощью полиномов

Приближая функцию $y(x)$ с помощью полинома $P(x)$, нужно прежде всего выбрать степень n этого полинома:

$$P(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

При выборе такой аппроксимации большую роль играет теорема Вейерштрасса, которая гласит, что "любую функцию на отрезке $[a, b]$ можно приблизить сколь угодно точно полиномом достаточно высокой степени". (аналогичные теоремы справедливы для функций двух и большего числа переменных).

Пусть заданы m точек и ордината y_i известна с какой-то ошибкой δ_i . Тогда

$$\begin{aligned} \delta_1 &= a_0 + a_1 x_1 + a_2 x_1^2 + \dots + a_n x_1^n - y_1 \\ \delta_2 &= a_0 + a_1 x_2 + a_2 x_2^2 + \dots + a_n x_2^n - y_2 \\ &\dots \dots \dots \\ \delta_m &= a_0 + a_1 x_m + a_2 x_m^2 + \dots + a_n x_m^n - y_m \end{aligned}$$

Формула (1) запишется в виде

$$S = \sum_{i=1}^m (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_n x_i^n - y_i)^2 = \min$$

Для решения этой задачи найдем частные производные от S по всем переменным a_0, a_1, \dots, a_n , приравняем их 0 и решим систему $(n + 1)$ уравнений с $(n + 1)$ неизвестными

Поскольку сплайновая кривая *Catmul IRom* является интерполяционной, то она проходит через каждую из базовых точек. Однако при построении составной сплайновой кривой каждый сегмент рассчитывается на участке между парой внутренних точек очередной четверки из набора базовых точек. Поэтому для построения составной сплайновой интерполяционной кривой, начинающейся в первой базовой точке и заканчивающейся в последней базовой точке, достаточно дополнить набор копиями первой и последней точек. Копия начальной точки при этом добавляется в начало набора, а копия последней точки — в конец набора.

Построить замкнутую сплайновую интерполяционную кривую можно, дополнив набор базовых точек из n штук точками:

$$P_{n+1} = P_0, P_{n+2} = P_1.$$

$$P_{n+3} = P_2.$$

Элементарная бета-сплайновая кривая

По заданному массиву точек P_0, P_1, P_2, P_3 бета-сплайновая кривая описывается уравнением:

$$R(t) = b_0(t)P_0 + b_1(t)P_1 + b_2(t)P_2 + b_3(t)P_3$$

$$0 \leq t \leq 1$$

Функциональные коэффициенты $b_0(t), b_1(t), b_2(t), b_3(t)$ задаются следующими формулами:

$$b_0(t) = \frac{2\beta_1^3}{\delta}(1-t)^3$$

$$b_1(t) = \frac{1}{\delta}(2\beta_1^3 t(t^2 - 3t + 3) + 2\beta_1^2(t^3 - 3t + 2) + 2\beta_1(t^3 - 3t + 2) + \beta_2(2t^3 - 3t^2 + 1))$$

$$b_2(t) = \frac{1}{\delta}(2\beta_1^2 t^2(-t + 3) + 2\beta_1 t(-t^2 + 3) + 2\beta_2 t^2(-2t + 3) + 2(-t^3 + 1))$$

$$b_3(t) = \frac{2t^3}{\delta}$$

где $\beta_1 > 0$ и $\beta_2 \geq 0$, $\delta = 2\beta_1^3 + 4\beta_1^2 + 4\beta_1 + \beta_2$ — числовые параметры β_1 и β_2 называются параметрами формы бета-сплайновой кривой, параметр β_1 называется параметром скоса (смещения), а параметр β_2 — параметром натяжения.

Свойства составной бета-сплайновой кривой:

- набор базовых функций однозначно определяет кривую, т.е. нет возможности регулировать ее форму;
- проходит внутри выпуклой оболочки, заданной опорными точками;
- является дважды геометрически непрерывной кривой;
- параметры β_1 и β_2 позволяют регулировать ее форму.

Составная бета-сплайновая кривая, как правило, не проходит ни через одну из базовых точек. Однако известно, что начальная точка кривой обязательно лежит в треугольнике, образованном тремя начальными базовыми точками, а конечная точка кривой — в треугольнике, образованном тремя конечными базовыми точками. Поэтому для построения составной бета-сплайновой кривой, начинающейся в первой базовой точке и заканчивающейся в последней базовой точке, достаточно дополнить набор двумя копиями первой точки и двумя копиями последней точки.

Построить замкнутую бета-сплайновую кривую можно, дополнив набор базовых точек и $3n$ штук точками: $P_{n+1}=P_0, P_{n+2}=P_1, P_{n+3}=P_2$



Рис. 8.31

РАЗДЕЛ 9. ФРАКТАЛЬНЫЕ КРИВЫЕ

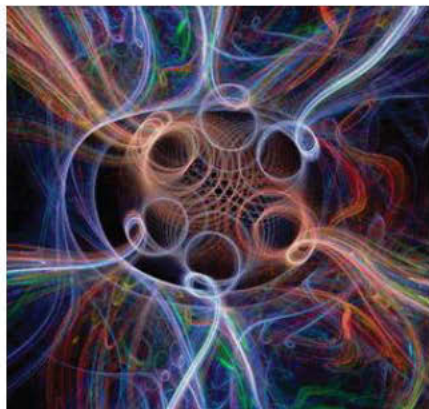
Математика, если на нее правильно посмотреть, отражает не только истину, но и несравненную красоту.

Бертран Рассел



*Сотворение Богом окружностей,
волн и фракталов*

*Бенуа Мальдерброт
«Фрактальная геометрия природы»*



Кажется, что все равнодушно к фракталам и действительно, многие рассматривают свою первую встречу с геометрией фракталов как с эстетической, так и с научной точки зрения.

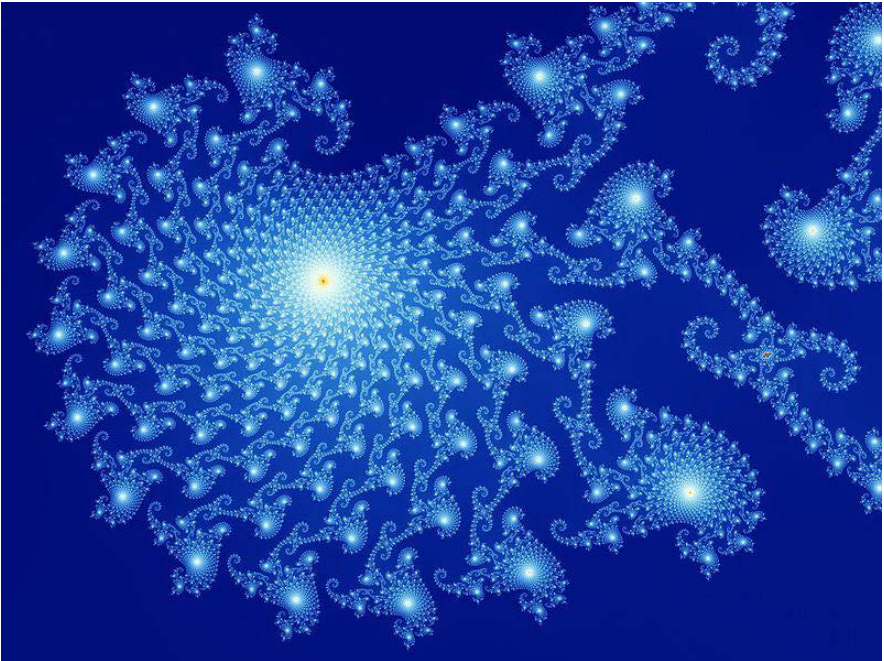
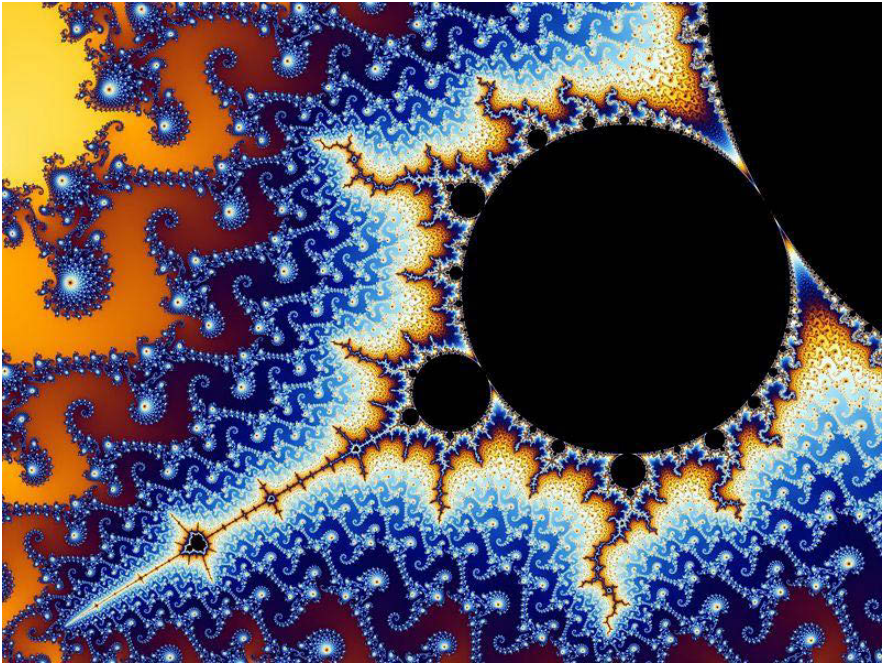
*Бенуа Мальдерброт
«Красота фракталов»*



*Paul Bourke, Горящий корабль.
Фрактал получен по модифициро-
ванной формуле множества Ман-
дельброта*

$$z_{n+1} = (|x_n| + i |y_n|)^2 - c = x_n^2 - y_n^2 + 2i |x_n y_n| - c$$





*Натуралисты знают — мир так плох,
Что блохи малые терзают своих блох,
А этих блохи меньшие грызут...
И нескончаем вечный этот зуд.*

Джонатан Свифт. Раисодия



Вся красота относительна... Мы не должны думать, что берега океана искажены и деформированы, потому что они не похожи на ровную стену; и мы не должны думать, что горы имеют неправильную форму, потому что они не являются правильными пирамидами или конусами; и мы не должны думать, что звезды неумело расположены на небе, раз они находятся на разном расстоянии от нас. Это не природные неточности — они кажутся такими только по нашему капризу.

Ричард Бентли, британский ученый XVIII века



БЕНУА МАНДЕЛЬБРОТ

Фрактальная геометрия принадлежит к таким немногим разделам математики, для которых можно указать как дату их появления, так и имя создателя. В 1975 г. вышла книга франко-американского математика Бенуа Мандельброта² "Фрактальная геометрия природы", послужившая толчком для возникновения новой науки. Работа Мандельброта продолжила исследования таких математиков как Пуанкаре, Жюлиа, Гильберт, Кантор, Хаусдорф. По определению самого Б.Мандельброта фрактальная геометрия располагается между теорией Евклида, исследующей "исключительно упорядоченные и гладкие фигуры (элементы кривых у Евклида всегда самоподобны, но тривиальным образом: все кривые являются локально прямыми, а прямая всегда самоподобна)" и фигурами произвольной сложности и неупорядоченности. "Сегодня эти фигуры заслуживают наименования "геометрически хаотичных" [Б.Мандельброт "Фракталы и возрождение теории итераций" в книге Пайтсен Х.-О., Рихтер П.Х. Красота фракталов. - М.: Мир, 1993].

Расшифруем приведенное высказывание. В прошлом, да и в основном и сейчас, основное внимание математики концентрируется на достаточно гладких функциях и множествах. С другой стороны негладкие функции (нерегулярные множества) обеспечивают лучшее представление многих природных явлений (турбулентность, протекание жидкости в порче, облака и многое другое).

Фрактальная геометрия связана с изучением именно нерегулярных множеств. Основной объект фрактальной геометрии – фрактал – находит применение в компьютерном дизайне, алгоритмах сжатия информации, радиосистемах и т.п. Фрактальность проявляется в искусстве и архитектуре, урбанистике, в т.н. математической истории, синергетических концепциях культуры, искусствоведении Европы, Африки, Индии. Говоря о фракталах, часто используются термины "компьютерное искусство", "художественный дизайн", "эстетический хаос". Главным результатом своих научных трудов Б. Мандельброт считал «возвращение глаголу «видеть» его исконного смысла, порядком подзабытого как в общепринятом употреблении, так и в лексике «твердой» (количественной) науки: видеть – значит, воспринимать *глазами*». Фрактальная оптика видения извлекает из старого хаоса форм новый порядок, создавая и новую образность, и новые правила смыслопорождения.

Бенуа Мандельброт (фр. *Benoit B. Mandelbrot*; 20 ноября 1924, Варшава — 14 октября 2010, Кембридж) — французский и американский математик, создатель фрактальной геометрии. Лауреат премии Вольфа по физике (1993).

Одни фрактальные множества представляют собой кривые или поверхности, другие – несвязную «пыль»; есть и такие, чья форма столь причудлива, что ни наука, ни искусство не в состоянии предложить подходящее для них название».

Ф. Манделъброт



9.1. Интуитивное понятие о фракталах

На приведенных выше фотографиях изображены примеры фрактальные образования в природе. Кривые, рассмотренные в приведенных примерах обладают весьма важным и не тривиальным свойством. Различные участки кривой могут быть совмещены друг с другом они самоподобны. Интуитивно это понятие означает, что эти формы выглядят «одинаково» при любом масштабировании: не имеет значение, насколько было увеличено изображение такой кривой – она имеет одну и ту же степень сложности. Одни кривые в точности самоподобны: какое бы сильное увеличение вы не делали, увеличенное изображение выглядит в точности как оригинал (за исключением поворота и сдвига). Другие кривые являются только статистически самоподобными: в этом



Рис. 9.1. Береговая линия

случае неправильности и изгибы кривой при любом увеличении картинки не изменяются только «в среднем». Наилучший пример береговая линия (рис. 9.1). При взгляде со спутника у нее имеется определенная степень изрезанности, образуемая заливами, бухтами, полуостровами. Если пролететь над ней ниже, увидим больше подробностей. Залив приобретает собственную неровность не видимую раньше. Этот процесс можно продолжать вплоть до разглядывания отдельных песчинок в микроскоп (а может быть и дальше?). Другие примеры: ветви дерева, поверхность губки. Определим длину береговой линии. Результат измерения зависит от длины звена измерения l . С уменьшением величины l длина береговой линии L не стремится к конечному пределу, а увеличивается по степенному закону $L \approx l \left(\frac{R}{l}\right)^D$

Фундаментальным предикатом описания категории фрактала является самоподобие. Самоподобие означает, что подсистемы нижних уровней фрактальной системы повторяют конфигурацию целой системы и в пределах общей формы заключен точно или с некоторыми изменениями «тиражируемый» (в предельном случае – бесконечно тиражируемый) паттерн. Иначе говоря, фрагмент фрактала, идентичный целостной форме, воспроизводится на каждом последующем уровне меньшего масштаба, образуя своего рода «вложенную» структуру. Природными фракталами являются, например, береговые линии, горы, деревья с их ветвистыми кронами и листьями, снежинки, кровеносная система человека и др.

Подобие может быть жестким (инвариантным), т.е. абсолютно точным рекурсивным воспроизведением паттерна (как в т.н. геометрических фракталах: снежинка Коха (рис. 9.2), треугольник Серпинского (рис. 9.3) и пр. или нежестким (ко-вариантным), т.е. относительным, когда элементы фрактала при увеличении масштаба рассмотрения не повторяют систему в целом, но происходит почти полное повторение базовой формы во все более и более уменьшенном виде последовательно через каждые несколько ступеней масштабного преобразования (например, знаменитый алгебраический фрактал – множество Мандельброта (рис. 9.4)). Наконец, случайные, или стохастические, фракталы, у которых на разных шагах итерации параметры меняются случайным образом, обладают статистическим подобием (например, Броуновское дерево).

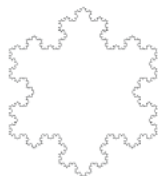


Рис. 9.2. Снежинка Коха



Рис. 9.3. Треугольник Серпинского

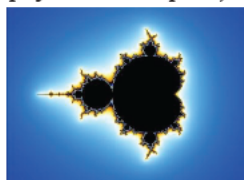


Рис. 9.4. Множество Мандельброта

При этом любой фрактал представляет собой визуализацию некоторого алгоритма, набора математических процедур, имеющих характер последовательных итераций. Общим для всех фрактальных структур является наличие рекурсивной процедуры их генерации, что означает бесконечную цепочку автопоэзиса, в которой каждый результат предыдущей итерации служит начальным значением нового цикла воспроизводства: $z_{n+1} = f(z_n)$.

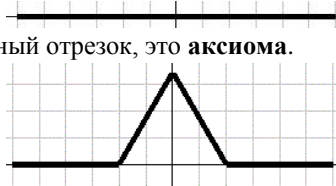
В итоге с помощью относительно несложных математических формул, включающих комплексные числа, «можно описать форму облака так же чётко и просто, как архитектор описывает здание с помощью чертежей, в которых применяется язык традиционной геометрии».

Рассмотрим ряд примеров построения фрактальных кривых

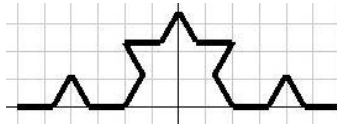


Кривая Коха.³ Эта кривая интересна тем, что она образует бесконечно длинную линию внутри области конечной площади. Кривая Коха является типичным детерминированным фракталом. Процесс ее построения выглядит следующим образом

Шаг 0 берем единичный отрезок, это **аксиома**.



Шаг 1 разделяем этот отрезок на три равные части и заменяем средний интервал равносторонним треугольником без этого сегмента. В результате образуется ломаная, состоящая из четырех звеньев длины $1/3$. Длина полученной кривой $4/3$ Это — **генератор**.



Шаг 2 повторяем операцию для каждого из четырех получившихся звеньев. Полученная ломаная состоит из $(4)^2$ звеньев длины $(1/3)^2$. Общая длина полученной линии равна $(4/3)^2$

На n -ом шаге получим ломаную, состоящую из $(4)^n$ звеньев длиной $(1/3)^n$ каждое, при этом длина всей линии будет $l=(4/3)^n$. Предельная кривая и будет

³Нильс Фабиан Хельге фон Кох (швед. *Niels Fabian Helge von Koch*, 25 января 1870 — 11 марта 1924) — шведский математик, фрактальная кривая, носящая его имя, описанная в статье «Об одной непрерывной кривой, не имеющей касательных...» в 1904 году

фракталом- кривой Коха при этом ее длина при $n \rightarrow \infty$ будет стремиться к бесконечности.

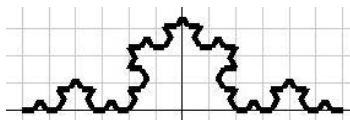


Рис. 9.5.Кривая Коха

На рис. 9.5 приведены четвертая итерации этой кривой Коха

Зададим произвольную ломаную с конечным числом звеньев, называемую генератором. Далее, заменим в ней каждый отрезок генератором (точнее, ломаной, подобной генератору, но уменьшенной в два раза). В получившейся ломаной вновь заменим каждый отрезок генератором, уменьшенным уже в четыре раза. Продолжая до бесконечности. На каждом шаге мы получаем предфрактал, а в пределе получим фрактальную кривую. Порождающий такие структуры механизм Б. Мандельброт назвал каскадом.

При каждом шаге длина кривой Коха $L(e)$ увеличивается на треть и при бесконечном числе шагов длина линии стремится к бесконечности. На первом шаге алгоритма длина отрезка составляет $e = 1/3$ от первоначальной. Тогда длина кривой Коха вычисляется просто:

$$L = 4 \cdot 1/3 = 4/3 = 1,33$$

На втором шаге алгоритма длина элементарного отрезка $e = 1/9$, соответственно длина кривой:

$$L = 16 \cdot 1/9 = 16/9 = 1,777$$

На третьем шаге алгоритма $e = 1/27$. И длина кривой:

$$L = 64 \cdot 1/27 = 64/27 = 2,370370$$

Процесс этот можно продолжить до бесконечности, заметив, что с увеличением числа шагов n длина элементарного отрезка e стремится к 0, а длина кривой L на каждом шаге изменяется (увеличивается) и в пределе стремится к бесконечности:

$$L = (4/3)^n$$

$$e = (1/3)^n$$

где $n = 1, 2, 3, \dots$ и из этих выражений получаем:

$$n = (1/\ln 3) \ln(1/e).$$

Подставляя n получим:

$$L = \exp[n \cdot \ln(4/3)] = \exp[(\ln(4/3)/\ln 3) \cdot \ln(1/e)]$$

Обозначив $D = \ln 4 / \ln 3$, получаем:

$$L(e) = e^{(1-D)}$$

Из последнего соотношения видно, что постоянным показателем при любом шаге остается только величина D , поскольку она не зависит от масштаба

измерения и является характеристикой линии "кривая Коха". Она называется фрактальной размерностью.

С геометрической точки зрения фрактальная размерность является показателем того, насколько плотно эта линия заполняет плоскость или пространство.

Б. Мандельброт обобщил и популяризировал исследования природы самоподобия. Он назвал различные формы самоподобных кривых фракталами. слово «фрактал» не является математическим термином и не имеет общепринятого строгого математического определения. оно может употребляться, когда рассматриваемая фигура обладает какими-либо из перечисленных ниже свойств:

- обладает нетривиальной структурой на всех масштабных шкалах в этом отличие от регулярных фигур (таких, как ОКРУЖНОСТЬ, ЭЛЛИПС, ГРАФИК ГЛАДКОЙ ФУНКЦИИ): если мы рассмотрим небольшой фрагмент регулярной фигуры в очень крупном масштабе, он будет похож на фрагмент прямой. для фрактала увеличение масштаба не ведёт к упрощению структуры, на всех шкалах мы увидим одинаково сложную картину.

- является самоподобной или приближённо самоподобной.

- обладает дробной метрической или фрактальной размерностью, превосходящей топологическую.

Рассмотрим более подробно понятия самоподобия и фрактальной размерности.

9.2. Самоподобие и фрактальная (дробная) размерность

Самоподобный объект (в математике) - объект, в точности или приближённо совпадающий с частью себя самого (то есть целое имеет ту же форму, что и одна или более частей). Многие объекты реального мира, например, береговые линии, обладают свойством статистического самоподобия: их части статистически однородны в разных шкалах измерения. Самоподобие есть характеристическое свойство фрактала.

Инвариантность относительно изменения масштабной шкалы является одной из форм самоподобия, при которой при любом приближении найдётся по крайней мере одна часть основной фигуры, подобная целой фигуре.

Самоподобие имеет важные приложения в построении компьютерных сетей, так как типичный сетевой поток обладает аналогичными свойствами. Например, в телефонии, потоки пакетных данных почти статистически самоподобны. Наличие данного свойства означает, что простые модели, использующие пуассоновское распределение, неточны, и сети, построенные без учёта самоподобия, могут функционировать в непредсказуемых режимах. Движение цен на фондовом рынке также демонстрирует самоподобие, так как представляется вполне обоснованным считать графики приближённо повторяющимися при изменении масштаба скважности.

Свойство точного самоподобия характерно только для регулярных, как правило построенных фракталов. Если в алгоритм построения включить элемент случайности, то в результате получим случайные фракталы. Их отличие состоит в том, что свойство самоподобия приблизительно выполняется только после усреднения по всем статистически независимым реализациям рассматриваемого объекта. При этом увеличенная часть фрактала не точно идентична исходному фрагменту, но их статистические характеристики совпадают

Математическое определение

Компактное топологическое пространство X самоподобно, если существует конечное множество S , индексирующее набор не сюръективных гооморфизмов $\{f_s\}_{s \in S}$, для которых $X = \cup_{s \in S} f_s(X)$. Если $x \subset y$, то X называется самоподобным, если оно является единственным непустым подмножеством Y , для которого вышеприведённое уравнение выполняется при заданном семействе $\{f_s\}_{s \in S}$. В таком случае $V = (X, S, \{f_s\}_{s \in S})$ называется самоподобной структурой.

Фрактальная (дробная) размерность

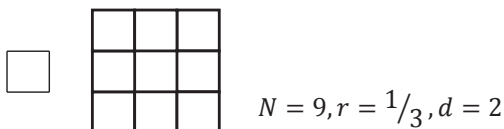
Семейство евклидовых размерностей 0, 1, 2, 3, 4, ... иногда называют топологическими размерностями, так как пространства с различными евклидовыми размерностями топологически различны, т. е. одно такое пространство нельзя перевести в другое непрерывной топологической деформацией. Точка имеет топологическую размерность 0. Гладкие, «хорошие» кривые - прямые, окружности, параболы и т. д. - имеют евклидову размерность 1. Поверхности имеют размерность 2, объёмные тела - размерность 3 и гипертела – более высокие размерности.

Чтобы понять, как вычисляются фрактальные размерности, рассмотрим сначала отрезок прямой (при достаточно большом увеличении малую дугу любой кривой мы будем видеть как прямую).

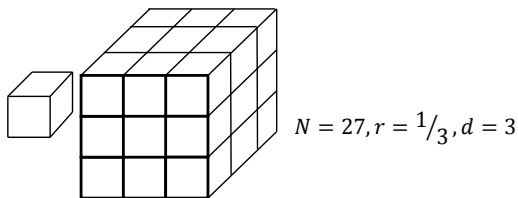
Разделим отрезок прямой на N равных частей. Тогда, каждую часть можно считать копией всего отрезка, уменьшенной в $1/r$ раз. Очевидно, N и r связаны соотношением $Nr = 1$.



Разобьем квадрат на N квадратов (с площадью в $1/r^2$ раз меньшей площади исходного) то соотношение запишется как $Nr^2 = 1$.



Разобьем куб на N равных кубов (с объемом в $1/r^3$ раз меньшим объема исходного), то соотношение примет вид $Nr^3 = 1$.



Заметим, что размерность d объекта, будь то одномерный отрезок, двумерный квадрат или трехмерный куб, появляется как степень r в соотношении между N , числом равных подобъектов, и коэффициентом подобия r . А именно:

$$Nr^d = 1 \quad (9.1)$$

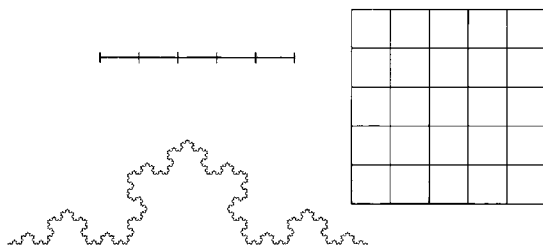
Рассмотренные выше множества обладают целой размерностью.

Множество, при разбиении которого на N непересекающихся подмножеств, полученных масштабированием оригинала с коэффициентом r называют самоподобным фракталом. Величину d называют фрактальной (дробной) размерностью или размерностью подобия.

Явное выражение для N и r находится логарифмированием (9.1)

$$d = \frac{\ln N}{\ln(1/r)} \quad (9.2)$$

Б. Мандельброт поясняет фрактальную размерность следующим образом:



«На рисунке показано, как, располагая некоторым целым числом (в данном случае $b = 5$), можно разбить прямолинейный отрезок единичной длины на $N = b$ подынтервалов, длина каждого из которых равна $r = 1/b$. Аналогичным образом мы можем разделить единичный квадрат на $N = b^2$ меньших квадратов с длиной стороны $r = 1/b$. И в том, и в другом случае величина $\ln N / \ln(1/r)$ представляет собой размерность подобия рассматриваемой фигуры, — величина, о которой школьная геометрия не считает нужным упоминать, так как ее значение сводится к евклидовой размерности. Нижняя фигура — это троичная кривая Коха или треть побережья острова Коха. Ее также можно разбить на подобные исходной кривой фигуры меньшего размера, при этом $N = 4$, а $r = 1/3$.

Размерность подобия $D = \ln N / \ln(1/r)$ в данном случае оказывается дробным числом (ее значение $\sim 1,2618$), не находя себе аналогов в стандартной геометрии. Хаусдорф показал, что величина D может быть весьма полезной в математике и что она совпадает с хаусдорфовой, или фрактальной, размерностью. Я же утверждаю, что без величины D не обойтись и в естественных науках.» («Фрактальная геометрия природы» с.75)



Понятие фрактальной (дробной) размерности появилось в 1919г. в работе Ф. Хаусдорфа. Размерность Хаусдорфа — естественный способ определить размерность множества в метрическом пространстве. Размерность Хаусдорфа согласуется с нашими обычными представлениями о размерности в тех случаях, когда эти обычные представления есть. Например, в трёхмерном евклидовом пространстве хаусдорфова размерность конечного множества равна нулю, размерность гладкой кривой — единице, размерность гладкой поверхности — двум и размерность множества ненулевого объёма — трём. Для фрактальных множеств размерность Хаусдорфа может принимать дробные значения. Определение размерности Хаусдорфа⁴ непросто и базируется на функциональном анализе, поэтому отметим только основные свойства размерности Хаусдорфа:

- с точностью до умножения на коэффициент: 1-мера Хаусдорфа для гладких кривых совпадает с их длиной;

- 2-мера Хаусдорфа для гладких поверхностей совпадает с их площадью;

- d -мера Хаусдорфа множеств R^d в совпадает с их d -мерным объёмом.

Размерность Хаусдорфа не более чем счётного объединения множеств равна максимуму из их размерностей. В частности, добавление счётного множества к любому множеству не меняет его размерности

Для самоподобных множеств размерность Хаусдорфа может быть вычислена явно. Неформально говоря, если множество разбивается на n частей, подобных исходному множеству с коэффициентами r_1^s, r_2^s, \dots то его размерность s является решением уравнения $r_1^s + r_2^s + \dots + r_n^s = 1$

Применим этот подход к снежинке фон Коха. Увеличивая линейные размеры какой-нибудь ее части в три раза, мы получаем четыре копии исходной снежинки. На каждом этапе построения длина штриховой линии ровно в $4/3$ раза превышает длину сплошной линии на предыдущем этапе, хотя каждый прямолинейный отрезок по длине составляет $1/3$ длины отрезка предыдущего этапа.

⁴Феликс Хаусдорф (нем. *Felix Hausdorff*; 8 ноября 1868, Бреслау — 26 января 1942, Бонн) — немецкий математик, один из основоположников современной топологии. Ввел и впервые исследовал важные топологические понятия хаусдорфова пространства (1914), топологического предела, частично упорядоченного множества, а также хаусдорфовой размерности (1919).

Следовательно, предельной кривой разумно сопоставить размерность Хаусдорфа D , или фрактальную размерность, равную

$$\ln 4 / \ln 3 = 1,261859$$

Размерность множества Кантора равна $\ln 2 / \ln 3$ (разбивается на две части, коэффициент подобия $1/3$), а размерность треугольника Серпинского — $\ln 3 / \ln 2$ (разбивается на 3 части, коэффициент подобия $1/2$).

Фрактальные кривые, одномерные в евклидовом смысле, могут иметь фрактальные размерности от 1 до 2, если кривая плоская, и больше, если кривая искривляется и скручивается в евклидовых пространствах более высоких размерностей. Если фрактальная кривая проходит через каждую точку области, заключенной внутри ее границы на плоскости, как извивающаяся змея (я имею в виду саму змею, а не ее границу) и другие кривые Пеано, то в пределе ее евклидова размерность и ее фрактальная размерность равны 2. Если фрактальная кривая проходит через каждую точку некоторой области пространства, то в пределе она достигает евклидовой и фрактальной размерности 3. Аналогично, фрактальная поверхность с евклидовой размерностью 2 может иметь фрактальную размерность, заключенную между 2 и 3, если кривая принадлежит трехмерному пространству, и выше, если кривая извивается в евклидовых пространствах более высокой размерности, и аналогичные утверждения справедливы относительно фрактальных структур более высоких топологических размерностей. Может ли сложность фракталов оцениваться числами, которые меньше единицы? Может, хотя структуры, о которых идет речь, топологически не эквивалентны отрезкам прямых. Например, возьмем отрезок прямой, удалим его центральную треть, затем удалим центральную треть каждого из двух оставшихся подотрезков и продолжим этот процесс выбрасывания средней трети до бесконечности. В результате мы получим то, что Мандельброт предпочитает называть канторовской пылью. В математической литературе такие множества принято называть дисконтинуумами Кантора, или канторовскими множествами, в честь Георга

Кантора, который впервые исследовал их. Применяя формулу Мандельброта для фрактальной размерности к канторовской пыли, получаем число, заключенное между 0 и 1. Точное значение фрактальной размерности канторовской пыли равно $D = \log 2 / \log 3 = 0,6309\dots$. Другие процедуры вырезания части отрезка приводят к другим фрактальным размерностям. Мандельброт называет их предфракталами, чтобы отличать от фракталов, размерность которых больше их евклидовой размерности.

В абстрактном мире чистой математики фрактальные структуры рассмотренного нами типа принято называть упорядоченными фракталами. Разумеется, в реальном мире никаких упорядоченных фракталов не существует. Береговые линии, деревья, реки, облака, молнии, траектории частиц в броуновском движении и тысячи других фракталоподобных явлений можно рассматривать как

несовершенные модели, которые в определенных пределах сверху и снизу являются фракталами в статистическом смысле. Самоподобие их проявляется в том, что они сохраняют статистическое подобие независимо от масштаба. Фрактальные размерности при различных масштабах подлежат усреднению, и, для того чтобы проводить такое усреднение, необходимо накопить достаточно обширный запас эмпирических данных. Такие фракталы называются случайными, или статистическими.

Пыль Кантора

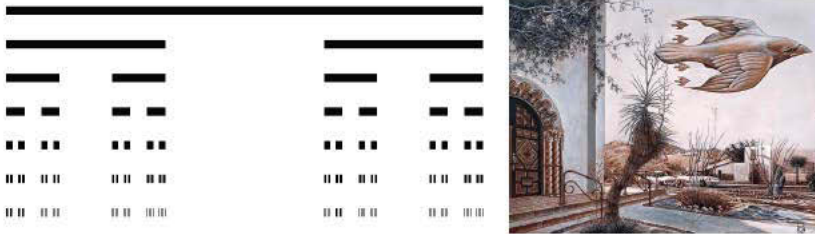


Рис. 9.6. Пыль Кантора

Способ построения этого множества следующий. Берется отрезок прямой единичной длины. Затем он делится на три равные части, и вынимается средний отрезок. Это первый шаг итерационной процедуры. На втором шаге подобной процедуре деления на три равные части и последующего удаления середины подвергается каждый из двух оставшихся отрезков. Так продолжая до бесконечности, получим множество Кантора. Нетрудно заметить, что суммарная длина получившихся в пределе отрезков равна нулю, так как мы исключили в результате длину, равную

$$1 = 1/3 + 2/9 + 4/27 + \dots = 1/3 \cdot (1 + 2/3 + 4/9 + \dots) = 1/3 \cdot 1 / (1 - 2/3) = 1.$$

Также нулю равна и топологическая размерность. Фрактальные свойства пыли Кантора имеют огромное значение, особенно учитывая тот факт, что многие известные фракталы являются близкими родственниками этого фрактала.

Треугольник (салфетка) Серпинского



Берется сплошной равносторонний треугольник, на первом шаге из центра удаляется перевернутый треугольник. На втором шаге удаляется три перевернутых треугольника из трех оставшихся треугольников. Продолжая этот процесс, на n -ом шаге удаляем 3^{n-1} перевернутых треугольников из центров

3^{n-1} оставшихся треугольников. По теории конца этому процессу не будет, и в треугольнике не останется живого места, но и на части он не распадется - получится объект состоящий из одних только дырок. Это и есть треугольник Серпинского. Этот фрактал описал в 1915 году польский математик Вацлав Серпинский⁵

«Точка кривой называется точкой ветвления, если граница сколь угодно малой ее окрестности содержит более чем две точки, принадлежащие той же кривой... Здравый смысл, судя по всему, настаивает на том, что никакая кривая просто не может состоять из одних лишь... точек ветвления. Это очевидное убеждение опровергается ... кривой Серпинского, все точки которой являются точками ветвления». У.Хан



«Я утверждаю, что (задолго до Коха, Пеано и Серпинского) в построенной Гюставом Эйфелем в Париже башне была осознанно воплощена идея фрактальной кривой, содержащей множество точек ветвления. В первом приближении Эйфелева башня состоит из четырех А-образных элементов. Согласно легенде, Эйфель выбрал букву А, чтобы выразить в своей башне слово Atour. Все четыре

А-образных элемента имеют общую вершину, а соседние А — общее ребро. Кроме того, на верхушке возвышается еще одна, прямая, башня. Обратите внимание, что и А-элементы, и верхняя башня сделаны не из цельных балок, а из колоссальных ферм. Фермой называется жестко скрепленная совокупность взаимосвязанных звеньев, каждое из которых не может быть деформировано без деформации, по крайней мере, одного из соседних звеньев.

При одинаковой прочности фермы оказываются значительно легче цельных цилиндрических балок. А Эйфель сообразил, что фермы, звенья которых сами являются фермами, еще легче. Бакминстер Фуллер открыл миру глаза на то, что секрет прочности скрыт в точках ветвления, однако умудренные опытом строители готических соборов знали об этом задолго до него. Чем дальше мы заходим в применении этого принципа, тем ближе подбираемся к идеалу Серпинского! Бывший ученик Безиковича Фримен Дайсон в поисках прочных и легких конструкций для своих межпланетных по-

⁵ **Вацлав Серпинский** *Sierpiński*; (14 марта 1882, Варшава, Польша — 21 октября 1969, Варшава) — выдающийся польский математик. континуум-гипотезе, теории чисел, теории функций, а также топологии. Автор 724 статей и 50 книг. Им предложен в 1915 г. один из двумерных аналогов множества Кантора, известный как салфетка (ковер) Серпинского.

строек описал однажды бесконечно экстраполированную Эйфелеву башню»
Б.Мандельброт

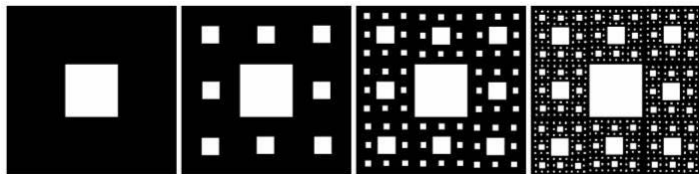


Александр **Гюстав Эйфель** (15.12.1832, Дижон — 28.12.1923, Париж) — французский инженер, специалист по проектированию металлических конструкций. Завоевал небывалую популярность после постройки в Париже к выставке 1889 года металлической башни, принадлежащей к замечательнейшим техническим сооружениям XIX века и названной в его честь. *«Почему такая странная форма? Ветровые нагрузки. Я считаю, что искривление четырех внешних краев монумента продиктовано и математическими расчётами, и эстетическими соображениями».*

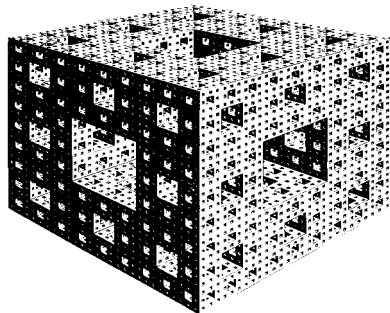
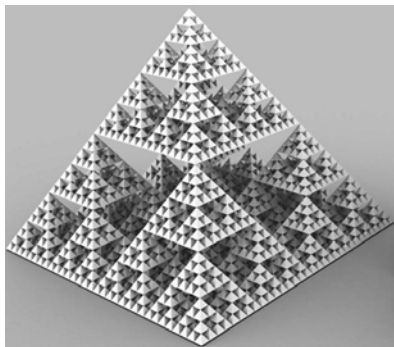
Эйфель

Ковер Серпинского и губка Менгера

Берется сплошной квадрат, на первом шаге из центра удаляется квадрат, уменьшенный в 9 раз.. На втором шаге удаляется, 8 квадратов, уменьшенных в 9^2 раз из оставшихся 8 квадратов. И т.д.



Обобщив плоские салфетки и ковер Серпинского, проведя построение в пространстве R^3 , получим пирамиду Серпинского и губку Менгера.



Кривые Гильберта и Серпинского

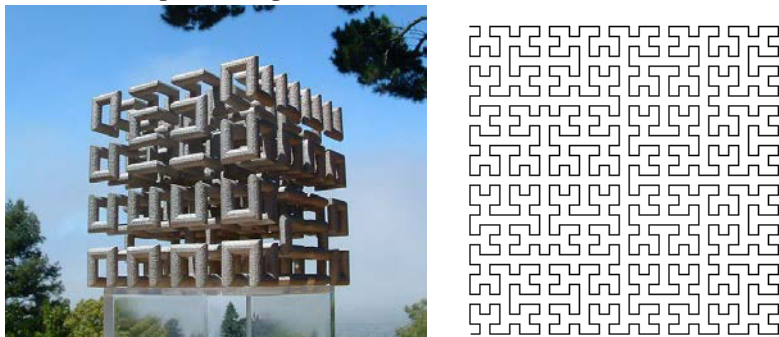


Рис. 9.7. Кривая Гильберта 5 порядка

Кривая Гильберта и кривая Серпинского связаны с любопытным понятием теории функций, а именно — всюду плотными кривыми. Кривая на плоскости называется всюду плотной в некоторой области, если она проходит через любую сколь угодно малую окрестность каждой точки этой области. Несколько упрощенно можно считать, что всюду плотные кривые целиком заполняют указанную область. Известные математики Гильберт и Серпинский построили примеры всюду плотных кривых. Хотя эти примеры различны, схема получения соответствующих кривых одинакова. По определенному правилу строятся кривые (соответственно Гильберта и Серпинского) первого, второго, ..., n -го порядка, вписанные в заданный квадрат. При неограниченном возрастании n они стремятся к некоторой предельной кривой, которая является всюду плотной в заданном квадрате.

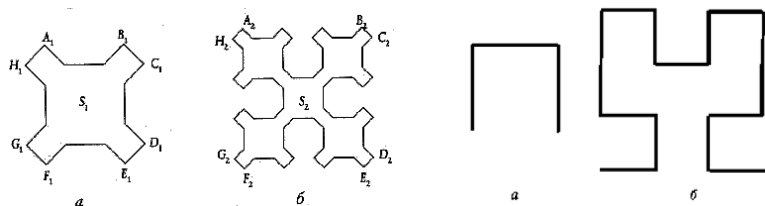


Рис. 9.8. Кривые Гильберта и Серпинского 1 и 2 порядка

Многие объекты в природе обладают фрактальными свойствами, например, побережья, облака, кроны деревьев, кровеносная система и система альвеол человека или животных.

Фракталы представляют собой математические модели сложных структур, пространственное изображение которых представляется в виде сломанных, морщинистых и нечетких форм. Фракталы (как математические абстракции)

обладают следующими характерными свойствами, отображающими их иррегулярную сущность:

- самоподобие (иерархический принцип организации);
- способность к развитию (принцип непрерывности формообразования);
- дробная метрическая размерность (принцип сингулярности меры);
- размытость, нечеткость контуров (принцип неопределенности границ);
- геометрическое представление хаотической динамики (принцип динамического хаоса).

Фракталы, особенно на плоскости, популярны благодаря сочетанию красоты с простотой построения при помощи компьютера.

Было предложено несколько определений фрактала:

■ Фрактал (лат. *fractus* — дробленный, сломанный, разбитый) — сложная геометрическая фигура, обладающая свойством самоподобия, то есть составленная из нескольких частей, каждая из которых подобна всей фигуре целиком. В более широком смысле под фракталами понимают множества точек в евклидовом пространстве, имеющие дробную метрическую размерность (в смысле Минковского или Хаусдорфа), либо метрическую размерность, строго большую топологической.

■ Фрактал — это бесконечно самоподобная геометрическая фигура, каждый фрагмент которой повторяется при уменьшении масштаба.

■ Фрактал — самоподобное множество нецелой размерности.

Первые примеры самоподобных множеств с необычными свойствами появились в XIX веке (например, множество Кантора).

Фракталы обычно делят на

■ конструктивные — множества получающиеся в результате линейных (аффинных) сжимающих отображений (преобразований) подобия. Результирующее сжимающее отображение обладает устойчивой неподвижной «точкой»_фракталом.

■ динамические — множества возникающие в нелинейных (в первую очередь дискретных) динамических системах. Такие множества могут обладать масштабной инвариантностью лишь приближенно.

9.3 Конструктивные (классические) фракталы

Предельного обобщения мы достигнем, выбирая какую-нибудь замкнутую кривую, допускающую разбиение на конгруэнтные сегменты, изменяя каждый сегмент каким угодно образом так, чтобы весь процесс можно было повторить на меньших сегментах последующих поколений и перейти к пределу. Аналогичные построения можно производить и на поверхностях пространственных тел. Разумеется, в результате таких операций могут получиться сложные кривые или поверхности с самопересечениями.

Гарднер

Очень сложные кривые можно получить рекурсивно (ситуация, когда преобразователь прямо или косвенно создает собственную копию, в теории алгоритмов называется рекурсией) посредством многократного «усложнения» простой кривой. Для описания и построения подобных кривых используются L -системы.

L – системы

Понятие L -системы тесно связанное с самоподобными фракталами.

L -системы использовались при изучении формальных языков и в биологических моделях селекции. Они нашли широкое применение в компьютерной графике для построения фрактальных деревьев, растений и различных фракталов.

Для графической реализации L -систем в качестве подсистемы ввода используется так называемая черепашня графика. Точка (черепашка) движется по экрану дискретными шагами, прочерчивая свой след (но может перемещаться и без рисования). Состояние черепашки описывается тремя параметрами (x, y, α) , где (x, y) - координаты черепашки, α - направление, в котором она смотрит (движется). Черепашка интерпретирует и выполняет последовательность команд, задаваемых кодовым словом, буквы которого читаются слева направо. Кодовое слово представляет собой результат работы L -системы и включает следующие буквы:

F - переместиться вперед на один шаг, прорисовывая след.

b - переместиться вперед на один шаг, не прорисовывая след (позволяет организовать разрыв в графике).

$+$ - увеличить угол α на величину θ (повернуть на θ по ходу ч.с.)

$-$ - уменьшить угол α на величину θ (повернуть на θ против хода ч.с.)

[- открыть ветвь

] - закрыть ветвь

X, Y - при генерировании инструкции для пера игнорируются.

Размер шага и величина приращения по углу θ задаются заранее и остаются неизменными для всех перемещений черепашки. Если начальное направление движения α (угол, отсчитываемый от положительного направления оси X) не указано, но полагают $\alpha = 0$.

Тогда строка " $F - F + +F - F$ " при значении $\theta = 60^\circ$ задает перу команду нарисовать первую генерацию K_1 кривой Коха.

Процесс превращения простой строки в более сложную основывается на наборе инструкций создания строк, который включается в подпрограмму *producestring()*. Правило для кривой Коха имеет вид:

$$'F' \rightarrow 'F - F + +F - F'' ,$$

где знак \rightarrow означает, что каждый встреченный символ F заменяется группой символов " $F - F + +F - F$ ". Для символов $' + '$ или $' - '$ инструкция отсутствует, поэтому эти символы передаются без изменений.

Формально, детерминированная L - система состоит из алфавита, слова инициации называемого аксиомой (инициатором, атомом) и набора порождающих правил, указывающих, как следует преобразовывать слово при переходе от уровня к уровню (от итерации к итерации).

Например, для снежинки Коха порождающее правило

$$newf = F - F + +F - F;$$

Аксиома: $F + +F + +F$ - равносторонний треугольник; $\theta = \frac{\pi}{3}$.

На первой стадии исходная строка называется атомом, в данном случае это " F ", «генерирует строку первого поколения

$$S1 = 'F - F + +F - F',$$

затем эта строка подается на вход того же самого процесса, который генерирует строку второго поколения

$$S2 = \underline{F - F + +F - F} - \underline{F - F + +F - F} + \underline{+F - F + +F - F} - \underline{F - F + +F - F}$$

В этой строке можно выделить четыре группы (кластера) вида

$$F - F + +F - F$$

разделенных символами $' - '$, затем $' + +'$ и наконец, снова $' - '$. Теперь перо нарисует вторую генерацию K_2 кривой Коха.

Для снежинки Коха: $\theta = \frac{\pi}{3}$; Аксиома: $F + +F + +F$ - равносторонний треугольник; правило: $newf = F - F + +F - F$.

Приведем алгоритм работы L - системы в псевдокодах.

Алгоритм L - системы.

Назначение: реализует правило $F = newf; b = newb/$

Вход:

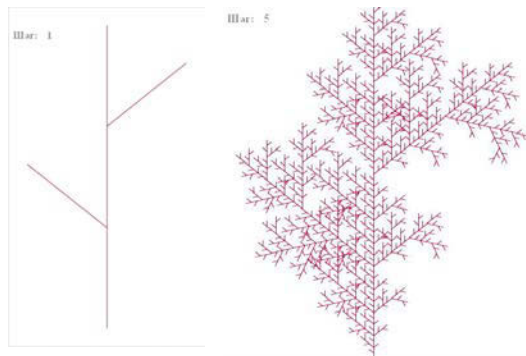
axiom - (слово инициации $F + +F + +F$)

newf - (порождающее правило) $newf = F - F + +F - F$

newb - (порождающее правило)

level - (число итераций, задается).

Выход:



word - (слово результат)

Инициация:

$W = axiom$

$n = length(W)$ (количество букв в слове W)

$T = \{\}$ - (пустое множество)

Шаги:

while $level > 0$

for $j = 1$ *to* n

if $W(j) = +, T = \{T +\},$ *end if*

if $W(j) = -, T = \{T -\},$ *end if*

if $W(j) = [, T = \{T [\},$ *end if*

if $W(j) =], T = \{T]\},$ *end if*

if $W(j) = F, T = \{T newf\},$ *end if*

if $W(j) = b, T = \{T newb\},$ *end if*

end for $W = T$

$level = level - 1$

end while

$word = W.$

Замечание: $W(j)$ – j -я буква в слове,

$\{T +\}$ - строка T , к которой присоединен знак +.

Ветвление

Для того, чтобы запомнить желаемое состояние, нужно сделать его копию и записать ее в удобном месте. Чтобы осуществить это можно работать со стеком. Если в строке встречается символ "[", то текущее состояние черепахи заталкивается в стек для дальнейшего использования. Если встречается символ "]", то из стека выталкивается верхнее значение, и состояние черепахи устанавливается равным вытолкнутому значению.

{ в *OpenGL* к оператору *switch* подпрограммы *produce String* () надо добавить две строки:

"[" : *save Turtle* () : *break* : - затолкнуть в стек текущее состояние

"]" : *restoreTurtle* () : *break* : - вытолкнуть из стека состояние }

«Куст». Атом " F ", $F \rightarrow F[-F]F[+F][F]$; $\theta = 51^\circ$.

На рисунке представлен первый и пятый шаг.

Конструкции генераторов (аксиом) фрактальных форм

1) Кривая Коха(треугольная)



$$D = \ln 4 / \ln 3 = 1.2618\dots$$

$$N=4, b=1/3$$

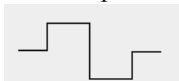
2) Видоизменённая кривая Коха (прямоугольная)



Единичный сегмент имеет три отрезка длиной $1/3$. два отрезка длиной $1/4$. Из соотношения $3(1/3)^D + 2(1/4)^D = 1$

$$D = 1.34.$$

3) Видоизменённая кривая Коха (меандр)



$$D = \ln 8 / \ln 4 = 3/2$$

$$N=8, b=1/4$$

4) Канторовская пыль



$$D = \ln 2 / \ln 3 = 0,6309\dots$$

$$N=2, b=1/3$$

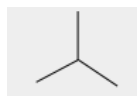
5) Дерево



$$D = \ln 5 / \ln 3 = 1,465\dots$$

$$N=5, b=1/3$$

6) Снежинка



$$D = \ln 3 / \ln(3)^{1/2} = 2$$

$$N=3, b=1/(3)^{1/2}$$

7) Прокладка Серпинского



$$D = \ln 3 / \ln 2 = 1,585\dots$$

$$N=3, b=1/2$$

Примерами таких кривых служат:



■ **кривая дракона;**

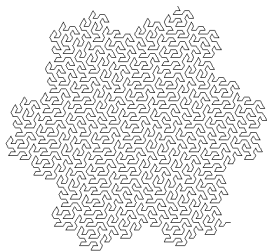
Аксиома: FX

Правила:

$X \rightarrow X+YF+$

$Y \rightarrow -FX-Y$

$\theta = \frac{\pi}{2}$



■ кривая Госпера

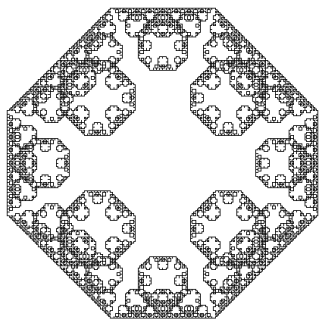
Аксиома: XF

Правила:

$X \rightarrow X+YF++YF-FX--FXFX-YF$

$Y \rightarrow -FX+YFYF++YF+FX--FX-Y$

$$\theta = \frac{\pi}{3}$$



Угол поворота: $\pi/6$

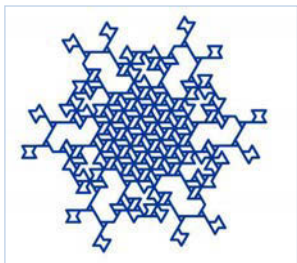
■ кривая Леви

Аксиома: F++F+

Правило: $F \rightarrow -F++F- \theta = \frac{\pi}{4}$

Аксиома: F++F++F++F

Правило: $F \rightarrow F-F++FF$



Один из фракталов случайного типа может начинаться с треугольника, лежащего в произвольной плоскости. Средние точки сторон треугольника соединены между собой, так что треугольник оказывается разделённым на четыре меньших треугольника. Затем каждая средняя точка сдвигается вверх или вниз на определённую, случайно выбираемую величину. Тот же процесс применяется к каждому из меньших треугольников, затем к ещё меньшим и так далее до бесконечности. После достаточно большого количества итераций начинает возникать всё более детализированная поверхность. В этом методе смещения средних точек случайные величины для перемещения средних точек вверх или вниз управляются определённым законом распределения, который тщательно подбирается, чтобы получить близкую аппроксимацию желаемой поверхности. Для того чтобы поверхность была относительно гладкой, в преобразованиях следует встроить правило, согласно которому величина смещения средних точек должна становиться очень малой уже после нескольких первых итераций. Такое правило позволяет

добавлять лишь небольшие «кочки» к общим очертаниям ландшафта. Для представления изрезанной поверхности, характерной, скажем, для горного хребта или береговой линии, более подходящим будет правило медленного уменьшения смещений после каждого шага итерационного процесса. У данного метода построения поверхностей существует много приложений. Он применялся, в частности, в качестве модели эрозии почвы, для анализа сейсмических явлений, чтобы лучше понять характер изменений в зоне разломов. Р. Восс, один из коллег Мандельброта по Исследовательскому центру корпорации ИВМ, воспользовался идеей метода, чтобы строить изображения планет, спутников, облаков и горных хребтов, которые выглядят весьма реалистично (рис. 9.9).

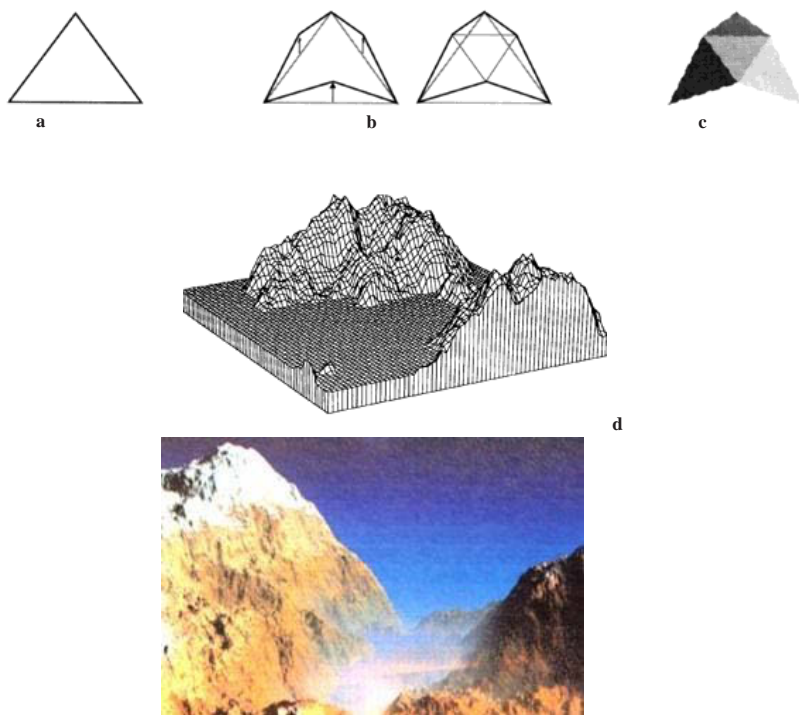


Рис. 9.9. Фрактальные ландшафты могут создаваться из фракталов методом случайного смещения средней точки. Средние точки сторон треугольника (а) смещаются вверх или вниз от плоскости изображения и соединяются с вершинами (b). При этом возникает четыре меньших треугольника, к которым повторно применяется та же процедура. Функция распределения вероятности определяет величину смещения и, следовательно, степень гладкости фрактального ландшафта. Затем графическая программа компьютера закрашивает треугольники, создавая различные оттенки (с). В результате получается весьма реалистичная картина (d).

Получение фрактальных изображений с помощью системы итерируемых функций.



Рассмотрим иной подход к получению фрактальных изображений на примере построения салфетки Серпинского. В качестве начального изображения I_0 возьмем равносторонний треугольник.

Применим к нему преобразование масштабирования с $s=0.5$. При этом матрица преобразования в однородных координатах имеет вид:

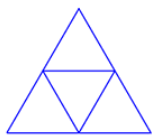


$$M_1 = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Применим к нему еще одно преобразование состоящее из масштабирования с $s=0.5$ и переноса по оси OX на 0.5 единиц. Матрица этого преобразования имеет вид:

$$M_2 = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Применим к нему еще одно преобразование состоящее из масштабирования с $s=0.5$, переноса по оси OX на 0.25 единиц и по оси OY на 0. Матрица этого преобразования имеет вид:



$$M_3 = \begin{bmatrix} 0.5 & 0 & 0.25 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{bmatrix}$$

Собрав полученные копии вместе получим новое изображение I_1 , т.е.

$$I_1 = T(M_1) \cup T(M_2) \cup T(M_3).$$

Чтобы получить изображение I_2 , применим к каждому из треугольников I_1 все три преобразования, а затем соберем все 9 копий. Повторив этот процесс бесконечно, получим последовательность изображений I_0, I_1, I_2, \dots которая называется *орбитой*. Барнсли назвал этот набор аффинных преобразований «системой итерированных функций».

Определение. Системой итерированных функций называется набор из N аффинных преобразований T_i , где $i = 1, 2, \dots, N$.

Иногда полезно давать имя каждому числу, появляющемуся в СИФ Назовем k -е число (изображение) d_k и предположим, что процесс начинается с $k = 0$ «введением» в систему начального значения. Тогда последовательность значе-

ний, создаваемых с помощью СИФ, имеет вид:

$$d_1 = f(d_0),$$

$$d_2 = f(f(d_0)),$$

$$d_3 = f(f(f(d_0)))$$

Так как d_3 формируется путем трехкратного применения функции $f(\cdot)$, и называется третьей итерацией функции $f(\cdot)$, примененной к начальному значению.

Обозначим k -ю итерацию $f(\cdot)$ как $\mathbf{d}_k = \mathbf{f}^{[k]}(\mathbf{d}_0)$

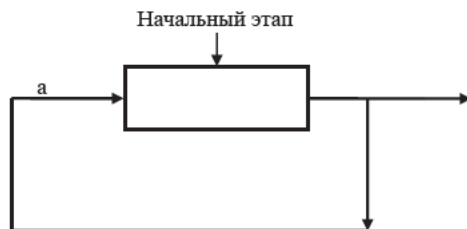


Рис. 9.10

На рис. 9.10 схематически представлен этот итерационный процесс.

Можно использовать рекуррентную форму и записать $\mathbf{d}_k = f(\mathbf{d}_{k-1})$ для $k = 1, 2, 3, \dots$, для заданной величины \mathbf{d}_0 . Последовательность величин $\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \dots$, называется «орбитой \mathbf{d}_0 » для данной системы.

В общем виде матрица аффинных преобразований имеет вид:

$$T = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

Ее удобно записывать в виде кортежа из шести элементов

$$T = \{m_{11}, m_{12}, m_{21}, m_{22}, m_{13}, m_{23}\}$$

В качестве начального изображения I_0 может быть использовано любое множество.

Так в рассмотренных примерах это ломаная линия для кривой Коха или треугольник, как для салфетки Серпинского. При этом можно наблюдать как изображение уменьшается при последующих итерациях. Однако при большом количестве итераций кривая все более и более уменьшается и изображение становится точечным. Поэтому в качестве начального изображения I_0 можно использовать изолированную точку. При этом каждая итерация будет представлять собой множество точек, которые нужно сохранять их в списке. Тогда если СИФ состоит из N сжимающих аффинных преобразований то I_1 состоит из N точек, I_2 - из N^2 , I_3 - из N^3 точек и т.д.

9.3.1. Сжимающие аффинные преобразования

Определение. Отображение $f: X \rightarrow X$ в метрическом пространстве (X, d) называется *сжимающим*, если существует такое число $\epsilon \in]0, 1[$, что для любых x_1 и x_2 из выполнено $d(f(x_1), f(x_2)) \leq \epsilon d(x_1, x_2)$. Константа ϵ называется *коэффициентом сжатия* отображения f . Преобразования вида $f(f(\dots f(x) \dots)) = f^n$ называются итерациями преобразования f . Точка a такая, что $f(a) = a$, называется неподвижной точкой отображения f .

Рассмотрим аффинное преобразование:

$$T = \begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0 \quad (9.3)$$

Тип неподвижной точки $0(0,0)$ определяют корни характеристического уравнения

$$\mu^2 - (a + d)\mu + \Delta = 0 \text{ где } \Delta = \det A \text{ или } \Delta = \det(A - \mu E) = 0$$

- при $|\mu_1| < 1, |\mu_2| < 1$ точка 0 – устойчивая;
- при $|\mu_1| < 1, |\mu_2| > 1$ – седловая (гиперболическая);
- при $|\mu_{1,2}| = 1$ – эллиптическая;
- при $|\mu_1| > 1, |\mu_2| > 1$ – неустойчивая.

Теорема. Существует линейное невырожденное преобразование

$(x, y) \rightarrow (\tilde{x}, \tilde{y})$ приводящее (1) при $\Delta = 1, |\mu_1| = |\mu_2| = 1$ к виду

$$\begin{cases} x' = x \cos \alpha - y \sin \alpha \\ y' = x \sin \alpha + y \cos \alpha \end{cases} \quad (9.4)$$

где α – параметр (угол поворота).

Поворот. Рассмотрим преобразование поворота вокруг точки $0(0,0)$ на угол α (против хода часовой стрелки) без растяжения (сжатия) имеет вид:

$$\begin{cases} x' = x \cos \alpha - y \sin \alpha \\ y' = x \sin \alpha + y \cos \alpha \end{cases} \quad (9.5)$$

$\Delta = 1$

Преобразование поворота вокруг произвольной точки (x_0, y_0) имеет вид

$$\begin{cases} x' = (x - x_0) \cos \alpha - (y - y_0) \sin \alpha + x_0 \\ y' = (x - x_0) \sin \alpha + (y - y_0) \cos \alpha + y_0 \end{cases} \quad (9.6)$$

Пример

Сжатие (растяжение).

Сжатие (расширение) характеризуется центром и показателем сжатия “с”:

$$C = \begin{cases} x' = c(x - x_0) + x_0 \\ y' = c(y - y_0) + y_0 \end{cases}$$

Рассмотрим аффинное преобразование T : $T(x) = x' = Ax + b$.

Оно будет сжимающим в \mathbb{R}^n , если для любых x_1 и x_2 выполнено, то $|T(x_1) - T(x_2)| < |x_1 - x_2|$. Откуда, подставляя выражение для T , получаем $|(Ax_1 + b) - (Ax_2 + b)| < |x_1 - x_2| \rightarrow |Ax_1 - Ax_2| < |x_1 - x_2| \rightarrow |A(x_1 - x_2)| < |x_1 - x_2|$, откуда, заменяя $x_1 - x_2$ на x , получаем $|Ax| < |x|$, для любых x .

Так как $|A(kx)| = |k||Ax|$ и $|kx| = |k||x|$, то мы можем ограничиться рассмотрением векторов x , имеющих единичную длину (единичная сфера S^n). Таким образом, матрица A должна удовлетворять следующему условию: $\max_{|x|=1} |Ax| < 1$. Заметим также, что

$$|Ax| = \sqrt{(Ax, Ax)} = \sqrt{(Ax)^t Ax} = \sqrt{x^t A^t Ax}.$$

Поэтому $\max_{|x|=1} |Ax| = \max_{|x|=1} \sqrt{x^t A^t Ax} = \sqrt{\max_{|x|=1} x^t A^t Ax}$.

Рассмотрим функцию $F(x) = x^t A^t Ax$. Для отыскания максимума функции $F(x)$ на единичной сфере воспользуемся методом Лагранжа. (Максимум существует, так как S^n есть компакт в \mathbb{R}^n , а функция $F(x)$ непрерывна на S^n . Функция Лагранжа для $F(x)$ примет вид:

$$L(x, \lambda) = F(x) - \lambda(x^t x - 1) = x^t A^t Ax - \lambda(x^t x - 1).$$

В точке максимума дифференциал функции должен равняться 0 при любом приращении аргумента.

$$\begin{aligned} \Delta L(x) &= L(x + \Delta x, \lambda) - L(x, \lambda) = \\ &((x + \Delta x)^t A^t A(x + \Delta x) - \lambda((x + \Delta x)^t(x + \Delta x) - 1)) - (x^t A^t Ax - \lambda(x^t x - 1)) = \\ &x^t A^t A \Delta x + \Delta x^t A^t Ax - \lambda(x^t \Delta x + \Delta x^t x) + o(|\Delta x|) = \\ &2x^t A^t A \Delta x - 2\lambda x^t \Delta x + o(|\Delta x|) = 2x^t (A^t A - E) \Delta x + o(|\Delta x|). \end{aligned}$$

Откуда получаем:

$$\begin{aligned} (A^t A - E)x &= 0, \\ |x| &= 1. \end{aligned}$$

Это означает, что x собственный вектор матрицы $A^t A$, имеющий единичную длину, λ собственное значение матрицы $A^t A$. Следовательно функция $F(x)$ достигает своего наибольшего значения на одном из собственных векторов матрицы $A^t A$, и ее максимальное значение равно собственному числу, которому соответствует этот вектор. Таким образом получаем, что

$$\max_{|x|=1} |Ax| = \sqrt{\max_i \lambda_i}.$$

Поэтому аффинное преобразование является сжимающим, когда корень из максимального собственного числа матрицы $A^t A$ меньше единицы

Примеры фракталов, построенных с помощью системы итерированных функций.

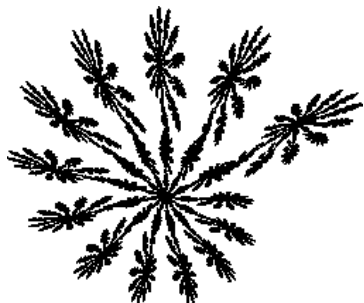
В этом случае рациональнее использовать простой рекурсивный метод создания изображения аттрактора известный как «Алгоритм случайных итераций». Рассмотрим его на примере рисования салфетки Серпинского. Точка, обозначенная буквой P, преобразуется в точку, лежащую посередине между самой этой точкой и одной из трех фиксированных точек $p[0], p[1], p[2]$, выбранных случайно. Затем рисуется эта новая точка, после чего процесс повторяется (наперед заданное число раз).

Выбирается случайным образом.

В рассматриваемом случае матрицу аффинных преобразований удобно записывать в виде кортежа из семи элементов $T_i = \{m_{11}, m_{12}, m_{21}, m_{22}, m_{13}, m_{23}, p[i]\}$, где $p[i]$ - вероятность применения i -го преобразования.

Dragon

{0.824074 0.281482 -0.212346 0.864198
-1.882290 -0.110607 0.787473};
{0.088272 0.520988 -0.463889 -0.377778 0.785360 8.095795 0.212527}.



Swirl

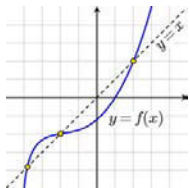
{0.745455 -0.459091 0.406061
0.887121 1.460279 0.691072 0.912675};
{-0.424242 -0.065152 -0.175758 -
0.218182 3.809567 6.741476 0.087325}



Dendrit

{0.5 0.5 0.5 -0.5 0.0 0.0 0.5};
{0.6 -0.2 -0.2 -0.6 0.4 0.2 0.5}.

Отображение с тремя неподвижными точками



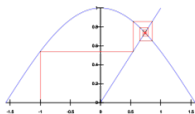
В математике, неподвижная точка отображения — точка, которую отображение переводит в неё же, иными словами, решение уравнения $f(x) = x$.

К примеру, отображение $f(x) = x^2 - 3x + 3$ имеет неподвижные точки $x = 1$ и $x = 3$, поскольку $f(1) = 1$ и $f(3) = 3$.

Неподвижные точки есть не у всякого отображения — скажем, отображение $f(x) = x + 1$ вещественной прямой в себя неподвижных точек не имеет.

Точки, возвращающиеся в себя после определённого числа итераций, то есть, решения уравнения $f(f(\dots f(x) \dots)) = x$, называются периодическими (в частности, неподвижные точки — это периодические точки периода 1).

Притягивающие неподвижные точки



Нахождение решения уравнения $x = \cos x$

Неподвижная точка $x = f(x)$ отображения f — притягивающая, если итерации любой начальной точки y , достаточно близкой к x , будут стремиться к x :

$$f(f(\dots f(x) \dots)) \rightarrow x, n \rightarrow \infty$$

(При этом, обычно, требуют, чтобы итерации y не покидали некоторой большей окрестности точки x — то есть, чтобы точка x была асимптотически устойчива.) В частности, достаточным условием, чтобы точка была притягивающей, является условие на производную: $|f'(x)| < 1$.

9.3.2. Метод Ньютона

Одним из применений идеи притягивающей неподвижной точки является метод Ньютона: искомое решение оказывается притягивающей неподвижной точкой построенного отображения, и потому может быть найдено как предел (очень быстро сходящейся) последовательности итераций.

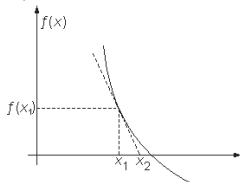
Наиболее известное применение этого метода — нахождение квадратного корня из числа $a > 0$ как последовательности итераций отображения

$$f(x) = \left(x + \frac{a}{x}\right) / 2$$

Бассейны Ньютона (проблема Кэли)

Области с фрактальными границами появляются при приближенном нахождении корней нелинейного уравнения $f(z) = 0$ алгоритмом Ньютона на комплексной плоскости. Для функции действительной переменной метод Ньютона часто называют методом касательных. Поясним суть этого метода.

Пусть нам задана функция $f(x)$, для которой известно приближенное значение ее корня x_1 , а также значение функции в этой точке $f(x_1)$ и значение ее первой производной $f'(x_1)$. Тогда, проводя касательную к графику функции $f(x)$ в этой точке и определяя ее пересечение с осью x , мы получим уточненное положение корня, равное x_2 (см. рис.).



Т.к. уравнение касательной к графику функции $f(x)$ в точке x_1 выглядит следующим образом

$$y = f'(x_1)(x - x_1) + f(x_1), \quad (9.7)$$

то, приравнявая y нулю, получаем, что уточненное значение корня x_2 связано с предыдущим значением x_1 соотношением

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \quad (9.8)$$

Беря теперь значение x_2 в качестве приближенного и повторяя этот алгоритм, находим следующее значение x_3 и т.д. При этом последовательность приближений быстро сходится к истинному значению корня (число верных десятичных знаков удваивается на каждом шаге).

Этот метод также эффективен и для комплексных чисел. Т.е. стартовав в непосредственной близости от значения корня уравнения $f(z) = 0$, и, используя итерационный алгоритм

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} \quad (9.9)$$

мы получаем последовательность комплексных чисел, быстро сходящуюся к этому корню. При этом возникает вполне закономерный вопрос, а что будет, если начальная точка z_0 в плоскости комплексных чисел выбрана не вблизи от корня, а произвольным образом?

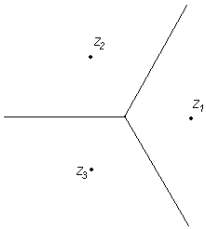
Примерно такой вопрос в 1977 году был задан молодому американскому математику Джону Хаббарду (John Hubbard) его студентами-первокурсниками, когда он преподавал им математику в Парижском университете Орсе́й. Хаббард довольно быстро доказал, что для уравнения второй степени данная последовательность всегда будет сходиться к ближайшему корню, за исключением случаев, когда начальная точка z_0 равноудалена от обоих корней. В этом случае последовательность итераций все время остается на этой прямой, совершая хаотическое движение.

Однако уже для уравнения 3-й степени такой простой ответ найти не удалось. В этой неудаче он был неодинок. Оказывается, еще в 1879 году эту проблему безуспешно пытался решить и Артур Кэли (Arthur Cayley). Им также был найден ответ для уравнения 2-й степени, а случай многочленов более высокой степени был анонсирован как будущая публикация, которой не суждено было появиться. Но в отличие от Кэли у Хаббарда было существенное преимущество, т.к. в его распоряжении был компьютер. Уже к концу семестра им и его студентами было получено несколько важных экспериментальных результатов.

Они начали с простейшего уравнения третьей степени $z^3 - 1 = 0$. В этом случае имеем три кубических корня из 1, а именно, $\omega_1 = 1$, $\omega_2 = -1/2 + i\sqrt{3}/2$ и $\omega_3 = -1/2 - i\sqrt{3}/2$. Нанесенные на комплексную плоскость, они образуют равносторонний треугольник, центр которого совпадает с началом координат. Можно легко показать, что эти корни действительно являются притягивающими неподвижными точками отображения (3) с $f(z) = z^3 - 1$

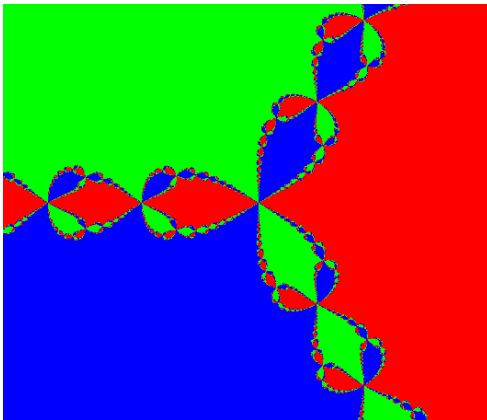
$$z_{n+1} = z_n - \frac{z_n^3 - 1}{3z_n^2} \quad (9.10)$$

Так какова же геометрия областей притяжения корней кубического уравнения? По аналогии с уравнением второй степени можно предположить, что ответ выглядит примерно так, как изображено на рисунке. Именно такую картину первоначально представлял себе Хаббард (и многие другие до него). Однако такое предположение является неверным. Более скрупулезное компьютерное исследование выявило, что геометрия границ областей притяжения имеет го-



раздо более сложную форму.

Так, если раскрашивать разные области притяжения разным цветом, то мы получим картинку, изображенную на рисунке (см. ниже). Из этого рисунка видно, что граница областей притяжения состоит из сильно переплетенных самоподобных структур. Оказывается, что на границе между любыми двумя цветами всегда расположена гирлянда островков третьего цвета. Границы этих островков, в свою очередь, состоят из гирлянд островков меньшего размера соответствующего дополнительного цвета и т.д. (см. рис.) Непрерывно уменьшаясь в размерах, детали границы постоянно воспроизводят самих себя. В результате оказывается, что каждая точка такой фрактальной границы соседствует сразу с тремя областями притяжения. Еще один тип динамических фракталов составляют фракталы (так называемые бассейны) Ньютона. Формулы для их построения основаны на методе решения нелинейных уравнений, который был придуман великим математиком еще в XVII веке. Применяя формулу метода Ньютона $z_{n+1} = z_n - f(z_n)/f'(z_n)$, $n = 0, 1, 2, \dots$ для решения уравнения $f(z) = 0$ к многочлену $z^k - a$, получим последовательность точек: $z_{n+1} = ((k-1)z_n^k - a)/kz_n^{k-1}$, $n = 0, 1, 2, \dots$. Выбирая в качестве начальных приближений различные комплексные числа z_0 , будем получать последовательности, которые сходятся к корням этого многочлена. Поскольку



корней у него ровно k , то вся плоскость разбивается на k частей — областей притяжения корней. Границы этих частей имеют фрактальную структуру. (Заметим в скобках, что если в последней формуле подставить $k = 2$, а в качестве начального приближения взять $z_0 = a$, то получится формула, которую реально используют для вычисления квадратного корня из a в компьютерах.) Наш фрактал получается из многочлена $f(z) = z^3 - 1$.

Построим теперь фрактал Ньютона на экране компьютера. Для этого рассмотрим отображение

$$z_{n+1} = z_n - \frac{z_n^3 - 1}{3z_n^2} = \frac{2z_n^3 + 1}{3z_n^2}$$

Полагая $z = x + yi$ и разделяя вещественную и мнимую части, приходим к двумерному вещественному отображению:

$$x_{n+1} = \frac{2}{3}x_n + \frac{x_n^2 - y_n^2}{3(x_n^2 + y_n^2)^2}$$
$$y_{n+1} = \frac{2}{3}y_n \left(1 - \frac{x_n}{(x_n^2 + y_n^2)^2}\right)$$

9.3.3 Последовательность Морса-Туэ



АКСЕЛЬ ТУЭ

Числа Фибоначчи и треугольник Паскаля далеко не исчерпывают самоподобных способностей целых чисел. Рассмотрим последовательность неотрицательных целых чисел

0, 1, 2, 3, 4, 5, 6, 7, ...

записанных в двоичной системе счисления,

0, 1, 10, 11, 100, 101, 110, 111, ...

и извлечем из каждого двоичного числа "цифровой корень" (т.е. найдем сумму его цифр по модулю 2). Мы получим последовательность

0, 1, 1, 0, 1, 0, 0, 1, ...

которая называется последовательностью Морса-Туэ в честь норвежского математика Акселя Туэ (1863-1922), который представил ее миру в 1906 году в качестве примера аперiodической рекурсивно вычислимой последовательности символов, и Марстона Морса из Пристона (1892-1977), обнаружившего, что эта последовательность играет важную роль в символической динамике при описании поведения некоторых нелинейных динамических систем в фазовом пространстве.



МАРСТОН МОРС

Интересно, что последовательность Морса-Туэ можно также получить, итерируя отображение $0 \rightarrow 01$ и $1 \rightarrow 10$, т.е. отображение, при котором вслед за каждой двоичной цифрой записывается ее дополнение. Начав с единственного нуля, мы получим следующее:

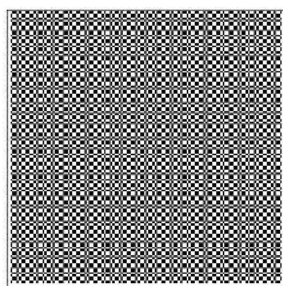


Рис. 9.11

0
01
0110
01101001

и т.д. Построенная таким образом последовательность называется самопорождающей последовательностью.

Следующий способ построения этой последовательности состоит в том, что каждое следующее поколение получается из предыдущего путем приписывания справа его дополнения. Этот способ – это лишь следствие того, что отображение $0 \rightarrow 01$, $1 \rightarrow 10$ непосредственно порождает отображение $01 \rightarrow 0110$, $10 \rightarrow 1001$ и т.д., где каждое отображение более высокого порядка следует исходному правилу "повтори двоичное число и припиши к нему справа его дополнение". Иначе говоря, исходное правило, порождающее отображение, наследуется всеми последующими поколениями. Наследование такого рода яв-

ляется важным следствием итерированных отображений и часто приводит к самоподобным структурам.

Последовательность Морса-Туэ самоподобна: сохраняя лишь нечетные члены этой бесконечной последовательности (подчеркнуты), мы вновь получаем последовательность Мора-Туэ:

$$\underline{0} \ 1 \ \underline{1} \ 0 \ \underline{1} \ 0 \ \underline{0} \ 0 \ \dots$$

Аналогично, сохраняя только нечетные пары двоичных чисел, мы также воспроизводим последовательность Морса-Туэ:

$$\underline{0} \ 1 \ 1 \ 0 \ \underline{1} \ 0 \ 0 \ 1 \ \dots$$

Последовательность Морса-Туэ воспроизводится и после замены каждой пары, четверки, восьмерки и т.д. на их самую левую цифру. Тот факт, что собственное подмножество может быть эквивалентно всему множеству, представляет собой хорошо известное свойство бесконечных множеств.

На рис. 9.11 изображена двумерная последовательность Морса-Туэ. Чёрным цветом обозначена единица, белым — ноль. Понять самоподобие последовательности Морса-Туэ очень легко. Сохранение только нечетных членов бесконечной последовательности эквивалентно умножению исходных чисел на 2 (основание двоичной системы счисления, в которой записаны члены последовательности). Так как в двоичной системе счисления умножение на 2 означает сдвиг цифр влево на один знак, то цифровые корни остаются неизменными, точнее, они и образуют по определению последовательность Морса-Туэ. (Если сохранять только четные члены, то последовательность Морса по аналогичным причинам перейдет в свое дополнение.)

Аналогичным свойством обладает и последовательность, построенная на троичной системе счисления. Цифровые корни (т.е. суммы цифр по модулю 3) целых неотрицательных чисел, записанных в троичной системе счисления,

$$0, 1, 2, 10, 11, 12, 20, 21, \dots$$

образуют самоподобную последовательность с коэффициентом подобия 3:

$$0, 1, 2, 1, 2, 0, 2, 0, \dots$$

Еще одним интересным свойством последовательности Морса-Туэ является ее аperiodичность. Хотя последовательность Морса-Туэ и аperiodична, но она не случайна. Более того, структура последовательности обладает ярко выраженным ближним и дальним порядком. Например, в этой последовательности не может быть групп, состоящих более чем из двух одинаковых членов подряд. И ее члены с индексами (если начинать с 0), различающимися на множитель 2^n , совпадают.

В физике последовательность Морса-Туэ была впервые обнаружена в символической динамике некоторых нелинейных динамических систем. Морс доказал, что траектории динамических систем, фазовые пространства которых имеют всюду отрицательную кривизну, полностью характеризуются дискретной последовательностью нулей и единиц. Это означает, что какая-нибудь

сложная кривая в \mathbb{R}^n , т.е. несчетно бесконечное множество в многомерном пространстве, может быть отображено на дискретную двоичную последовательность. Пользуясь последовательностью, Морс доказал возможность существования бесконечно долгих шахматных партий (при определенных ограничениях).

9.3.4 Множество Жюлиа и Мандельброта

Компьютерная графика предоставляет мощный инструмент для изучения множеств, которые считаются одними из самых сложных объектов в математике. В 1918 г. Г. Жюлиа опубликовал работу по теории камфорных преобразований (отображений, которые «оставляют углы» неизменными).

Рассмотрим аналитическое преобразование

$$z_{n+1} = z_n^2 + c, \quad c = \text{const}$$

которое после разделения вещественной и мнимой частей ($z = x + iy, c = a + bi$) запишется в виде

$$\begin{aligned} x_{n+1} &= x_n^2 - y_n^2 + a \\ y_{n+1} &= 2x_n y_n + b \end{aligned}$$

$$a \in [-2,5; 1,5]; \quad b \in [-2,2].$$

Множества Жюлиа квадратичных отображений и множество Мандельброта появляются из последовательностей комплексных чисел, определяемых по индукции с помощью соотношения $z_{n+1} = z_n^2 + c$.

Система генерирует каждый выход посредством возведения входа s в квадрат и добавляем константы c . Данный процесс начинается со стартовой величины s тогда система генерирует следующую последовательность значений или орбиту:

$$\begin{aligned} d_1 &= (s)^2 + c \\ d_2 &= (d_1)^2 + c = ((s)^2 + c)^2 + c \\ d_3 &= (d_2)^2 + c = (((s)^2 + c)^2 + c)^2 + c \\ &\dots \dots \\ d_n &= (d_{n-1})^2 + c. \end{aligned}$$

Очевидно, что эта орбита зависит от двух параметров: стартовой точки s и заданного значения константы c . Неподвижные точки - такие комплексные числа z , которые отображаются сами в себя:

$$\begin{aligned} z^2 + c &= z \quad \mathbf{P} \\ z^2 + c - z &= 0 \quad \mathbf{P} \\ P_{1,2} &= 0.5 \pm \sqrt{0.25 - c} \end{aligned}$$

Если орбита «пролегает» вблизи от неподвижной точки P , то следующая точка вдоль той орбиты будет вынуждена находиться

- ближе к точке P , если P - неподвижная точка притяжения (аттрактор),
- дальше от точки P , если P - неподвижная точка отталкивания (репеллер).

Любая точка z комплексной плоскости имеет свой характер поведения (остается конечной, стремится к бесконечности, принимает фиксированные значения) при итерациях функции $f(z)$, а вся плоскость делится на части. При этом точки, лежащие на границах этих частей, обладают таким свойством: при сколь угодно малом смещении характер их поведения резко меняется (такие точки называют **точками бифуркации**). При этом множества точек, имеющих один конкретный тип поведения, а также множества бифуркационных точек часто имеют фрактальные свойства. Это и есть множества Жюлиа для функции $f(z)$.

Нулевая орбита $0, c, c^2 + c, (c^2 + c)^2 + c, ((c^2 + c)^2 + c)^2 + c, \dots$

$c \in M$ (c принадлежит M), если 0 орбита притягивается,

$c \notin M$, если 0 орбита взрывается.

Определение: множество Мандельброта — это множество всех комплексных чисел c , которым соответствуют конечные 0 орбиты.

Плотным множеством Жюлиа K_C в точке c называется множество всех стартовых точек, орбиты которых конечны.

При $c = 0$ при старте из произвольной точки орбита s, s^2, \dots, s^n всасывается в 0 при $|s| < 1$ и взрывается при $|s| > 1$.

При $c = -2$ K_C состоит из всех точек, лежащих на оси между -2 и 2

K_C - связное множество, состоит из одного «куска»,

K_C - множество Кантора (пыль).

Вопрос в том уходит или нет орбита в бесконечность.

Пусть κ_{max} - максимальное число итераций (выбирается заранее). Если орбита за κ_{max} не покидает круга $x^2 + y^2 < 4$, точка (a, b) принадлежит множеству Мандельброта (обозначается черным цветом).

Если $\kappa < \kappa_{max}$ то точка уходит в бесконечность (цвет точки (a, b) выбирается, например по формуле $\kappa \bmod 256$ (если используется 256 цветов).

$$\{ |f_c^n(z)| < 4; n = 20, \text{лучше } 50, 100 \text{ и более} \}$$

Существует всего два типа множества Жюлиа. Каждое множество Жюлиа функции $f_c(z) = z^2 + c$ либо связно, либо вполне несвязно.

[Множество A связно, если его нельзя представить в виде объединения двух непустых непересекающихся множеств B и C , причем $B \cap \bar{C} = \emptyset$ и $C \cap \bar{B} = \emptyset$.

Здесь \bar{B}, \bar{C} дополнения множеств B и C соответственно, (т.е. каждое из множеств B и C не содержит предельных точек другого. Точка P предельная точка точечного множества S , если каждая окрестность точки P содержит точки множества S , отличные от P т.е. P - граничная точка. Множество A вполне несвязно (вполне разрывно) если все компоненты - одиночные точки, например множество Кантора].

В случае $0 < c < 0.25$ связные множества Жюлиа выглядят как простые замкнутые кривые, которые являются фракталами. Все вполне связные множества Жюлиа представляют собой «канторову» пыль.

Поведение этой последовательности чисел зависит от параметра c и начальной точки z_0 в поле комплексных чисел. Если зафиксировать c и варьировать z_0 получим множество Жюлиа, а если зафиксировать $z_0 = 0$ и изменять параметр c , то получим множество Мандельброта.

Каждая точка в множестве Мандельброта представляет значение c , для которого множество Жюлиа $S(f_c)$ связно. Каждая точка из дополнения к множеству Мандельброта представляет значение c для которого $S(f_c)$ вполне несвязно. Пусть задана пара значений s и c . как поведут себя точки d_k вдоль орбиты по мере того, как k становится все больше и больше? Остается ли эта орбита конечной (все точки на орбите остаются на конечном расстоянии от 0) или она взорвется (умчится в бесконечность)?

Те орбиты, которые останутся конечными, располагаются внутри соответствующего множества Жюлиа или Мандельброта, взрывающиеся орбиты лежат вне этого множества. Пример:

1) Пусть $c = 0$. Тогда

при $s = 1$ - орбита равна $1, 1, 1, \dots$ - является конечной.

при $s = -1$ - орбита равна $1, -1, 1, -1, \dots$

при $s = 0.5$ - орбита равна $0.5, 0.25, 0.0625, \dots$, сходится к 0.

при $|s| < 1$ - орбита конечна; при $|s| > 1$ - орбита взрывается.

2) Пусть $c = 0$. Тогда

при $c = 1$ - орбита равна $0, 1, 2, 5, 26, \dots$ - происходит взрыв.

при $c = -1$ - орбита равна $0, -1, 0, -1, \dots$ - конечна

при $c = -1, 3$ - орбита равна $0, -1, 3$ и примерно после 50 шагов закидывается в периодическую последовательность из 4 значений до бесконечности.

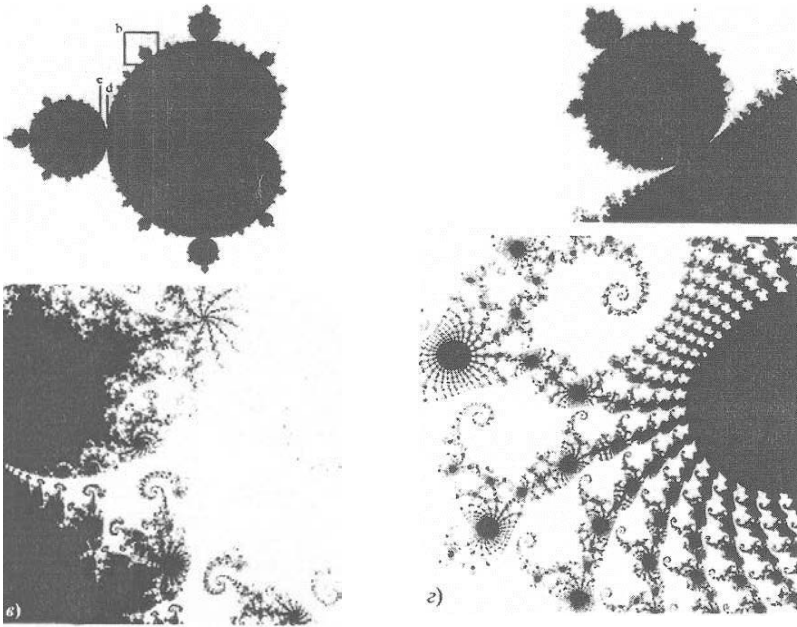


Рис 9.12. а) Множество Мандельброта; б), в) и г) увеличенные (в 11,6; 168,9 и 1042 раза) изображения некоторых участков, позволяющие увидеть чрезвычайно сложную структуру множества. Количество шагов итерации: 300 а), 300 б), 200 в), 200 г).

Каждая точка на рис. 9.12 затенена или раскрашена в соответствии с результатом экспериментального прогона в системе *IFS*, в которой использована следующая функция $\mathbf{f}(\mathbf{z}) = z^2 + c$, где c - некоторая константа, $z = a + ib$ - комплексная переменная. Вещественная часть z равна a , а мнимая - вещественному числу b . Модуль $|z|$ определяется как евклидова длина $\text{ра}[\mathbf{a}, \mathbf{b}]^T$, т.е. $|z| = \sqrt{a^2 + b^2}$.

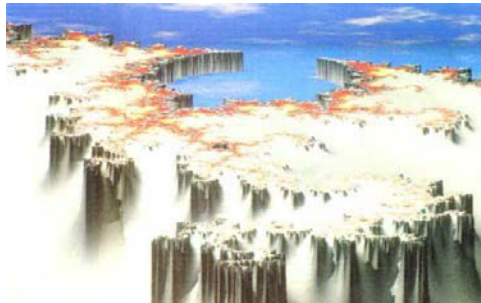


Рис. 9.13 Трёхмерное представление множества Мандельброта. Используется для изучения этой сложнейшей и интереснейшей фрактальной структуры. На рисунке показан электрический потенциал, окружающий заряженное множество Мандельброта

Рассмотрим функции, представляющие собой полиномы одного комплексного переменного. Пусть

$$f(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0, \quad a_n \neq 0$$

полином степени $n > 2$, коэффициенты a_0, a_1, \dots, a_n - комплексные числа (в частном случае, вещественные).

Множество Жюлиа $J(f)$ функции f , определяется как

$$J(f) = \partial\{z: f^{(n)} \rightarrow \infty \text{ при } n \rightarrow \infty\}$$

т.е. множество Жюлиа функции f есть граница множества точек z , стремящихся к бесконечности при генерировании $f(z)$.

Простейшее множество Жюлиа соответствует случаю $f(z) = z^2$. Т.к. $f^{(n)}(z) = z^{2^n}$ то $f^{(n)} \rightarrow \infty$ при $n \rightarrow \infty$ тогда и только тогда, когда $|z| > 1$. Границей этого множества (т.е. множеством Жюлиа), является единичная окружность $\partial\{z: |z| = 1\}$, которая фракталом не является, хотя в общем случае множество Жюлиа - фрактал.

Функция $f(z) = z^2$ хаотична на своем множестве Жюлиа (на единичной окружности).

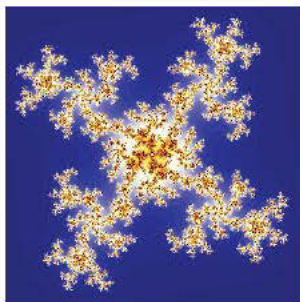
Рассмотрим множество Жюлиа квадратичной функции

$$f_c(z) = z^2 + c$$

где c - константа.

Теорема. Предположим, что $|c| < 2$. Пусть $z_0 = c$ и пусть $z_n = f_c^{(n)}(z)$ для $n = 1, 2, 3,$

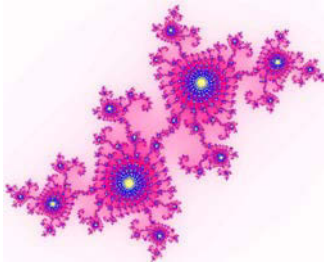
Если существует такой n_0 , что $|z_{n_0}| \geq 2$, то имеет место $\lim_{n \rightarrow \infty} z_n = \infty$, то есть орбита $\{f_c^n(z)\}_{n=1}^{\infty}$ стремится к бесконечности и z не принадлежит множеству Жюлиа $J(f_c)$.



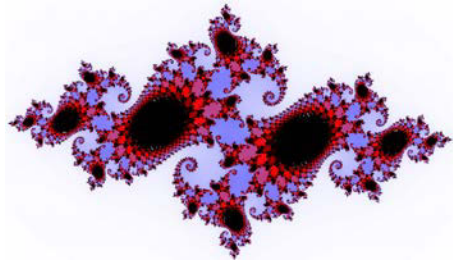
$f_c(z) = z^2 + c; c = 0.6 + 0.551i$



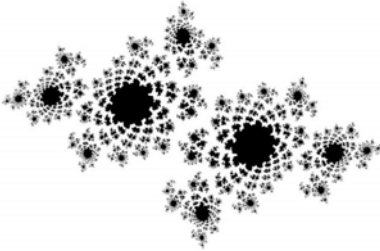
$f_c(z) = z^2 + c; c = -0.8 + 0.167i$



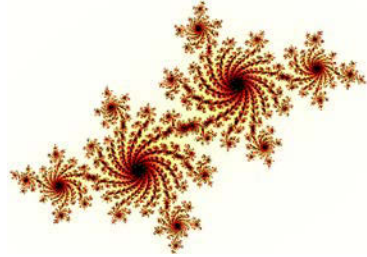
$$z = z^2 + c \quad C = -0,74524 + 0,10412i$$



$$Z = Z^2 + c \quad C = -0,21044 - 0,657i$$



$$Z = Z^2 + c \quad C = -0,41929 - 0,58579i$$



$$C = -0,7 + 0,299i$$

9.3.5 Фрактальное сжатие

Фрактальное сжатие изображений — это алгоритм сжатия изображений потерями, основанный на применении систем итерируемых функций (IFS, как правило являющимися аффинными преобразованиями) к изображениям. Данный алгоритм известен тем, что в некоторых случаях позволяет получить очень высокие коэффициенты сжатия (лучшие примеры — до 1000 раз при приемлемом визуальном качестве) для реальных фотографий природных объектов, что недоступно для других алгоритмов сжатия изображений в принципе. Из-за сложной ситуации с патентованием широкого распространения алгоритм не получил. Основа метода фрактального кодирования — это обнаружение самоподобных участков в изображении. Впервые возможность применения теории систем итерируемых функций (IFS) к проблеме сжатия изображения была исследована Майклом Барнсли и Аланом Слоуном. Они запатентовали свою идею в 1990 и 1991 гг (U.S. Patent 5,065,447). А. Жакен представил метод фрактального кодирования, в котором используются системы доменных и ранговых блоков изображения (*domain and range subimage blocks*), блоков квадратной формы, покрывающих все изображение. Этот подход стал основой для большинства методов фрактального кодирования, применяемых сегодня. Он был усовершенствован Ювалом Фишером и рядом других исследователей.

В соответствии с данным методом изображение разбивается на множество неперекрывающихся ранговых подизображений (*range subimages*) и опре-

деляется множество перекрывающихся доменных подизображений (англ. *domain subimages*). Для каждого рангового блока алгоритм кодирования находит наиболее подходящий доменный блок и аффинное преобразование, которое переводит этот доменный блок в данный ранговый блок. Структура изображения отображается в систему ранговых блоков, доменных блоков и преобразований.

Идея заключается в следующем: предположим что исходное изображение является неподвижной точкой некоего сжимающего отображения. Тогда можно вместо самого изображения запомнить каким-либо образом это отображение, а для восстановления достаточно многократно применить это отображение к любому стартовому изображению.

По теореме Банаха, такие итерации всегда приводят к неподвижной точке, то есть к исходному изображению. На практике вся трудность заключается в отыскании по изображению наиболее подходящего сжимающего отображения и в компактном его хранении. Как правило, алгоритмы поиска отображения (то есть алгоритмы сжатия) в значительной степени переборные и требуют больших вычислительных затрат. В то же время, алгоритмы восстановления достаточно эффективны и быстры.

Вкратце метод, предложенный Барнсли, можно описать следующим образом. Изображение кодируется несколькими простыми преобразованиями (в нашем случае аффинными), то есть определяется коэффициентами этих преобразований (в нашем случае A, B, C, D, E, F). Например, изображение кривой Коха можно закодировать четырьмя аффинными преобразованиями, мы однозначно определим его с помощью всего 24-х коэффициентов. Далее, поставив чёрную точку в любой точке картинки мы будем применять наши преобразования в случайном порядке некоторое (достаточно большое) число раз (этот метод ещё называют фрактальным пинг-понгом).

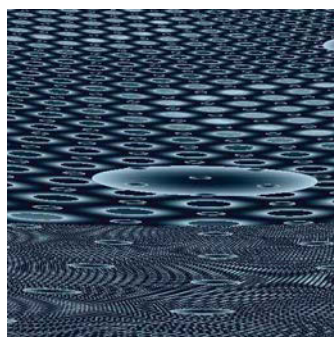
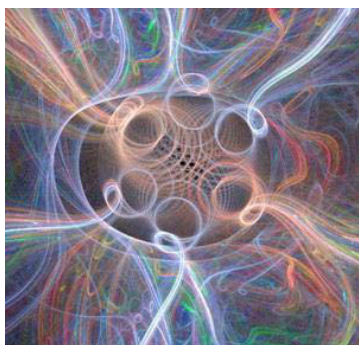
В результате точка обязательно перейдёт куда-то внутрь чёрной области на исходном изображении. Прделав такую операцию много раз мы заполним все чёрное пространство, тем самым восстановив картинку.

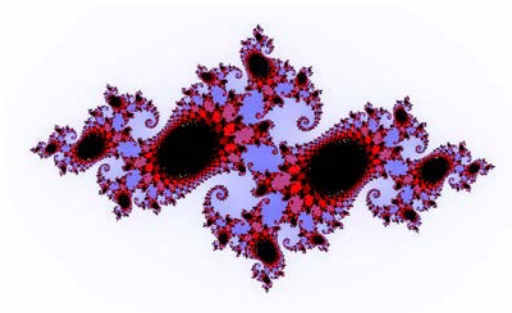
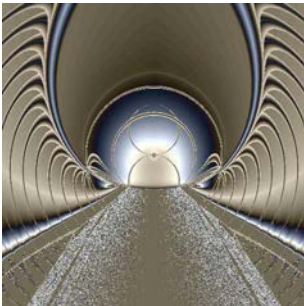
Основная сложность фрактального сжатия заключается в том, что для нахождения соответствующих доменных блоков вообще говоря требуется полный перебор. Поскольку при этом переборе каждый раз должны сравниваться два массива, данная операция получается достаточно длительной. Сравнительно простым преобразованием её можно свести к операции скалярного произведения двух массивов, однако даже скалярное произведение вычисляется сравнительно длительное время.

На данный момент известно достаточно большое количество алгоритмов оптимизации перебора, возникающего при фрактальном сжатии. Наиболее эффективными оказались два направления исследований: метод выделения осо-

бенностей (feature extraction) и метод классификации доменов (classification of domains).

9.4 Галерея фракталов

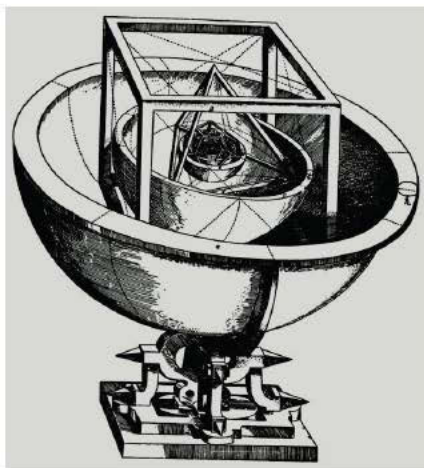






ФОМЕНКО А.Т. ПОЛИЭДР

В рисунке использованы разнообразные топологические объекты: полиэдры, особые точки поверхностей и т. п.



Модель Кеплера солнечной системы из 5 платоновых тел

Из всех языков, которыми мы обладаем, математика – единственный, не отягощенный предубеждениями, порожденными содержанием. Лишенная содержания и независимая от личного опыта, математика является единственным «космополитическим» языком человека. Она дает возможность соединять теории с очень различным содержанием, но близкие по логическим структурам. Она, таким образом, в высшей степени подходит для того, чтобы быть языком всей науки.

А. Панопорт (1911-?)

10.1. Способы задания поверхностей

В теории геометрического моделирования поверхность определяется как совокупность последовательного положения различных прямых и кривых линий, расположенных в пространстве по определенному закону. Поверхность считается заданной, если относительно любой точки пространства можно однозначно и сколь угодно точно решить вопрос о ее принадлежности данной поверхности.

Существуют три основных способа задания поверхностей:

- Аналитический
- Кинематический
- Задание поверхности каркасом.

Аналитический способ задания поверхности

Определение. Поверхностью называется непрерывное двупараметрическое множество точек.

Аналитическое задание поверхности может быть представлено

- В неявном виде: $F(x, y, z) = 0$
- В явном виде: $z = f(x, y)$,
- В функциях от независимого переменного (векторно-численном виде)

$$\vec{F}(u, v) = [x(u, v), y(u, v), z(u, v)]$$

- В параметрическом виде $x = f_1(t), y = f_2(t), z = f_3(t)$.

В зависимости от вида функции $F(x, y, z) = 0$ поверхности делятся на алгебраические и трансцендентные.

Наибольшая степень многочлена относительно x, y, z указывает на порядок алгебраической поверхности.

Недостатки алгебраического задания поверхности:

1. В понятии не заложены принципы конструирования поверхности.
2. Чрезвычайно трудно получить уравнение поверхности по первоначально заданному дискретному каркасу.
3. Уравнение поверхности громоздко даже для невысокого порядка.

Геометрическое определение поверхности: множество точек, которые могут быть склеены (состыкованы) из конечного (считанного) множества простых кусков.

Простым куском называется множество точек топологически (взаимно-однозначно и непрерывно) эквивалентное множеству круга (квадрата).

Уравнение поверхности состоит из уравнений простых кусков поверхности.

Кинематический способ образования поверхности

При кинематическом способе поверхность образуется движением линии – называемой образующей. Закон перемещения в пространстве образующей задается неподвижными кривыми – направляющими в процессе движения образующая линия может менять свою форму, опять таки по определенному закону. При этом конкретную поверхность из всего многообразия выделяет определитель поверхности.

Определение. *Определителем поверхности называется совокупность условий задающих конкретную поверхность. Определитель состоит из двух частей:*

■ *Геометрической, в которой задаются некоторые постоянные величины (точки, линии, прямые, плоскости и т.п.)*

■ *Алгебраической, задающей законы изменения формы образующей и ее перемещения в пространстве.*

Определитель поверхности обозначается $\Phi\{(G), [A]\}$.

Конкретным поверхностям соответствует конкретное содержание геометрической (G) и алгебраической $[A]$ части определителя.

Пример. Определитель поверхности с криволинейной образующей a , направляющей m , законом изменения формы A_1 и законом движения образующей A_2

$$\Phi\{(a, m), [A_1, A_2]\} \Rightarrow$$

Пусть $m: x^2 + y^2 = R^2, l$ – прямая, s – точка.

Тогда $\Phi\{(m, l, s), [s \in l, M = l \times m, \forall M \in m]\}$ - определитель кругового конуса.

Определение. Каркасом поверхности называется множество точек и линий определяющих поверхность. Как правило каркас поверхности совпадает с силовым каркасом.

Исходная информация о конструируемой поверхности может быть представлена

■ Точечным базисом. Аппроксимация (интерполяция) точечного представления поверхности осуществляется с помощью выбранной базовой кривой в виде декартова тензорного представления.

■ Линейным каркасом. (каркасом линий одного семейства). Каждая линия каркаса аппроксимируется одним и тем же параметрическим уравнением. При этом линии каркаса поверхности имеют различные значения одного и того же параметра

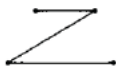
■ Сетчатым каркасом (трансфинитное представление). При этом используется интерполяция семейством параметрических кривых. Метод обеспечивает заданную точность представления поверхности и гибкость в управлении формой поверхности.

Иерархию математического и компьютерного представления поверхности можно представить следующим образом

1. Алгебраическое уравнение поверхности
2. Уравнение каркаса поверхности
3. Обобщенный метод Кунса
4. Представление поверхности в виде декартова произведения.

10.2 Моделирование поверхностей полигональными сетками.

Основные определения



Ломаная



Полигон



Простой полигон



Выпуклый полигон

Рис. 10.1

Ломаная линия, или просто линия (*polyline*) – последовательность соединенных между собой прямых линий.

Ломаная не обязана образовывать замкнутую фигуру, но если последняя и первая точки соединяются ребром, то такая ломаная становится многоугольником или полигоном (*polygon*). Если, кроме того никакие два ребра не пересекаются, то такой полигон называется простым.

Полигон называется выпуклым, если прямая, соединяющая две его любые точки, целиком лежит внутри его.

Набор полигонов, или «граней», которые в совокупности формируют «оболочку» объекта, называется полигональной сеткой.

Полигональная сетка это совокупность полигонов вместе с нормальными векторами, связанными с каждой вершиной этих полигонов.

Популярность полигональных сеток в графике

основана на простоте использования полигонов:

- Их легко представлять (последовательностью вершин)
- Преобразовывать;
- Единственный нормальный вектор;
- Четко определяемые внутренняя и внешняя области.

Рассмотрим трехмерное моделирование полигональными сетками.

Определение. *Объект считается монолитным, если его полигональные грани плотно примыкают друг к другу.*

Полигональная сетка задается списком полигонов и информацией о направлении, куда обращен каждый полигон, которое, как правило, задается в виде нормали к плоскости его грани.

Нормали в вершинах и нормали к поверхностям.

Выгоднее связывать нормальный вектор с каждой вершиной грани, чем задавать одну нормаль для целой грани.

$\overline{V_1}, \overline{V_2}, \overline{V_3}, \overline{V_4}$ ассоциируются с одной и той же нормалью $\overline{n_1}$; $\overline{V_5}, \overline{V_2}$. (рис. 10.2а)

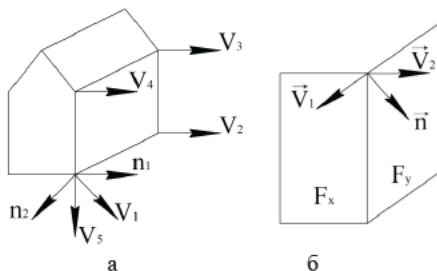


Рис. 10.2

Для гладких поверхностей (например, цилиндра) используется другой подход.

Обе вершины $\overline{V_1}$ с гранью $\overline{F_1}$ и $\overline{V_1}$ с гранью $\overline{F_2}$ используют в одну и ту же нормалью $\overline{n_1}$, являющуюся перпендикуляром к будущей гладкой поверхности (рис. 10.2б).

Задание полигональной сетки

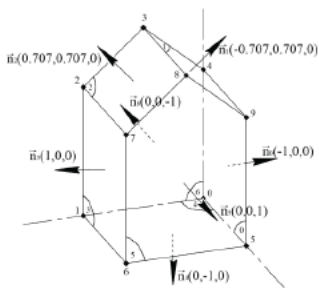


Рис. 10.3. «Базовый сарай»

Рассмотрим рис. 10.3 «Базовый сарай» имеет 7 полигональных граней и 10 вершин, каждая из которых одновременно принадлежит 3 граням.

Можно использовать список из 7 полигонов и для каждого из них список, содержащий вершины полигона и нормали в каждой из этих вершин: всего 30 вершин и 30 нормалей. Структура избыточна, т.к. всего 10 различных вершин и 7 нормалей.

Более эффективно использовать 3 различных списка:

- список вершин (содержатся координаты различных вершин сетки) – геометрия;
- список нормалей (описываются направления различных нормалей) – ориентация;
- список граней (является просто индексом для первых двух списков – топология (связность)).

Пример. Составим перечисленные списки для «базового сарая»:

Список вершин

Вершина	x	y	z
0	0	0	0
1	1	0	0
2	1	1	0
3	0,5	1,5	0
4	0	1	0
5	0	0	1
6	1	0	1
7	1	1	1
8	0,5	1,5	1
9	0	1	1

Список нормалей

Нормал	n_x	n_y	n_z
0	-1	0	0
1	-0,7	0,7	0
2	0,7	0,7	0
3	1	0	0
4	0	-1	0
5	0	0	1
6	0	0	-1

Список граней

грань	Вершины	Ассоциир. норм.
0	0,5,9,4	0,0,0,0
1(крыша левая)	3,4,9,8	1,1,1,1
2(крыша правая)	2,3,8,7	2,2,2,2
3(правая)	1,2,7,6	3,3,3,3
4(нижняя)	0,1,6,5	4,4,4,4
5(передняя)	5,6,7,8,9	5,5,5,5,5
6(задняя)	0,4,3,2,1	6,6,6,6,6

Обходим полигон против часовой стрелки, если смотреть на объект снаружи (это соглашение, но его надо выполнять).

Нахождение нормальных векторов.

Возьмем на грани три соседние точки V_1, V_2, V_3 и вычислим нормаль как векторное произведение:

$$\vec{m} = (V_1 - V_2) \times (V_3 - V_2)$$

После чего нормализуем до единичной длины.

Если $(V_1 - V_2)$ и $(V_3 - V_2)$ - почти параллельны, векторное произведение очень мало и возникают большие погрешности.

Если полигон не является полностью плоским – используется усредненная величина.

Метод Ньюэлла

Рассмотрим полигональную грань

$$P = \{P_0, P_1, \dots, P_{N-1}\}$$

созданную своими N трехмерными вершинами. Покажем, что формулы:

$$\begin{aligned} m_x &= \sum_{i=0}^{N-1} (y_i - y_{next(i)})(z_i + z_{next(i)}), \\ m_y &= \sum_{i=0}^{N-1} (z_i - z_{next(i)})(x_i + x_{next(i)}), \\ m_z &= \sum_{i=0}^{N-1} (x_i - x_{next(i)})(y_i + y_{next(i)}), \end{aligned} \quad (10.1)$$

где $next(j) = (j + 1) \bmod N$ - индекс следующей после j вершины при обходе граней, что обеспечивает «круговой» переход от $(N - 1)$ вершины к нулевой. Формулы (10.1) обеспечивают точное вычисление нормального вектора $\vec{m}(m_x, m_y, m_z)$ к грани P в случае, когда P является плоской, и предлагают хо-

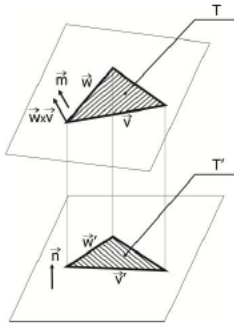


Рис. 10.4

рошее направление для «усредненной» нормали, когда P - неплоская грань. Для простоты рассмотрим случай, когда P является треугольником, как показано на рис. 10.4. Обозначим этот треугольник T , а его единичный нормальный вектор \vec{n} . Пусть $\Delta T'$ - проекция ΔT на плоскость с единичной нормалью \vec{n} . Покажем, что площадь $\Delta T' \equiv S(T')$ является определенной частью от площади всего $\Delta T \equiv S(T)$ и эта часть является скалярным произведением:

$$S(T') = (\vec{n} \cdot \vec{n})S(T)$$

Пусть ребра ΔT представляют собой два вектора \vec{V} и \vec{W} , как показано на рис.4. Из геометрической интерпретации векторного произведения следует:

$$|\vec{W} \times \vec{V}| = 2S(T) \Rightarrow \vec{V} \times \vec{W} = 2S(T)\vec{n}$$

Рассмотрим проекцию T' треугольника T : $\vec{V}' \times \vec{W}' = 2S(T')\vec{n}$

Но \vec{W}' - проекция вектора \vec{W} , а \vec{V}' - проекция вектора \vec{V} на плоскость с нормалью \vec{n} .

Разложим вектор \vec{V} на составляющие в направлении \vec{n} и \vec{n}^\perp :

$$\vec{V} = \left(\frac{\vec{V} \cdot \vec{n}}{|\vec{n}|^2} \right) \vec{n} + \left(\frac{\vec{V} \cdot \vec{n}^\perp}{|\vec{n}^\perp|^2} \right) \vec{n}^\perp$$

Второе слагаемое это проекция вектора \vec{V} на плоскость с нормалью \vec{n} , т.е.

$$\vec{V}' = \left(\frac{\vec{V} \cdot \vec{n}^\perp}{|\vec{n}^\perp|^2} \right) \vec{n}^\perp$$

Учтем, что \vec{n} - единичный вектор. Окончательно получим:

$$\vec{V} = (\vec{V} \cdot \vec{n})\vec{n} + \vec{V}' \Rightarrow \vec{V}' = \vec{V} - (\vec{V} \cdot \vec{n})\vec{n}$$

Аналогично получим: $\vec{W}' = \vec{W} - (\vec{W} \cdot \vec{n})\vec{n}$.

$$\begin{aligned} \vec{V}' \times \vec{W}' &= (\vec{V} - (\vec{V} \cdot \vec{n})\vec{n}) \times (\vec{W} - (\vec{W} \cdot \vec{n})\vec{n}) \\ &= \vec{V} \times \vec{W} - (\vec{W} \cdot \vec{n})\vec{n} \times \vec{V} - (\vec{V} \cdot \vec{n})\vec{n} \times \vec{W} + (\vec{V} \cdot \vec{n})(\vec{W} \cdot \vec{n})\vec{n} \times \vec{n}; \end{aligned}$$

Очевидно, что последнее слагаемое равно 0, ($\vec{n} \times \vec{n} = 0$)

$$2S(T')\vec{n} = \vec{V} \times \vec{W} - (\vec{W} \cdot \vec{n})(\vec{V} \times \vec{n}) - (\vec{V} \cdot \vec{n})(\vec{n} \times \vec{W}) \quad (10.2)$$

Умножим скалярно обе части уравнения (2) на вектор \vec{n} , учтем $(\vec{V} \times \vec{n})\vec{n} = 0$ и $(\vec{W} \times \vec{n})\vec{n} = 0$, т.к. $(\vec{V} \times \vec{n}) \perp \vec{n}$ и $(\vec{W} \times \vec{n}) \perp \vec{n}$ окончательная формула имеет вид:

$$2S(T') = (\vec{V} \times \vec{W})\vec{n} = 2S(T)(\vec{n} \cdot \vec{n}). \quad (10.3)$$

Полученный результат справедлив для площади любого плоского полигона P и площади его проекции P' , т.к. в этом случае справедлив принцип суперпозиции, а любой плоский полигон можно представить, как объединение плоских треугольников.

Т.к. \vec{m} и \vec{n} единичные векторы, то $\vec{m} \cdot \vec{n} = \cos \psi$, где $\angle \psi$ это угол, отсчитываемый против хода часовой стрелки между положительными направлениями векторов \vec{m} и \vec{n} . Окончательно имеем:

$$S(T') = S(T) \cos \psi$$

Спроецируем полигон P на координатные плоскости YOZ ($x = 0$), XOZ ($y = 0$), XOY ($z = 0$), а A_x, A_y, A_z - площади проекций полигона P на эти плоскости. Тогда

$$A_x = S(T)m_x, A_y = S(T)m_y, A_z = S(T)m_z.$$

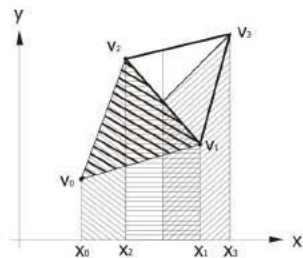


Рис. 10.5.

Чтобы найти вектор \vec{m} , нужно вычислить

вектор $\vec{A}(A_x, A_y, A_z)$ и нормировать его к единичной длине. Каждая трехмерная вершина $P_i =$

(x_i, y_i, z_i) проецируется на плоскость XOY как $V_i = (x_i, y_i)$. Каждое ребро полигона P' определяет трапециевидную область, лежащую между этим ребром и осью x . Площадь этой трапеции равна произведению длины ее основания на

высоту средней точки ребра $\left[x_0 V_0 \bar{V} x_1 - \text{прямоугольник}, = (x_1 - x_0) y_0 \right]$,

а $x_0 \bar{V} V_1$ - прямоугольный треугольник, $S_{\Delta} = \frac{1}{2} (x_1 - x_0) \cdot (y_1 - y_0) =$

$\frac{1}{2} [(x_1 - x_0)(y_1 + y_0)] \left] A_0 = 0,5(x_1 - x_0)(y_1 + y_0) \right.$. Эта величина отрицательна,

если x_0 лежит левее x_1 , и положительна в противном случае. Применим полученную формулу для остальных ребер. Тогда для i -го ребра имеем

$$A_i = 0,5(x_i - x_{next(i)})(y_i + y_{next(i)})$$

где $next(i) = 0$ при $i = N - 1$ равно $i + 1$ в остальных случаях.

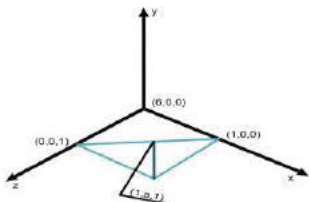
Вектор \vec{m} является внешней нормалью, при условии, что вершины пронумерованы против часовой стрелки, если смотреть снаружи.

Рассмотрим четырехугольник приведенный на рис.5 с вершинами:

$$N_0(0,0,0) N_1(1,0,0) N_2(1, a, 1) N_3(0,0,1)$$

Определим нормаль по методу Ньюэлла

$$m_x = \sum_{i=0}^{N-1} (y_i - y_{next(i)})(z_i + z_{next(i)})$$



$$m_y = \sum_{i=0}^{N-1} (z_i - z_{next(i)})(x_i + x_{next(i)})$$

$$m_z = \sum_{i=0}^{N-1} (x_i - x_{next(i)})(y_i + y_{next(i)})$$

$$m_x = 0 \cdot 0 + (-a) \cdot 1 + a(2) + 0 \cdot 1 = a$$

$$m_y = 0 \cdot 1 + (-1) \cdot 2 + 0 \cdot 1 + 1 \cdot 0 = -2$$

$$m_z = -1 \cdot 0 + 0 \cdot a + 1 \cdot a + 0 \cdot 0 = a$$

$$\bar{m}\{a, -2, a\} \text{ при } a = 0 \quad m\{0, -2, 0\}$$

$$|m| = \sqrt{2a^2 + 4} = 2\sqrt{0,5a^2 + 1}$$

$$\sin \varphi = \frac{a\sqrt{2}}{2\sqrt{0,5a^2 + 1}}$$

$$0,5 \cdot a^2 \ll 1 \Rightarrow \sin \varphi \approx 0,7a \sim \varphi.$$

Свойства сеток:

■ Монолитность – сетка представляет монолитный объект, если совокупность его граней заключает в себе некоторое конечное пространство;

■Связность. Сетка называется связной, если между любыми двумя вершинами существует непрерывный путь вдоль ребер полигона (сетка обычно не является связной, если она представляет более одного объекта);

■Простота. Сетка называется простой, если отображаемый ею объект является монолитным и не содержит отверстий. Это означает, что объект может быть деформирован в сферу, не подвергаясь разрезанию;

■ Плоскостность. Сетка называется плоской, если каждая грань представляемого ею объекта является плоским полигоном, т.е. вершины каждой грани лежат в одной плоскости. Треугольники по определению являются плоскими, поэтому это обстоятельство используется достаточно часто в программах. 4-х угольник может и не быть плоским;

■ Выпуклость. Сетка представляет выпуклый объект, если прямая соединяющая любые две точки внутри этого объекта, целиком лежит внутри него.

10.3 Многогранники.

Часто бывает удобно ограничить данные в сетках так, чтобы она изображала многогранник (полиэдр).

Определение. *Полиэдром называется связная сетка из простых плоских полигонов, которая ограничивает конечный объем пространства.*

Согласно определению, полиэдр представляет собой единый монолитный объект. Из этого следует:

■ Каждое ребро принадлежит ровно двум граням;

■ В каждой вершине встречается не менее трех ребер;

■ Грани не являются взаимопроникающими: две грани или не имеют общих точек, или пересекаются только вдоль их общего ребра.

Полиэдр простой, если он не имеет отверстий.

Формула Эйлера устанавливает фундаментальное соотношение между количеством граней Γ , ребер P и вершин B простого многогранника:

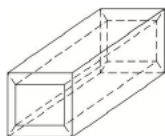
$$B + \Gamma - P = 2 \quad (10.4)$$

Обобщение этой формулы на непростой полиэдр выглядит так:

$$B + \Gamma - P = 2 + H - 2G, \quad (10.5)$$

Где H – общее число отверстий, имеющих в гранях (не сквозных)

G – число отверстий в самом полиэдре.



На рис. 10.6 параллелепипед с отверстием в форме другого параллелепипеда. Две его торцевые грани выдвинуты из ос нового параллелепипеда. Для этого объекта $B=16$, $\Gamma=16$, $P=32$, $H=0$, $G=1$:

$$16 + 16 - 32 = 2 + 0 - 2 \cdot 1 \Rightarrow 0 = 0.$$

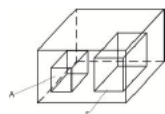


Рис. 10.6.

На рисунке 10.6 изображен полиэдр с частично проникающим в него отверстием А и сквозным отверстием Б: $B=24$, $\Gamma=15$, $P=36$, $G=1$, $H=3$

$$24 + 15 - 36 = 2 + 3 - 2 \cdot 1 \Rightarrow 3 = 3.$$

«структуру полиэдра удобно описывать с помощью диаграммы Шлегеля, которая основана на взгляде на полиэдр из точки, расположенной на некотором расстоянии от центра одной из его граней.

Несколько примеров диаграмм Шлегеля

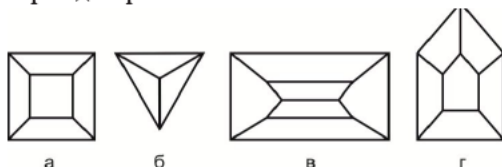


Рис. 10.7

а - параллелепипед, б - пирамида, в и г – базовый сарай.

Свойства многогранника, связанные лишь с общей схемой соединения его граней, называются комбинаторными или топологическими свойствами полиэдра. Остальные свойства называются метрическими свойствами (величина поверхности или объема многогранника, величины его линейных или двугранных углов)

Более точно комбинаторные свойства полиэдра могут быть охарактеризованы при помощи изоморфизма. Два многогранника называются изоморфными,

если между вершинами, ребрами и гранями одного многогранника и соответственно вершинами, ребрами и гранями другого можно установить взаимно однозначное соответствие, при котором инцидентным между собой вершинам, ребрам и граням первого многогранника отвечают инцидентные вершины ребра и грани второго.

(Вершины, ребра и грани называются инцидентными, если они примыкают друг к другу. Ребро и грань инцидентны, если это ребро является одной из сторон рассматриваемой грани). Треугольная призма, треугольная усеченная пирамида и «клин» изоморфны между собой. О двух изоморфных многогранниках говорят, что они одного топологического (комбинаторного) типа.

Отношение инцидентности:

- Симметрично: т.е. если $a \sim A \Rightarrow A \sim a$;
- Транзитивно: если $A \sim a, a \sim \alpha \Rightarrow A \sim \alpha$
- Если $A \sim \alpha$, то существуют два и только два ребра, инцидентных как вершине A , так и грани α ;
- Если каждая из двух заданных граней инцидентна каждой из двух вершин, то существует одно и только одно ребро, инцидентное обеим граням и обеим вершинам;
- Каждое ребро инцидентно 2 и только 2 вершинам;
- Каждое ребро инцидентно 2 и только 2 граням.

Призмы и антипризмы

Призма (рис.8) это частный случай полиэдра. Она обладает симметрией и поэтому довольно проста в описании. Призма получается в результате заметания, или выдавливания (*extrudig* – экструзия), полигона вдоль прямой линии, что превращает двумерный полигон в трехмерный полиэдр.

Призма прямая, если боковая поверхность перпендикулярна основанию.

Призма правильная, если в основании лежит правильный полигон, а в качестве боковых граней квадраты (прямоугольники).

У антипризмы верхний n –угольник повернут на $\frac{180}{n}$ градусов и соединен с нижним n –угольником таким образом, что получаются грани в виде равно-сторонних (равнобедренных) треугольников.

Правильная призма и антипризма являются примерами так называемых «полуправильных» многоугольников.

Платоновы тела

Если все грани полиэдра одинаковы и каждая из них является правильным многоугольником, то такой объект называется правильным многогранником. Требование данной симметрии является настолько жестким, что существует всего 5 таких объектов: Платоновы тела.

Пусть V , G . и P – число вершин, граней и ребер многогранника.

Пусть каждая грань многогранника имеет n вершин и в каждой вершине сходятся S граней. Т.к. каждая грань имеет n ребер, то все Γ граней имеют $n\Gamma$ ребер, но при этом каждое ребро многогранника учитывается дважды (т.к. оно принадлежит двум граням). Следовательно

$$n\Gamma = 2P. \quad (10.6)$$

Аналогично, подсчитывая ребра, примыкающие к каждой вершине, получим соотношение:

$$SB = 2P \quad (10.7)$$

Согласно теореме Эйлера

$$B + \Gamma - P = 2$$

Выражая B и Γ через P по формулам (10.6) и (10.7) и подставляя в формулу Эйлера, приходим к уравнению:

$$\frac{2P}{n} + \frac{2P}{S} - P = 2 \Rightarrow \frac{1}{S} + \frac{1}{n} = \frac{1}{2} + \frac{1}{P}$$

Это уравнение позволяет найти все возможные (для топологически правильного многогранника) значения S , n и P .

Действительно: $\frac{1}{S} + \frac{1}{n} > \frac{1}{2}$ (справа отбросим $\frac{1}{P}$)

Очевидно, что хотя бы одно из чисел $\frac{1}{S}$ и $\frac{1}{n}$ должно быть больше $\frac{1}{4}$ (пусть $\frac{1}{S} = \frac{1}{n} \Rightarrow \frac{2}{n} > \frac{1}{2} \Rightarrow \frac{1}{n} > \frac{1}{4}$), т.е. хотя бы одно из чисел S и n должно быть меньше 4.

По геометрическому смыслу ни одно из чисел S и n не может быть меньше 3, то должно быть либо $S = 3$, либо $n = 3$.

Т.к. $\frac{1}{S} \leq \frac{1}{3}$ то $\frac{1}{n} > \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$, т.е. $n < 6$ (аналогично $S < 6$).

Таким образом, имеется лишь пять возможностей: либо $n = 3, S = 3, 4, 5$; либо $S = 3, n = 4, 5$. Проверка показывает, что во всех случаях, число P определяемое соотношением (10.4), а значит и числа B и Γ , определяемые (10.6) и (10.7) являются целыми числами. Находя эти числа, мы определяем все возможные характеристики правильных многогранников (рис. 10.8).

тела	$B \ n \ \Gamma \ S \ P$	Символ Шлефля
а) Тетраэдр	4 3 4 3 6	(3,3)
б) Гексаэдр	8 4 6 3 12	(4,3)
в) Октаэдр	6 3 8 4 12	(3,4)
г) Икосаэдр	12 3 20 5 30	(3,5)
д) додекаэдр	20 5 12 3 30	(5,3)

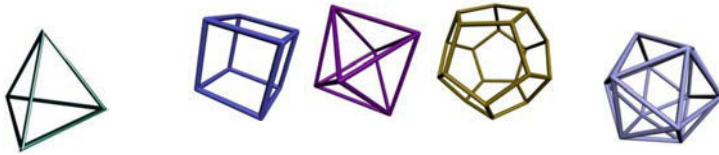


Рис. 10.8

Символ Шлефля (p, q) означает, что каждая грань является p –угольником и что q из них сходятся в каждой вершине.

Икосаэдр состоит из антипризмы и двух пятиугольных пирамид в его вершине и основании.

Двойственные многогранники

Каждому Платонову телу P соответствует двойственный многогранник D . Вершинами полиэдра D являются центры граней полиэдра P , так что ребра полиэдра D соединяют средние точки смежных граней полиэдра P . (рис. 10.8)

- а) двойственным для тетраэдра является тетраэдр.
- в) куб и октаэдр двойственны.
- г) икосаэдр и додекаэдр – двойственны.

Двойственные многогранники имеют то же число ребер E , а параметр V для одного из них является параметром F для другого. Если (p, q) - символ Шлефля для одного из двойственных тел, то для другого символ Шлефля будет равен (q, p) .

Если известен список вершин для одного Платонова тела P , то легко создать такой же список для двойственного ему тела D , поскольку вершина k полиэдра D располагается в центре грани k полиэдра P . Т.е. фактически создается многогранник D , вписанный в многогранник P .

Чтобы следить за нумерацией вершин и граней используют модель (развертку), образуемую путем разрезания полиэдра вдоль определенных ребер и «разворачивания» (совмещения) его в плоскую фигуру таким образом, что все его грани видны с внешней стороны.

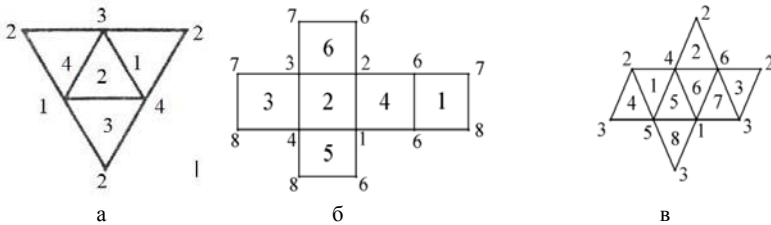


Рис. 10.9

а – тетраэдр; б – гексаэдр; в – октаэдр

Рассмотрим двойственную пару куб-октаэдр. Грань 4 куба окружена вершинами 1,5,6,2. Из двойственности следует, что вершина 4 октаэдра окружена гранями 1,5,6,2. Отметим, что поскольку тетраэдр двойственен самому себе, то его список вершин, окружающих k -ю грань, совпадает со списком граней, окружающих k -ю вершину.

Вершина 4 октаэдра является центром грани 4 куба. (координаты центра грани вычисляются как среднее арифметическое координат вершин, принадлежащих этой грани).

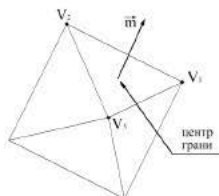
Если нам известны вершины V_1, V_5, V_6, V_2 для грани 4 куба, то получим

$$V_4 = \frac{1}{4}(V_1 + V_5 + V_6 + V_2).$$

Нормальные векторы к Платоновым телам.

Высокая степень симметричности (допускает) позволяет достаточно просто вычислить для каждой грани нормальный вектор. Пусть тело центрировано относительно начала координат. Тогда нормальный вектор к каждой грани – это вектор из начала координат (радиус вектор) к центру грани, представляющему собой среднее значение вершин. (Заметим, что этот вектор совпадает с вектором, проведенным из начала координат к соответствующей вершине двойственного к данному Платонову телу) (рис.10.10)

Для октаэдра



$$\vec{m} = \frac{V_1 + V_2 + V_3}{3}$$

Существует бесконечное разнообразие многогранников (табл. 10.1).

Многогранник называется (метрически) равноугольно правильным или архимедовым, если все его грани – правильные многоугольники, а все многогранные углы равны между собой. Их всего 14, не считая двух бесконечных серий призм и антипризм. Такие многогранники имеют в качестве граней несколько различных правильных многоугольников.

Архимедовы тела все еще обладают достаточной симметрией, так что нормальный вектор к каждой грани может быть найден с использованием координат середины этой грани.

Изображения Архимедовых тел приведены на рис.10.11, а основные характеристики в таблице. На (рис. 10.11 г) изображено одно из Архимедовых тел, представляющее особый

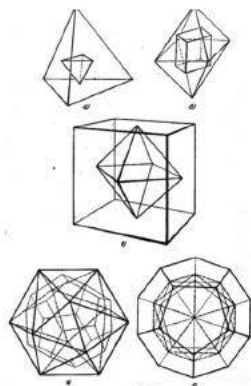


Рис. 10.10.

интерес: оно представляет собой усеченный икосаэдр $5 \cdot 6^2$ и состоит из правильных гексагонов и пентагонов. Этот многогранник известен во всем мире – именно такую форму имеет футбольный мяч. В последнее время он получил новое название – бакибол – в честь Бакминстера Фуллера из-за его интереса к подобным геофизическим структурам.

Недавно кристаллографы обнаружили, что 60 атомов углерода, расположенные в вершинах усеченного икосаэдра, образуют новый вид молекулы углерода, отличный от графита и алмаза. Этот материал обладает многими замечательными свойствами, например, устойчивостью к высоким температурами и сверхпроводимостью. Он получил имя фуллерен.

Таблица 10.1

№ п/п	s	s_1	s_2	s_3	n_1	n_2	n_3	Γ_1	Γ_2	Γ_3	V	Γ	p	№ рис	
1	3	2	1	-	4	n	-	n	2	-	$2n$	$n+2$	$3n$	а	Правильная n -угольная призма с квадратными боковыми гранями($n=3, 5, 6, 7, \dots$)
2	3	2	1	-	6	3	-	4	4	-	12	8	18	б	Усеченный тетраэдр
3	3	2	1	-	6	4	-	8	6	-	24	14	36	в	Усеченный октаэдр
4	3	2	1	-	6	5	-	20	12	-	60	32	90	г	Усеченный икосаэдр
5	3	2	1	-	8	3	-	6	8	-	24	14	36	д	Усеченный куб
6	3	2	1	-	10	3	-	12	20	-	60	32	90	е	Усеченный додекаэдр
7	3	1	1	1	4	6	8	12	8	6	48	26	72	ж	Усеченный кубоктаэдр
8	3	1	1	1	4	6	10	30	20	12	120	62	180	з	Усеченный икосододекаэдр
9	4	3	1	-	3	n	-	$2n$	2	-	$2n$	$2n+2$	$4n$	и	n -угольная анти-призма ($n=4, 5, 6, \dots$)
10	4	3	1	-	4	3	-	18	8	-	24	26	48	к	Ромбукубоктаэдр
11	4	3	1	-	4	3	-	18	8	-	24	26	48	л	
12	4	2	2	-	3	4	-	8	6	-	12	14	24	м	Кубоктаэдр
13	4	2	2	-	3	5	-	20	12	-	30	32	60	н	Икосододекаэдр
14	4	2	1	1	4	3	5	30	20	12	60	62	120	о	Ромбикосододекаэдр
15	5	4	1	-	3	4	-	32	6	-	24	38	60	п	Плосконосый куб
16	5	4	1	-	3	5	-	80	12	-	60	92	150	р	Плосконосый додекаэдр

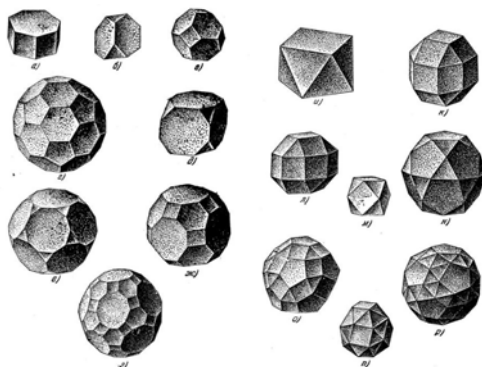


Рис. 10.11. Архимедовы тела

Геодезические купола

Понятие геодезических куполов ввел Бакминстер. Эти тела составляют интересный класс полиэдров, обладающих многими полезными свойствами. Геодезический купол, созданный из реальных материалов, чрезвычайно прочен для своего веса. Геодезический купол может принимать различные формы, но все они расположением своих граней (обычно Δ) аппроксимируют сферу, нижняя часть которой удалена.

Рассмотрим пример, основанный на икосаэдре. Для определения граней каждое ребро икосаэдра разбивается на $3F$ равных частей, где F - (величина введенная Фауллером) – частота купола. В примере $F = 3$, т.е. каждое ребро делится на три части для образования меньших Δ граней. Геодезическую конструкцию изобрел и запатентовал в 1951 году американский изобретатель Ричард Бакминстер Фуллер (12.07.1895 - 01.07.1983). Над этой идеей изобретатель работал с 1947 года, искренне надеясь, что применение геодезических куполов в строительстве поможет решить послевоенный жилищный кризис. Конструкция геодезического купола заинтересовала Фуллера прежде всего благодаря малой массе при большом внутреннем пространстве. Купола нашли свое применение в различных архитектурных строениях — больших оранжереях, планетариях, аудиториях, складах, ангарах. Жилые же купола не оправдали своих надежд из-за высокой стоимости. Купола обладают рядом преимуществ, которые делают их уникальными архитектурными сооружениями. Купола обладают большой несущей способностью, причем чем больше купол, тем она выше. Простые сооружения создаются очень быстро из достаточно лёгких элементов силами небольшой строительной группы: структуры до 50 метров собираются даже без строительного крана. Купола также обладают идеальной аэродинамической формой, благодаря чему их можно возводить в ветреных и ураганных районах. Однако есть и недостатки. Обычно современные материалы для создания граней купола имеют прямоугольное сечение, их приходится дополнительно обрабатывать, придавая треугольное сечение, из-за чего появляется много лишних отходов.



Монреальская Биосфера (бывший Павильон США на Экспо-67), созданная архитектором Ричардом Фуллером. Элементы конструкции геодезического купола

10.4. Экструзивные формы

Обширный класс форм может быть создан путем выдавливания (экструзии) двумерной формы в пространство.

Примером форм, образованных посредством «линейной экструзии» (выдавливания по прямой линии), являются призмы.

Тетраэдр, октаэдр – примеры экструзии формы в пространство определенным способом. Поверхности вращения также могут быть аппроксимированы путем «вытеснения» полигона в пространство при слегка расширенном определении экструзии.

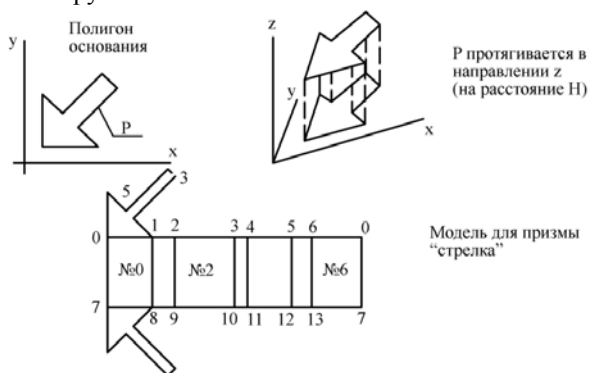


Рис. 10.12. Модель для призмы «стрелка»

Полигон P протянут на расстояние H вдоль оси z , в результате чего образовалась призма «стрелка» (рис. 10.12). (В более общем случае экструзия может проводиться вдоль вектора \vec{d} , достаточно произвольно направленного). В результате протягивания полигона P из каждой вершины образуется ребро в заданном направлении. Этот процесс приводит к возникновению еще 7 вершин, всего у призмы их 14. Вершины появляются парами: если $(x_i, y_i, 0)$ – одна из вершин полигона P , то у призмы имеется вершина (x_i, y_i, H) . Призма имеет 7 боковых граней, основание и крышку. Поскольку данная призма имеет плоские грани, с каждой вершиной грани связан один и тот же нормальный вектор – нормальный вектор к самой грани.

Инструмент, создающий сетку для призмы на основе произвольного полигона.

Пусть основанием призмы является полигон с N вершинами (x_i, y_i) . Занумеруем вершины основания числами $0, 1, \dots, N - 1$, а вершины крышки числами $N, \dots, 2N - 1$ так, чтобы ребро соединяло вершины с номерами i и $i + N$. Тогда список вершин содержит точки $(x_i, y_i, 0)$ и (x_i, y_i, H) для $i = 0, 1, \dots, N - 1$.

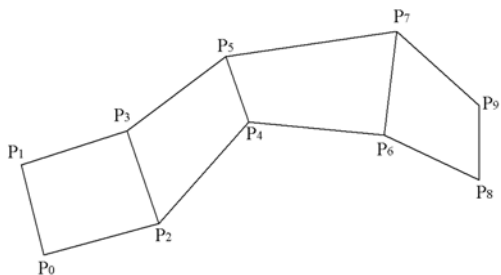


Рис. 10.13

В список граней запишем вначале «боковые» грани или «стены», а затем добавим основание и крышку. Для j -той стенки ($j = 0, \dots, N - 1$) получим грань с четырьмя вершинами с индексами $j, j + N, next(j) + N, next(j)$, где $next(j) = \begin{cases} j + 1, & \text{для } j = N - 1. \\ 0 \end{cases}$

Частный случай: экструзия полос из четырехугольников.

Это призмы, полигон основания которых может быть представлен виде полосы четырехугольников (*quad - strip*) – совокупность четырехугольников, соединенных в цепочку (подобно кирпичам, выложенным в ряд), таким образом, что соседние грани полностью совпадают. Полоса описывается последовательностью вершин

$$quad - strip = \{P_0, P_1, \dots, P_{M-1}\}.$$

Пусть вершины берутся парами, причем нечетные вершины формируют одно «ребро» полосы, а четные – другое. Очевидно, что не каждый полигон можно представить в виде полосы.

Когда сетка сформирована как экструзивная полоса, в список вершин помещаются только $2M$ вершин, а в список граней помещаются только «внешние стенки», которых всего $2(M - 1)$ и при визуализации лишние стенки вообще не рисуются. Для экструзивных полос параметрами сетки является массив двумерных точек и вектор экструзии.

Экструзия с поворотом

Экструзия передвигает полигон основания в новое положение. Определяющее полигон крышки. Обобщим эту концепцию.

В качестве полигона крышки можно взять увеличенный или уменьшенный, и даже повернутый полигон крышки.

Пусть полигон основания P задан вершинами $\{P_0, P_1, \dots, P_{N-1}\}$, тогда полигон P' крышки будет задан:

$$P' = \{MP_0, MP_1, \dots, MP_{M-1}\}$$

где M - некоторая матрица 4×4 , соответствующая аффинному преобразованию.

$$M = \begin{bmatrix} 0,7 & 0 & 0 & 0 \\ 0 & 0,7 & 0 & 0 \\ 0 & 0 & 1 & H \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ матрица содержит масштабный множитель } 0,7 \text{ и}$$

перемещение на H вдоль оси z .

$$M = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & H \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ крышка перед перемещением была повернута}$$

на угол θ вокруг оси z и перемещена на H вдоль оси z .

Трубки и змейки.

Семейство объектов может быть создано путем применения последовательности экструзий, каждая со своим преобразованием и укладкой их непрерывной цепью в форме трубки. Первый сегмент имеет концевые полигоны M_0P и M_1P , где M_0 позиционирует и ориентирует начальный конец трубы. Второй сегмент концевые полигоны M_1P и M_2P и т.д. Различающиеся преобразованные полигоны называются перетяжками трубы. В рассматриваемом примере список вершин сетки содержит 16 вершин:

$$M_0P_0, M_0P_1, M_0P_2, M_0P_3, M_1P_0, M_1P_1, M_1P_2, \dots, M_3P_0, M_3P_1, M_3P_2, M_3P_3.$$

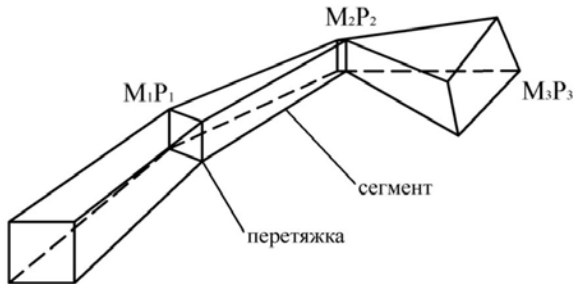


Рис. 10.14

Конструирование трубок, на базе трехмерных кривых.

Представим трубу, окружающую некоторую кривую, называемую хребтом (*spine*) трубы и которая изгибается в пространстве определенным образом. Представим эту кривую параметрически в виде $C(t)$. Например, винтовая линия: $C(t) = (\cos t, \sin t, bt)$, где $b = const$.

Для формирования различных перетяжек трубы, вычислим $C(t)$ для значений параметра $\{t_0, t_1, \dots\}$ и построим преобразованный полигон в каждой точке $C(t_i)$ в плоскости, перпендикулярной к кривой. Это равносильно созданию локальной системы координат в каждой выбранной точке вдоль хребта.

Локальная ось z показывает направление вдоль кривой, а оси x и y направлены перпендикулярно оси z и перпендикулярно друг другу. Полигон перетяжки при этом располагается в xy -плоскости и необходимо задать только вершины каждой перетяжки.

Удобнее всего предоставлять самой кривой $C(t)$ определять локальную систему координат: базис Френе. При каждом значении t_i вычисляется:

$\vec{T}(t_i)$ -единичный вектор, касательный к кривой;

$\vec{B}(t_i)$ -единичный вектор бинормали;

$\vec{N}(t_i)$ - единичный вектор главной нормали.

$$\vec{T}(t) = \frac{\vec{c}(t)}{|\vec{c}(t)|} \rightarrow \bar{T}(t) = -\frac{1}{\sqrt{1+b^2}}(-\sin t, \cos t, b)$$

$$\vec{B}(t) = \frac{\vec{c}(t)}{|\vec{c}(t)|} \rightarrow \bar{B}(t) = -\frac{1}{\sqrt{1+b^2}}(b\sin t, -b\cos t, 1)$$

$$\rightarrow \bar{N}(t) = (-b\cos t, -\sin t, 0)$$

10.5 Линейчатые поверхности

Поверхность называется линейчатой, если через каждую ее точку проходит хотя бы одна прямая, целиком лежащая на этой поверхности.

Линейчатая поверхность порождается семейством прямых и описывается уравнением:

$$\vec{r} = \vec{r}(u, v) = \vec{r}_0(u) + v\vec{n}(u),$$

Где $\vec{r}_0(u)$ - заданная точка на прямой с параметром u ,

А $\vec{n}(u)$ -направляющий вектор этой прямой. Параметр v дает расстояние между точками $\vec{r}(u, v)$ и $\vec{r}_0(u)$. Линейчатая поверхность, образующие которой соединяют соответствующие точки двух пространственных кривых $\vec{r} = \vec{r}_0(u)$ и $\vec{r} = \vec{r}_1(u)$ может быть представлена в виде

$$\vec{r} = \vec{r}(u, v) = (1 - v)\vec{r}_0(u) + v\vec{r}_1(u) \quad (10.8)$$

Кривые $\vec{r} = \vec{r}_0(u)$ и $\vec{r} = \vec{r}_1(u)$ называются направляющими, а прямые, соединяющие соответствующие точки направляющих, называются образующими линейчатой поверхности.

В точечном представлении уравнение (10.8) имеет вид:

$$P(v) = (1 - v)P_0 + vP_1,$$

где P_0 и P_1 - точка, а v - единственный параметр. Однако для линейчатых поверхностей точки P_0 и P_1 становятся функциями второго параметра u : P_0 становится $P_0(u)$, а $P_1 \Rightarrow P_1(u)$. И параметрическая форма исследуемых линейчатых поверхностей имеет вид

$$P(u, v) = (1 - v)P_0(u) + vP_1(u)$$

Функции $P_0(u)$ и $P_1(u)$ определяют пространственные кривые, причем каждая из них описывается тремя функциями компонент:

$P_0(u) = (x_0(u), y_0(u), z_0(u)), P_1(u) = (x_1(u), y_1(u), z_1(u))$, определенных на одном и том же интервале u (обычно от 0 до 1).

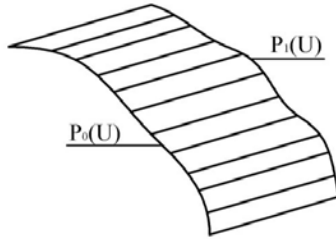


Рис. 10.15

Линейчатая поверхность (рис. 10.15) состоит из прямых линий, соединяющих каждую пару соответствующих точек $P_0(u')$ и $P_1(u')$. Прямую для $u = u'$ называют образующей в точке u' .

Для конкретного фиксированного значения v', v' - контур является аффинной комбинацией двух кривых $P_0(u)$ и $P_1(u)$, причем первая из них имеет вес $(1 - v')$, а вторая v' .

Если v ничем не ограничено, то каждая прямая будет неограниченно продолжаться в обоих направлениях, и поверхность будет напоминать неограниченный криволинейный «лист».

Линейчатый лоскут формируется посредством ограничения обеих величин u и v некоторыми значениями, например $[0,1]$.

Линейчатая поверхность может быть легко «полигонизирована». Выбирается множество значений u_i, v_i и для каждого из них вычисляется точка $P(u_i, v_i)$ и нормаль $\vec{n}(u_i, v_i)$. На основе этого формируются списки.

Примеры линейчатых поверхностей: конусы, цилиндры, билинейный лоскут.

Конус – линейчатая поверхность, одна из кривых которой, $P_0(u)$, представляет собой единственную точку – вершину конуса.

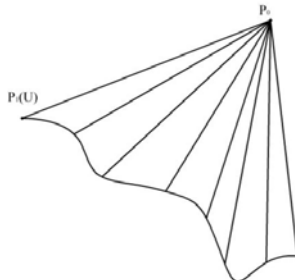


Рис. 10.16

Параметрическое уравнение конуса имеет вид:

$$P(u, v) = (1 - v)P_0 + vP_1(u),$$

Где P_0 - вершина. Все прямые проходят через $(\cdot)P_0$ при $v = 0$ и через точку $P_1(u)$ при $v = 1$.

Частный случай: круговой конус, если $P_1(u)$ - окружность.

Цилиндр – линейчатая поверхность, у которой кривая $P_1(u)$ представляет собой смещенную кривую $P_0(u)$: $P_1(u) = P_0(u) + \vec{d}$ где \vec{d} - некоторый вектор. Параметрическая форма цилиндра имеет вид:

$$P(u, v) = P_0(u) + \vec{d}v$$



Билинейные лоскуты получаются в случае, когда обе кривые $P_0(u)$ и $P_1(u)$ являются отрезками прямых из одного и того же интервала для u , например $[0,1]$. Обозначим концевые точки прямой $P_0(u)$ через P_{00} и P_{01} , (тогда $P_0(u) = (1 - u)P_{00} + uP_{01}$), а концевые точки $P_1(u)$ через P_{10} и P_{11} . В этом случае параметрическое уравнение билинейного лоскута имеет вид:

$$P(u, v) = (1 - v)(1 - u)P_{00} + (1 - v)uP_{01} + v(1 - u)P_{10} + uvP_{11}.$$

Поверхность называется билинейной потому, что она ли-

нейно зависит и от u и от v . Если прямые $P_0(u)$ и $P_1(u)$ располагаются в одной плоскости – то билинейный лоскут плоский (т.е. лоскут плоскости). В общем случае это – косяя плоскость.

Существует много других линейчатых поверхностей: лента Мёбиуса,

Три базовые формы: сфера, цилиндр и конус

Сфера $F(x, y, z) = x^2 + y^2 + z^2 - 1$ единичного радиуса с центром в начале координат называется «базовой сферой».

Другая запись: $F(P) = |P^2| - 1$

В параметрическом виде (который получается из стандартного описания в сферических координатах):

$$P(u, v) = (\cos v \cos u, \cos v \sin u, \sin v)$$

$$0 < u \leq 2\pi, -\frac{\pi}{2} < v < \frac{\pi}{2}.$$

Из географии:

u - контуры – называются меридианами (долгота)

v - контуры – параллелями.

Параметрам u, v можно придать другой геометрический смысл и получить для сферы другую параметрическую форму. Пусть u по-прежнему долгота, а v - высота точки P над плоскостью xy тогда

$$P_2(u, v) = (\sqrt{1 - v^2} \cos u, \sqrt{1 - v^2} \sin u, v).$$

$$0 < u \leq 2\pi, -1 < v < 1.$$

Нормальный вектор $\vec{n}(u, v)$ к поверхности сферы всегда направлен по радиусу наружу, т.е. параллелен радиус-вектору из начала координат в данную точку.

Градиент $\nabla F = 2(x, y, z)$, т.е. пропорционален P .

В параметрической форме $\vec{n}(u, v) = -\cos v \cdot \vec{p}(u, v)$.

Цилиндр $F(x, y, z) = x^2 + y^2 - 1$, ось которого совпадает с осью z , поперечное сечение является окружностью единичного радиуса, а сам цилиндр простирается вдоль оси z от 0 до 1 называется «образцовым цилиндром».

Удобно рассматривать такой цилиндр в качестве представителя семейства конических цилиндров

$$F(x, y, z) = x^2 + y^2 - (1 + (s - 1)z)^2 \text{ при } 0 < z < 1.$$

При $s = 0 \Rightarrow F(x, y, z) = x^2 + y^2 - (1 - z)^2$ - базовый конус с $s(0,0,1)$

При $s = 1 \Rightarrow F(x, y, z) = x^2 + y^2 - 1$ - базовый цилиндр.

$$P(u, v) = ((1 + (s - 1)v) \cos u, (1 + (s - 1)v) \sin u, v) \\ \begin{cases} 0 < u \leq 2\pi \\ 0 < v < 1 \end{cases}$$

Если необходимо представить конический цилиндр в виде монолитного объекта, на его концы добавляют два круга, основание – единичного радиуса ($z = 0$) и крышку – радиуса $s(z = 1)$.

Нормальные векторы к базовым поверхностям

Поверхность	$n(u, v)$ в точке $P(u, v)$	$F(x, y, z)$
Сфера	$p(u, v)$	(x, y, z)
Конический цилиндр	$(\cos u, \sin u, 1 - s)$	$(x, y, -(s - 1)(1 + (s - 1)z))$
Цилиндр	$(\cos u, \sin u, 0)$	$(x, y, 0)$
конус	$(\cos u, \sin u, 1)$	$(x, y, 1 - z)$

$$\vec{p} \cdot \vec{a} = 0 \Rightarrow \vec{a} \perp \vec{p}$$

Пусть A фиксированная точка с радиус-вектором \vec{a} .

P - произвольная точка \vec{p} .

$$\begin{aligned} \vec{p} \cdot \vec{a} &= |\vec{p}||\vec{a}| \cos \theta \\ |\vec{p} \times \vec{a}| &= |\vec{p}||\vec{a}| \sin \theta \\ \vec{p} \cdot \vec{a} &= |a| \\ |\vec{p} \times \vec{a}| &= |a| \\ \vec{p} \cdot \vec{a} &= \vec{p} \cdot \vec{p} \\ \vec{p} \cdot \vec{a} &= |a| \cdot \vec{p} / 2 \end{aligned}$$

Поверхности вращения

Поверхности вращения образуются вращательным перемещением образующей вокруг неподвижной оси. Таким образом получают: сфера – при вращении окружности вокруг своего диаметра, тор – если образующая окружность вращается вокруг оси, лежащей в плоскости окружности и не проходящей через ее центр, а также эллипсоиды, гиперболоиды, параболоиды, конусы и цилиндры вращения, псевдосфера.

Пусть образующую поместили в плоскость XOZ и задали параметрически: $C(v) = (x(v)y(v))$. Для создания поверхности этот профиль будет поворачиваться вокруг оси z , меняя параметр u , где u определяет угол, на который каждая точка образующей повернута относительно оси. Различные положения кривой C называются меридианами. Когда точка $(x(v), 0, z(v))$ поворачивается на u радиан, она становится точкой $(x(v) \cos u, y(v) \sin u, z(v))$. Полный поворот кривой образует полный круг. Следовательно, контуры при постоянном v являются окружностями и называются параллелями. Параллель для каждого значения v имеет радиус $x(v)$ и располагается на высоте $z(v)$ над плоскостью xOy . При этом, произвольная точка поверхности задается выражением

$$P(u, v) = (x(v) \cos u, y(v) \sin u, z(v)).$$

Чтобы получить уравнение поверхности вращения (в неявном виде) образуемой вращением кривой $f(x, z) = 0$ вокруг оси z надо заменить в уравнении этой линии x на $\sqrt{x^2 + y^2}$. рассмотрим пример.

Окружность $l: (x - R)^2 + z^2 = r^2$, лежащая в плоскости XOZ вращается вокруг оси $i \equiv Oz$.

Заменяем $x \Rightarrow \sqrt{x^2 + y^2}$. Получаем

$$\begin{aligned} & (\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2 \Rightarrow \\ & x^2 + y^2 + z^2 + R^2 - r^2 = 2R\sqrt{x^2 + y^2} \Rightarrow \\ & (x^2 + y^2 + z^2 + R^2 - r^2)^2 = 4R^2(x^2 + y^2) - \text{уравнение тора.} \end{aligned}$$

Если: $R > r$ - открытый тор (рис.)

$$R < r$$

$R = 0$ - закрытый тор

$R = r$ - сфера.

Прямая $l: \frac{z}{H} + \frac{x}{R} = 1$, лежащая в плоскости XOZ вращается вокруг оси $i \equiv Oz$.

$$\begin{aligned} & \text{Заменяем } x \Rightarrow \sqrt{x^2 + y^2}, \text{ получим } \frac{z}{H} + \frac{\sqrt{x^2 + y^2}}{R} = 1 \Rightarrow \left(\frac{\sqrt{x^2 + y^2}}{R}\right)^2 - \\ & \left(1 - \frac{z}{H}\right)^2 = 0 \\ & \frac{x^2 + y^2}{R^2} - \frac{(H-z)^2}{H^2} = 0 - \text{уравнение прямого кругового конуса.} \end{aligned}$$

Псевдосфера – образуется вращением трактрисы вокруг ее базы. Если трактриса задана в плоскости параметрическими уравнениями $x = a \sin u$, $y = 0$, $z = a \left(\ln \tan \frac{u}{2} + \cos u \right)$, то параметрическое уравнение псевдосферы будет: $x = a \sin u \cos v$, $y = a \sin u \sin v$, $z = a \left(\ln \tan \frac{u}{2} + \cos u \right)$. Кривизна у нее постоянная, отрицательная и равна $-1/a^2$.

Нормальный вектор в этом случае задается выражением:

$$\vec{n}(u, v) = x(v)(\dot{z}(v) \cos u, \dot{z}(v) \sin u, -x(v)).$$

Для тора: профиль $C(v) = (D + A \cos v, A \sin v)$;

$$P(u, v) = ((D + A \cos v) \cos u, (D + A \cos v) \sin u, A \sin v);$$

$$\vec{n}(u, v) = (\cos v \cos u, \cos v \sin u, \sin v)(D + A \cos v)$$

Чтобы построить поверхность вращения в случае профиля, заданного дискретными точками. Определим -ю вершину:

$$P_{ij} = (x_j \cos u_i, y_j \sin u_i, z_j)$$

Нормальные векторы к поверхностям 2-го порядка.

Уравнение поверхности второго порядка – квадратичная форма. Получение градиента для прохождения нормалей достаточно просто. В параметрической форме достаточно подставить $x(u, v)$ вместо x . следовательно, градиент от $F(x, y, z)$ для эллипсоида $F(2x, 2y, 2z)$.

$$\vec{n}(u, v) = (\cos v \cos u, \cos v \sin u, \sin v)$$

Суперквадрики (табл. 10.2)

Таблица 10.2

Суперэллипсоид	$(x^n + y^n)^{\frac{m}{n}} + z^m - 1$	$\cos^{\frac{2}{m}} v \cos^{\frac{2}{n}} u, \cos^{\frac{2}{m}} v \sin^{\frac{2}{n}} u, \sin^{\frac{2}{m}} v$
Однополосный супергиперboloид	$(x^n + y^n)^{\frac{m}{n}} - z^m - 1$	$\sec^{\frac{2}{m}} v \cos^{\frac{2}{n}} u, \sec^{\frac{2}{m}} v \sin^{\frac{2}{n}} u, \operatorname{tg}^{\frac{2}{m}} v$
Двухполосный супергиперboloид	$(x^n - y^n)^{\frac{m}{n}} - z^m - 1$	$\sec^{\frac{2}{m}} v \sec^{\frac{2}{n}} u, \sec^{\frac{2}{m}} v \operatorname{tg}^{\frac{2}{n}} v, \operatorname{tg}^{\frac{2}{m}} v$
Супертороид	$\left((x^n - y^n)^{\frac{1}{n}} - d \right)^{\frac{m}{n}} + z^m - 1$	$(d + \cos^{\frac{2}{m}} v) \cos^{\frac{2}{n}} u,$ $(d + \cos^{\frac{2}{m}} v) \sin^{\frac{2}{n}} u, \sin^{\frac{2}{m}} v$

Поверхности на базе явных функций двух переменных

Многие формы поверхности однозначны в одном направлении (измерении). Можно сказать, что высота поверхности в точке (x, z) есть некоторая $f(x, z)$ - поле высот, и задают в виде формулы следующего типа

$$f(x, z) = e^{-ax^2 - bz^2}$$

$a, b = \text{const}$; или в виде

$$f(x, z) = \frac{\sin \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}$$

С круговой симметрией

В параметрическом виде $P(u, v) = (u, f(u, v), v)$, а нормальный вектор имеет вид:

$$\vec{n}(u, v) = \left(-\frac{\partial f}{\partial u}, 1, \frac{\partial f}{\partial v} \right)$$

Здесь u - контуры располагаются в плоскостях с постоянным x , а v - контуры - в плоскостях с постоянным z . Каждая линия является следом пересечения поверхности плоскостью $x = k$ или $z = k$.

Используются при иллюстрации поведения математических формул.

10.6 Поверхности второго порядка

Таблица 10.3

Эллипсоид	$x^2 + y^2 + z^2 - 1$	$\cos v \cos u, \cos v \sin u, \sin v$	$-\frac{\pi}{2}, \frac{\pi}{2}; -\pi, \pi$
Однополосный Гиперболоид	$x^2 + y^2 - z^2 - 1$	$\sec v \cos u, \sec v \sin u, \operatorname{tg} v$	$-\frac{\pi}{2}, \frac{\pi}{2}; -\pi, \pi$
Двуполосный гиперболоид	$x^2 - y^2 - z^2 - 1$	$\sec v \cos u, \sec v \operatorname{tg} u, \operatorname{tg} v$	$-\frac{\pi}{2}, \frac{\pi}{2}; -\pi, \pi$ $-\frac{\pi}{2}, \frac{3\pi}{2}$
Эллиптический Конус	$x^2 + y^2 - z^2$	$v \cos u, v \sin u, v$	$-\infty, \infty; (-\pi, \pi)$
Эллиптический Параболоид	$x^2 + y^2 + z^2$	$v \cos u, v \sin u, v^2$	$v \geq 0; (-\pi, \pi)$
Гиперболический параболоид	$-x^2 + y^2 - z^2$	$v \operatorname{tg} u, v \sec u, v^2$	$v \geq 0; (-\pi, \pi)$

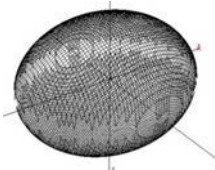
Замена знака одного из членов внешне-внутренней функции приводит к замене $\cos(\cdot)$ на $\sin(\cdot)$, а $\sin(\cdot)$ на $\operatorname{tg}(\cdot)$ в параметрических формах.

Базовые поверхности промасштабированы вдоль осей x, y, z масштабными множителями a, b, c соответственно.

След – кривая, образующаяся при сечении поверхности плоскостью.

Главный след – сечение поверхности плоскостями уровня

$$z = k; y = k; x = k; k = \text{const.}$$



Эллипсоид – все следы эллипсы. Если два параметра (из a, b, c) равны, то – поверхность вращения, если $a = b = c$ - сфера.

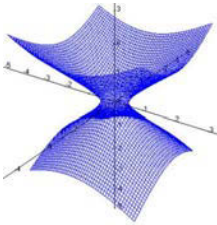
Однополосный гиперboloид – если $a = b$ - поверхность вращения гиперболаы вокруг z . Главные следы для $z = k$ - эллипсы, а для $z = k, y = k$ - гиперболаы. Однополосный гиперboloид – линейчатая поверхность.

$$x^2 + y^2 - z^2 - 1 = (x^2 - z^2) + (y^2 - 1) \Rightarrow (x - z)(x + z) = (1 - y)(1 + y)$$

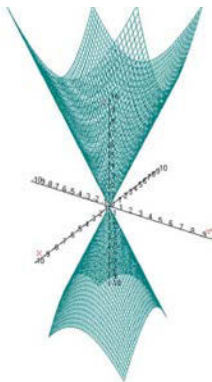
$$\Rightarrow$$

$$x - z = A(1 - y) \Rightarrow A(x + z) = 1 + y \text{ где } A = \text{const.}$$

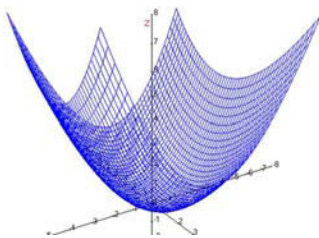
Такая поверхность может получиться, если переплести нитью два параллельных эллипса



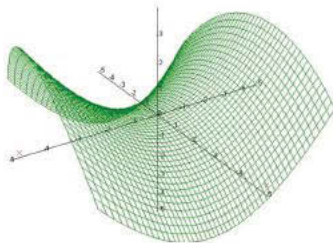
Двуполосный гиперboloид. Между $-a, a$ поверхность отсутствует. При $a = b$ - поверхность вращения. Следы при $x = k, |k| > a$ - эллипсы, остальные главные следы гиперболаы.



Эллиптический конус – линейчатая поверхность. Главные следы для $z = k$ - эллипсы. При $a = b$ - поверхность вращения – прямой круговой конус.



Эллиптический параболоид. Следы для $z = k > 0$ - эллипсы, остальные – параболы. При $a = b$ - поверхность вращения.



Гиперболический параболоид – линейчатая поверхность. Следы $z = k (k \neq 0)$ – гиперболы; $x = k, y = k$ – параболы
 $z = 0 \Rightarrow (x + y)(y - x) = 0 \Rightarrow x + y = 0, y - x = 0$, где x прямые.

10.7. Билинейно сопряженные поверхности (лоскуты Кунса)

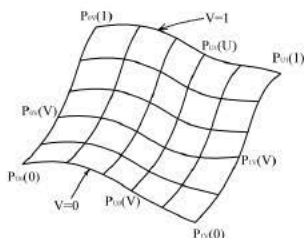


Рис. 10.18

Обобщение линейчатой поверхности интерполирующей две граничные кривые $P_0(u)$ и $P_1(u)$ является билинейно сопряженный лоскут, интерполирующий четыре граничные кривые: лоскуты Кунса (рис. 10.18). Семейство таких лоскутов впервые было исследовано Кунсом.

Обозначим граничные кривые $P_{uv}, P_{u1}, P_{0v}, P_{1v}$. Эти кривые пересекаются в углах лоскута (где u и v принимают значения 0 или 1). При других значениях параметров u и v кривые могут иметь любую форму. Следовательно,

данный пример является обобщением билинейного лоскута, границы которого прямые линии. Выведем формулу для поверхности $P(u, v)$, осуществляющей плавный переход от одной граничной кривой к другой, по мере изменения значений параметров u и v .

Скомбинируем линейчатый лоскут, построенный из $P_{u0}(u)$ и $P_{u1}(u)$ с линейчатой поверхностью, построенной из $P_{0v}(v)$ и $P_{1v}(v)$. Простое сложение этих поверхностей не даст результата (не удастся правильно интерполировать четыре кривые, кроме того, эта операция некорректна, поскольку дает не аффинную комбинацию точек). Выход заключается в том, чтобы сложить эти поверхности, а затем вычесть билинейный лоскут, построенный на четырех углах граничных кривых. Формула для данного лоскута имеет вид:

$$P(u, v) = [P_{0v}(v)(1 - u) + P_{1v}(v)u] + [P_{u0}(u)(1 - v) + P_{u1}(u)v] - [(1 - u)(1 - v)P_{0v}(0) + u(1 - v)P_{1v}(0) + v(1 - u) + uvP_{1v}(1)]. \quad 10.9$$

Пусть необходимая сетка кривых построена каким-либо методом. Для большей общности будем сначала применять параметрическое представление, введя семейство продольных и поперечных кривых, параметры этих кривых обозначим через u и v соответственно. Затем, полученные результаты применим к непараметрическому случаю.

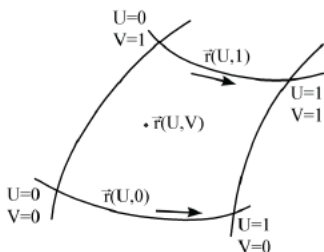


Рис.10.19. Параметрическое задание порции поверхности

Сетка кривых делит поверхность на совокупность топологически прямоугольных порций, каждая из которых ограничена парой u -кривых и парой v -кривых (рис. 10.19). Пусть u и v изменяются в пределах $[0,1]$ вдоль соответствующих границ; тогда $\vec{r}(u, v)$ $0 < u, v < 1$, представляет внутренность лоскута поверхности, а $\vec{r}(u, 0)$, $\vec{r}(1, v)$, $\vec{r}(u, 1)$, $\vec{r}(0, v)$ - представляют четыре известные граничные кривые. Тем самым задача

определения лоскута поверхности сводится к нахождению функции $\vec{r}(u, v)$ с подходящим типом «хорошего поведения», которая при $u = 0, u = 1, v = 0, v = 1$ представляет соответствующую кривую.

Упростим задачу и рассмотрим случай, когда заданы две границы лоскута $\vec{r}(0, v)$ и $\vec{r}(1, v)$. Применяя линейную интерполяцию в u -направлении, получим линейчатую поверхность (рис. 10.20 а)

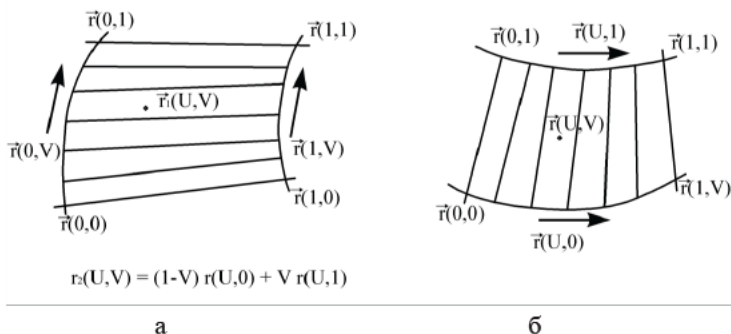
$$\vec{r}_1(u, v) = (1 - u)\vec{r}(0, v) + u\vec{r}(1, v)$$

Или в точечном представлении.

$$P_1(u, v) = (1 - u)P_{0v}(v) + uP_{1v}(u)$$

Линейная интерполяция в v -направлении дает линейчатую поверхность (рис. 10.20 б), удовлетворяющую двум другим граничным условиям:

$$\vec{r}_2(u, v) = (1 - u)\vec{r}(u, 0) + v\vec{r}(u, 1)$$



$$r(U, V) = (1 - V)r(U, 0) + V r(U, 1)$$

а

б

Рис.10.20.

Сумма $\vec{r}_1 + \vec{r}_2$ представляет лоскут поверхности, каждая из границ которой является суммой заданной граничной кривой и прямолинейного отрезка, соединяющего концевые точки этой кривой. Например, граница, соответствующая $v = 0$, определяется не вектор-функцией $\vec{r}(u, 0)$, а выражением

$$\vec{r}(u, 0) + [(1 - u)\vec{r}(0, 0) + u\vec{r}(1, 0)].$$

Если каким-либо способом определить лоскут поверхности $\vec{r}_3(u, v)$, границами которого будут вышеупомянутые прямолинейные отрезки, то можно восстановить первоначальные граничные кривые, составив выражение $\vec{r}_1 + \vec{r}_2 - \vec{r}_3$. Построить \vec{r}_3 не сложно, его границы, отвечающие

$v = 0$ и $u = 0$, будут определяться выражениями

$$(1 - u)\vec{r}(0,0) + u\vec{r}(1,0), (1 - u)\vec{r}(0,1) + u\vec{r}(1,1)$$

Соответственно, и последующая интерполяция в $-$ направлении даст

$$\vec{r}_3(u, v) = (1 - u)(1 - v)\vec{r}(0,0) + u(1 - v)\vec{r}(1,0) + (1 - u)v\vec{r}(0,1) + uv\vec{r}(1,1),$$

Вектор $\vec{r} = \vec{r}_1 + \vec{r}_2 - \vec{r}_3$, определяющий искомую поверхность, представим в матричном виде

$$\vec{r}(u, v) = [(1 - u)u] \begin{bmatrix} \vec{r}(0, v) \\ \vec{r}(1, v) \end{bmatrix} + [\vec{r}(u, 0)\vec{r}(u, 1)] \begin{bmatrix} 1 - v \\ v \end{bmatrix} - [(1 - u)u] \begin{bmatrix} \vec{r}(0,0) & \vec{r}(0,1) \\ \vec{r}(1,0) & \vec{r}(1,1) \end{bmatrix} \begin{bmatrix} 1 - v \\ v \end{bmatrix} \quad (10.10)$$

Последовательные подстановки $u = 0, u = 1, v = 0, v = 1$ подтверждают, что лоскут поверхности, определенный вектором $\vec{r}(u, v)$ (10.21), имеет своими границами четыре первоначальные кривые.

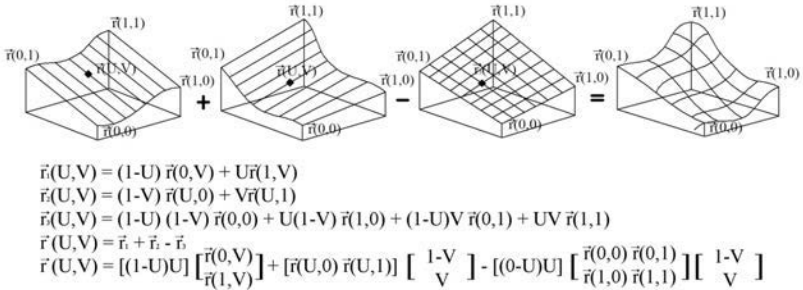


Рис. 10.21

Эта конструкция, использующая только информацию о границах поверхности и некоторые вспомогательные скалярные функции переменных u и v , определяет наиболее простой класс поверхностей – поверхности Кунца. Вспомогательные функции $u, (1 - u), v, (1 - v)$ называются функциями смешивания, т.к. они соединяют воедино четыре отдельные граничные кривые и в результате дают одну корректно определенную поверхность.

Линейные функции смешения в уравнении лоскута возникают в результате использования равномерной линейной интерполяции. Заменяем $(1 - u)u$ на $\alpha_0(u), \alpha_1(u)$. При этом функции смешения будут любыми, но для них должны выполняться соотношения

$$1 - \alpha_1 = \alpha_0 \text{ и } \begin{cases} \alpha_0(0) = 1 & \alpha_0(1) = 0 \\ \alpha_1(0) = 0 & \alpha_1(1) = 1 \end{cases}$$

При такой замене интерполяция в $-$ направлении останется линейной, но скорость движения конца вектора \vec{r} вдоль линии интерполяции не останется постоянной при равномерном возрастании u от 0 до 1. Соответствующая замена $(1 - v)v$ на $\alpha_0(v), \alpha_1(v)$ приводит к формуле

$$\begin{aligned} \vec{r}(u, v) = & \\ \alpha_0(u), \alpha_1(u) \begin{bmatrix} \vec{r}(0, v) \\ \vec{r}(1, v) \end{bmatrix} + [\vec{r}(u, 0)\vec{r}(u, 1)] \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \end{bmatrix} - & \\ - [\alpha_0(u), \alpha_1(u)] \begin{bmatrix} \vec{r}(0, 0) & \vec{r}(0, 1) \\ \vec{r}(1, 0) & \vec{r}(1, 1) \end{bmatrix} \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \end{bmatrix} & \end{aligned} \quad (10.11)$$

Функции смещения α_0 и α_1 выбираются, как правило, непрерывными и монотонными (непрерывно возрастающими или непрерывно убывающими) на интервале $0 \leq u \leq 1$. На практике, по аналитическим и вычислительным соображениям, используются полиномиальные функции смещения. Условие $1 - \alpha_1 = \alpha_0$ можно опустить, если α_0 и α_1 удовлетворяют другим наложенным условиям. Однако в этом случае интерполяции между противоположными границами $\vec{r}_1(u, v)$ и $\vec{r}_2(u, v)$ являются нелинейными, а лоскут, представленный формулой (3), не становится плоским, когда все его границы компланарны (лежат в одной плоскости).

Имея сетку кривых, можно сконструировать составную поверхность из лоскутов описанного типа. Однако на границах лоскутов эта поверхность будет только непрерывной, но не гладкой. Непрерывность градиента (гладкость), существенная для большинства приложений, достигается более сложным определением лоскута не только через его граничные кривые, но также четырех граничные наклоны в направлениях, трансверсальных границам, т.е. через $\vec{r}_v(u, 0), \vec{r}_u(1, 0), \vec{r}_v(u, 1), \vec{r}_u(0, v)$.

Можно автоматически обеспечить непрерывность градиента поперек границ элементарных лоскутов подходящим выбором функций смещения. Но это уменьшает гибкость, а искомая поверхность имеет нежелательные плоские области.

Используя вместо обобщенной линейной интерполяции обобщенную интерполяцию Эрмита, выведем уравнение для лоскута поверхности:

$$\vec{r}_1(u, v) = \alpha_0(u)\vec{r}(0, v) + \alpha_1(u)\vec{r}(1, v) + \beta_0(u)\vec{r}_u(0, v) + \beta_1(u)\vec{r}_u(1, v).$$

Поверхность интерполирует две граничные кривые $\vec{r}(0, v)$ и $\vec{r}(1, v)$ и имеет заданные наклоны поперек границ, если четыре функции смещения удовлетворяют соотношениям

$$\begin{aligned} & \begin{cases} \alpha_0(0) = 1 & \alpha_0(1) = 0 \\ \alpha_1(0) = 0 & \alpha_1(1) = 1 \end{cases} \\ & \alpha_0'(0) = \alpha_0'(1) = \alpha_1'(0) = \alpha_1'(1) = 0 \end{aligned}$$

$$\begin{aligned} \beta_0(0) = \beta_0(1) = \beta_1(0) = \beta_1(1) = 0 \\ \begin{cases} \beta'_0(0) = 1 & \beta'_0(1) = 0 \\ \beta'_1(0) = 0 & \beta'_1(1) = 1 \end{cases} \end{aligned} \quad (10.11)$$

Две кривые $\vec{r}(u, 0)$ и $\vec{r}(u, 1)$ с соответствующими данными о наклонах интерполируются формулой:

$$\vec{r}_2(u, v) = \alpha_0(v)\vec{r}(u, 0) + \alpha_1(v)\vec{r}(u, 1) + \beta_0(v)\vec{r}_v(u, 0) + \beta_1(v)\vec{r}_v(u, 1) \quad (10.12)$$

Очевидно, как и в случае линейной, $\vec{r}_1 + \vec{r}_2$ не дает требуемой поверхности. Чтобы получить ее необходимо вычесть дополнительные слагаемые, получаемое с помощью того же метода интерполяции в обоих направлениях при использовании только информации от угловых точках. Искомое уравнение имеет вид:

$$\begin{aligned} \vec{r}(u, v) = & [\alpha_0(u)\alpha_1(u)\beta_0(u)\beta_1(u)] \begin{bmatrix} \vec{r}(0, v) \\ \vec{r}(1, v) \\ \vec{r}(0, v) \\ \vec{r}(1, v) \end{bmatrix} + [\vec{r}(u, 0)\vec{r}(u, 1)\vec{r}_v(u, 0)\vec{r}_v(u, 1)] \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix} - \\ & - [\alpha_0(u), \alpha_1(u), \beta_0(u), \beta_1(u)] \begin{bmatrix} \vec{r}(0,0) & \vec{r}(0,1) & \vec{r}_v(0,0) & \vec{r}_v(0,1) \\ \vec{r}(1,0) & \vec{r}(1,1) & \vec{r}_v(1,0) & \vec{r}_v(1,1) \\ \vec{r}_u(0,0) & \vec{r}_u(0,1) & \vec{r}_{uv}(0,0) & \vec{r}_{uv}(0,1) \\ \vec{r}_u(1,0) & \vec{r}_u(1,1) & \vec{r}_{uv}(1,0) & \vec{r}_{uv}(1,1) \end{bmatrix} \times \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix} \end{aligned} \quad (10.13)$$

Для построения \vec{r}_3 необходимы значения смешанных производных в углах \vec{r}_{uv} . Поверхность имеет заданные граничные кривые и градиенты в направлениях, перпендикулярных границам. Используя такие лоскуты, можно сделать направления нормалей к ним непрерывными на границах и построить всюду гладкую составную поверхность.

Установим, какой геометрический смысл имеет смешанная производная \vec{r}_{uv} , входящая в уравнение (10.14). Сначала рассмотрим поверхность $z = z(x, y)$ в декартовых координатах. Т.к. $z_{xy} = \left(\frac{\partial}{\partial x}\right)\left(\frac{\partial z}{\partial y}\right)$, то очевидно, что смешанная производная измеряет скорость изменения в x -направлении наклона поверхности в y направлении, т.е. z_{xy} является мерой закрученности поверхности. По аналогии, параметрическая смешанная производная \vec{r}_{uv} часто называется вектором кручения.

Значение \vec{r}_{uv} в точке параметрически заданной поверхности зависит не столько от геометрических свойств этой поверхности, сколько от способа ее параметризации. Рассмотрим простой пример. Пусть $\vec{a}, \vec{b}, \vec{c}$ -постоянные векторы, такие, что $\vec{b} \times \vec{c} \neq 0$, а u, v -любые скаляры. Тогда $\vec{r}_{uv} = \vec{a} + u\vec{b} + v\vec{c}$ определяет плоскость, проходящую через точку \vec{a} и параллельную векторам \vec{b} и \vec{c} . Плоскость незакрученная поверхность, $\vec{r}_{uv} = 0$. Рассмотрим теперь выражение

$\vec{r}_{uv} = \vec{a} + u\vec{b} + uv\vec{c}$. Эта новая параметризация определяет ту же плоскость, но $\vec{r}_{uv} = \vec{c} \neq 0$.

Т.е. $\vec{r}_{uv} \neq 0$ необязательно означает закрученность поверхности. Можно существенно упростить уравнение лоскута по Кунцу подходящим определением граничных кривых и градиентов в ортогональных к ним направлениях. Представим сегмент кривой с помощью положений его концевых точек и значений касательных в них:

$$\vec{r}(u) = \alpha_0(u)\vec{r}(0) + \alpha_1(u)\vec{r}(1) + \beta_0(u)\vec{r}'(0) + \beta_1(u)\vec{r}'(1)$$

Это позволяет записать граничные кривые и поперечные градиенты в виде:

$$\begin{aligned} \vec{r}(i, v) &= \alpha_0(v)\vec{r}(i, 0) + \alpha_1(v)\vec{r}(i, 1) + \beta_0(v)\vec{r}'_v(i, 0) + \beta_1(v)\vec{r}'_v(i, 1); \\ \vec{r}'_u(i, v) &= \alpha_0(v)\vec{r}'_u(i, 0) + \alpha_1(v)\vec{r}'_u(i, 1) + \beta_0(v)\vec{r}'_{uv}(i, 0) + \beta_1(v)\vec{r}'_{uv}(i, 1); \\ \vec{r}(u, j) &= \alpha_0(u)\vec{r}(0, j) + \alpha_1(u)\vec{r}(1, u) + \beta_0(u)\vec{r}'_u(0, j) + \beta_1(u)\vec{r}'_u(1, j); \\ \vec{r}'_v(u, j) &= \alpha_0(u)\vec{r}'_v(0, j) + \alpha_1(u)\vec{r}'_v(1, u) + \beta_0(u)\vec{r}'_{uv}(0, j) + \beta_1(u)\vec{r}'_{uv}(1, j); \end{aligned}$$

где $i, j = 0, 1$. После подстановки в () этих выражений оказывается, что все три члена одинаковые за исключением того, что третий имеет отрицательный знак. В результате получим:

$$\vec{r}(u, v) = [\alpha_0(u)\alpha_1(u)\beta_0(u)\beta_1(u)] \times \begin{bmatrix} \vec{r}(0,0) & \vec{r}(0,1) & \vec{r}'_v(0,0) & \vec{r}'_v(0,1) \\ \vec{r}(1,0) & \vec{r}(1,1) & \vec{r}'_v(1,0) & \vec{r}'_v(1,1) \\ \vec{r}'_u(0,0) & \vec{r}'_u(0,1) & \vec{r}'_{uv}(0,0) & \vec{r}'_{uv}(0,1) \\ \vec{r}'_u(1,0) & \vec{r}'_u(1,1) & \vec{r}'_{uv}(1,0) & \vec{r}'_{uv}(1,1) \end{bmatrix} \cdot \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix}$$

Поверхности Фергюсона

Простейший набор полиномиальных функций смешения, удовлетворяющий соотношениям:

$$\begin{cases} \alpha_0(0) = 1 & \alpha_0(1) = 0 \\ \alpha_1(0) = 0 & \alpha_1(1) = 1 \\ \alpha_0'(0) = \alpha_0'(1) = \alpha_1'(0) = \alpha_1'(1) = 0 \\ \beta_0(0) = \beta_0(1) = \beta_1(0) = \beta_1(1) = 0 \\ \beta'_0(0) = 1 & \beta'_0(1) = 0 \\ \beta'_1(0) = 0 & \beta'_1(1) = 1 \end{cases}$$

есть следующий набор полиномов третьей степени

$$\begin{aligned} \alpha_0(u) &= 1 - 3u^2 + 2u^3, \alpha_1(u) = 3u^2 - 2u^3 \\ \beta_0(u) &= u - 2u^2 + u^3, \beta_1(u) = -u^2 + u^3 \end{aligned}$$

подставив их в уравнение

$$\vec{r}(u) = \alpha_0(u)\vec{r}(0) + \alpha_1(u)\vec{r}(1) + \beta_0(u)\vec{r}'(0) + \beta_1(u)\vec{r}'(1)$$

найдем, что получается в точности сегмент Фергюсона кубической кривой

$$\vec{r}(u) = (1 - 3u^2 + 2u^3)\vec{r}(0) + (3u^2 - 2u^3)\vec{r}(1) + (u - 2u^2 + u^3)\dot{\vec{r}}(0) + (-u^2 + u^3)\dot{\vec{r}}(1)$$

Использование этого набора функций смещения приводит к тому, что составная поверхность строится на сетке кривых, составленной из сегментов Фергюсона. Можно записать векторы функций смещения в следующем виде:

$$F(u) = [\alpha_0(u)\alpha_1(u)\beta_0(u)\beta_1(u)] = [1 \ u \ u^2 \ u^3] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -3 & 3 & -2 & -1 \\ 2 & -2 & 1 & 1 \end{bmatrix} = UC$$

$$F^T(v) = \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix} = \begin{bmatrix} 1 & 0 & -3 & 2 \\ 0 & 0 & 3 & -2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ v \\ v^2 \\ v^3 \end{bmatrix} = C^T V$$

и лоскут поверхности по Фергюсону задается формулой

$$\vec{r}(u, v) = UCQC^T V$$

где матрица Q имеет вид:

$$Q = \begin{bmatrix} \vec{r}(0,0) & \vec{r}(0,1) & \vec{r}_v(0,0) & \vec{r}_v(0,1) \\ \vec{r}(1,0) & \vec{r}(1,1) & \vec{r}_v(1,0) & \vec{r}_v(1,1) \\ \vec{r}_u(0,0) & \vec{r}_u(0,1) & \vec{r}_{uv}(0,0) & \vec{r}_{uv}(0,1) \\ \vec{r}_u(1,0) & \vec{r}_u(1,1) & \vec{r}_{uv}(1,0) & \vec{r}_{uv}(1,1) \end{bmatrix}$$

CQC^T - суть линейные комбинации значений векторов положения и векторов производных в углах лоскута поверхности.

Условие $\vec{r}(u, v) = 0$ во всех узлах сетки может привести к местным уплощениям синтезируемой поверхности вблизи углов лоскута. Однако этот эффект сводится к минимуму, если линии сетки проводить примерно по линиям кривизны поверхности, т.к. поверхность вдоль таких линий не закручивается.

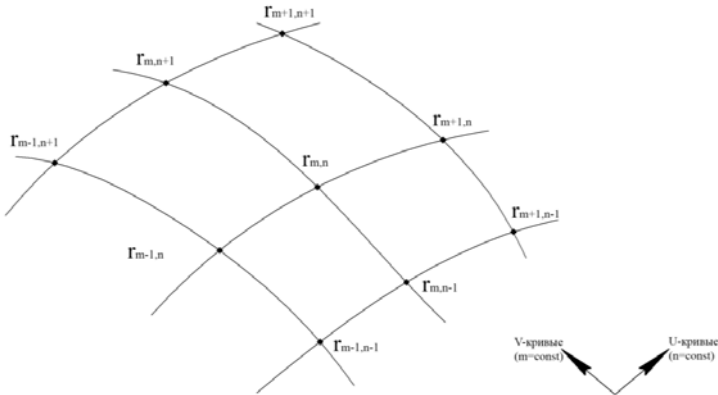


Рис. 10.22

Если предположить, что имеются нулевые векторы кручения в узлах лоскута поверхности, то возникает проблема, как их определить.

Предположим, что узлы сетки кривых имеют векторы положения \vec{r}_{mn} , где $m = 0, 1, \dots, M; n = 0, 1, \dots, N$.

Если первые производные согласованы по величине и направлению в узловых точках, то вторые производные u -кривых будут непрерывными, если касательные векторы $\vec{r}_{u,mn}$ вычислены из системы уравнений:

$$\vec{r}_{u,m-1,n} + 4\vec{r}_{u,mn} + \vec{r}_{u,m+1,n} = 3(\vec{r}_{m+1,n} - \vec{r}_{m-1,n})$$

$$m = 1, 2, \dots, M - 1$$

Аналогично вычисляются вторые производные для v кривых:

$$\vec{r}_{v,m,n-1} + 4\vec{r}_{v,mn} + \vec{r}_{v,m,n+1} = 3(\vec{r}_{m,n+1} - \vec{r}_{m,n-1})$$

$$n = 1, 2, \dots, N - 1$$

Рассмотрим условия непрерывности $\vec{r}(u, v)$ на границе двух лоскутов, уравнения которых записаны в виде:

$$\vec{r}^1(u, v) = UCQ^1C^TV; \vec{r}^2(u, v) = UCQ^2C^TV$$

имеющих общую границу: $\vec{r}^1(1, v) = \vec{r}^2(0, v)$.

Дифференцируя дважды уравнения лоскутов по u , получим

$$\vec{r}_{uu}^1(1, v) = [0, 0, 2, 6]CQ^1C^TV$$

$$\vec{r}_{uu}^2(0, v) = [0, 0, 2, 0]CQ^2C^TV$$

Правые части являются кубическими полиномами от v . Для непрерывности \vec{r}_{uu} на границе они должны быть равны при $0 \leq v \leq 1$. Приравняв коэффициенты, получим

$$[0, 0, 2, 6]CQ^1C^T = [0, 0, 2, 0]CQ^2C^T$$

Т.к. C^T —неособая квадратная матрица, умножив справа на $(C^T)^{-1}$ получим

$$[6, -6, 2, 4]Q^1 = [-6, 6, -4, -2]Q^2$$

В скалярной форме это выражение соответствует системе четырех уравнений, содержащих значения $\vec{r}_0, \vec{r}_u, \vec{r}_v, \vec{r}_{uv}$ в узлах двух лоскутов. Эта система упрощается, т.к. $\vec{r}^1(1, 0) = \vec{r}^2(0, 0), \vec{r}^1(1, 1) = \vec{r}^2(0, 1)$,

$\vec{r}_u^1(1, 0) = \vec{r}_u^2(0, 0), \vec{r}_u^1(1, 1) = \vec{r}_u^2(0, 1)$ —условия непрерывности поверхности и ее градиента на границе двух лоскутов. Оказывается, что два из этих уравнений сводятся к условиям гладкости второго порядка составных граничных кривых поверхности, образованной лоскутами. Остающиеся два уравнения суть дополнительные условия, обеспечивающие гладкость второго порядка поверхности в целом в u -направлениях. Если они налагаются на всех v -границах большой составной поверхности, то их можно представить в виде:

$$\vec{r}_{uv,m-1,n} + 4\vec{r}_{uv,mn} + \vec{r}_{uv,m+1,n} = 3(\vec{r}_{v,m+1,n} - \vec{r}_{v,m-1,n})$$

$$m = 1, 2, \dots, M - 1$$

причем градиенты \vec{r}_0 в правой части известны. Можно добиться непрерывности \vec{r}_{uv} в -направлении для всей поверхности, потребовав выполнения аналогичных условий:

$$\vec{r}_{vu,m,n-1} + 4\vec{r}_{vu,mn} + \vec{r}_{vu,m,n+1} = 3(\vec{r}_{u,m,n+1} - \vec{r})$$

$$n = 1, 2, \dots, N - 1$$

Выполнение условия $\vec{r}_{uv} = \vec{r}_{vu}$, не очевидно. Если оно не выполняется, то гладкость второго порядка может быть получена только в одном направлении. Но эти два уравнения дают одно и то же множество векторов кручения, поэтому достижима общая гладкость второго порядка.

Построение всей поверхности с гладкостью второго порядка требует определения касательных векторов во всех узлах сетки на границах, но векторов кручения только в четырех угловых точках. Вся необходимая информация может быть записана в виде таблицы 10.4

Таблица 10.4

X_{00}	$S_{00}S_{01} \dots \dots \dots S_{0N}$	X_{0N}
t_{00}	$r_{00}r_{01} \dots \dots \dots r_{0N}$	t_{0N}
t_{10}	$r_{10}r_{11} \dots \dots \dots r_{1N}$	t_{1N}
...		...
...		...
...		...
t_{M0}	$r_{M0}r_{M1} \dots \dots \dots r_{MN}$	t_{MN}
X_{M0}	$S_{M0}S_{M1} \dots \dots \dots S_{MN}$	X_{MN}

Здесь $\vec{s} = \vec{r}_u, \vec{t} = \vec{r}_v, \vec{X} = \vec{r}_{uv}$.

Поверхность, получаемая в результате описанной процедуры, - частный случай параметрической сплайновой поверхности.

При формировании поверхности задаются позиции вершин и при помощи уравнения определяется форма поверхности.

Бикубическая ($M = 3, L = 3$) поверхность Безье.

$P(u, v)$

$$= (u^3, u^2, u, 1) \begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{pmatrix} \begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & -3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v^3 \\ v^2 \\ v \\ 1 \end{pmatrix}$$

Если необходимо составить сложную форму из нескольких фрагментов поверхности Безье так, чтобы эти лоскуты плавно стыковались на их общих границах, то для этого достаточно, чтобы вдоль границы совпадали их контрольные полигоны. Если каждые два ребра полиэдров, встречающихся на границе, коллинеарны, то этим обеспечивается непрерывность касательной, т.е. гарантирована непрерывность нормали к поверхности в месте соединения.

При проектировании поверхностей для обеспечения большего локального контроля вместо полиномов Берштейна можно использовать В-сплайн функции. При этом уравнение имеет следующий вид:

$$P(u, v) = \sum_{i=0}^M \sum_{k=0}^L P_{ik} N_{im}(u) N_{kn}(v),$$

где $N_{im}(u)$ и $N_{kn}(v)$ -базисные В-сплайн функции (возможно различных порядков), задаваемые уравнением

$$N_{km}(t) = \left(\frac{t - t_k}{t_{k+m-1} - t_k} \right) N_{t, m-1} + \left(\frac{t_{k+m} - t}{t_{k+m} - t_{k+1}} \right) N_{t+1, m-1}(t).$$

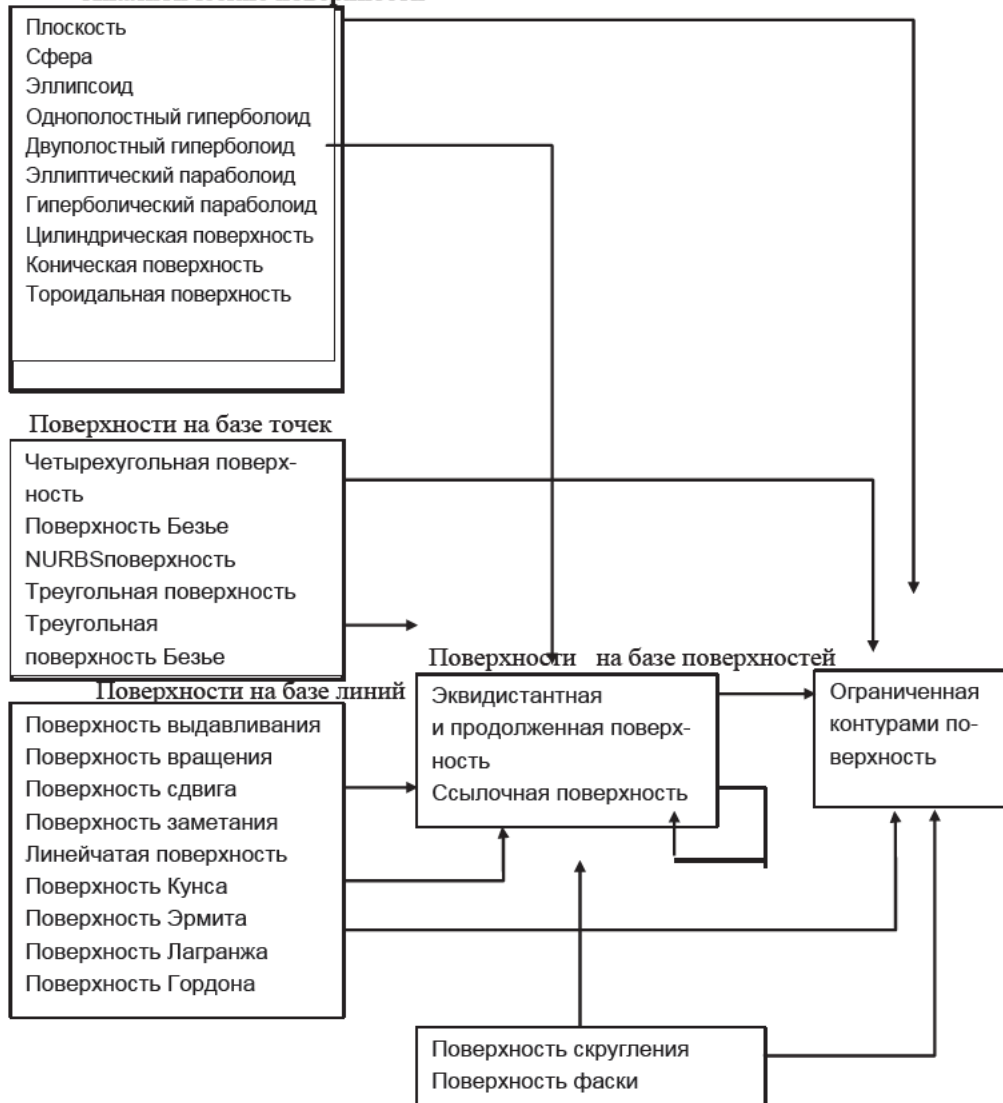
Обычно для обеих В-сплайн форм берется стандартный узловой вектор, так что углы полиэдра интерполируются точно. Удачным выбором являются кубические В-сплайны ($m = n = 4$). в этом случае нет ограничений на число контрольных точек (для кривых Безье число контрольных точек влияет на порядок полиномов). Для создания поверхности желаемой формы необходимо выбрать узловой полиэдр.

Распространим понятие *NURBS*- кривых на *NURBS* поверхности:

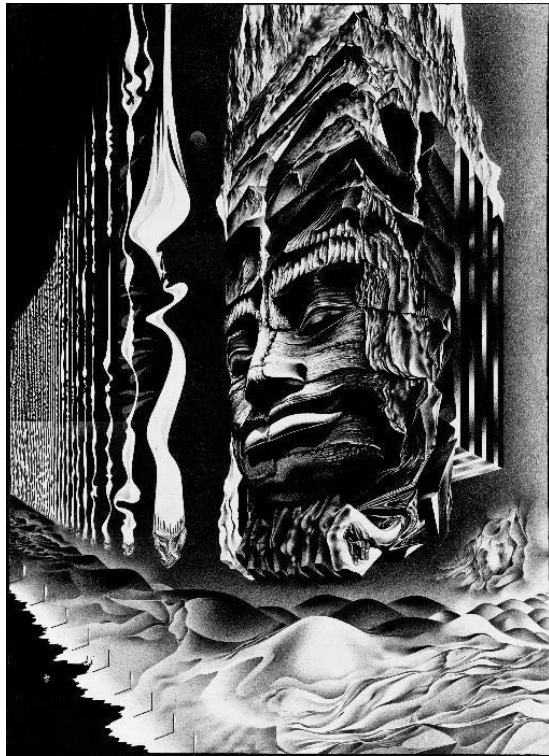
$$P(u, v) = \frac{\sum_{i=0}^M \sum_{k=0}^L \omega_{ik} P_{ik} N_{im}(u) N_{kn}(v)}{\sum_{i=0}^M \sum_{k=0}^L \omega_{ik} N_{im}(u) N_{kn}(v)}$$

Для контроля над формой поверхности задают контрольный полиэдр на базе контрольных точек P_{ik} , а также веса ω_{ik} . Если все веса одинаковы, то поверхность упрощается и становится В-сплайн поверхностью.

Аналитические поверхности



ЗАКЛЮЧЕНИЕ. ГЕОМЕТРИЧЕСКИЕ ОБЪЕКТЫ В ПРИРОДЕ,
МАТЕМАТИКЕ И ИСКУССТВЕ
(КРАТКИЕ ИСТОРИЧЕСКИЕ СВЕДЕНИЯ)



*Число живет в искусстве.
Августин Аврелий (354-430)
Геометрия есть прообраз красоты мира.
Иоганн Кеплер (1571-1630)*

*Кто из богов придумал этот знак?
Какое исцеленье от унынья
Дает мне сочетанье этих линий!
Расходится томивший душу мрак.
Всё проясняется как на картине.
И вот мне кажется, что сам я- бог
И вижу, символ мира разбирая,
Вселенную от края и до края
И.Ф. Гёте «Фауст».*

*Природа подражает себе. Зерно брошенное
в хорошую землю, производит плод. Мысль,
брошенная в здравый ум, тоже дает плод.
Числа подражают пространству, хотя они
столь непохожи на него по природе*

Паскаль

Одни геометрические формы представляют особый интерес, поскольку они отображают какой либо реальный объект, существующий в природе или спроектированный человеком на основе сложного сочетания целесообразности, простоты изготовления и интуитивного чувства, что это понравится потребителям, поскольку они способны доставлять эстетическое удовольствие. Если мы хотим написать подпрограмму для их рисования, нам необходимы методы их математического описания, методы преобразования этих форм из одного представления в другое, более подходящее для решения стоящей задачи. Другим видом являются «свободные формы», которые в большей степени базируются на данных, чем на математических выражениях. Но их тоже необходимо представлять в виде, удобном для машинной обработки, если они используются в программе.

Одним из основных элементов, образующих и описывающих эти геометрические формы являются линии. Поэтому, одним из основных объектов геометрического исследования и моделирования является линия. Это вызвано, прежде всего, тем, что само понятие линии возникло в процессе практической деятельности человека: при определении границ земельных участков, изучении траекторий движения тел (брошенного камня, струи воды, луча света), изготовлении чертежей и т.п. Очертания цветов и листьев растений, извилистая линия берега реки и моря, рисунки – все это линии. Ученые стремились точно определить, что такое линия как математическое понятие, т.е. что общего есть у всех вещей (в том числе и перечисленных выше) которые на практике называют линией.

Появление собственно человека происходило одновременно с зарождением искусства, религии, математических представлений

Лики С. Б. Происхождение человека.
«Курьер ЮНЕСКО» -1972, № 5

Геометрическое подобие, представляющее фундаментальную основу эволюции жизни и метод конструирования ею форм, стало инструментом и методом работы человеческого сознания.

Шевелев И. Ш. «О формообразовании в природе и искусстве»
«Золотое сечение: Три взгляда на природу гармонии».-1990

Интерес к кривым линиям начал проявляться еще у первобытного человека. Люди неолита и раннего бронзового века не располагали какими-либо современными математическими знаниями. Тем не менее, трудно не признать их

увлечения геометрией. Спирали, полукружья и другие фигуры на стоячих камнях и могильниках свидетельствуют о таком интересе. Особый интерес представляют каменные круги и кольца. «Нередко это правильные круги, аккуратно размеченные и аккуратно построенные. Такие каменные кольца бывают хорошо вычерченными эллипсами, и очень часто это модифицированные круги, либо сознательно уплощенные, так что фигура напоминает слегка приплюснутый апельсин, либо вытянутые, как яйцо».

Даже очень сложные фигуры можно создавать с помощью самых простых методов. «Примером фигуры, которую легче нарисовать, чем описать математически, может служить спираль, которая часто встречается в орнаментах, вырезанных на камнях. Хотя уравнение для спирали с постоянным шагом не очень сложно, понимание его требует знания аналитической геометрии. Однако нарисовать такую спираль очень просто. Возьмите кол, вбейте его в землю, и привяжите к нему веревку так, чтобы она по нему не скользила. К другому концу веревки привяжите кольцо и идите вокруг кола так, чтобы веревка наматывалась на него, держа ее все время натянутой. Острый кольцо опишет на земле спираль, витки которой постепенно будут становиться все меньше, пока веревка не намотается вся».

Уже на заре своего существования человек использовал кривые, во-первых, для украшения, то есть в искусстве и, во-вторых, для организации наблюдения за Солнцем, Луной и звездами, то есть в научных целях.

*Возьму линейку, проведу прямую,
И мигом круг квадратом обернется
Посередине рынок мы устроим.
А от него уж улицы пойдут-
Ну как на Солнце! Хоть оно само
И круглое, а ведь лучи прямые!...*

Аристофан “Птицы”

Изучение кривых продолжили древнегреческие ученые, заложивших основы математики как науки в современном смысле этого слова.

«Существование в доэллинический период развитой математики сегодня не может быть подвергнуто сомнению... Важнейшая самобытная особенность греков состоит именно в их сознательном стремлении расположить математические доказательства в такие цепочки, чтобы переход от одного звена к следующему не оставлял никакого места сомнениям и заставлял всех с ним согласиться. О том, что греческие математики, совсем как современные, в своих изысканиях пользовались скорее «эвристическими» рассуждениями, чем доказательствами, свидетельствует (если в том есть нужда), например «Трактат о методе» Архимеда» (Н. Бурбаки «теория множеств»).

Для математического изучения какого-либо объекта необходимо задать его аксиоматические, т.е. указать ряд свойств этого объекта, из которых можно было бы логически выводить другие его свойства.

В конце 4-го века до н.э. вся математика была собрана в трудах Евклида (около 365-300 г. до н.э.), прежде всего в «Началах», и притом так, что этому суждено было остаться образцом на целые тысячелетия.

Евклид определяет линию:

- как длину без ширины («Начала», определение 2)
- как границу поверхности (определение 6).

Здесь линия определяется через другие понятия, которые сами, в свою очередь, нуждаются в определении. Это, прежде всего понятие размерности, интуитивно очевидно, что точка имеет 0 размерность, линия 1, поверхность 2, тело 3. Однако, само понятие размерности пространства потребовало своего развития и уточнения. Более подробно мы рассмотрим этот вопрос при изучении фракталов.

«Первоначально занимались математикой, имея одну основную цель - понять, какое место во вселенной занимает человек в рамках некоторой рациональной схемы. Математика помогла найти порядок в хаосе, связать идеи в логические цепочки, обнаружить основные принципы. Она была наиболее теоретической из всех наук.» Стройк Д.Я. [7.с.48]

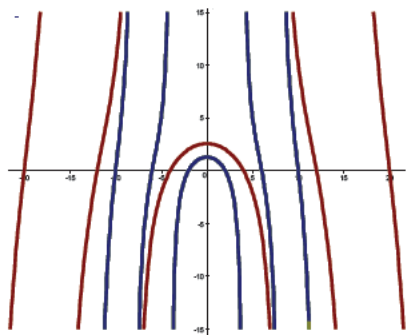


Рис.11.1. Квадратриса Динострата

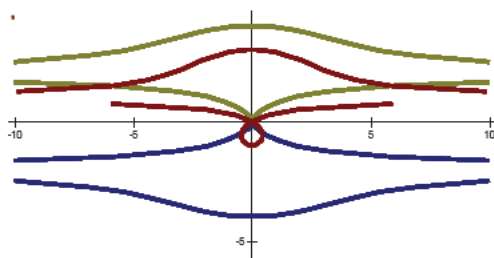


Рис.11.2. Конхоида Никомеда

Ряд математиков в поисках решения «трех великих задач древности - о трисекции угла (делении угла на три равные части),

- об удвоении куба (заданы сторона и объем куба, найти сторону куба, объем которого в два раза больше),
- о квадратуре круга (нахождение такого квадрата, площадь которого была бы равна площади круга) изобрели ряд замечательных кривых.

Так софист Гипсий из Элиды (420г.до н.э.), решавший задачу о делении угла в заданном отношении, открыл *квадратрису Динострата* (рис. 11.1).

Никомед (3 век до н.э.) исследовал конхоиду и применил её для решения задачи о трисекции угла (рис. 11.2).

Следует иметь в виду, что указанные задачи нельзя точно решить геометрически с помощью конечного числа построений прямых линий и окружностей, - это можно сделать только приближенно. Архимед исследовал кривую,

получившую его имя-спираль Архимеда (рис. 11.3), определил способ построения касательной к ней и выполнил её квадратуру.

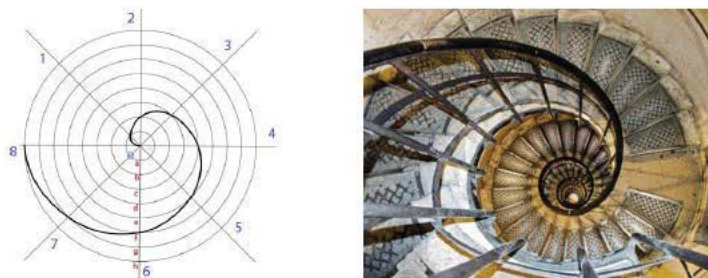


Рис. 11.3. Спираль Архимеда

Греки создали теорию конических сечений – линий, имеющих особенно большое значение в науке и технике. Открытие их приписывают Менехму (IV век до н.э.), как полагают, учителю Александра Македонского. Полагают, что поводом к этому открытию послужили поиски решения знаменитой делосской задачи об удвоении куба, или практический вопрос: насколько должен быть вытянут овал, находящийся в виде архитектурного украшения на фронтоне здания, чтобы с известного места перед зданием он казался кругом. Аполлоний Пергский (около 260-170г. до н.э.) в сочинении «Конические сечения» установил характеристические свойства коник и их уравнения (рис. 11.4).



$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

Рис. 11.4 – Коники и их уравнение (в декартовой системе координат)

Развитие математики, а значит и развитие теории кривых линий известный российский математик С.Новиков представляет следующим образом

«XVI в.: развилась алгебра многочленов; решили алгебраические уравнения 3-й и 4-й степени; как главный продукт было кардинально усовершенствовано учение о числе, ввели и начали использовать отрицательные и комплекс-

ные числа - отрицательные числа прижились сразу, а вот борьба за комплексные числа была долгой, до нашего времени.

XVII в.: появились координаты, позволившие перевести геометрию на язык алгебраических формул и расширить ее предмет; стал развиваться анализ; были сформулированы математические законы, лежащие в основе многих явлений природы, - вариационный принцип Ферма для световых лучей, принцип Галилея, закон Гука, универсальный закон гравитации, общие законы Ньютона. Возникли первые значительные прецеденты математического вывода законов природы из фундаментальных принципов (недостаточно оцененный современниками вывод закона преломления света на границе двух сред из вариационного принципа Ферма и вывод законов Кеплера Ньютоном, ставший основой современного научного метода). Появились идеи теории вероятностей.

XVIII в.: развитие анализа превратилось в мощный поток, включая линейные дифференциальные уравнения и метод собственных колебаний, вариационное исчисление и многое другое. Возникли дифференциальная геометрия, теория чисел, развилась теория вероятностей. Механика, включая небесную механику, стала зрелой далеко развитой наукой. Возникла гидродинамика.

XIX в.: математический поток, включая теорию вероятностей, продолжает набирать силу. Возникает комплексный анализ; проблема разрешимости алгебраических уравнений порождает теорию римановых поверхностей и теорию групп; создается линейная алгебра; углубляется изучение симметрии и возникают алгебры Ли; геометрия, теория чисел, теория римановых поверхностей, теория дифференциальных уравнений, теория рядов Фурье и др. превращаются в мощные развитые дисциплины. Появились новые разделы физики со своими математическими законами: электричество и магнетизм, рожденная техникой термодинамика, затем - статистическая физика и кинетика. В конце XIX в. возникли первые ростки абстрактных разделов математики - такие, как теория множеств и функций действительного переменного. Возникли качественно-топологические разделы математики (качественная теория динамических систем и топология). Появились первые идеи математической логики»

Новиков С.П. – Математика на пороге XXI века (Историко-математические исследования.(интернет исследование)

Вновь к кривым математическая наука обратилась только в XVII веке, в связи с созданием аналитической геометрии. В 1637 г. Рене Декарт (1596-1650), великий французский ученый – математик, физик, биолог и философ – в книге «Геометрия» изложил основы метода координат. Открытие этого метода для исследования кривых стало фактом первостепенного значения. Главный прием этого метода-изображение точек пространства с помощью трех числовых координат. Это позволило заменить исследование геометрических объектов, представляемых чертежами, исследованием соответствующих алгебраических выражений и дал возможность ввести в рассмотрение объекты, не допускающие наглядного представления.

Предложенный Декартом метод координат впервые позволил сформулировать понятие кривой в достаточно общей форме. Так, плоской кривой, задаваемой уравнением $F(x,y)=0$, стали называть множество точек на плоскости, ко-

ординаты которых удовлетворяют этому уравнению. Из механики пришло представление о кривой как о траектории движущейся точки с координатами, зависящими от времени t . Систематическое изучение плоских алгебраических кривых высших порядков привело к расширению известных свойств кривых, их особенностей, таких как точки перегиба, кратные точки, точки возврата. Сам Декарт исследовал немало кривых, одна из них, «листок жасмина», известна теперь как «Декартов лист» (рис. 11.5).

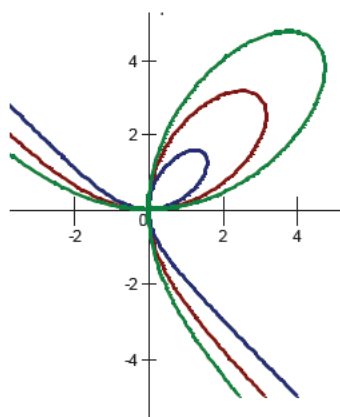


Рис.11.5. Лист Декарта

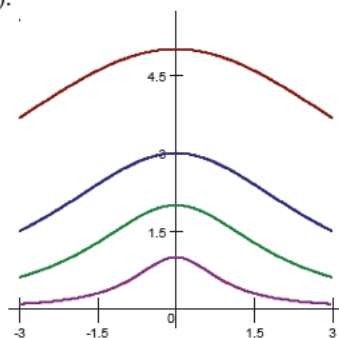


Рис. 11.6. Верзиера

Итальянский геометр Г. Гранди (1671-1742) исследовал так называемые «лепестковые» кривые. Некоторые из этих кривых имеют свои названия, например, «локон Аньези» или «верзиера» (рис. 11.6) в честь известного математика Марии Гаэтано Аньези (1718-1799) работавшей в Болонье. Другой пример; «розы» (рис. 11.7). Немецкий натуралист Б. Хабенихт использовал полярные координаты для описания контуров цветов (рис. 11.8).

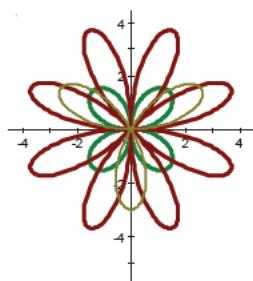


Рис. 11.7. Розы

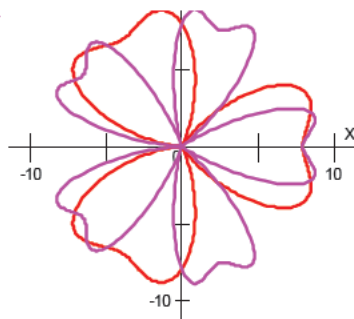


Рис. 11.8.Трилисник

В это же время эстетические свойства именно кривых начинают рассматриваться художниками. В1753 году выхолит книга выдающегося английского живописца Уильяма Хогарта (1697-1764) «Анализ красоты», в которой он писал:

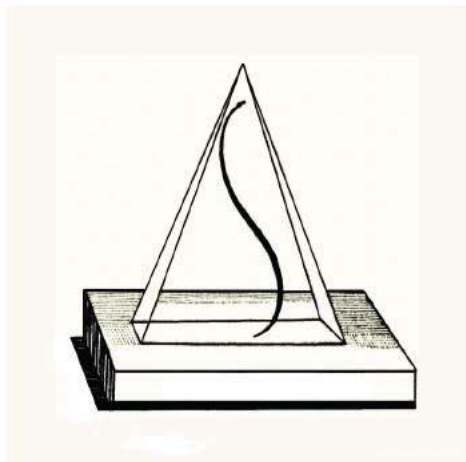


Рис. 11.9. Кривая Красоты

«Постоянное употребление линий при изображении предметов на бумаге, как математиками, так и художниками создало установившееся предпоследнее, будто бы эти линии существуют и на самих предметах. Исходя их этой предпосылки, мы сначала скажем вообще, что прямая и округлая линии, в различных комбинациях и вариантах, могут ограничить и описать любой видимый предмет, и они производят благодаря этому, такое бесконечное разнообразие форм, которое делает необходимым распределить эти линии и разбить их на определенные группы, оставив промежуточные формы и связи дальнейшим наблюдениям самого читателя».

Предлагаемую им «линию красоты» Хогарт изобразил на своем автопортрете с собакой(1745). Она приведена также на титульном листе первого издания «Анализа красоты» (рис.9).

Василий Кандинский (1886-1944)-один из крупнейших художников 20 века, с живописных «импровизаций» и «композиций» которого началась история абстрактного искусства, в 1926 году написал свое знаменитое исследование «Точка и линия на плоскости». В ней он в частности писал:

«Геометрическая линия-это невидимый объект. Она след перемещающейся точки, то есть ее произведение. Она возникла из движения -а именно вследствие уничтожения высшего, замкнутого в себя покоя точки. здесь произошел скачок из статики в динамику. Таким образом, линия - величайшая противоположность живописного первоэлемента - точки.»



Рис. 11.10. Г. Эрни «Рисовальщик»

Мотивы линий использовали в своем творчестве многие современные художники. В качестве примера приведем работу известносовременного швейцарского художника Ганса Эрни «Рисовальщик» (рис. 11.10). Широко использовал геометрические мотивы в своем творчестве М.К. Эшер (1898-1972).

Использование аналитического метода в геометрии встретилося с проблемами, внутренне ему присущими, но чуждыми изучаемым геометрическим объектам. Аналитический метод не обеспечивает единообразия изображения геометрического объекта в алгебраическом виде. Поэтому алгебраические выражения могут различаться, несмотря на изображаемое ими общее геометрическое содержание.

Дальнейшее развитие методов исследования кривых связано с дифференциальной геометрией, основы которой заложили Леонард Эйлер (1707-1783) и Гаспар Монж (1746-1818), алгебраической геометрией развитой немецкими математиками Л. Гессе (1811-1874) и Р. Клебшем (1833-1872). Использование векторного анализа при исследовании свойств линий связано с Германом Грассманом (1809-1877). К. Жордан (1838-1922) в 1874 году в работе «О теории кривых в пространстве n -измерений» определил кривую как образ непрерывного отображения отрезка $[a, b]$ в пространство. Это определение казалось вполне соответствующим наглядному представлению о кривой, но в 1890 г. Д. Пеано (1858-1932) построил непрерывную кривую, которая полностью заполняла квадрат. Только в 1921 г. П. С. Урысон (1898-1924) дал наиболее общее типологически инвариантное определение линии.

Литература

1. Савелов А. А. Плоские прямые. Систематика, свойства, применение. (Справочное руководство) / А. А. Савелов. - М. : Физматгиз, 1960. - 294 с.
2. Смогоржевский А. С. Справочник по теории плоских кривых 3-го порядка / А. С. Смогоржевский, Е. С. Стопова. - М. : Физматгиз, 1961. - 426 с.
3. Гильберт Д. Наглядная геометрия / Д Гильберт, С. Кон-Фоссен. - М. : Наука, 1981. - 344 с.
4. Фокс А. Вычислительная геометрия. Применение в проектировании и на производстве. Пер. с англ. / А. Фокс, М. Пратт. - М. : Мир, 1982. - 304 с.
5. Иванов Г. С. Конструирование технических поверхностей (математическое моделирование на основе нелинейных преобразований) / Г. С. Иванов. - М. : Машиностроение, 1987. - 192 с.
6. Стародетко Е. А. Элементы вычислительной геометрии / Е. А. Стародетко - Минск. : Наука и техника, 1986. - 240 с.
7. Погорелов А. В. Дифференциальная геометрия / А. В. Погорелов. - М. : Наука; 1974. - 176 с.
8. Постников М. М. Лекции по геометрии. Семестр 1. Аналитическая геометрия / М. М. Постников. - М. : Наука, 1979. - 336 с.
9. Постников М. М. Лекции по геометрии. Семестр 2. Линейная алгебра и дифференциальная геометрия / М. М. Постников. - М. : Наука, 1979. - 312 с.
10. Постников М. М. Лекции по геометрии. Семестр 3. Гладкие многообразия / М. М. Постников. - М. : Наука, 1979. - 478 с.
11. Брукс Дж. Кривые и особенности. Геометрическое введение в теорию особенностей / Дж. Брукс, П. Джиблино. - М. : Мир, 1988. - 262 с.
12. Клеменс Г. Мозаика теории комплексных кривых / Г. Клеменс. - М. : Мир, 1984. - 160 с.
13. Рид М. Алгебраическая геометрия для всех / М. Рид. - М : Мир, 1991. - 151 с.
14. Залгаллер В. А. Теория огибающих / В. А. Залгаллер. - М. : Наука, 1975. -104 с.
15. Пахоменко А. С. Что такое линия / А. С. Пахоменко. - М. : ГИТТЛ, 1954. -140 с.
16. Shikin E. V. Handbook and atlas of curves / Eugene V. Shikin. - Boca Raton etc. : CRC press, Cop. 1995. XIV. - 545 s.
17. Von Seggern Devid H. Handbook of Matematikal Curves and Surfaces / D. Von Seggern. - 1990. - 226 p. (Макет - справочник по математическим кривым и поверхностям).
18. Хилл Ф. Open GL Программирование компьютерной графики / Ф. Хилл. -СПб. : Питер, 2002. - 1088 с.
19. Бадаев Ю. И. Поликоординатный метод в прикладной геометрии / Ю. И. Бадаев. - Киев, Техника, 2006. - 173 с.
20. Альберг Дж. Теория сплайнов и ее применение / Дж. Альберг, Э. Нильсон, Дж. Уолш. - М.: Мир, 1972. - 317 с.
21. Завьялов Ю. С. Методы сплайн - функций / Ю. С. Завьялов, Б. И. Квасов, В. Л. Мирошниченко. - М.: Наука, 1980. - 350 с.
22. Макаров В. Л. Сплайн-аппроксимация функций / В. Л. Макаров, В. В. Хлобыстов. - М. : Наука, 1983. - 80 с.
23. Мандельброт Б. Фрактальная геометрия природы / Б. Мандельброт. - М. : Институт компьютерных исследований, 2002. - 650 с.
24. Пайтген Х. О. Красота фракталов / Х. О. Пайтген, П. Х. Рихтер. - М. : Мир, 1993. - 176 с.
25. Федер Е. Фракталы / Е. Федер. - М. : Мир, 1991. - 254 с.
26. Фоменко А. Т. Наглядная геометрия и топология / А. Т. Фоменко. - М. : МГУ, 1993. - 416 с.
27. Фракталы в физике. //Труды 6-го международного симпозиума по фракталам в физике, 1985. - М. : Мир, 1988. - 672 с.

28. Шредер М. Фракталы, хаос, степенные законы. Миниатюры из бесконечного рая / М. Шредер. - Ижевск.: РХД, 2001. - 528 с.
29. Кроновер Р. М. Фракталы и хаос в динамических системах / Р. М. Кроновер. - М. : Техносфера, 2006. - 528 с.
30. Морозов А. Д. Введение в теорию фракталов / А. Д. Морозов. - Москва-Ижевск: Институт компьютерных исследований, 2002. - 160 с.
31. Арнольд В. И. Гюйгенс и Барроу, Ньютон и Гук. Первые шаги математического анализа и теории катастроф. От эвольвент до квазикристаллов / В. И. Арнольд. - М. ГРФМЛ, 1989. - 94 с.
32. Кольман Э. История математики в древности /Э. Кольман. - М. : ГИФМЛ, 1961. - 236 с.
33. Юшкевич А. П. История математики в средние века / А. П. Юшкевич. - М. : ГИФМЛ, 1961. - 448 с.
34. Хрестоматия по истории математики. Математический анализ. Теория вероятностей. - М. : Просвещение, 1977. - 224 с.
35. Цейтен Г. Г. История математики в XVI и XVII веках. Пер. с нем. / Г. Г. Цейтен. - М. : ОНТИ, 1938. - 456 с.
36. Бурбаки Н. Очерки по истории математики. Пер. с франц./ Н. Бурбаки. - М. : ИЛ, 1963. - 292 с.
37. Стройк Д. Я. Краткий очерк по истории математики. Пер. с англ. / Д. Я. Стройк. - М. : Наука, 1969. - 256 с.
38. Вилейтнер Г. История математики от Декарта до середины XIX столетия. Пер. с нем. / Г. Вилейтнер. - М. : ГИФМЛ, 1960. - 468 с.
39. Даан-Дельмедико А. Пути и лабиринты. Очерки по истории математики. Пер. с франц./ А. Даан-Дельмедико, Ж. Пейффер. - М. : Мир, 1986. - 432 с.
40. Ван дер Варден. Пробуждающаяся наука. Математика древнего Египта, Вавилона и Греции. Перевод с голландского. / Ван дер Варден. - М.: ГИФМЛ, 1959. - 460 с.
41. Пиндоу Д. Геометрия и искусство. Пер с англ. / Д. Пиндоу. - М. : Мир, 1979. - 332 с.
42. Аталай Б. Математика и «Мона Лиза». Искусство и наука в творчестве Леонардо да Винчи. Пер. с англ. / Б. Аталай. - М. : Техносфера, 2007. - 304 с.
43. Вуд Дж. Солнце, Луна и древние камни. Пер. с англ. /Дж. Вуд. - М. : Мир, 1981. - 268 с.
44. Яглом И. М. Математические структуры и математическое моделирование / И. М. Яглом. - М. : Сов. радио, 1980. - 144 с.
45. Яглом И. М. Современная культура и компьютеры. - М. : Знание, 1990. – 48 с.
46. Antoniou I. Dynamical systems and irreversibility. Proc. XXI Solvay congress in physics. (2002, eng, 378 p)
47. Cox B. L. , Wang J. S. Y. Fractal surfaces measurement and applications in the Earth sciences // Fractals: An interdisciplinary of the complex geometry of nature. 1993. - V. 1, №1. pp. 87-115.
48. Galor O. Discrete Dynamical Systems (Springer, 2007)(158s)
49. Pilgrim K. M. Combinations of complex dynamical systems (Springer LNM1.)
50. Shiah A., Ong K-K., Musielak Z. E. Fractal images of generalized mandelbrot sets // Fractals: An interdisciplinary of the complex geometry of nature. 1994. - V. 2, №1. pp 111-121.
51. Peitgen H. O., Jürgens H., Saupe D. Fractals for the classroom. NY, 1992.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ. Геометрические модели и моделирование	3
РАЗДЕЛ 1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ МАТЕМАТИКИ.....	11
1.1. Матрицы и операции над ними	12
1.1.1 Действия над матрицами.....	12
1.1.2 Умножение двух матриц	13
1.1.3 Разбиение матриц на блоки	14
1.1.4 Определитель матрицы.....	14
1.1.5 Обращение матриц	15
1.2. Элементы векторной алгебры.....	16
1.2.1 Операции с векторами	16
1.2.2 Линейные операции векторов	20
1.2.3 Декартовы компоненты вектора	22
1.2.4 Скалярное произведение	29
1.2.5 Векторное уравнение прямой	32
1.2.6 Векторное произведение	34
1.2.7 Векторное уравнение плоскости	38
1.2.8 Смешанное произведение.....	46
1.2.9 Тройное векторное произведение	50
РАЗДЕЛ 2. ПРЕОБРАЗОВАНИЯ	52
2.1. Графический конвейер.....	53
2.2. Преобразование точек и объектов.....	53
2.3. Аффинные преобразования.....	54
2.4. Однородные координаты.....	55
2.5. Элементарные преобразования плоскости.....	60
2.6. Элементарные трехмерные преобразования.....	66
2.7. Плоские проекции трехмерных объектов.....	75
2.7.1 Косоугольное параллельное проецирование	76
2.7.2 Ортогональное проецирование	79
2.7.3 Перспектива.....	87
РАЗДЕЛ 3. КРИВЫЕ ЛИНИИ, ЗАКОНОМЕРНЫЕ КРИВЫЕ	91
3.1. Общие сведения о кривых линиях.....	92
3.1.1 Способы образования кривых.....	92
3.1.2 Классификация кривых.....	93
3.2. Способы задания кривой	94

3.3. Основные характеристики кривой.....	98
3.3.1 Касательная и нормаль	98
3.3.2 Соприкасающаяся плоскость к кривой	100
3.3.3 Кривизна и кручение кривой.....	101
3.3.4 Формулы Френе.....	103
3.3.5 Кривизна плоской кривой.....	104
3.3.6 Соприкасающаяся окружность к плоской кривой	106
3.3.7 Определение непрерывности кривой.....	107
3.3.8 Особые точки плоских кривых	107
3.3.9 Эволюта и эвольвента	111
3.3.10 Огибающая семейства кривых.....	116
РАЗДЕЛ 4. АЛГЕБРАИЧЕСКИЕ КРИВЫЕ.....	120
4.1. Общие свойства алгебраических кривых.....	121
4.2. Понятие о тангенциальных координатах.....	123
4.3. Класс алгебраической кривой	124
4.4. Формулы Плюккера	125
4.5. Род алгебраической кривой.....	125
4.6. Циклические точки и циркулярные кривые	127
РАЗДЕЛ 5. ПРЕОБРАЗОВАНИЯ ПЛОСКИХ КРИВЫХ ЛИНИЙ.....	129
5.1. Точечные преобразования	130
5.2. Применение аффинных преобразований к плоским кривым	130
5.3. Перспективные преобразования.....	131
5.3.1 Трилинейные (проективные) координаты	131
5.3.2 Проективные преобразования	134
5.3.3 Частные случаи проективного преобразования	134
5.4. Преобразование инверсии	136
5.5. Подерное преобразование	140
5.6. Катакаустика	146
5.7. Конхоидное преобразование.....	148
РАЗДЕЛ 6. КРИВЫЕ ВТОРОГО ПОРЯДКА.....	150
6.1. Классификация кривых 2-го порядка.....	150
6.2. Приведение уравнения кривой 2-го порядка к стандартному (каноническому) виду.....	153
6.3. Геометрическое определение невырожденной кривой 2-го порядка ..	154
6.4. Параметризация кривых второго порядка.....	156
6.5. Суперконики.....	156

РАЗДЕЛ 7. КРИВЫЕ ТРЕТЬЕГО ПОРЯДКА	159
7.1. Классификация прямых третьего порядка	159
7.2. Примеры важнейших кривых третьего и высших порядков.....	163
РАЗДЕЛ 8. КОНСТРУИРОВАНИЕ КРИВЫХ (ОБВОДЫ И СПЛАЙНЫ)	169
8.1. Основные понятия и определения.....	170
8.2. Построение обводов методом интерполяции	172
8.3. Интерполяция полиномными кривыми.....	180
8.4. Интерполяционный сплайн.....	189
8.5. Интерполирование параметрическими кубическими кривыми.....	194
8.5.1 Параметрические кривые	194
8.6. Кусочно-полиномиальные кривые	204
8.7. B – сплайны	208
8.8. Рациональные сплайны и NURBS – кривые	213
8.9. Аппроксимация	215
РАЗДЕЛ 9. ФРАКТАЛЬНЫЕ КРИВЫЕ.....	220
9.1. Интуитивное понятие о фракталах	224
9.2. Самоподобие и фрактальная (дробная) размерность	228
9.3. Конструктивные (классические) фракталы	238
9.3.1 Сжимающие аффинные преобразования.....	245
9.3.2 Метод Ньютона.....	249
9.3.3 Последовательность Морса – Туэ.....	253
9.3.4 Множество Жюлиа и Мандельброта	255
9.3.5 Фрактальное сжатие	260
9.4. Галерея фракталов	262
РАЗДЕЛ 10. ПОВЕРХНОСТИ.....	264
10.1. Способы задания поверхностей	265
10.2. Моделирование поверхностей полигональными сетками	267
10.3. Многогранники	272
10.4. Экструзивные формы.....	280
10.5. Линейчатые поверхности.....	283
10.6. Поверхности 2-го порядка.....	289
10.7. Билинейно сопряженные поверхности (лоскуты Кунса)	291
ЗАКЛЮЧЕНИЕ. ГЕОМЕТРИЧЕСКИЕ ОБЪЕКТЫ В ПРИРОДЕ, МАТЕМАТИКЕ И ИСКУССТВЕ. (КРАТКИЕ ИСТОРИЧЕСКИЕ СВЕДЕНИЯ)	302
ЛИТЕРАТУРА.....	313

Навчальне видання

Чермних Ігор Олексійович
Журило Алла Григорівна
Краєвська Олена Олександрівна
Адашевська Ірина Юріївна

ГЕОМЕТРИЧНЕ МОДЕЛЮВАННЯ В КОМП'ЮТЕРНІЙ ГРАФІЦІ

Навчальний посібник
для студентів вищих навчальних закладів

Російською мовою

Роботу до друку рекомендував

О. І. Пономаренко

В авторській редакції

План 2017 р., поз.74

Підписано до друку 23.03. 2017. Формат 60х84 1/16. Папір офсетний.

Друк – ризографія. Гарнітура Times New Roman. Ум. друк. арк. 19,25.

Наклад 100 прим. Зам. № 03/17-19. Ціна договірна.

Видавництво "НТМТ".

Свідоцтво про внесення до Державного реєстру видавців ДК № 1748 від 15.04.2004 р.
61072, м. Харків, пр. Науки, б. 58, к. 106.

E-mail: jornal_2016@ukr.net.
