



Н.Н. Голованов

УЧЕБНОЕ ПОСОБИЕ

ГЕОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Н.Н. Голованов

ГЕОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Учебное пособие

Допущено Учебно-методическим объединением вузов Российской Федерации по университетскому политехническому образованию в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлениям 09.03.01 «Информатика и вычислительная техника» (уровень бакалавриата), 09.04.01 «Информатика и вычислительная техника» (уровень магистратуры)

Москва
КУРС
ИНФРА-М
2016

УДК 004.925.8(075.8)

ББК 22.151я73

Г61

Рецензенты:

В.А. Мартынюк — доцент кафедры систем автоматизированного проектирования МГТУ им. Баумана, канд. техн. наук;

Г.В. Носовский — доцент кафедры дифференциальной геометрии МГУ им. М.В. Ломоносова, канд. физ.-мат. наук.

Н.Н. Голованов

Г61 Геометрическое моделирование. Учебное пособие. — М.: КУРС: ИНФРА-М, 2016. — 400 с.

ISBN 978-5-905554-76-6 (КУРС)

ISBN 978-5-16-011357-9 (ИНФРА-М, print)

ISBN 978-5-16-103551-1 (ИНФРА-М, online)

Изложены методы построения кривых, поверхностей и твёрдых тел. Описан состав геометрических моделей, приведены принципы взаимодействия их элементов, рассмотрены применения геометрических моделей.

Для студентов высших учебных заведений и специалистов по системам автоматизированного проектирования.

УДК 004.925.8(075.8)

ББК 22.151я73

ISBN 978-5-905554-76-6 (КУРС)

ISBN 978-5-16-011357-9 (ИНФРА-М, print)

ISBN 978-5-16-103551-1 (ИНФРА-М, online)

© Голованов Н.Н., 2016

© КУРС, 2016

Оглавление

Предисловие	6
Введение	7
Глава 1. Точки	10
1.1. Радиус-вектор	10
1.2. Аффинные координаты	13
1.3. Модификации векторов и точек	16
1.4. Однородные координаты	18
Глава 2. Кривые	22
2.1. Кривая	22
2.2. Аналитические кривые	27
2.3. Кривые, построенные по набору точек	30
2.4. Кривые Безье	38
2.5. Кривые Безье и конические сечения	42
2.6. Рациональные кривые Безье	51
2.7. Разделенные разности	54
2.8. <i>B</i> -сплайны	59
2.9. <i>B</i> -кривые	67
2.10. Алгоритм Де Бура	72
2.11. Вставка точки и узла <i>B</i> -кривой	74
2.12. Примеры <i>B</i> -кривых	77
2.13. <i>B</i> -кривые и кривые Безье	80
2.14. Частные случаи <i>B</i> -кривых	84
2.15. Кривые, построенные на базе кривых	92
2.16. Составная кривая	96
Глава 3. Поверхности	99
3.1. Поверхность	99
3.2. Аналитические поверхности	111
3.3. Поверхности, полученные движением кривой	114
3.4. Поверхности, построенные на семействе кривых	120
3.5. Поверхности, построенные на сетке кривых	126
3.6. Поверхность на сетке точек	133
3.7. Поверхность Безье	136
3.8. <i>B</i> -поверхность	139
3.9. <i>T</i> -поверхности	144
3.10. Поверхности треугольной формы	148
3.11. Треугольные поверхности Безье	152
3.12. Однородные разделенные разности	158
3.13. Симплексные сплайны	166
3.14. <i>S</i> -поверхности	172

3.15. Поверхности, построенные на базе поверхностей	174
3.16. Деформируемая поверхность.....	177
3.17. Поверхность на базе триангуляции	179
3.18. Поверхность с произвольной границей	181
Глава 4. Проекция и пересечения	185
4.1. Проекция точки на кривую	185
4.2. Проекция точки на поверхность	188
4.3. Точки пересечения кривых.....	192
4.4. Точки пересечения поверхности и кривой	195
4.5. Точки пересечения трех поверхностей	198
4.6. Кривые на поверхностях.....	199
4.7. Кривая пересечения поверхностей	205
4.8. Алгоритм построения кривых пересечения	209
4.9. Поверхности сопряжения	215
4.10. Поверхность фаски.....	224
4.11. Положение точки относительно поверхности	226
4.12. Поиск начальных приближений.....	227
4.13. Точность геометрических построений	230
Глава 5. Тела	234
5.1. Оболочка	234
5.2. Характеристики оболочки	236
5.3. Однородная оболочка	242
5.4. Тело	243
5.5. Структуры данных.....	244
5.6. Элементарные тела.....	246
5.7. Тела, полученные движением кривой.....	251
5.8. Тела, построенные по сечениям	256
5.9. Тело, построенное по поверхности	258
Глава 6. Методы построения тел.....	261
6.1. Методы геометрического моделирования	261
6.2. Булевы операции над телами	262
6.3. Алгоритм булевых операций.....	271
6.4. Разрезанное тело	276
6.5. Симметричное тело	277
6.6. Тело с достраиваемыми элементами	280
6.7. Эквидистантное тело.....	282
6.8. Тонкостенное тело	286
6.9. Скругление ребер тела.....	288
6.10. Алгоритм скругления ребер тела.....	293
6.11. Построение фаски ребер тела	295
6.12. Прямое моделирование.....	296
6.13. Деформирование оболочек	299

Глава 7. Геометрические ограничения	302
7.1. Наложение геометрических ограничений	302
7.2. Объекты геометрических ограничений.....	305
7.3. Позиционирование набора тел.....	308
7.4. Решение уравнений геометрических ограничений	312
7.5. Консервативный метод	313
7.6. Метод декомпозиции	315
Глава 8. Геометрическая модель	320
8.1. Состав геометрической модели	320
8.2. Применение геометрической модели	321
8.3. Построение векторного изображения.....	323
8.4. Моделирование оптических свойств.....	327
8.5. Построение точечного изображения.....	334
8.6. Триангуляция	340
8.7. Триангуляция поверхности	347
8.8. Триангуляция оболочки.....	350
8.9. Инерционные характеристики	351
8.10. Вычисление инерционных характеристик.....	354
8.11. Главные моменты инерции.....	359
Приложение. Криволинейные системы	366
A.1. Криволинейные координаты	366
A.2. Дифференцирование в криволинейных координатах	371
A.3. Тензоры в криволинейных координатах	374
A.4. Примеры тензоров.....	380
A.5. Ортогональные криволинейные координаты	387
Список литературы	395
Предметный указатель	397

Предисловие

Предлагаемая книга предназначена для тех, кто изучает, использует в работе или разрабатывает системы геометрического моделирования. Она посвящена математическому аппарату этих систем и содержит методы построения численных моделей, моделей геометрии реальных и воображаемых объектов.

В книге изложены методы геометрического моделирования, включающие методы построения кривых, поверхностей и тел, приведены алгоритмы и структуры данных геометрических объектов, изложены принципы, обеспечивающие взаимосвязь элементов модели, рассмотрен состав геометрической модели и некоторые области ее применения.

В основу книги положен опыт работы автора над геометрическим ядром C3D.

Автор благодарит Г.Ф. Горшкова, В.А. Мартынюка, Г.В. Носовского, С.И. Роткова, К. Версприлла, Дж. Орра и коллег по разработке геометрического ядра C3D за полезные советы и поддержку.

Введение

С развитием компьютеров появилась возможность создавать численные модели различных объектов и экспериментировать с ними. Численные модели используются в системах, выполняющих проектирование (Computer Aided Design), расчеты (Computer Aided Engineering) и производство (Computer Aided Manufacturing) моделируемых объектов. Во всех этих системах с той или иной точностью требуется описать геометрическую форму моделируемых объектов. В процессе развития этих систем сформировалась область математики, которая называется геометрическим моделированием.

Геометрическое моделирование изучает методы построения численных моделей геометрии реальных и воображаемых объектов, а также методы управления этими моделями. Геометрическая модель содержит описание формы моделируемого объекта и описание связей элементов модели. Для возможности редактирования и создания подобных моделей в геометрическую модель часто включают дерево построения, хранящее последовательность и способы построения модели. Элементы геометрической модели, как правило, наделяют атрибутами, которые несут информацию о физических и других свойствах этих элементов.

Геометрическое моделирование начало свое развитие с систем компьютерного черчения. Позже появились системы каркасного и поверхностного моделирования. Компьютерные системы параметрического твердотельного моделирования кардинально изменили технологию работы конструктора. Они позволили фиксировать конструкторскую мысль не в виде плоского чертежа, а в виде трехмерной модели. Для описания связей элементов модели стали применяться вариационные методы. Еще одним шагом в развитии геометрического моделирования являются методы прямого моделирования, расширяющие возможности конструктора и дизайнера.

Окружающие нас объекты занимают некоторый конечный объем пространства. Для моделирования этих объектов нужно описать занимаемую ими часть пространства. В определенных случаях это можно сделать с помощью элементов объема моделируемого объекта. Элементы объема называют вокселями. Элементами объема могут служить кубы, призмы и пирамиды. **Воксельные модели** используются в случаях, когда атрибуты элементов объема модели имеют большее значение, чем геометрическая форма модели.

С определенной степенью точности геометрическую форму объектов можно описать, используя плоские грани. Описание геометрической формы моделируемого объекта плоскими гранями называют **полигональным представлением**. Криволинейные поверхности полигональное представление аппроксимирует набором пластин треугольной или четырехугольной формы. Использование плоских граней значительно упрощает работу с моделью. Полигональную модель обычно строят на основе замеров реальных объектов или на основе другой модели. Полигональное представление широко применяется для визуализации геометрических моделей и при передаче информации для 3D-печати.

Многие объекты можно получить, используя поступательное и вращательное движения. Элементы поверхности таких объектов можно задать плоскостью, поверхностью сферы, поверхностью цилиндра, поверхностью конуса, поверхностью тора. Все перечисленные поверхности делят пространство на две части и для них можно указать, с какой стороны поверхности находится внутренний объем моделируемого объекта. Используя для моделирования такие поверхности, можно построить геометрическую модель путем выполнения операций над примитивами, к которым, как правило, относят прямоугольную призму, треугольную призму, сферу, цилиндр, конус, тор. Описанным образом устроена **конструктивная твердотельная геометрия** (Constructive Solid Geometry). В настоящее время конструктивная твердотельная геометрия практически не используется.

Наиболее общий подход к описанию геометрической формы моделируемого объекта использует поверхности произвольной формы, представленные в явном виде. Этот подход состоит в представлении поверхности моделируемого объекта совокупностью граней, стыкующихся по ребрам и содержащих информацию о своих границах и связях с соседями. Грани стыкуются так, что внешняя сторона одной грани переходит во внешнюю сторону соседней грани. Грани могут иметь произвольную форму. Описание геометрической формы моделируемого объекта криволинейными гранями, стыкующимися друг с другом по общим краям, называется **граничным представлением** (Boundary Representation). Оно дает возможность выполнять над моделями множество операций, сохраняя при этом единый способ их внутреннего устройства. Граничное представление содержит точное описание границы моделируемого объекта, отделяющей его от остальной части пространства. Граничное представление используют все современные CAD-системы.

Между элементами геометрической модели устанавливают вариационные связи, поддерживающие совпадение точек, совпадение осей, совпадение плоскостей, принадлежность объекта плоскости, касание

поверхностей и многие другие зависимости. Взаимозависимости элементов модели называют **геометрическими ограничениями** (Geometric Constraints). Связанными могут быть различные размеры одного объекта и нескольких объектов. Через такие связи легко редактировать модель и создавать подобные модели. Как правило, геометрические ограничения могут быть удовлетворены не единственным образом. Большинство САД-систем имеют модули, устанавливающие геометрические ограничения на элементы геометрической модели. Геометрическое моделирование изучает методы поиска решения, которое соответствует поставленной практической задаче.

В разработке своих методов геометрическое моделирование опирается на разные области математики, в первую очередь на дифференциальную геометрию и численные методы. Геометрическое моделирование тесно связано с программированием, оно в полной мере востребовало особенности объектно-ориентированного подхода: сокрытие данных (encapsulation), наследование (inheritance), переопределение семантики операций (polymorphism). Например, объектно-ориентированное программирование позволяет создавать кривые и поверхности, обладающие некоторым множеством общих методов, скрыв реализацию этих методов каждой конкретной кривой и поверхности и используемые для этого данные.

Геометрическая модель используется для визуализации моделируемого объекта, проверки возможности корректно собрать моделируемый объект из его элементов, кинематической проверки, вычисления инерционных характеристик, расчета траектории режущего инструмента, проектирования оснастки и других этапов подготовки производства моделируемого объекта. С помощью геометрической модели выполняются численные эксперименты и производство моделируемого объекта. Для этого используются атрибуты элементов модели, описывающие физические и другие свойства моделируемого объекта.

Геометрическое моделирование позволяет сократить время и материальные затраты на производство проектируемых объектов и повысить их качество. Геометрическое моделирование автоматизирует труд дизайнеров, конструкторов, архитекторов, технологов, позволяет им уйти от рутинных работ и сосредоточиться на творчестве.

Глава 1. Точки

Мы будем моделировать геометрические свойства реальных и вообразимых объектов. Нашей конечной целью является построение геометрических моделей этих объектов. Эти модели нужны для проведения исследований, для принятия решений, для производства моделируемых объектов. Геометрическая модель состоит из множества тел, которые описываются кривыми и поверхностями. Кривые и поверхности в свою очередь описываются точками, векторами и скалярными величинами. Преобразования над геометрической моделью сводятся к соответствующим преобразованиям над точками, векторами и скалярами. В данной главе мы введем обозначения и рассмотрим преобразования координат точек и компонент векторов в аффинных системах координат при изменении положения описываемых ими объектов.

1.1. Радиус-вектор

Для построения геометрических моделей будем использовать декартовы координаты с фиксированным ортонормированным базисом, считая, что пространство является евклидовым.

Скалярные величины будем обозначать строчными буквами латинского алфавита, выделенными курсивом, а также строчными буквами греческого алфавита, например h , r , β , μ .

Для обозначения столбца координат точек и компонент векторов в трехмерном пространстве будем использовать строчные буквы латинского алфавита, выделенные полужирным шрифтом. Так, положение точки с координатами p_1 , p_2 , p_3 будем описывать, например, символом \mathbf{p} , за которым стоят три координаты:

$$\mathbf{p} \equiv \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}.$$

Кроме того, будем использовать и такое обозначение точки с координатами x_i , y_i , z_i :

$$\mathbf{p}_i \equiv \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}$$

или $\mathbf{p}_i \equiv [x_i, y_i, z_i]^T$.

Аналогичным способом будем обозначать столбцы компонент векторов в трехмерном пространстве. Так, вектор с компонентами v_1, v_2, v_3 будем описывать, например, символом \mathbf{v} , за которым стоят три компоненты:

$$\mathbf{v} \equiv \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}.$$

Для обозначения компонент матриц и тензоров будем использовать прописные буквы латинского алфавита, выделенные полужирным шрифтом, например,

$$\mathbf{M} \equiv \begin{bmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \end{bmatrix}.$$

Во многих формулах важны не конкретные координаты, а действия над их совокупностью, поэтому в формулах будем оперировать принятыми обозначениями векторов и точек. Многие формулы приобретают компактный вид при использовании векторных обозначений. Например, равенство $\mathbf{d} = \mathbf{a} + \mathbf{b}$ обозначает три скалярных равенства. Скалярное произведение векторов \mathbf{a} и \mathbf{b} будем обозначать через $\mathbf{a} \cdot \mathbf{b}$. Векторное произведение векторов \mathbf{a} и \mathbf{b} будем обозначать через $\mathbf{a} \times \mathbf{b}$.

При описании сложных геометрических объектов нам придется работать с точками и векторами как в трехмерном, так и в двумерном пространствах, например, в области параметров поверхности.

Для обозначения столбца координат точек и столбца компонент векторов в двумерном пространстве будем использовать строчные буквы латинского алфавита, выделенные полужирным наклонным шрифтом. Так, положение двумерной точки с координатами u_i и v_i будем описывать, например, символом \mathbf{p}_i , за которым стоят две координаты:

$$\mathbf{p}_i \equiv \begin{bmatrix} u_i \\ v_i \end{bmatrix}.$$

Кроме того, будем использовать и такое обозначение двумерной точки с координатами p_1 и p_2 , как $\mathbf{p} \equiv [p_1 p_2]^T$.

Аналогичным способом будем обозначать столбцы компонент векторов в двумерном пространстве. Так, двумерный вектор с компонентами α_k и β_k будем описывать, например, символом \mathbf{v}_k , за которым стоят две компоненты:

$$\mathbf{v}_k \equiv \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix}.$$

Вектор определяет направление в выбранной системе координат, точка определяет положение в выбранной системе координат. В аффинной системе координат точку пространства можно описать с помощью радиуса-вектора.

Радиус-вектор определяет преобразование переноса, переводящее начальную точку системы координат в заданную точку пространства. Пусть векторы $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ описывают оси некоторой аффинной системы координат. Рассмотрим точку пространства, имеющую в этой системе координаты p_1, p_2, p_3 . Построим вектор

$$\mathbf{p} = p_1 \mathbf{e}_1 + p_2 \mathbf{e}_2 + p_3 \mathbf{e}_3, \quad (1.1.1)$$

выходящий из начала аффинной системы координат. Тогда вектор (1.1.1) будет оканчиваться в рассматриваемой точке. Вектор (1.1.1) представляет собой радиус-вектор рассматриваемой точки, рис. 1.1.1. В криволинейной системе координат базисные векторы $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ в разных точках пространства различны, поэтому выражением (1.1.1) можно описывать только векторы в указанной точке, а не положение точки в пространстве. Запись вектора в виде правой части (1.1.1) в криволинейной системе координат используется для векторных полей, так как при дифференцировании позволяет учитывать изменение базисных векторов $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ при переходе из одной точки пространства в другую точку.

Радиус-вектор описывается координатами заданной точки пространства. Радиус-вектор будем обозначать так же, как и столбцы координат описываемой им точки, то есть строчными буквами латинского алфавита, выделенными полужирным шрифтом. Оси глобальной системы координат и отдельные радиусы-векторы на рисунках, как правило, не показывают, а отображают относительное положение точек в пространстве или изменение радиуса-вектора. Так, на рис. 1.1.2. можно показать только вектор \mathbf{v} перемещения точки \mathbf{p} из исходного положения в точку \mathbf{q} , а все остальное можно скрыть.

Мы будем использовать радиус-вектор для описания кривых и поверхностей. Радиус-вектор дает положение точки пространства и в отличие от вектора связан с началом координат. Эта разница сказывается на формулах при переходе от одной системы координат к другой системе координат и на формулах изменения положения в пространстве. Рассмотрим различия между преобразованиями координат точки и преобразованиями компонент вектора при переходе от одной аффинной системы к другой аффинной системе координат.

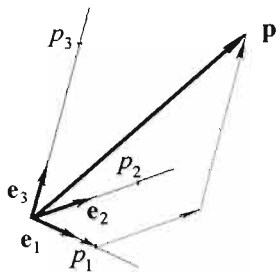


Рис. 1.1.1

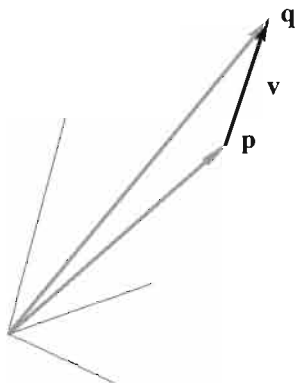


Рис. 1.1.2

1.2. Аффинные координаты

Пусть мы выбрали в трехмерном евклидовом пространстве некоторую точку и три взаимно ортогональных вектора единичной длины. Построим на выбранной точке и векторах декартову систему координат, которую будем называть глобальной системой. Пусть в выбранной глобальной системе координат построена локальная аффинная система координат с фиксированными базисными векторами

$$\mathbf{i}_1 = \begin{bmatrix} i_{11} \\ i_{12} \\ i_{13} \end{bmatrix}, \quad \mathbf{i}_2 = \begin{bmatrix} i_{21} \\ i_{22} \\ i_{23} \end{bmatrix}, \quad \mathbf{i}_3 = \begin{bmatrix} i_{31} \\ i_{32} \\ i_{33} \end{bmatrix}$$

и началом в точке начала глобальной системы. Векторы \mathbf{i}_1 , \mathbf{i}_2 , \mathbf{i}_3 должны быть линейно независимыми. Положим, что векторы \mathbf{i}_1 , \mathbf{i}_2 , \mathbf{i}_3 могут быть не ортогональными друг другу и иметь произвольную длину. Пусть положение некоторой точки \mathbf{q} в глобальной системе описывается координатами q_1 , q_2 , q_3 , а в локальной системе положение этой же точки описывается координатами y_1 , y_2 , y_3 . Тогда из равенства

$$\mathbf{q} = y_1 \mathbf{i}_1 + y_2 \mathbf{i}_2 + y_3 \mathbf{i}_3$$

следует соотношение для глобальных q_1 , q_2 , q_3 и локальных y_1 , y_2 , y_3 координат рассматриваемой точки:

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = [\mathbf{i}_1 \quad \mathbf{i}_2 \quad \mathbf{i}_3] \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad (1.2.1)$$

где столбцы матрицы преобразования \mathbf{A} составлены из компонент базисных векторов $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ локальной системы координат:

$$\mathbf{A} = [\mathbf{i}_1 \quad \mathbf{i}_2 \quad \mathbf{i}_3] = \begin{bmatrix} i_{11} & i_{21} & i_{31} \\ i_{12} & i_{22} & i_{32} \\ i_{13} & i_{23} & i_{33} \end{bmatrix}.$$

Обратное преобразование имеет вид

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathbf{A}^{-1} \cdot \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}. \quad (1.2.2)$$

Если базис глобальной системы координат и базис $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ ортонормирован, то определитель матрицы \mathbf{A} равен единице, а обратная матрица равна транспонированной матрице:

$$\mathbf{A}^{-1} = \mathbf{A}^{-T} = \begin{bmatrix} i_{11} & i_{21} & i_{31} \\ i_{12} & i_{22} & i_{32} \\ i_{13} & i_{23} & i_{33} \end{bmatrix}.$$

Сдвинем начало локальной системы координат в точку \mathbf{r} . Пусть положение некоторой точки \mathbf{p} в глобальной системе описывается координатами p_1, p_2, p_3 , а в локальной системе положение этой же точки описывается координатами x_1, x_2, x_3 , рис. 1.2.1.

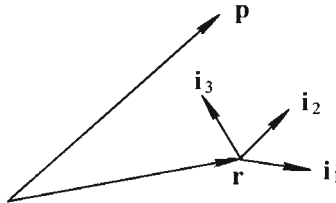


Рис. 1.2.1

Тогда из равенства

$$\mathbf{p} = \mathbf{r} + x_1 \mathbf{i}_1 + x_2 \mathbf{i}_2 + x_3 \mathbf{i}_3$$

следует соотношение для глобальных p_1, p_2, p_3 и локальных x_1, x_2, x_3 координат рассматриваемой точки:

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}. \quad (1.2.3)$$

Обратное преобразование имеет вид

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{A}^{-1} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} - \mathbf{A}^{-1} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}. \quad (1.2.4)$$

Если базис глобальной системы координат и базис $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ ортонормирован, то обратная матрица равна транспонированной матрице, а преобразование (1.2.4) примет вид

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix},$$

где компоненты o_1, o_2, o_3 и r_1, r_2, r_3 связаны соотношением:

$$\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix} = - \begin{bmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{bmatrix} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}.$$

Вектор в аффинной системе координат определяет перемещение от одной точки пространства к другой точке, поэтому вектор не связан с началом координат. Компоненты вектора u_1, u_2, u_3 в локальной системе координат и q_1, q_2, q_3 в глобальной системе координат связаны соотношениями (1.2.1) и (1.2.2). Радиус-вектор определяет преобразование переноса начальной точки системы координат в заданную точку пространства, поэтому в формулах преобразования координат радиуса-вектора присутствуют координаты начала локальной системы. Координаты точки x_1, x_2, x_3 в локальной системе координат и p_1, p_2, p_3 в глобальной системе координат связаны соотношениями (1.2.3) и (1.2.4).

1.3. Модификации векторов и точек

Простейшей модификацией точки является ее **сдвиг в пространстве** на вектор сдвига \mathbf{w} . Положение точки до модификации будем описывать радиусом-вектором \mathbf{r}_0 , а положение точки после модификации будем описывать радиусом-вектором \mathbf{r} . Положение точки после сдвига на вектор \mathbf{w} будет описываться радиусом-вектором, равным сумме радиуса-вектора ее исходного положения \mathbf{r}_0 и вектора сдвига \mathbf{w} :

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{w}.$$

Координаты радиуса-вектора \mathbf{r} равны сумме соответствующих координат радиуса-вектора \mathbf{r}_0 и компонент вектора \mathbf{w} .

В отличие от точки сдвиг вектора в пространстве не изменяет вектор, так как вектор определяет относительное положение двух точек, которое не меняется при их одновременном перемещении.

Рассмотрим **вращение точки** вокруг оси. Пусть начальное положение рассматриваемой точки описывается радиусом-вектором \mathbf{r}_0 , а ось вращения определяется точкой \mathbf{q} и ортом \mathbf{v} . Выполним поворот рассматриваемой точки вокруг оси на угол α против часовой стрелки, если взгляд направить навстречу вектору \mathbf{v} , рис. 1.3.1.

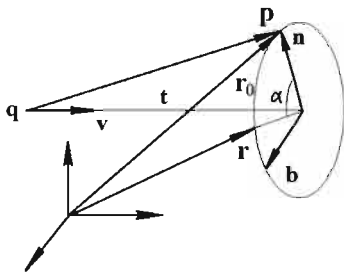


Рис. 1.3.1

Построим вектор $\mathbf{p} = \mathbf{r}_0 - \mathbf{q}$. Разложим вектор \mathbf{p} на две составляющие:

$$\mathbf{p} = \mathbf{t} + \mathbf{n}, \quad (1.3.2)$$

где вектор $\mathbf{t} = (\mathbf{p} \cdot \mathbf{v})\mathbf{v}$ параллелен вектору \mathbf{v} , а вектор $\mathbf{n} = \mathbf{p} - (\mathbf{p} \cdot \mathbf{v})\mathbf{v}$ ортогонален вектору \mathbf{v} . При вращении вектор \mathbf{t} не изменится, а вектор \mathbf{n} повернется на угол α в сторону вектора

$$\mathbf{b} = \mathbf{v} \times \mathbf{n} = \mathbf{v} \times (\mathbf{p} - (\mathbf{p} \cdot \mathbf{v})\mathbf{v}) = \mathbf{v} \times \mathbf{p}. \quad (1.3.3)$$

Так как вектор \mathbf{v} имеет единичную длину, то вектор \mathbf{b} будет иметь длину, равную длине вектора \mathbf{n} . Кроме того, он ортогонален векторам \mathbf{v} и \mathbf{n} . По-

сле поворота на угол α вектор \mathbf{n} станет равным вектору $\mathbf{n}\cos\alpha + \mathbf{b}\sin\alpha$. Следовательно, после поворота рассматриваемая точка будет определяться радиусом-вектором

$$\begin{aligned}\mathbf{r} &= \mathbf{q} + \mathbf{t} + \mathbf{n} \cos\alpha + \mathbf{b} \sin\alpha = \\ &= \mathbf{q} + (\mathbf{p}\cdot\mathbf{v})\mathbf{v} + (\mathbf{p} - (\mathbf{p}\cdot\mathbf{v})\mathbf{v}) \cos\alpha + \mathbf{v} \times \mathbf{p} \sin\alpha.\end{aligned}\quad (1.3.4)$$

С учетом равенства $(\mathbf{p}\cdot\mathbf{v})\mathbf{v} = (\mathbf{v}\mathbf{v})\cdot\mathbf{p}$ выражение (1.3.4) примет вид

$$\mathbf{r} = \mathbf{q} + (\mathbf{v}\mathbf{v})\cdot\mathbf{p} + \cos\alpha (\mathbf{E} - \mathbf{v}\mathbf{v})\cdot\mathbf{p} + \sin\alpha \mathbf{v} \times \mathbf{p} = \mathbf{q} + \mathbf{A}\cdot(\mathbf{r}_0 - \mathbf{q}).\quad (1.3.5)$$

Матрица поворота \mathbf{A} определяется равенством

$$\mathbf{A} = (1 - \cos\alpha) \mathbf{v}\mathbf{v} + \cos\alpha \mathbf{E} + \sin\alpha \mathbf{v}\times,$$

где $\mathbf{v}\mathbf{v} = \begin{bmatrix} v_1v_1 & v_1v_2 & v_1v_3 \\ v_2v_1 & v_2v_2 & v_2v_3 \\ v_3v_1 & v_3v_2 & v_3v_3 \end{bmatrix}$, $\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\mathbf{v}\times = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$

v_1, v_2, v_3 – компоненты орта \mathbf{v} .

Матрица \mathbf{A} является ортогональной. При транспонировании матрицы \mathbf{A} изменится только знак перед последним ее слагаемым, что соответствует повороту точки на угол $-\alpha$.

Формулу для поворота свободного вектора в пространстве получим из формулы (1.3.5), положив в ней $\mathbf{q} = \mathbf{0}$.

Определим координаты точки \mathbf{r} , симметричной точке \mathbf{r}_0 относительно плоскости. Пусть плоскость симметрии определяется точкой \mathbf{q} и ортом \mathbf{n} , рис 1.3.2.

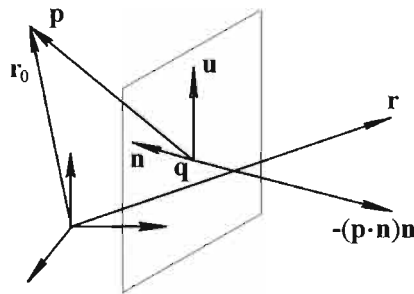


Рис. 1.3.2

Построим вектор $\mathbf{p} = \mathbf{r}_0 - \mathbf{q}$ и представим его в виде суммы двух векторов – проекции на орт \mathbf{n} и перпендикулярной орту \mathbf{n} составляющей \mathbf{u} :

$$\mathbf{p} = (\mathbf{p}\cdot\mathbf{n})\mathbf{n} + \mathbf{u},$$

где $\mathbf{u} = \mathbf{p} - (\mathbf{p} \cdot \mathbf{n})\mathbf{n}$. После зеркального отражения вектора \mathbf{p} его нормальная к плоскости составляющая изменит знак на противоположный. Положение симметричной точки будет описываться радиусом-вектором

$$\begin{aligned} \mathbf{r} &= \mathbf{q} - (\mathbf{p} \cdot \mathbf{n}) \mathbf{n} + (\mathbf{p} - (\mathbf{p} \cdot \mathbf{n}) \mathbf{n}) = \mathbf{q} + (\mathbf{E} - 2\mathbf{n}\mathbf{n}^\top) \cdot \mathbf{p} = \\ &= \mathbf{q} + \mathbf{A} \cdot (\mathbf{r}_0 - \mathbf{q}), \end{aligned} \quad (1.3.6)$$

где $\mathbf{A} = \begin{bmatrix} 1 - 2n_1n_1 & -2n_1n_2 & -2n_1n_3 \\ -2n_2n_1 & 1 - 2n_2n_2 & -2n_2n_3 \\ -2n_3n_1 & -2n_3n_2 & 1 - 2n_3n_3 \end{bmatrix}$ – матрица симметрии;

n_1, n_2, n_3 – компоненты орта \mathbf{n} .

Формулу для преобразования симметрии свободного вектора в пространстве получим из формулы (1.3.6), положив в ней $\mathbf{q} = \mathbf{0}$.

Рассмотрим **масштабирование** радиуса-вектора точки. Пусть даны три взаимно ортонормированных вектора $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ и требуется увеличить размер масштабируемого объекта в направлении вектора \mathbf{i}_1 – в m_1 раз, в направлении вектора \mathbf{i}_2 – в m_2 раз, в направлении вектора \mathbf{i}_3 – в m_3 раз. Рассмотрим изменение положения точки \mathbf{r}_0 масштабируемого объекта, при котором точка \mathbf{q} не меняет своего положения. Положение рассматриваемой точки после масштабирования будет описываться радиусом-вектором

$$\mathbf{r} = \mathbf{q} + \mathbf{A} \cdot (\mathbf{r}_0 - \mathbf{q}), \quad (1.3.7)$$

где $\mathbf{A} = \begin{bmatrix} m_1\mathbf{i}_1 & m_2\mathbf{i}_2 & m_3\mathbf{i}_3 \end{bmatrix} = \begin{bmatrix} m_1i_{11} & m_2i_{21} & m_3i_{31} \\ m_1i_{12} & m_2i_{22} & m_3i_{32} \\ m_1i_{13} & m_2i_{23} & m_3i_{33} \end{bmatrix}$ – матрица масштабирования.

Формулу для масштабирования свободного вектора в пространстве получим из формулы (1.3.7), положив в ней $\mathbf{q} = \mathbf{0}$.

1.4. Однородные координаты

Формулы для пересчета координат радиуса-вектора точки при ее повороте вокруг оси (1.3.5), симметрии относительно плоскости (1.3.6) и масштабирования (1.3.7) имеют единый вид:

$$\mathbf{r} = \mathbf{q} + \mathbf{A} \cdot (\mathbf{r}_0 - \mathbf{q}) = \mathbf{A} \cdot \mathbf{r}_0 + (\mathbf{q} - \mathbf{A} \cdot \mathbf{q}).$$

В общем случае преобразование радиуса-вектора точки \mathbf{r}_0 можно записать в виде

$$\mathbf{r} = \mathbf{A} \cdot \mathbf{r}_0 + \mathbf{t}, \quad (1.4.1)$$

где \mathbf{A} – матрица трансформации (1.3.5) – (1.3.7), $\mathbf{t} = \mathbf{q} - \mathbf{A} \cdot \mathbf{q}$ – вектор перемещения. Сдвиг точки \mathbf{r}_0 формально описывается формулой (1.4.1) с единичной матрицей \mathbf{A} . Аналогичный вид имеет формула преобразования координат (1.2.3). Всем перечисленным преобразованиям можно придать единый вид:

$$\mathbf{r} = \mathbf{M} \cdot \mathbf{r}_0,$$

если использовать **однородные координаты** и расширенные матрицы.

Расширенная матрица \mathbf{M} представляет собой матрицу \mathbf{A} , окаймленную снизу нулями, а справа – компонентами вектора перемещения \mathbf{t} , и имеет вид:

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}, \quad (1.4.2)$$

При работе с расширенной матрицей полагается, что точки имеют четыре координаты, например, радиус-вектор точки \mathbf{r} с координатами r_1, r_2, r_3 описан с помощью матрицы-столбца

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ 1 \end{bmatrix}, \quad (1.4.3)$$

а вектор \mathbf{v} с компонентами v_1, v_2, v_3 описан с помощью матрицы-столбца

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \end{bmatrix}. \quad (1.4.4)$$

Однородные координаты и расширенные матрицы строятся на базе обычных координат и матриц. Для получения однородных координат к трем координатам точки добавляют четвертую координату, равную единице, а к трем компонентам вектора добавляют четвертую компоненту, равную нулю. Для получения расширенной матрицы увеличивают размерность обычной матрицы на единицу. В терминах однородных координат равенство (1.2.3) имеет вид

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ 1 \end{bmatrix} = \begin{bmatrix} i_{11} & i_{12} & i_{13} & r_1 \\ i_{21} & i_{22} & i_{23} & r_2 \\ i_{31} & i_{32} & i_{33} & r_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}. \quad (1.4.5)$$

Обратное преобразование (1.2.4) имеет вид:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & o_1 \\ a_{21} & a_{22} & a_{23} & o_2 \\ a_{31} & a_{32} & a_{33} & o_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ 1 \end{bmatrix}, \quad (1.4.6)$$

$$\text{где } \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \mathbf{A}^{-1}, \quad \begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix} = -\mathbf{A}^{-1} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}.$$

Если базис $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ ортонормирован, то обратное преобразование (1.2.4) имеет вид

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} i_{11} & i_{12} & i_{13} & o_1 \\ i_{21} & i_{22} & i_{23} & o_2 \\ i_{31} & i_{32} & i_{33} & o_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ 1 \end{bmatrix}. \quad (1.4.7)$$

Расширенная матрица в двумерном пространстве имеет вид аналогичный матрице (1.4.2) с той разницей, что трехмерный вектор \mathbf{t} заменен двумерным вектором \mathbf{t} . При работе с расширенной матрицей в двумерном пространстве полагается, что точки имеют три координаты, например, радиус-вектор двумерной точки \mathbf{r} с координатами r_1, r_2 описан с помощью матрицы-столбца

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ 1 \end{bmatrix},$$

а двумерный вектор \mathbf{v} с компонентами v_1, v_2 описан с помощью матрицы-столбца

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix}.$$

Расширенные матрицы и однородные координаты используются также при работе с рациональными кривыми и поверхностями. Для построения **рациональных** кривых и поверхностей наряду с координатами для точек необходим дополнительный параметр — так называемый **вес** точки. Вес точки выступает в роли дополнительной координаты. При

наличии у точки веса w запись ее радиуса-вектора в однородных координатах имеет вид

$$\mathbf{p} = \begin{bmatrix} wp_1 \\ wp_2 \\ wp_3 \\ w \end{bmatrix}. \quad (1.4.8)$$

При построении рациональных кривых в двумерном пространстве запись радиуса-вектора точки в однородных координатах имеет вид

$$\mathbf{p} = \begin{bmatrix} wp_1 \\ wp_2 \\ w \end{bmatrix}. \quad (1.4.9)$$

При использовании однородных координат вычисления производятся в векторном виде. Пространственные координаты точки вычисляются на конечном этапе вычислений делением однородных координат на последнюю координату w , равную весу.

Упражнения

1. Чем отличается радиус-вектор от вектора?
2. Составьте матрицу преобразования координат двумерной точки в двумерную локальную аффинную систему координат.
3. Составьте расширенную матрицу вращения точки в двумерном пространстве.

Глава 2. Кривые

Геометрическая модель содержит описание формы моделируемого объекта. Геометрическая форма описывается поверхностями. Для построения поверхностей используются кривые. С кривых начинается геометрический дизайн. В геометрическом моделировании широко применяются кривые, которыми легко управлять. Управление кривыми осуществляется путем изменения данных, на основе которых они построены. Кривые могут быть построены с помощью аналитических функций, по набору точек, на базе других кривых и на базе поверхностей.

2.1. Кривая

Кривой будем называть векторную функцию

$$\mathbf{r}(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \\ r_3(t) \end{bmatrix} \quad (2.1.1)$$

скалярного параметра t , принимающего значения на отрезке $[t_{\min}, t_{\max}]$. Пусть координаты $r_1(t)$, $r_2(t)$, $r_3(t)$ точки кривой $\mathbf{r}(t)$ являются однозначными непрерывными функциями параметра t . Такое описание кривой называется *параметрическим*.

Точку кривой будем называть *обыкновенной*, если в этой точке не обращается в ноль длина первой производной кривой по параметру. В противном случае точку кривой будем называть *особой*.

Кривую будем называть *периодической*, если существует $p > 0$ такое, что $\mathbf{r}(t \pm kp) = \mathbf{r}(t)$, где k – целое число. Для устранения неоднозначности область определения периодической кривой должна лежать в пределах одного периода. Периодическую кривую будем называть *циклически замкнутой*, если $p = t_{\max} - t_{\min}$.

Область изменения параметра кривой есть отрезок $[t_{\min}, t_{\max}]$ в одномерном пространстве. Кривая представляет собой непрерывное отображение некоторого участка числовой оси в трехмерное пространство.

Введем обозначения

$$\mathbf{r}' = \frac{d\mathbf{r}}{dt}, \quad \mathbf{r}'' = \frac{d^2\mathbf{r}}{dt^2}, \quad \mathbf{r}''' = \frac{d^3\mathbf{r}}{dt^3}$$

для производных кривой. В обыкновенной точке производная кривой есть вектор, направленный по касательной к кривой в сторону возрастания параметра. Единичный вектор

$$\mathbf{t} = \frac{\mathbf{r}'}{s'}, \quad (2.1.2)$$

где $s' = \sqrt{\mathbf{r}' \cdot \mathbf{r}'}$ — длина первой производной кривой, называется *касательным вектором*. Длина первой производной зависит от способа параметризации кривой. Если в качестве параметра используется длина дуги кривой, отсчитываемая от некоторой ее точки, то длина производной равна единице. Векторная функция $\mathbf{r}(s)$, где s — длина дуги, называется кривой с *натуральной параметризацией*.

Предположим, что нам дана кривая с натуральной параметризацией. В этом случае касательный вектор равен первой производной кривой:

$$\mathbf{t} = \frac{d\mathbf{r}}{ds}.$$

Построим в рассматриваемой точке кривой плоскость, перпендикулярную ее первой производной. Такая плоскость называется *нормальной плоскостью кривой*. Плоскость, в которой лежат и первая производная кривой, и ее вторая производная, называется *соприкасающейся плоскостью*. Если вторая производная кривой параллельна первой производной или ее длина равна нулю, то в качестве соприкасающейся плоскости можно взять любую плоскость, в которой лежит первая производная кривой. Точка кривой, в которой векторы первой и второй производных кривой коллинеарны, называется *точкой распрямления*. Точки распрямления не зависят от способа параметризации кривой. Название соприкасающейся плоскости обусловлено тем, что она проходит через заданную точку кривой с наивысшим порядком касания, и ее можно определить как предельное положение плоскости, построенной по трем бесконечно близким точкам кривой. Плоскость, перпендикулярная нормальной и соприкасающейся плоскостям, называется *спрямляющей плоскостью*.

Единичный вектор, ортогональный касательному вектору, лежащий в соприкасающейся плоскости и направленный в сторону второй производной (в сторону вогнутости кривой), называется *главной нормалью кривой*. Главную нормаль обозначим через \mathbf{n} . Векторы первой и второй производной кривой с натуральной параметризацией ортогональны, так как $\frac{d(\mathbf{t} \cdot \mathbf{t})}{ds} = 2\mathbf{t} \cdot \frac{d\mathbf{t}}{ds} = 2\mathbf{t} \cdot \frac{d^2\mathbf{r}}{ds^2} = 0$. Следовательно, вторая производная кривой с натуральной параметризацией направлена по главной нормали

$$\frac{dt}{ds} = \frac{d^2 \mathbf{r}}{ds^2} = k\mathbf{n}.$$

Коэффициент k называется *кривизной* кривой. Обратная кривизне кривой величина равна радиусу соприкасающейся с кривой в рассматриваемой точке окружности.

Единичный вектор, направленный вдоль линии пересечения нормальной и спрямляющей плоскостей и образующий с касательным и нормальным вектором правую тройку векторов, называется *бинормалью* кривой. Бинормаль обозначим через \mathbf{b} . Касательный вектор \mathbf{t} , главная нормаль \mathbf{n} , бинормаль \mathbf{b} кривой приведены на рис. 2.1.1.

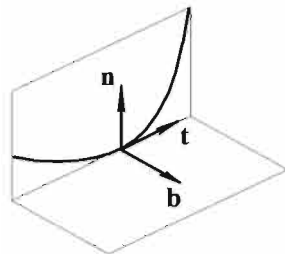


Рис. 2.1.1

Бинормаль по определению ортогональна касательному вектору и главной нормали. Из этого следует, что

$$\begin{aligned} \frac{d(\mathbf{b} \cdot \mathbf{t})}{ds} &= \mathbf{t} \cdot \frac{d\mathbf{b}}{ds} + \mathbf{b} \cdot \frac{d\mathbf{t}}{ds} = \mathbf{t} \cdot \frac{d\mathbf{b}}{ds} = 0, \\ \frac{d(\mathbf{b} \cdot \mathbf{b})}{ds} &= 2\mathbf{b} \cdot \frac{d\mathbf{b}}{ds} = 0. \end{aligned}$$

Таким образом, производная бинормали по длине дуги кривой ортогональна векторам \mathbf{t} и \mathbf{b} и, следовательно, она параллельна главной нормали. Это принято записывать в виде

$$\frac{d\mathbf{b}}{ds} = -\chi\mathbf{n}.$$

Коэффициент χ называется *кручением* кривой.

Используя производные касательного вектора и бинормали кривой с натуральной параметризацией, найдем производную нормали по длине дуги:

$$\frac{d\mathbf{n}}{ds} = \frac{d(\mathbf{b} \times \mathbf{t})}{ds} = \frac{d\mathbf{b}}{ds} \times \mathbf{t} + \mathbf{b} \times \frac{d\mathbf{t}}{ds} = -\chi\mathbf{n} \times \mathbf{t} + k\mathbf{b} \times \mathbf{n} = \chi\mathbf{b} - k\mathbf{t}.$$

Формулы Френе–Серре

$$\frac{d\mathbf{t}}{ds} = k\mathbf{n}, \quad \frac{d\mathbf{n}}{ds} = \chi\mathbf{b} - k\mathbf{t}, \quad \frac{d\mathbf{b}}{ds} = -\chi\mathbf{n}$$

выражают дифференциальные зависимости между касательным вектором \mathbf{t} , главной нормалью \mathbf{n} и бинормалью \mathbf{b} . Векторы \mathbf{t} , \mathbf{n} , \mathbf{b} связаны соотношениями

$$\mathbf{b} = \mathbf{t} \times \mathbf{n}, \quad \mathbf{t} = \mathbf{n} \times \mathbf{b}, \quad \mathbf{n} = \mathbf{b} \times \mathbf{t}$$

и могут использоваться в качестве ортов локальной системы координат, связанной с кривой.

Совокупность нормальной, соприкасающейся, спрямляющей плоскостей и векторов \mathbf{t} , \mathbf{n} , \mathbf{b} называется *сопровождающим трехгранником кривой*. Кручение кривой равно угловой скорости вращения сопровождающего трехгранника вокруг касательного вектора при его движении вдоль кривой. Если кручение равно нулю, то кривая является плоской. Полный вектор угловой скорости вращения сопровождающего трехгранника по отношению к пути, проходимому по кривой, называется **вектором Дарбу**. Он равен

$$\boldsymbol{\omega} = k\mathbf{b} + \chi\mathbf{t}.$$

Вектор Дарбу придает механический смысл формулам Френе–Серре, с использованием которого последние имеют вид

$$\frac{d\mathbf{t}}{ds} = \boldsymbol{\omega} \times \mathbf{t}, \quad \frac{d\mathbf{n}}{ds} = \boldsymbol{\omega} \times \mathbf{n}, \quad \frac{d\mathbf{b}}{ds} = \boldsymbol{\omega} \times \mathbf{b}.$$

Кривизну и кручение кривой с натуральной параметризацией вычислим по формулам

$$k = \left| \frac{d^2\mathbf{r}}{ds^2} \right|, \quad \chi = \frac{1}{k^2} \left(\frac{d\mathbf{r}}{ds} \times \frac{d^2\mathbf{r}}{ds^2} \right) \cdot \frac{d^3\mathbf{r}}{ds^3}.$$

Положение точки кривой зависит от параметра s , который является ее внутренней координатой. Параметр s своими значениями однозначно определяет точки кривой. Но так как геометрически он ничем не связан с кривой, то для нее можно использовать другую параметризацию. Для этого достаточно ввести новый параметр t , который связан с параметром s функциональной зависимостью $s = s(t)$. Будем полагать, что функция $s(t)$ является взаимно однозначной и монотонно возрастающей. Кривая $\mathbf{r}(s)$, выраженная через параметр t , будет иметь вид

$$\mathbf{r}(t) = \mathbf{r}(s(t)),$$

при этом форма кривой останется прежней. Как параметр s , так и параметр t одинаково годятся, чтобы характеризовать точки рассматриваемой кривой, и выбор параметра зависит от нас. В частном случае параметром кривой может служить длина ее дуги, отсчитываемая от начальной точки. В общем случае параметр t , как любая координатная система, может быть определен удобным для нас способом.

Как правило, параметризация кривой не является натуральной. Используя приведенные формулы для кривой с натуральной параметризацией, получим формулы, связывающие главную нормаль, бинормаль, кривизну и кручение с производными кривой при произвольной параметризации. Производные кривой с различной параметризацией связаны соотношениями:

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= \frac{d\mathbf{r}}{ds} \frac{ds}{dt}, \\ \frac{d^2\mathbf{r}}{dt^2} &= \frac{d^2\mathbf{r}}{ds^2} \left(\frac{ds}{dt}\right)^2 + \frac{d\mathbf{r}}{ds} \frac{d^2s}{dt^2}, \\ \frac{d^3\mathbf{r}}{dt^3} &= \frac{d^3\mathbf{r}}{ds^3} \left(\frac{ds}{dt}\right)^3 + 3 \frac{d^2\mathbf{r}}{ds^2} \frac{ds}{dt} \frac{d^2s}{dt^2} + \frac{d\mathbf{r}}{ds} \frac{d^3s}{dt^3}. \end{aligned}$$

В общем случае кривизну кривой вычислим по формуле

$$k = \frac{|\mathbf{r}' \times \mathbf{r}''|}{|\mathbf{r}'|^3}. \quad (2.1.3)$$

В точках кривой с отличной от нуля кривизной главная нормаль, бинормаль, радиус кривизны ρ и кручение кривой вычислим по формулам

$$k\mathbf{n} = \frac{\mathbf{r}''}{|\mathbf{r}'|^2} - \frac{\mathbf{r}' \cdot \mathbf{r}''}{|\mathbf{r}'|^4} \mathbf{r}', \quad \mathbf{b} = \frac{\mathbf{r}' \times \mathbf{r}''}{k |\mathbf{r}'|^3}, \quad \rho = \frac{|\mathbf{r}'|^3}{|\mathbf{r}' \times \mathbf{r}''|}, \quad \chi = \frac{(\mathbf{r}' \times \mathbf{r}''') \cdot \mathbf{r}''}{|\mathbf{r}' \times \mathbf{r}''|^2}. \quad (2.1.4)$$

Из приведенных формул видно, что кривизна всегда неотрицательна, а кручение может иметь любой знак. Если кривизна равна нулю, то направление главной нормали, бинормали и кручение не определены. Если кривизна равна нулю в каждой точке кривой, то кривая является отрезком прямой. Главная нормаль в этом случае может иметь произвольное направление в нормальной плоскости. Если первая, вторая и третья производные кривой лежат в одной плоскости, то кручение кривой равно нулю и кривая является плоской.

При описании поверхностей нам придется работать с точками и векторами в двумерном пространстве параметров поверхности. Для описания границы области определения поверхности потребуются двумерные кривые.

Двумерной кривой будем называть векторную функцию

$$\mathbf{r}(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \quad (2.1.5)$$

скалярного параметра t , изменяющегося в пределах $t_{\min} \leq t \leq t_{\max}$. Напомним, что для обозначения столбца координат точек в двумерном пространстве используются строчные буквы латинского алфавита, выделенные полужирным наклонным шрифтом. Пусть координаты $u(t)$, $v(t)$ точки двумерной кривой являются однозначными непрерывными функциями параметра t . Двумерные кривые, так же как и трехмерные, могут быть периодическими и циклически замкнутыми.

Большинство формул для двумерных кривых можно получить из соответствующих формул для трехмерных кривых, положив в них третью координату равной нулю.

Для двумерной кривой $\mathbf{r}(s)$ формулы Френе–Серре имеют вид

$$\frac{d\mathbf{t}}{ds} = k\mathbf{n}, \quad \frac{d\mathbf{n}}{ds} = -k\mathbf{t},$$

где $\mathbf{t} = d\mathbf{r}/ds$ — касательный вектор кривой; \mathbf{n} — нормаль кривой; s — длина дуги кривой. Кривизна двумерной кривой определяется равенством

$$k = \left| \frac{d^2\mathbf{r}}{ds^2} \right|.$$

Методы построения двумерных кривых аналогичны методам построения соответствующих трехмерных кривых.

2.2. Аналитические кривые

Аналитическими кривыми будем называть кривые, координаты которых в некоторой локальной системе координат можно описать с помощью аналитических функций, не используя точки, векторы и другие кривые.

Для аналитических кривых используются локальные системы координат, в которых кривые имеют канонический вид.

Построим локальную декартову прямоугольную систему координат с началом в точке \mathbf{p} и базисными векторами $\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z$. Кривая, координаты которой в локальной системе равны соответственно $x(t)$, $y(t)$, $z(t)$, будет описываться векторной функцией

$$\mathbf{r}(t) = \mathbf{p} + x(t) \mathbf{i}_x + y(t) \mathbf{i}_y + z(t) \mathbf{i}_z. \quad (2.2.1)$$

Пусть p_i – координаты начала \mathbf{p} локальной системы координат, x_i – компоненты базисного вектора \mathbf{i}_x , y_i – компоненты базисного вектора \mathbf{i}_y , z_i – компоненты базисного вектора \mathbf{i}_z , $i = 1, 2, 3$. Тогда аналитическая кривая будет представлять собой функцию

$$\begin{bmatrix} r_1(t) \\ r_2(t) \\ r_3(t) \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \cdot \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix}.$$

Координаты кривой (2.2.1) равны

$$r_i(t) = p_i + x(t) x_i + y(t) y_i + z(t) z_i.$$

При изменении положения или ориентации подобным образом описанной аналитической кривой изменяются координаты начала местной системы координат и ее базисные векторы, а аналитические функции кривой остаются неизменными, сохраняя канонический вид. Рассмотрим примеры аналитических кривых.

Дугу эллипса опишем векторной функцией

$$\mathbf{r}(t) = \mathbf{p} + a \cos t \mathbf{i}_x + b \sin t \mathbf{i}_y, \quad (2.2.2)$$

$$t_{\min} \leq t \leq t_{\max},$$

где a и b – полуоси эллипса, $0 \leq t_{\min} < 2\pi$ – начальный параметр дуги $t_{\min} < t_{\max} \leq \alpha + 2\pi$ – конечный параметр дуги. Начало \mathbf{p} локальной системы координат мы расположили в центре эллипса, базисные векторы \mathbf{i}_x и \mathbf{i}_y направили вдоль полуосей. Скалярные функции $x(t) = a \cos t$, $y(t) = b \sin t$ эллипса связаны уравнением

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1.$$

При области определения параметра $t_{\max} = t_{\min} + 2\pi$ получим **эллипс**. Эллипс является циклически замкнутой кривой. Эллипс показан на рис. 2.2.1.

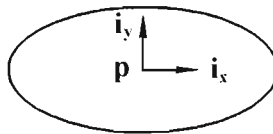


Рис. 2.2.1

Если $t_{\max} - t_{\min} < 2\pi$, то получим дугу эллипса. При $a = b$ и $t_{\max} = t_{\min} + 2\pi$ мы получим окружность.

Дугу гиперболы опишем векторной функцией

$$\mathbf{r}(t) = \mathbf{p} + a \operatorname{ch} t \mathbf{i}_x + b \operatorname{sh} t \mathbf{i}_y, \\ t_{\min} \leq t \leq t_{\max},$$

где a и b — действительная и мнимая полуоси гиперболы, t_{\min} — начальный параметр дуги, $t_{\min} < t_{\max}$ — конечный параметр дуги. Скалярные функции $x(t) = a \operatorname{ch} t$, $y(t) = b \operatorname{sh} t$ гиперболы связаны уравнением

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1.$$

Дугу параболы опишем векторной функцией

$$\mathbf{r}(t) = \mathbf{p} + f t^2 \mathbf{i}_x + 2ft \mathbf{i}_y, \\ t_{\min} \leq t \leq t_{\max},$$

где f — фокусное расстояние параболы; t_{\min} — начальный параметр дуги; $t_{\min} < t_{\max}$ — конечный параметр дуги. Скалярные функции $x(t) = f t^2$, $y(t) = 2ft$ параболы связаны уравнением

$$4fx - y^2 = 0.$$

Фокус параболы располагается в точке с радиусом-вектором $\mathbf{p} + f \mathbf{i}_x$. Директрисой параболы является прямая $\mathbf{r}(w) = \mathbf{p} - f \mathbf{i}_x + w \mathbf{i}_y$.

Спираль радиусом r , шагом h и параметрической длиной $t_{\max} - t_{\min}$ опишем векторной функцией

$$\mathbf{r}(t) = \mathbf{p} + r \cos(t) \mathbf{i}_x + r \sin(t) \mathbf{i}_y + \frac{ht}{2\pi} \mathbf{i}_z. \quad (2.2.3)$$

Начало локальной системы координат спирали мы расположили в точке пересечения оси и торцевой плоскости спирали, базисный вектор \mathbf{i}_z направили вдоль оси спирали. При равенстве длин и ортогональности базисных векторов локальной системе координат мы получим цилиндрическую спираль, рис. 2.2.2.

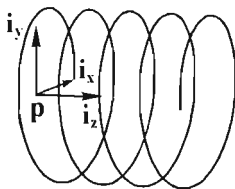


Рис. 2.2.2

При $h=0$ данная формула (2.2.3) будет описывать радиус-вектор окружности или ее дуги в плоскости базисных векторов \mathbf{i}_x и \mathbf{i}_y .

2.3. Кривые, построенные по набору точек

Рассмотрим построение кривых, которые при значениях параметра t_0, t_1, \dots, t_n проходят через заданные точки $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n$. Точки $\mathbf{p}_i, i = 0, 1, \dots, n$ называют *контрольными точками* (control points) кривой, а параметры $t_i, i = 0, 1, \dots, n$ называют *узлами*. Узловое значение параметра для каждой последующей опорной точки \mathbf{p}_{i+1} должно быть больше узлового значения параметра для предыдущей опорной точки \mathbf{p}_i : $t_i < t_{i+1}, i = 0, \dots, n - 1$. Кривые, построенные на множестве равноотстоящих узлов, называются *однородными* (uniform). Для однородных кривых выполняется равенство: $t_i - t_{i-1} = t_{i+1} - t_i$. Кривые, построенные на множестве неравноотстоящих узлов, называются *неоднородными* (non-uniform).

Ломаная линия является простейшей кривой, построенной по набору точек. Она состоит из отрезков, последовательно соединяющих заданные точки. Радиус-вектор ломаной линии определяется равенством

$$\mathbf{r}(t) = \mathbf{p}_i(1-w) + \mathbf{p}_{i+1}w, \quad (2.3.1)$$

где $w = \frac{t-t_i}{t_{i+1}-t_i}$, а $t_i \leq t \leq t_{i+1}$. Параметр w будем называть местным параметром на участке кривой между точками \mathbf{p}_i и \mathbf{p}_{i+1} .

Первая производная ломаной линии в контрольных точках \mathbf{p}_i терпит разрыв по длине и по направлению. В некоторых случаях значение параметра для точки \mathbf{p}_i можно принять равным номеру точки: $t_i = i$. Ломаная линия приведена на рис. 2.3.1.

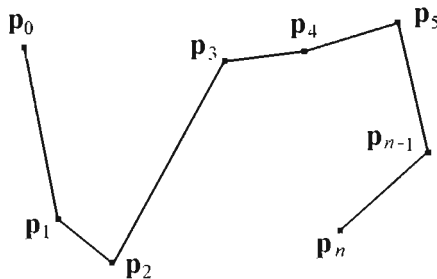


Рис. 2.3.1

Ломаная может быть замкнутой, в этом случае первая опорная точка одновременно является и последней. Ломаная обладает рядом полезных свойств: работа с ней требует минимум вычислений, проекция ломаной линии также будет ломаной линией.

Отрезок прямой мы получим в частном случае, когда ломаная линия содержит один участок. Радиус-вектор отрезка, построенного по точкам \mathbf{p}_0 и \mathbf{p}_1 , вычислим по формуле

$$\mathbf{r}(t) = \mathbf{p}_0(1-t) + \mathbf{p}_1 t.$$

Кривые, построенные по набору точек и обладающие определенной гладкостью, называют **сплайнами**. Термин «сплайн» для кривых заимствован у названия чертежного инструмента – упругой гибкой линейки, которая может изгибаться так, чтобы проходить через заданные точки.

Интерполяционная формула Лагранжа позволяет построить кривую по заданным точкам в виде

$$\mathbf{r}(t) = \sum_{i=0}^n L_i^n(t) \mathbf{p}_i, \quad t_0 \leq t \leq t_n, \quad (2.3.2)$$

где коэффициенты $L_i^n(t)$ равны единице при $t = t_i$ и равны нулю в остальных узлах $t_j, j \neq i$. Коэффициенты $L_i^n(t)$ представляют собой многочлены степени $n - 1$

$$L_i^n(t) = \lambda_i(t - t_0) \dots (t - t_{i-1})(t - t_{i+1}) \dots (t - t_n).$$

Из условия $L_i^n(t_i) = 1$ следует, что $\frac{1}{\lambda_i} = (t_i - t_0) \dots (t_i - t_{i-1})(t_i - t_{i+1}) \dots (t_i - t_n)$ и

$$L_i^n(t) = \frac{\prod_{j=0, j \neq i}^n (t - t_j)}{\prod_{j=0, j \neq i}^n (t_i - t_j)}. \quad (2.3.3)$$

Итак, радиус-вектор **сплайна Лагранжа** описывается формулой

$$\mathbf{r}(t) = \sum_{i=0}^n \frac{\prod_{j=0, j \neq i}^n (t - t_j)}{\prod_{j=0, j \neq i}^n (t_i - t_j)} \mathbf{p}_i. \quad (2.3.4)$$

Сплайн Лагранжа можно записать в виде

$$\mathbf{r}(t) = \sum_{i=0}^n \frac{W_{0,n}(t)}{(t - t_i)W_{0,n}'(t_i)} \mathbf{p}_i,$$

где $W_{0,n}(t) = (t - t_0)(t - t_1) \dots (t - t_{n-1})(t - t_n)$,

$$W_{0,n}'(t_i) = \frac{dW_{0,n}}{dt} \Big|_{t_i} = (t_i - t_0) \dots (t_i - t_{i-1})(t_i - t_{i+1}) \dots (t_i - t_n).$$

На форму сплайна оказывает влияние не только расположение контрольных точек, но и значения узлов. На рис. 2.3.2 приведен сплайн Лагранжа, изменение узловых значений параметра которого пропорционально расстоянию между соседними точками:

$$\frac{t_{i+1} - t_i}{t_i - t_{i-1}} \approx \frac{|\mathbf{p}_{i+1} - \mathbf{p}_i|}{|\mathbf{p}_i - \mathbf{p}_{i-1}|}.$$

Тонкой штриховой линией на рис. 2.3.2 приведен сплайн Лагранжа с однородной параметризацией ($t_i = i$).

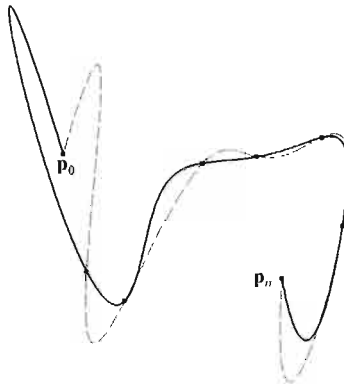


Рис. 2.3.2

На рис. 2.3.2 можно видеть, что на форму кривой влияет не только положение контрольных точек, но и соответствующие им значения параметров, то есть параметризация кривой.

Сплайн Ньютона, проходящий через заданные точки, ищем в форме

$$\mathbf{r}(t) = \mathbf{a}_0 + \mathbf{a}_1(t - t_0) + \mathbf{a}_2(t - t_0)(t - t_1) + \dots + \mathbf{a}_n(t - t_0)(t - t_1) \dots (t - t_{n-1}). \quad (2.3.5)$$

Векторы \mathbf{a}_i определим из условия прохождения векторной функции $\mathbf{r}(t)$ через точки \mathbf{p}_i при значениях параметра t_i , $i = 0, 1, \dots, n$

$$\begin{aligned} \mathbf{p}_0 &= \mathbf{a}_0, \\ \mathbf{p}_1 &= \mathbf{a}_0 + \mathbf{a}_1(t_1 - t_0), \\ \mathbf{p}_2 &= \mathbf{a}_0 + \mathbf{a}_1(t_2 - t_0) + \mathbf{a}_2(t_2 - t_0)(t_2 - t_1), \\ &\dots \dots \dots \\ \mathbf{p}_n &= \mathbf{a}_0 + \mathbf{a}_1(t_n - t_0) + \mathbf{a}_2(t_n - t_0)(t_n - t_1) + \dots + \mathbf{a}_n(t_n - t_0)(t_n - t_1) \dots (t_n - t_{n-1}). \end{aligned} \quad (2.3.6)$$

Матрица этой системы уравнений является треугольной, и система решается за один прямой проход. Первое уравнение системы (2.3.6) дает \mathbf{a}_0 . Подставляя этот результат во второе уравнение, получим \mathbf{a}_1 . Из третьего уравнения найдем \mathbf{a}_2 и так далее до \mathbf{a}_n . Вставляя дополнительную точку \mathbf{p}_{i+1} в сплайн, мы увеличиваем систему (2.3.6) на одно уравнение, а остальные уравнения остаются неизменными. Это является преимуществом сплайна Ньютона по сравнению со сплайном Лагранжа при построении кривых с увеличивающимся числом точек. Формулы (2.3.2) и (2.3.5) описывают одну и ту же кривую. Сплайны Лагранжа и Ньютона редко используются в геометрическом моделировании из-за их поведения, в том числе и из-за сильной зависимости от значений параметров в контрольных точках.

Построим на заданной совокупности контрольных точек **кубический сплайн**, который бы имел непрерывными первые и вторые производные радиуса-вектора. На каждом участке между соседними контрольными точками будем описывать радиус-вектор кривой кубическим полиномом. Вторые производные радиуса-вектора сплайна в контрольных точках t_i обозначим через \mathbf{s}_i . Вторая производная радиуса-вектора сплайна на участке $t_i \leq t \leq t_{i+1}$ по определению является линейной функцией параметра t :

$$\frac{d^2 \mathbf{r}}{dt^2} = \mathbf{s}_i \frac{t_{i+1} - t}{t_{i+1} - t_i} + \mathbf{s}_{i+1} \frac{t - t_i}{t_{i+1} - t_i}.$$

После двукратного интегрирования получим

$$\mathbf{r}(t) = \mathbf{s}_i \frac{(t_{i+1} - t)^3}{6(t_{i+1} - t_i)} + \mathbf{s}_{i+1} \frac{(t - t_i)^3}{6(t_{i+1} - t_i)} + \mathbf{c}_1 t + \mathbf{c}_2.$$

Постоянные интегрирования \mathbf{c}_1 и \mathbf{c}_2 определим из условий на концах участка $\mathbf{r}(t_i) = \mathbf{p}_i$ и $\mathbf{r}(t_{i+1}) = \mathbf{p}_{i+1}$. После вычислений получим

$$\mathbf{r}(t) = \mathbf{s}_i \frac{(t_{i+1} - t)^3}{6(t_{i+1} - t_i)} + \mathbf{s}_{i+1} \frac{(t - t_i)^3}{6(t_{i+1} - t_i)} + \left(\frac{\mathbf{p}_i}{t_{i+1} - t_i} - \mathbf{s}_i \frac{t_{i+1} - t_i}{6} \right) (t_{i+1} - t) + \left(\frac{\mathbf{p}_{i+1}}{t_{i+1} - t_i} - \mathbf{s}_{i+1} \frac{t_{i+1} - t_i}{6} \right) (t - t_i). \quad (2.3.7)$$

Выражение (2.3.7) описывает кубический полином на отрезке $t_i \leq t \leq t_{i+1}$ и содержит две неизвестные величины \mathbf{s}_i и \mathbf{s}_{i+1} . Для их определения приравняем первую производную сплайна на правом конце отрезка $t_{i+1} \leq t \leq t_i$ первой производной сплайна на левом конце отрезка $t_i \leq t \leq t_{i+1}$. После дифференцирования (2.3.7) и подстановки $t = t_i$ получим

$$\left. \frac{d\mathbf{r}}{dt} \right|_{t_i} = -\frac{(2\mathbf{s}_i + \mathbf{s}_{i+1})(t_{i+1} - t_i)}{6} + \frac{\mathbf{p}_{i+1} - \mathbf{p}_i}{t_{i+1} - t_i}.$$

После замены в (2.3.7) i на $i-1$, дифференцирования и подстановки $t=t_i$, получим

$$\left. \frac{d\mathbf{r}}{dt} \right|_{t_i} = \frac{(2\mathbf{s}_i + \mathbf{s}_{i-1})(t_i - t_{i-1})}{6} + \frac{\mathbf{p}_i - \mathbf{p}_{i-1}}{t_i - t_{i-1}}.$$

Приравняв правые части последних двух выражений, получим следующее уравнение:

$$\mathbf{s}_{i-1}(t_i - t_{i-1}) + 2\mathbf{s}_i(t_{i+1} - t_{i-1}) + \mathbf{s}_{i+1}(t_{i+1} - t_i) = 6\frac{\mathbf{p}_{i+1} - \mathbf{p}_i}{t_{i+1} - t_i} - 6\frac{\mathbf{p}_i - \mathbf{p}_{i-1}}{t_i - t_{i-1}}. \quad (2.3.8)$$

Таких уравнений можно составить для $n-1$ контрольных точек. Подлежат определению $n+1$ неизвестных векторов \mathbf{s}_i . Два недостающих уравнения для определения всех неизвестных составим, исходя из условий на концах кривой. Например, если считать концы кривой свободными, то можно положить $\mathbf{s}_0 = \mathbf{s}_n = 0$. В некоторых случаях можно принять $\mathbf{s}_0 = \mathbf{s}_1$, $\mathbf{s}_n = \mathbf{s}_{n-1}$, тогда на концевых участках сплайн будет иметь постоянную кривизну. Дополнительные условия зависят от геометрических условий в каждом конкретном случае. Если требуется построить замкнутую кривую, то мы будем иметь равное число неизвестных и уравнений. Таким образом, неизвестные векторы вторых производных \mathbf{s}_i в контрольных точках найдем из системы линейных уравнений. Матрица этой системы линейных уравнений трехдиагональная, что упрощает решение системы. Правую часть выражения (2.3.7) можно представить в виде

$$\mathbf{r}(t) = (1-w)\mathbf{p}_i + w\mathbf{p}_{i+1} + ((-2w+3w^2-w^3)\mathbf{s}_i + (-w+w^3)\mathbf{s}_{i+1})\frac{(t_{i+1}-t_i)^2}{6}, \quad (2.3.9)$$

где $w = \frac{t-t_i}{t_{i+1}-t_i}$, а $t_i \leq t \leq t_{i+1}$. Полученная кривая называется **кубическим сплайном**.

На форму сплайна оказывает влияние расположение контрольных точек и значения узлов. Если есть свобода выбора узловых значений параметра, то для избежания петель и необоснованных изгибов сплайнов желательно использовать параметризацию, пропорциональную натуральной параметризации сплайна. При натуральной параметризации кривой за параметр принимается длина кривой, отсчитываемая от некоторой начальной точки. Хорошие результаты дает также параметризация, при которой изменение узловых значений параметра пропорционально расстоянию между соседними точками.

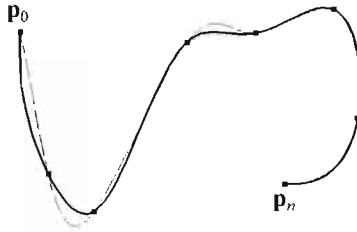


Рис. 2.3.3

На рис. 2.3.3 приведен кубический сплайн, изменение узловых значений параметра которого пропорционально расстоянию между соседними точками. Тонкой штриховой линией на рис. 2.3.3 приведен кубический сплайн с однородной параметризацией ($t_i = i$). При однородной параметризации изменение формы сплайна будет тем больше, чем менее равномерно расположены контрольные точки.

По аналогии с кубическим сплайном могут быть построены сплайны более высокой степени, когда производные радиуса-вектора третьего и более высокого порядка непрерывны в контрольных точках. В предельном случае, сохраняющем непрерывность производных n -го порядка при $n + 1$ контрольных точках, мы построим сплайн Лагранжа или сплайн Ньютона.

Если задана последовательность $m + 1$ точек, через которую должна пройти кривая, и производные ее радиуса-вектора в этих точках, то по этим данным можно построить сплайн, описываемый полиномом степени $2m + 1$ и носящий имя Эрмита. Мы рассмотрим частный случай **сплайна Эрмита** для $m = 1$.

Ломаную линию можно рассматривать как составную кривую, построенную из отрезков прямой линии. По аналогии можно построить составную кубическую кривую, состоящую из сплайнов Эрмита третьей степени, гладко стыкующихся между собой. Построим составной сплайн Эрмита, проходящий через заданную последовательность точек и имеющий в этих точках заданные производные. Пусть радиусы-векторы этих точек равны \mathbf{p}_i , векторы производных кривой в этих точках равны \mathbf{q}_i , а значения параметра в этих точках равны t_i ($t_i < t_{i+1}$), где $i = 0, 1, 2, \dots, n - \text{номера точек}$. На участке между точками \mathbf{p}_i и \mathbf{p}_{i+1} составной сплайн Эрмита является полиномом третьей степени местного параметра

$$w = \frac{t - t_i}{t_{i+1} - t_i};$$

$$\mathbf{r}_i(w) = \mathbf{a}_0 + \mathbf{a}_1 w + \mathbf{a}_2 w^2 + \mathbf{a}_3 w^3.$$

На рис. 2.3.4 приведен участок сплайна Эрмита.



Рис. 2.3.4

Местный параметр w изменяется от 0 до 1. Векторы $\mathbf{a}_j, j = 0, 1, 2, 3$ найдем из условий на границе участка кривой:

$$\mathbf{r}_i(0) = \mathbf{p}_i, \mathbf{r}_i(1) = \mathbf{p}_{i+1}, \mathbf{r}_i'(0) = \mathbf{q}_i, \mathbf{r}_i'(1) = \mathbf{q}_{i+1}.$$

После решения этой системы уравнений и подстановки значений $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ получим зависимость радиуса-вектора для сплайна Эрмита:

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{p}_i (1 - 3w^2 + 2w^3) + \mathbf{p}_{i+1} (3w^2 - 2w^3) + \\ &+ \mathbf{q}_i (w - 2w^2 + w^3)(t_{i+1} - t_i) + \mathbf{q}_{i+1} (-w^2 + w^3)(t_{i+1} - t_i) = \\ &= \alpha_0(w) \mathbf{p}_i + \alpha_1(w) \mathbf{p}_{i+1} + \beta_0(w) (t_{i+1} - t_i) \mathbf{q}_i + \beta_1(w) (t_{i+1} - t_i) \mathbf{q}_{i+1}, \end{aligned} \quad (2.3.10)$$

где $w = \frac{t - t_i}{t_{i+1} - t_i}$ и $t_i \leq t \leq t_{i+1}$. В (2.3.10) введены обозначения для функций

$$\begin{aligned} \alpha_0(w) &= 1 - 3w^2 + 2w^3, \\ \alpha_1(w) &= 3w^2 - 2w^3, \\ \beta_0(w) &= w - 2w^2 + w^3, \\ \beta_1(w) &= -w^2 + w^3. \end{aligned} \quad (2.3.11)$$

Сплайн Эрмита можно строить в отсутствии производных в контрольных точках. Если производные \mathbf{q}_i неизвестны, то их можно вычислить, построив векторный полином Лагранжа второй степени по трем соседним точкам $\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1}$

$$\mathbf{p}(t) = \mathbf{p}_{i-1} \frac{(t - t_i)(t - t_{i+1})}{(t_{i-1} - t_i)(t_{i-1} - t_{i+1})} + \mathbf{p}_i \frac{(t - t_{i-1})(t - t_{i+1})}{(t_i - t_{i-1})(t_i - t_{i+1})} + \mathbf{p}_{i+1} \frac{(t - t_{i-1})(t - t_i)}{(t_{i+1} - t_{i-1})(t_{i+1} - t_i)}$$

и взяв в качестве \mathbf{q}_i его производную при $t = t_i$:

$$\mathbf{q}_i = \mathbf{p}_{i-1} \frac{t_i - t_{i+1}}{(t_{i-1} - t_i)(t_{i-1} - t_{i+1})} + \mathbf{p}_i \frac{2t_i - t_{i-1} - t_{i+1}}{(t_i - t_{i-1})(t_i - t_{i+1})} + \mathbf{p}_{i+1} \frac{t_i - t_{i-1}}{(t_{i+1} - t_{i-1})(t_{i+1} - t_i)}.$$

Производные на краях сплайна можно получить из условия, что в этих точках обращаются в нуль вторые производные радиуса-вектора. Для этого вычислим по (2.3.10) вторые производные для соответствующих участков и подставим в них соответствующие значения параметра, в результате получим

$$\mathbf{q}_0 = \frac{3}{2} \frac{\mathbf{p}_1 - \mathbf{p}_0}{t_1 - t_0} - \frac{1}{2} \mathbf{q}_1, \quad \mathbf{q}_n = \frac{3}{2} \frac{\mathbf{p}_n - \mathbf{p}_{n-1}}{t_n - t_{n-1}} - \frac{1}{2} \mathbf{q}_{n-1}.$$

На рис. 2.3.5 приведен составной сплайн Эрмита третьей степени, построенный данным способом, изменение узловых значений параметра которого пропорционально расстоянию между соседними контрольными точками. Тонкой штриховой линией на рис. 2.3.5 приведен составной сплайн Эрмита третьей степени с равномерной параметризацией ($t_i = i$).

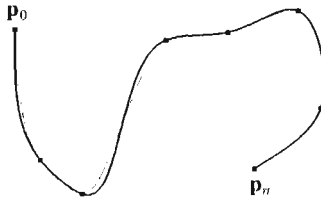


Рис. 2.3.5

Составной сплайн Эрмита дает приемлемую аппроксимацию и требует меньше вычислений, чем рассмотренные выше сплайны. Вторые производные в контрольных точках составного сплайна Эрмита в общем случае не сохраняют непрерывность.

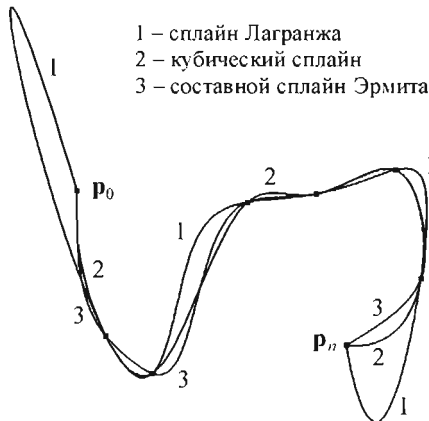


Рис. 2.3.6

На рис. 2.3.6 приведены сплайн Лагранжа (1), кубический сплайн (2) и составной сплайн Эрмита (3), построенные по одним и тем же данным. Мы видим, что сплайны ведут себя по-разному. Сплайны Эрмита и ку-

бические сплайны строятся для аппроксимации. Для эргономических целей используются кривые, которые мы рассмотрим далее.

2.4. Кривые Безье

Де Кастелье (De Casteljau F.) и независимо от него Безье (Bezier P.) предложили строить кривые, каждая точка которых вычисляется как взвешенная сумма заданных контрольных точек \mathbf{p}_i , $i = 0, 1, \dots, n$. Эти кривые получили название кривых Безье.

Кривая Безье первой степени ($n = 1$) представляет собой отрезок

$$\mathbf{r}(t) = (1 - t) \mathbf{p}_0 + t \mathbf{p}_1.$$

Радиус-вектор кривой Безье второй степени ($n = 2$) описывается зависимостью

$$\mathbf{r}(t) = (1 - t)^2 \mathbf{p}_0 + 2t(1 - t)\mathbf{p}_1 + t^2 \mathbf{p}_2.$$

Радиус-вектор кубической кривой Безье ($n = 3$) описывается зависимостью

$$\mathbf{r}(t) = (1 - t)^3 \mathbf{p}_0 + 3t(1 - t)^2 \mathbf{p}_1 + 3t^2(1 - t) \mathbf{p}_2 + t^3 \mathbf{p}_3.$$

На рис. 2.4.1 приведены кубическая кривая Безье и ломаная линия, построенные по одним и тем же контрольным точкам.

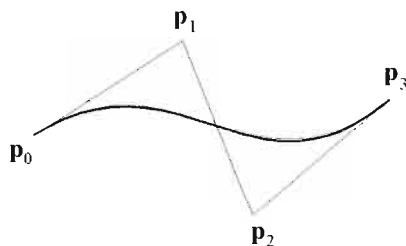


Рис. 2.4.1

Кривая Безье в общем случае описывается формулой

$$\mathbf{r}(t) = \sum_{i=0}^n B_i^n(t) \mathbf{p}_i = \sum_{i=0}^n \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i} \mathbf{p}_i, \quad t \in [0, 1]. \quad (2.4.1)$$

По точкам \mathbf{p}_i строится *контрольная ломаная* кривой Безье. На рис. 2.4.2 приведена кривая Безье, построенная по восьми точкам, и ее контрольная ломаная.

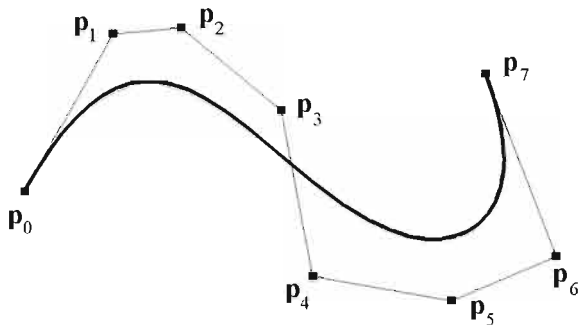


Рис. 2.4.2

Область изменения параметра кривой Безье любой степени принадлежит единичному отрезку: $0 \leq t \leq 1$. Кривая Безье не проходит через свои контрольные точки за исключением крайних точек. Контрольная ломаная касается кривой Безье в крайних точках. Гладкость Кривой Безье определяется числом точек. Путем перемещения одной или нескольких контрольных точек кривой Безье можно придать желаемую форму.

Форрест (A.R. Forrest) установил связь между коэффициентами при контрольных точках кривой Безье и полиномами Бернштейна. Коэффициентами при контрольных точках кривой Безье являются *функции Бернштейна* (Сергей Натанович Бернштейн)

$$B_i^n(t) = \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i}. \quad (2.4.2)$$

Совокупность функций Бернштейна для некоторого заданного n называется *базисом Бернштейна*. Коэффициенты при $t^i(1-t)^{n-i}$ в (2.4.2) равны коэффициентам $C_n^i = \frac{n!}{i!(n-i)!}$ бинома Ньютона: $(a+b)^n = \sum_{i=0}^n C_n^i a^{n-i} b^i$. Из

этого следует, что базис Бернштейна представляет собой разложение единицы:

$$\sum_{i=0}^n B_i^n(t) = (t + (1-t))^n = 1^n = 1. \quad (2.4.3)$$

На рис. 2.4.3 и рис. 2.4.4 приведены базисы Бернштейна третьей и четвертой степеней соответственно.

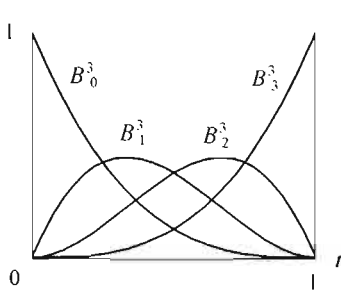


Рис. 2.4.3

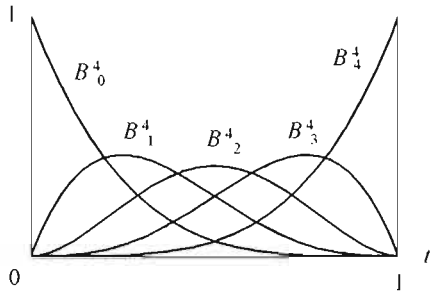


Рис. 2.4.4

Из полиномов Бернштейна только нулевой и последний принимают максимально возможные значения $B_0^n(0) = 1$, $B_n^n(1) = 1$, поэтому кривая Безье проходит только через начальную \mathbf{p}_0 и конечную \mathbf{p}_n точки.

Функции Бернштейна удовлетворяют рекуррентному соотношению:

$$B_i^n(t) = t B_{i-1}^{n-1}(t) + (1-t) B_{i+1}^{n-1}(t). \quad (2.4.4)$$

Это соотношение доказывается непосредственной подстановкой:

$$\begin{aligned} & t B_{i-1}^{n-1}(t) + (1-t) B_{i+1}^{n-1}(t) = \\ &= \frac{i}{n} \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i} + \frac{n-i}{n} \frac{n!}{i!(n-i)!} t^i (1-t)^{n-i} = B_i^n(t). \end{aligned}$$

Используя это рекуррентное соотношение, можно вычислить все функции Бернштейна. Вычисление начнем с функции $B_0^n(t) = 1$, далее получим $B_1^n(t) = t$, $B_2^n(t) = 1-t$, ..., $B_{n-1}^n(t) = (1-t)^n$, $B_n^n(t) = t^n$. При вычислении считается, что функции, один из индексов которых равен отрицательному числу, равны нулю.

Подставим рекуррентное соотношение (2.4.4) в выражение (2.4.1), выделив крайние точки и учитывая, что $B_0^n(t) = (1-t)^n$, $B_n^n(t) = t^n$, получим

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{p}_0(1-t)^n + \mathbf{p}_n t^n + \sum_{i=1}^{n-1} (t B_{i-1}^{n-1}(t) + (1-t) B_{i+1}^{n-1}(t)) \mathbf{p}_i = \\ &= (1-t) \left(\mathbf{p}_0(1-t)^{n-1} + \sum_{i=1}^{n-1} B_i^{n-1}(t) \mathbf{p}_i \right) + t \left(\mathbf{p}_n t^{n-1} + \sum_{i=1}^{n-1} B_i^{n-1}(t) \mathbf{p}_i \right) = \\ &= (1-t) \sum_{i=0}^{n-1} B_i^{n-1}(t) \mathbf{p}_i + t \sum_{i=0}^{n-1} B_i^{n-1}(t) \mathbf{p}_{i+1} = \\ &= (1-t) \mathbf{r}_0^{(n-1)}(t) + t \mathbf{r}_1^{(n-1)}(t), \end{aligned}$$

где $\mathbf{r}_0^{(n-1)}(t) = \sum_{i=0}^{n-1} B_i^{n-1}(t) \mathbf{p}_i$, $\mathbf{r}_1^{(n-1)}(t) = \sum_{i=0}^{n-1} B_i^{n-1}(t) \mathbf{p}_{i+1}$. Продолжив описанный процесс разложения для $\mathbf{r}_0^{(n-1)}$ и $\mathbf{r}_1^{(n-1)}$, в конце придем к равенствам

$$\mathbf{r}_i^{(0)} = \sum_{j=0}^0 B_j^0(t) \mathbf{p}_j = \mathbf{p}_i, \quad i = 0, 1, \dots, n.$$

Обозначив $\mathbf{r}(t)$ через $\mathbf{r}_0^{(n)}(t)$, а \mathbf{p}_i — через $\mathbf{r}_i^{(0)}$, получим рекуррентное соотношение для вычисления точки кривой Безье на k -ой итерации рекурсии:

$$\mathbf{r}_i^{(k)} = (1-t) \mathbf{r}_i^{(k-1)} + t \mathbf{r}_{i+1}^{(k-1)}, \quad i+k \leq n. \quad (2.4.5)$$

Алгоритм, описываемый соотношением (2.4.5), называется *алгоритмом Де Кастелье*. Алгоритм Де Кастелье позволяет вычислить любую точку кривой Безье по контрольным точкам, ничего не зная о функциях Бернштейна. Кривую Безье можно определить как линию, точки которой определяются рекуррентным соотношением (2.4.5).

Проиллюстрируем алгоритм Кастелье на примере квадратичной кривой Безье, рис. 2.4.5.

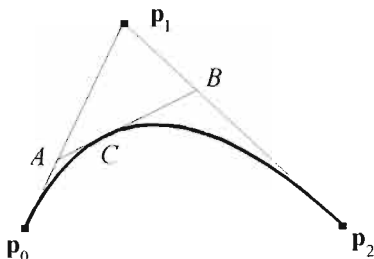


Рис. 2.4.5

Точка с произвольным параметром t квадратичной кривой Безье обладает тем свойством, что касательная в ней делит векторы $\mathbf{p}_1 - \mathbf{p}_0$ и $\mathbf{p}_2 - \mathbf{p}_1$ в отношении $\frac{t}{1-t}$. Это следует из записи кривой Безье второй степени в виде

$$\mathbf{r}(t) = (1-t) (\mathbf{p}_1 + (\mathbf{p}_0 - \mathbf{p}_1)(1-t)) + t (\mathbf{p}_1 + (\mathbf{p}_2 - \mathbf{p}_1)t).$$

В этом же отношении точкой кривой делится и часть касательной, заключенная внутри контрольной ломаной. Таким образом, $\mathbf{r}_0^{(0)} = \mathbf{p}_0$, $\mathbf{r}_1^{(0)} = \mathbf{p}_1$, $\mathbf{r}_2^{(0)} = \mathbf{p}_2$, точка A имеет радиус-вектор $\mathbf{r}_0^{(1)}$, точка B имеет радиус-вектор $\mathbf{r}_1^{(1)}$, точка C имеет радиус-вектор $\mathbf{r}_0^{(2)} = \mathbf{r}(t)$.

Для кривых, представимых в виде $\mathbf{r}(t) = \sum_{i=0}^n f_i(t) \mathbf{p}_i$, определим *оператор правого сдвига* E и *оператор левого сдвига* L , действие которых описывается формулами

$$E\mathbf{r}(t) = \sum_{i=0}^{n-1} f_i(t) \mathbf{p}_{i+1}, \quad L\mathbf{r}(t) = \sum_{i=1}^n f_i(t) \mathbf{p}_{i-1}.$$

Операторы сдвига используются для компактной записи алгоритма Де Кастелье. Если действие оператора повторяется i раз, то будем использовать степени операторов E и L , например, E^i и L^i . Представим контрольные точки \mathbf{p}_i в виде кривых $\mathbf{p}_i = \sum_{j=0}^n \delta_j^i \mathbf{p}_j$, где $\delta_j^i = 1$ при $i = j$ и $\delta_j^i = 0$ при $i \neq j$.

Кривая Безье, построенная по контрольным точкам \mathbf{p}_i , $i = 0, 1, \dots, n$, может быть записана в виде

$$\mathbf{r}(t) = ((1-t) + tE)^n \mathbf{p}_0 \text{ или } \mathbf{r}(t) = ((1-t)L + t)^n \mathbf{p}_n.$$

Действительно, $((1-t) + tE)^n \mathbf{p}_0 = \sum_{i=0}^n C_n^i t^i (1-t)^{n-i} E^i \mathbf{p}_0 = \sum_{i=0}^n B_n^i(t) \mathbf{p}_i$, поскольку $E^i \mathbf{p}_0 = \mathbf{p}_i$. Аналогично, $((1-t)L + t)^n \mathbf{p}_n = \sum_{i=0}^n C_n^i t^i (1-t)^{n-i} L^{n-i} \mathbf{p}_n = \sum_{i=0}^n B_n^i(t) \mathbf{p}_i$, поскольку $L^{n-i} \mathbf{p}_n = \mathbf{p}_i$.

2.5. Кривые Безье и конические сечения

Квадратичная кривая Безье

$$\mathbf{r}(t) = (1-t)^2 \mathbf{p}_0 + 2t(1-t) \mathbf{p}_1 + t^2 \mathbf{p}_2 \quad (2.5.1)$$

является плоской линией и представляет собой полином второй степени параметра. Эллипс, парабола, гипербола записываются в виде функции второй степени координат. Возникает вопрос: **можно ли некоторую часть конического сечения описать кривой Безье второй степени?**

Пусть $l_1 = a_1 x + b_1 y + c_1$ и $l_2 = a_2 x + b_2 y + c_2$, где x и y – координаты некоторой декартовой системы в двумерном пространстве, $a_1, b_1, c_1, a_2, b_2, c_2$ – числа. Тогда равенства $l_1 = 0$ и $l_2 = 0$ будут представлять собой уравнения прямых линий в двумерном пространстве. Пусть $l_1 = 0$ и $l_2 = 0$ – уравнения двух несовпадающих прямых линий. Тогда уравнение $l_1 l_2 = 0$ будет описывать некоторое коническое сечение на этой плоскости. Возьмем еще одну пару прямых линий, описываемых уравнениями $l_3 = 0$ и $l_4 = 0$ ($l_3 = a_3 x + b_3 y + c_3, l_4 = a_4 x + b_4 y + c_4$), пересекающихся прямые первой пары, и составим уравнение

$$(1-\lambda)l_1 l_2 + \lambda l_3 l_4 = 0, \quad (2.5.2)$$

где λ – некоторый параметр. Уравнение (2.5.2) описывает семейство конических сечений, проходящих через четыре точки пересечения двух пар линий, рис. 2.5.1.

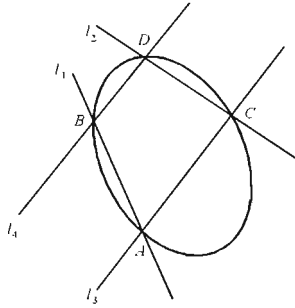


Рис. 2.5.1

По мере приближения прямой l_3 к l_4 точка A приближается к точке B , точка C приближается к точке D , хорды AB и CD стремятся к касательным линиям конических сечений семейства. Когда прямые линии l_3 и l_4 совпадут, уравнение (2.5.2) примет вид

$$(1-\lambda)l_1l_2 + \lambda l_3^2 = 0 \quad (2.5.3)$$

и будет представлять семейство конических сечений, касающихся линий l_1 и l_2 в точках их пересечения с линией l_3 , рис. 2.5.2.

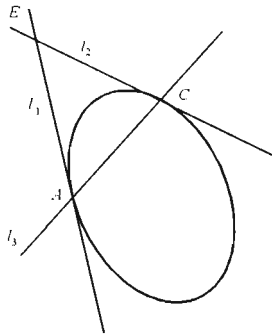


Рис. 2.5.2

Если задать еще одну точку R , не совпадающую с точками A и C , то мы определим коническое сечение и параметр λ в (2.5.3). Коническое сечение в данном случае определяется четырьмя точками: A , C , R и точкой E пересечения прямых линий l_1 и l_2 .

Попробуем описать часть конического сечения, лежащую внутри треугольника AEC , квадратичной кривой Безье, построенной по контрольной ломаной с вершинами в точках A , E и C . Пусть радиусы-векторы точек A , E , C равны соответственно \mathbf{p}_0 , \mathbf{p}_1 , \mathbf{p}_2 . Введем косоуголь-

ную систему координат на плоскости, начало которой лежит в точке \mathbf{p}_1 , а координатным базисом являются векторы $\mathbf{p}_0 - \mathbf{p}_1$ и $\mathbf{p}_2 - \mathbf{p}_1$. Координаты этой системы обозначим через u и v . В косоугольной системе uv вектор \mathbf{p}_0 имеет координаты $u = 1, v = 0$, вектор \mathbf{p}_1 имеет координаты $u = 0, v = 0$, вектор \mathbf{p}_2 имеет координаты $u = 0, v = 1$. Положение произвольной точки с координатами u и v будет описывать радиус-вектор

$$\mathbf{r}(u, v) = \mathbf{p}_1 + u(\mathbf{p}_0 - \mathbf{p}_1) + v(\mathbf{p}_2 - \mathbf{p}_1). \quad (2.5.4)$$

В косоугольной системе прямые линии l_1, l_2, l_3 описываются уравнениями $v = 0, u = 0, u + v - 1 = 0$, соответственно. Подставим их в (2.5.3) и получим уравнение

$$(1 - \lambda)uv + \lambda(u + v - 1)^2 = 0, \quad (2.5.5)$$

которое описывает семейство конических сечений в косоугольной системе координат. Левую часть уравнения (2.5.5) обозначим через $S(u, v)$.

Предположим, что мы выбрали из семейства (2.5.3) некоторое коническое сечение, то есть нам известен в (2.5.3) параметр λ . Возьмем некоторую точку \mathbf{r} конического сечения внутри треугольника с вершинами в точках $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2$, рис. 2.5.3.

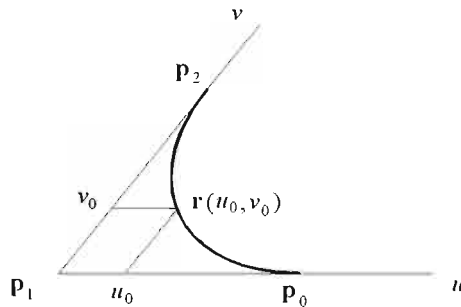


Рис. 2.5.3

Пусть ее координаты равны $0 < u_0 < 1$ и $0 < v_0 < 1$. В общем случае можно считать, что коническое сечение описывается функцией (2.5.4), в которой параметры u и v связаны соотношением (2.5.5). Запишем радиус-вектор конического сечения в виде функции одного параметра, в свою очередь являющегося функцией u и v . Для этого нам потребуется касательная к коническому сечению в точке $\mathbf{r}(u_0, v_0)$ и точки пересечения этой касательной с осями косоугольной системы координат.

Касательная линия к коническому сечению (2.5.5), записанному в виде $S(u, v) = 0$, в точке с координатами $u = u_0, v = v_0$ описывается уравнением

$$\left. \frac{\partial S}{\partial u} \right|_{u_0, v_0} (u - u_0) + \left. \frac{\partial S}{\partial v} \right|_{u_0, v_0} (v - v_0) = 0.$$

Производные $\partial S/\partial u$ и $\partial S/\partial v$ в точке $\mathbf{r}(u_0, v_0)$ с учетом того, что u_0, v_0 удовлетворяют уравнению (2.5.5), равны

$$\left. \frac{\partial S}{\partial u} \right|_{u_0, v_0} = (1 - \lambda)v_0 + 2\lambda(u_0 + v_0 - 1) = \lambda \frac{(u_0 + v_0 - 1)(u_0 - v_0 + 1)}{u_0},$$

$$\left. \frac{\partial S}{\partial v} \right|_{u_0, v_0} = (1 - \lambda)u_0 + 2\lambda(u_0 + v_0 - 1) = \lambda \frac{(u_0 + v_0 - 1)(v_0 - u_0 + 1)}{v_0}.$$

Таким образом, уравнение касательной к коническому сечению в точке $\mathbf{r}(u_0, v_0)$ имеет вид

$$(u_0 - v_0 + 1)v_0(u - u_0) + (v_0 - u_0 + 1)u_0(v - v_0) = 0.$$

Она пересекает координатные оси u и v в точках

$$u_1 = u_0 + u_0 \frac{v_0 - u_0 + 1}{u_0 - v_0 + 1} = \frac{2u_0}{u_0 - v_0 + 1}, \quad v_1 = 0,$$

$$v_2 = v_0 + v_0 \frac{u_0 - v_0 + 1}{v_0 - u_0 + 1} = \frac{2v_0}{v_0 - u_0 + 1}, \quad u_2 = 0,$$

что показано на рис. 2.5.4.

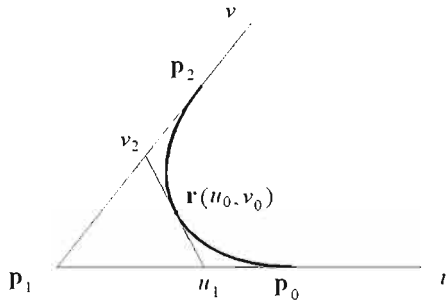


Рис. 2.5.4

Обозначим отношения длин отрезков, на которые касательная к коническому сечению делит векторы $\mathbf{p}_0 - \mathbf{p}_1$ и $\mathbf{p}_1 - \mathbf{p}_2$, через

$$z_1 = \frac{u_1}{1 - u_1} = \frac{2u_0}{1 - u_0 - v_0}, \quad z_2 = \frac{v_2}{1 - v_2} = \frac{2v_0}{1 - u_0 - v_0}. \quad (2.5.6)$$

Точку $\mathbf{r}(u_0, v_0)$ конического сечения можно характеризовать координатами u_0, v_0 , или координатами u_1, v_2 , или отношениями z_1, z_2 . Остановимся

на последнем способе. Выразим u_0 и v_0 через z_1 и z_2 , используя соотношения (2.5.6)

$$u_0 = \frac{z_1}{2 + z_1 + z_2}, \quad v_0 = \frac{z_2}{2 + z_1 + z_2}. \quad (2.5.7)$$

Подставив (2.5.7) в (2.5.4), получим зависимость радиуса-вектора конического сечения в виде функции параметров z_1 и z_2 :

$$\mathbf{r}(z_1, z_2) = \frac{z_1 \mathbf{p}_0 + 2\mathbf{p}_1 + z_2 \mathbf{p}_2}{z_1 + 2 + z_2}. \quad (2.5.8)$$

Параметры z_1 и z_2 связаны соотношением

$$z_1 z_2 = \frac{4u_0 v_0}{(1 - u_0 - v_0)^2} = -\frac{4\lambda}{1 - \lambda},$$

то есть произведение $z_1 z_2 = \text{const}$ для заданного конического сечения. Фактически радиус-вектор конического сечения зависит от одного параметра.

Введем параметр t для конического сечения, от которого будут зависеть параметры z_1 и z_2 . Из (2.5.5) следует

$$\frac{u_0 v_0}{(1 - u_0 - v_0)^2} = \frac{\lambda}{\lambda - 1}.$$

Подставим это равенство в (2.5.6) и получим

$$z_1 = \frac{1}{w} \sqrt{\frac{u_0}{v_0}}, \quad z_2 = \frac{1}{w} \sqrt{\frac{v_0}{u_0}},$$

где $w = \sqrt{\frac{\lambda - 1}{4\lambda}}$ – постоянный коэффициент. В качестве параметра t ко-

нического сечения примем величину $t = \frac{\sqrt{v_0}}{\sqrt{u_0} + \sqrt{v_0}}$, тогда

$$z_1 = \frac{1}{w} \left(\frac{1-t}{t} \right), \quad z_2 = \frac{1}{w} \left(\frac{t}{1-t} \right). \quad (2.5.9)$$

Подставим последние равенства в (2.5.8) и получим зависимость радиуса-вектора конического сечения (2.5.4) как функцию одного параметра t :

$$\mathbf{r}(t) = \frac{(1-t)^2 \mathbf{p}_0 + 2t(1-t)w\mathbf{p}_1 + t^2 \mathbf{p}_2}{(1-t)^2 + 2t(1-t)w + t^2}. \quad (2.5.10)$$

Коэффициент

$$w = \frac{1}{k} = \sqrt{\frac{\lambda - 1}{4\lambda}} = \frac{1 - u_0 - v_0}{\sqrt{4u_0v_0}} = \frac{1}{\sqrt{z_1z_2}} \quad (2.5.11)$$

называется *весом* точки \mathbf{p}_1 .

Сравним выражение для радиуса-вектора кривой Безье второй степени (2.5.1) и выражение для радиуса-вектора конического сечения (2.5.10). Они совпадают в случае, когда $w = 1$, что справедливо для одной из кривых семейства конических сечений, которой соответствует $\lambda = -1/3$. Коническое сечение (2.5.5) при $\lambda = -1/3$ описывается уравнением $u^2 - 2uv + v^2 - 2u - 2v + 1 = 0$ и является параболой. Действительно, если ввести новые переменные $x = u - v$, $y = 2u + 2v - 1$, то они будут связаны параболическим уравнением $y = x^2$. Будем рассматривать выражение (2.5.10) как обобщение квадратичной кривой, позволяющее в частных случаях получить как кривую Безье второй степени, так и коническое сечение.

Теперь мы можем ответить на поставленный выше вопрос: **кривой Безье второй степени в общем случае нельзя описать часть конического сечения, но его можно описать похожей кривой (2.5.10)**. Кривая Безье второй степени также может быть описана как частный случай кривой (2.5.10).

Построим с помощью формулы (2.5.10) дугу окружности радиуса ρ и углом раствора α , показанную на рис. 2.5.5. Через \mathbf{p}_0 и \mathbf{p}_2 обозначены радиусы-векторы крайних точек дуги, а через \mathbf{p}_1 обозначена точка пересечения касательных линий к дуге, построенных в крайних ее точках.

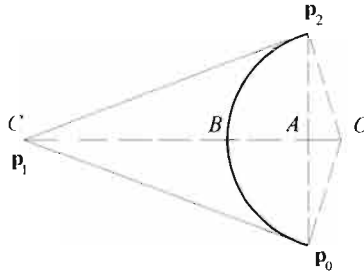


Рис. 2.5.5

Используя симметрию дуги относительно линии OC , найдем вес w по формуле (2.5.11). Из подобия треугольников следует, что

$$\frac{1}{w} = z_1 = z_2 = \frac{BC}{AB}.$$

Длина отрезков BC и AB вычисляются через радиус ρ и угол раствора дуги α

$$AB = \rho(1 - \cos(\alpha/2)), \quad BC = CO - BO = \frac{\rho}{\cos(\alpha/2)} - \rho = \rho \frac{1 - \cos(\alpha/2)}{\cos(\alpha/2)}.$$

Тогда

$$w = \frac{AB}{BC} = \cos\left(\frac{\alpha}{2}\right).$$

Таким образом, дуга окружности радиуса ρ с углом раствора α может быть построена как квадратичная кривая (2.5.10), заданная точками \mathbf{p}_0 , \mathbf{p}_2 и \mathbf{p}_1 с весом средней точки $w = \cos(\alpha/2)$. Радиус-вектор дуги описывается функцией

$$\mathbf{r}(t) = \frac{(1-t)^2 \mathbf{p}_0 + 2t(1-t) \cos(\alpha/2) \mathbf{p}_1 + t^2 \mathbf{p}_2}{(1-t)^2 + 2t(1-t) \cos(\alpha/2) + t^2}, \quad (2.5.12)$$

$$0 \leq t \leq 1,$$

где точки связаны соотношениями

$$|\mathbf{p}_0 - \mathbf{p}_1| = |\mathbf{p}_2 - \mathbf{p}_1| = \rho \operatorname{tg}(\alpha/2),$$

$$(\mathbf{p}_0 - \mathbf{p}_1) \cdot (\mathbf{p}_2 - \mathbf{p}_1) = |\mathbf{p}_0 - \mathbf{p}_1| |\mathbf{p}_2 - \mathbf{p}_1| \cos \alpha.$$

Параметрическая длина дуги равна единице. Формула (2.5.12) справедлива в диапазоне углов $0 < \alpha < 2\pi$. На рис. 2.5.6 и 2.5.7 приведены дуги окружности с углами $2/3\pi$ и $4/3\pi$ соответственно. Обе дуги имеют одни и те же контрольные точки, но в первом случае вес средней точки равен $0,5$, а во втором случае вес средней точки равен $-0,5$.

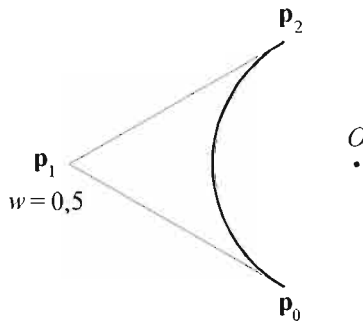


Рис. 2.5.6

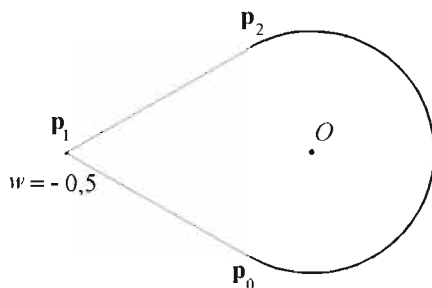


Рис. 2.5.7

Построим с помощью формулы (2.5.10) четверть эллипса по трем вершинам p_0 , p_1 и p_2 , рис. 2.5.8.

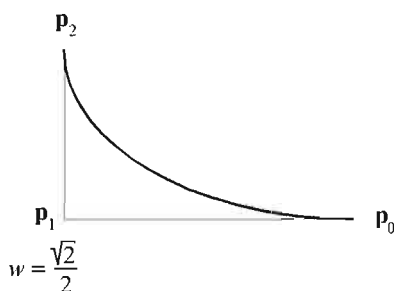


Рис. 2.5.8

Векторы $p_0 - p_1$ и $p_2 - p_1$ ортогональны. Полуоси эллипса определяются длинами этих векторов: $a = |p_0 - p_1|$, $b = |p_2 - p_1|$. В косоугольных координатах u и v , приведенных на рис. 2.5.4, уравнение эллипса имеет вид уравнения окружности: $(u - 1)^2 + (v - 1)^2 = 1$, которому в (2.5.5) соответствует $\lambda = -1$. В соответствии с (2.5.11) вес точки p_1

$$w = \frac{\sqrt{2}}{2},$$

то есть такой же, как и у дуги в четверть окружности.

Рассмотрим общий случай построения конического сечения. Пусть известны три вершины контрольной ломаной кривой p_0 , p_1 , p_2 и четвертая точка r , принадлежащая коническому сечению. Точка r должна лежать в одной плоскости с точками p_0 , p_1 , p_2 , рис. 2.5.9.

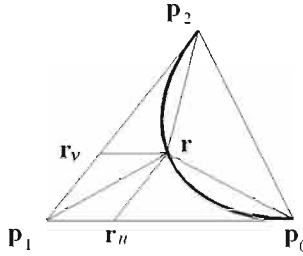


Рис. 2.5.9

Построим коническое сечение в виде (2.5.10), считая вес вершины \mathbf{p}_1 функцией положения четвертой точки \mathbf{r} относительно первых трех:

$$w(\mathbf{r}) = \frac{1 - u_0 - v_0}{\sqrt{4u_0v_0}}.$$

Косоугольные координаты u_0, v_0 точки \mathbf{r} , через которые вычисляется вес, равны отношениям длин векторов $|\mathbf{r} - \mathbf{r}_u| / |\mathbf{p}_0 - \mathbf{p}_1|$ и $|\mathbf{r} - \mathbf{r}_v| / |\mathbf{p}_2 - \mathbf{p}_1|$, соответственно. Обозначим через S_1 – площадь треугольника $\mathbf{p}_1\mathbf{p}_0\mathbf{r}$, через S_2 – площадь треугольника $\mathbf{p}_1\mathbf{p}_2\mathbf{r}$, через S – площадь треугольника $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$, тогда

$$u_0 = \frac{|\mathbf{r} - \mathbf{r}_v|}{|\mathbf{p}_0 - \mathbf{p}_1|} = \frac{S_2}{S}, \quad v_0 = \frac{|\mathbf{r} - \mathbf{r}_u|}{|\mathbf{p}_2 - \mathbf{p}_1|} = \frac{S_1}{S}.$$

Это следует из соотношений

$$S = 0,5|\mathbf{p}_0 - \mathbf{p}_1||\mathbf{p}_2 - \mathbf{p}_1|\sin\alpha, \quad S_1 = 0,5|\mathbf{p}_0 - \mathbf{p}_1||\mathbf{r} - \mathbf{r}_v|\sin\alpha, \quad S_2 = 0,5|\mathbf{p}_2 - \mathbf{p}_1||\mathbf{r} - \mathbf{r}_u|\sin\alpha,$$

где α – угол $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$. В принятых обозначениях вес средней точки выразится через площади треугольников формулой

$$w = \frac{S - S_1 - S_2}{\sqrt{4S_1S_2}}.$$

Вес w точки \mathbf{p}_1 можно найти, если $S_1S_2 \neq 0$, то есть если точка \mathbf{r} не лежит на линиях $\mathbf{p}_0 - \mathbf{p}_1$ и $\mathbf{p}_2 - \mathbf{p}_1$. В общем случае вес может быть как положительным, так и отрицательным. Как видно из рис. 2.5.6 и 2.5.7, если точка \mathbf{r} лежит внутри треугольника $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$, то вес точки \mathbf{p}_1 положителен, если же точка \mathbf{r} лежит вне треугольника $\mathbf{p}_0\mathbf{p}_1\mathbf{p}_2$, то вес точки \mathbf{p}_1 отрицателен.

Знак второго инварианта $I = \frac{3\lambda^2 - 2\lambda - 1}{4}$ конического сечения (2.5.5)

характеризует тип кривой: если $I > 0$, то коническое сечение имеет эллиптический тип, если $I < 0$, то коническое сечение имеет гиперболический тип.

ский тип, если $I = 0$, то коническое сечение имеет параболический тип. Из соотношения (2.5.11) следует, что при $|w| < 1$ кривая (2.5.10) описывает участок эллипса, при $|w| > 1$ кривая (2.5.10) описывает участок гиперболы, при $|w| = 1$ кривая (2.5.10) описывает участок параболы.

2.6. Рациональные кривые Безье

Формула (2.5.10) имеет несимметричный вид, так как коэффициент w приписан только точке \mathbf{p}_1 . Умножим числитель и знаменатель правой части (2.5.10) на некоторое число w_0 и получим

$$\mathbf{r}(t) = \frac{(1-t)^2 w_0 \mathbf{p}_0 + 2t(1-t) w_1 \mathbf{p}_1 + t^2 \mathbf{p}_2 w_0}{(1-t)^2 w_0 + 2t(1-t) w_1 + t^2 w_0},$$

где $w_1, w_2 = w_0 w$. Заменяем w_0 на некоторое число w_2 у последних слагаемых числителя и знаменателя правой части равенства. После таких изменений контрольные точки $\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2$ получают коэффициенты w_0, w_1, w_2 соответственно, а формула (2.5.10) примет симметричный вид. В результате получим кривую

$$\mathbf{r}(t) = \frac{(1-t)^2 w_0 \mathbf{p}_0 + 2t(1-t) w_1 \mathbf{p}_1 + t^2 w_2 \mathbf{p}_2}{(1-t)^2 w_0 + 2t(1-t) w_1 + t^2 w_2} = \frac{\sum_{i=0}^2 B_i^2 w_i \mathbf{p}_i}{\sum_{i=0}^2 B_i^2 w_i}. \quad (2.6.1)$$

Коэффициенты w_0, w_1, w_2 называются *весами* контрольных точек.

Используем полученный результат для обобщения кривой Безье произвольной степени. Аналогично зависимости (2.6.1) запишем формулу для вычисления радиуса-вектора кривой Безье, построенной по произвольному числу n вершин, в виде

$$\mathbf{r}(t) = \frac{\sum_{i=0}^n B_i^n(t) w_i \mathbf{p}_i}{\sum_{i=0}^n B_i^n(t) w_i}, \quad t \in [0, 1], \quad (2.6.2)$$

где $B_i^n(t)$ – функции Бернштейна (2.4.2). В общем случае каждая контрольная точка \mathbf{p}_i кривой (2.6.2) имеет свой вес w_i . Кривые, контрольные точки которых обладают весами, называются *рациональными*. Кривая (2.6.2) называется рациональной кривой Безье степени n .

В общем случае мы всегда можем перейти от рациональной кривой Безье степени n к рациональной кривой Безье степени $n + 1$, умножив числитель и знаменатель правой части равенства (2.6.2) на равное единице число $t + (1 - t)$ и перегруппировав слагаемые.

Покажем это на примере квадратичной рациональной кривой Безье. Умножим числитель и знаменатель правой части равенства (2.6.1) на равное единице число $t + (1 - t)$, перегруппируем слагаемые по степеням произведений t и $(1 - t)$ и, в результате, придем к выражению

$$r(t) = \frac{(1-t)^3 w_0 \mathbf{p}_0 + t(1-t)^2 (2w_1 \mathbf{p}_1 + w_0 \mathbf{p}_0) + t^2(1-t)(2w_1 \mathbf{p}_1 + w_2 \mathbf{p}_2) + t^3 w_2 \mathbf{p}_2}{(1-t)^3 w_0 + t(1-t)^2 (2w_1 + w_0) + t^2(1-t)(2w_1 + w_2) + t^3 w_2}.$$

Придадим правой части этого равенства форму кубической кривой Безье. Для этого введем обозначения для весов и вершин в правой части последнего выражения:

$$w_0' = w_0, \quad w_1' = \frac{2w_1 + w_0}{3}, \quad w_2' = \frac{2w_1 + w_2}{3}, \quad w_3' = w_2,$$

$$\mathbf{p}_0' = \mathbf{p}_0, \quad \mathbf{p}_1' = \frac{2w_1 \mathbf{p}_1 + w_0 \mathbf{p}_0}{3w_1'}, \quad \mathbf{p}_2' = \frac{2w_1 \mathbf{p}_1 + w_2 \mathbf{p}_2}{3w_2'}, \quad \mathbf{p}_3' = \mathbf{p}_2.$$

Теперь квадратичная кривая (2.6.1) примет форму кубической кривой (опустим звездочки у контрольных точек и их весов):

$$r(t) = \frac{(1-t)^3 w_0 \mathbf{p}_0 + 3t(1-t)^2 w_1 \mathbf{p}_1 + 3t^2(1-t)w_2 \mathbf{p}_2 + t^3 w_3 \mathbf{p}_3}{(1-t)^3 w_0 + 3t(1-t)^2 w_1 + 3t^2(1-t)w_2 + t^3 w_3} = \frac{\sum_{i=0}^3 B_i^3 w_i \mathbf{p}_i}{\sum_{i=0}^3 B_i^3 w_i}.$$

Полученная кубическая кривая полностью совпадает с исходной квадратичной кривой, но ее контрольная ломаная и веса контрольных точек стали другими. Умножив числитель и знаменатель правой части последнего равенства на неравное нулю число $t + (1 - t)$, перейдем к рациональной кривой Безье четвертой степени, которая описывает все то же коническое сечение. Этот процесс можно продолжать и дальше.

Если считать, что в знаменателе правой части (2.6.2) стоит вес $w(t)$ радиуса-вектора кривой $\mathbf{r}(t)$, то его можно считать дополнительной координатой, так как формула его вычисления через веса вершин совпадает с формулой вычисления координат радиуса-вектора через координаты вершин. При работе с точками, имеющими вес, удобно использовать **однородные координаты** (1.4.8), где p_1, p_2, p_3 — координаты точки \mathbf{p} , w — вес точки. В терминах однородных координат радиус-вектор рациональной кривой Безье описывается равенством, по форме совпадающим с равенством для кривой Безье (2.4.1):

$$\mathbf{r}(t) = \sum_{i=0}^n B_i^n(t) \mathbf{p}_i, \quad t \in [0, 1].$$

Вес точки подвергается тем же преобразованиям, что и ее координаты, поэтому с однородными координатами радиуса-вектора можно работать

без выделения декартовых координат. Декартовы координаты точки рациональной кривой получим делением однородных координат на дополнительную координату.

Радиус-вектор рациональной кривой вычисляется как частное от деления двух функций параметра кривой, поэтому при вычислении производной рациональной кривой правую часть следует рассматривать как сложную функцию. Если условно обозначим радиус-вектор рациональной кривой как $\mathbf{r} = \frac{w\mathbf{r}}{w}$, то производная радиуса-вектора рациональной

кривой определится формулой $\frac{d\mathbf{r}}{dt} = \frac{1}{w} \frac{d(w\mathbf{r})}{dt} - \frac{w\mathbf{r}}{w^2} \frac{dw}{dt}$.

Если веса всех контрольных точек равны, то их можно вынести из-под знаков суммирования и сократить. Тогда в силу свойства полиномов Бернштейна (2.4.3) рациональная кривая Безье (2.6.2.) становится равной обычной кривой Безье (2.4.1).

Наличие весов у контрольных точек дает дополнительные возможности при редактировании кривой. Заменим вес w_j на $w_j + \Delta w$, тогда кривая (2.6.2) изменится на кривую

$$\mathbf{r}'(t) = \frac{\sum_{i=0}^n B_i^n(t) w_i \mathbf{p}_i + B_j^n(t) \Delta w \mathbf{p}_j}{\sum_{i=0}^n B_i^n(t) w_i + B_j^n(t) \Delta w} = \frac{\mathbf{r}(t) + \delta(t) \mathbf{p}_j}{1 + \delta(t)},$$

где $\delta(t) = \frac{B_j^n(t) \Delta w}{\sum_{i=0}^n B_i^n(t) w_i}$. Следовательно, для любого $\Delta w > 0$ точка исходной

кривой смещается к точке \mathbf{p}_j и ее новое положение $\mathbf{r}(t)'$ делит в отношении $1 : \delta(t)$ отрезок с началом в точке $\mathbf{r}(t)$ и концом в точке \mathbf{p}_j . На рис. 2.6.1 приведены рациональные кривые Безье пятой степени, построенные по одним и тем же контрольным точкам с весами, равными единице, для всех точек, кроме точки \mathbf{p}_3 . Мы видим, что чем больше относительный вес точки, тем ближе к этой точке располагается рациональная кривая Безье. Если вес точки отрицательный, то точка как бы отталкивает кривую от себя.

Несмотря на большие возможности, кривые Безье обладают некоторыми ограничениями, а именно: порядок кривых жестко связан с числом точек; область определения параметра кривых предопределена; кривые не могут быть циклически замкнутыми.

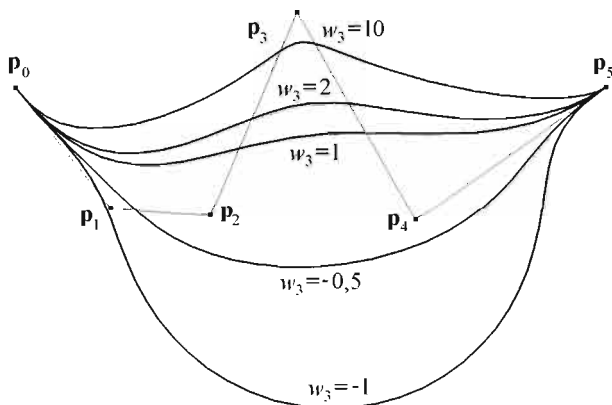


Рис. 2.6.1

2.7. Разделенные разности

Разделенные разности используются для построения сплайнов. На основе разделенных разностей строятся скалярные функции, являющиеся обобщением функций Берштейна, которые широко применяются в геометрическом моделировании.

Пусть нам известны значения $f(t_i)$ некоторой функции $f(t)$ при значениях параметра t_i . Разделенными разностями первого порядка называются отношения

$$f[t_0, t_1] = \frac{f(t_1) - f(t_0)}{t_1 - t_0}. \quad (2.7.1)$$

По разделенным разностям первого порядка вычисляются разделенные разности второго порядка

$$f[t_0, t_1, t_2] = \frac{f[t_1, t_2] - f[t_0, t_1]}{t_2 - t_0}.$$

Разделенные разности m -го порядка вычисляются по разделенным разностям $(m - 1)$ -го порядка

$$f[t_0, t_1, \dots, t_m] = \frac{f[t_1, t_2, \dots, t_m] - f[t_0, t_1, \dots, t_{m-1}]}{t_m - t_0}. \quad (2.7.2)$$

Параметры t_i , $i = 0, 1, \dots, m$ называют *узлами*. Разделенные разности будем обозначать символом функции, по которой она построена, с квадратны-

ми скобками, в которых перечислены узлы. В литературе для разделенной разности используется также обозначение $[t_1, t_2, \dots, t_m]f$. Порядок разделенной разности на единицу меньше числа узлов в квадратных скобках. Разделенными разностями нулевого порядка считаются значения функций в узлах $f[t_i] = f(t_i)$.

Разделенная разность m -го порядка выражается через значения функции $f(t)$ в узлах соотношением

$$\begin{aligned} f[t_0, t_1, \dots, t_m] &= \sum_{j=0}^m \frac{f(t_j)}{(t_j - t_0)(t_j - t_1) \dots (t_j - t_{j-1})(t_j - t_{j+1}) \dots (t_j - t_m)} = \\ &= \sum_{j=0}^m \frac{f(t_j)}{W_{0^*m}(t_j)}, \end{aligned} \quad (2.7.3)$$

где $W_{0^*m}(t_j)$ – производная функции $W_{0^*m}(t) = (t - t_0)(t - t_1) \dots (t - t_{m-1})(t - t_m)$, вычисленная при $t = t_j$. Индексы в обозначении функции $W_{0^*m}(t)$ говорят о номере начального и конечного узлов. Равенство (2.7.3) докажем методом индукции. Для разделенной разности первого порядка формула (2.7.3) совпадает с (2.7.1), то есть для $m = 1$ равенство (2.7.3) справедливо. Предположим, что равенство (2.7.3) справедливо для разделенных разностей $(m - 1)$ -го порядка, и убедимся, что из этого предположения следует его верность для разделенных разностей m -го порядка. Для этого найдем разность равенств

$$\begin{aligned} f[t_1, t_2, \dots, t_m] &= \sum_{j=1}^m \frac{(t_j - t_0)f(t_j)}{(t_j - t_0)W_{1^*m}(t_j)} \text{ и} \\ f[t_0, t_1, \dots, t_{m-1}] &= \sum_{j=0}^{m-1} \frac{(t_j - t_m)f(t_j)}{(t_j - t_m)W_{0^*m-1}(t_j)}, \end{aligned}$$

в правых частях которых соответствующие слагаемые приведены к общему знаменателю. В результате получим равенство

$$\begin{aligned} & f[t_1, t_2, \dots, t_m] - f[t_0, t_1, \dots, t_{m-1}] = \\ &= \frac{(t_m - t_0)f(t_0)}{W_{0^*m}(t_0)} + \sum_{j=1}^{m-1} \frac{(t_j - t_0)f(t_j) - (t_j - t_m)f(t_j)}{W_{0^*m}(t_j)} + \frac{(t_m - t_0)f(t_m)}{W_{0^*m}(t_m)} = \\ &= (t_m - t_0) \sum_{j=0}^m \frac{f(t_j)}{W_{0^*m}(t_j)}, \end{aligned}$$

из которого после деления на $(t_m - t_0)$ на основании (2.7.2) следует (2.7.3).

Из равенства (2.7.3) видно, что **разделенная разность не зависит от порядка, в котором следуют узлы t_0, t_1, \dots, t_m в списке аргументов.**

Разделенные разности могут быть построены на последовательности с совпадающими узлами. Узлы, значения которых совпадают, называются

ся *кратными узлами*. Так, если $t_1 = t_2 = \dots = t_k$, то говорят, что узел обладает кратностью k . Если разделенная разность первого порядка построена на кратном узле, то она равна производной функции в данной точке

$$f[t_i, t_i] = \lim_{t \rightarrow t_i} \frac{f(t_i) - f(t)}{t_i - t} = \left. \frac{df}{dt} \right|_{t_i} = f'(t_i).$$

Разделенная разность m -го порядка, построенная на узле кратности $m + 1$, равна m -й производной функции, вычисленной в данном узле, деленной на $m!$:

$$f[t_i, t_i, \dots, t_i] = \frac{1}{m!} \frac{d^m f}{dt^m} \Big|_{t_i} = \frac{f^{(m)}(t_i)}{m!}. \quad (2.7.4)$$

Если $p(t)$ является полиномом m -ой степени параметра t и нам известны значения полинома $p(t_i)$ в узлах t_i , то разделенная разность первого порядка

$$p[t, t_1] = \frac{p(t_1) - p(t)}{t_1 - t}$$

является полиномом $(m - 1)$ -й степени. В самом деле, функция $p(t_1) - p(t)$ имеет корень t_1 и, следовательно, согласно теореме Безу, делится без остатка на $t_1 - t$. Разделенная разность второго порядка

$$p[t, t_1, t_2] = \frac{p[t_1, t_2] - p[t, t_1]}{t_2 - t}$$

является полиномом $(m - 2)$ -ой степени. Действительно, функция $p[t_1, t_2] - p[t, t_1]$ имеет корень t_2 и, следовательно, согласно теореме Безу, делится без остатка на $t_2 - t$. С помощью аналогичных рассуждений приходим к тому, что разделенная разность m -го порядка есть полином нулевой степени

$$p[t, t_1, t_2, \dots, t_m] = \text{const},$$

а разделенная разность $(m + 1)$ -го порядка полинома степени m равна нулю:

$$p[t, t_1, t_2, \dots, t_{m+1}] = 0.$$

Пусть $p(t)$ является интерполяционным полиномом функции $f(t)$, совпадающим с ней в $m + 1$ узлах t_i , $p(t_i) = f(t_i)$, $i = 0, 1, \dots, m$, тогда разделенные разности, построенные по узлам t_i , для функций $p(t)$ и $f(t)$ будут равны. Из определения разделенной разности для функции $p(t)$, построенной на последовательности узлов t, t_0, t_1, \dots, t_m ,

$$p[t, t_0, t_1, \dots, t_m] = \frac{p[t_0, t_1, \dots, t_m] - p[t, t_0, \dots, t_{m-1}]}{t_m - t}$$

и равенств разделенных разностей $p[t_0, t_1, \dots, t_m] = f[t_0, t_1, \dots, t_m]$ следуют равенства:

$$\begin{aligned} p[t, t_0, t_1, \dots, t_{m-1}] &= f[t_0, t_1, \dots, t_m] + (t - t_m) p[t, t_0, t_1, \dots, t_m] = f[t_0, t_1, \dots, t_m], \\ p[t, t_0, \dots, t_{m-2}] &= f[t_0, t_1, \dots, t_{m-1}] + (t - t_{m-1}) p[t, t_0, t_1, \dots, t_{m-1}], \\ &\dots \dots \dots \dots \dots \dots \dots, \\ p[t, t_0] &= f[t_0, t_1] + (t - t_1) p[t, t_0, t_1], \\ p[t] &= f[t_0] + (t - t_0) p[t, t_0]. \end{aligned}$$

Так как $p(t) = p[t]$, то, подставляя последовательно эти равенства – каждое последующее в предыдущее, начиная с последнего, получим интерполяционную формулу Ньютона для функции $f(t)$:

$$\begin{aligned} p(t) = f[t_0] + f[t_0, t_1] (t - t_0) + f[t_0, t_1, t_2] (t - t_0) (t - t_1) + \dots \\ + f[t_0, t_1, \dots, t_m] (t - t_0) (t - t_1) \dots (t - t_{m-1}). \end{aligned} \quad (2.7.5)$$

Таким образом, коэффициенты в полиноме Ньютона являются соответствующими разделенными разностями интерполируемой функции. Если интерполяцию (2.7.5) выполнить на совокупности $m + 1$ совпадающих узлов $t_0 = t_1 = \dots = t_m$, то в соответствии с (2.7.4) разделенные разности пропорциональны производным соответствующего порядка интерполируемой функции, и мы получим усеченный ряд Тейлора.

Известно, что по заданной совокупности значений функции $f(t_i)$ в узлах $t_i, i = 0, 1, \dots, m$ можно построить единственный полином степени m , совпадающий в узлах t_i с заданной функцией. Из интерполяционной формулы Ньютона (2.7.5) видим, что разделенная разность $f[t_0, t_1, \dots, t_m]$ m -го порядка на заданной последовательности $m + 1$ узлов равна коэффициенту при наивысшей степени аргумента t^m полинома степени m , значения которого в заданных узлах согласуются с функцией $f(t)$. Это свойство может быть принято в качестве определения разделенной разности.

Свойство 1. Итак, если $f(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_m t^m$ (полином степени m), то

$$f[t_0, t_1, \dots, t_m] = a_m, f[t_0, t_1, \dots, t_m, t_{m+1}] = 0, \quad (2.7.6)$$

для любого дополнительного узла t_{m+1} .

Свойство 2. В силу единственности такого полинома, а также в силу (2.7.3), любая разделенная разность является симметричной функцией своих аргументов, то есть значение $f[t_0, t_1, \dots, t_m]$ не зависит от порядка, в котором следуют узлы t_0, t_1, \dots, t_m в списке аргументов. Разделенная разность выражается через значения функции $f(t)$ в узлах с помощью формулы (2.7.3). Если некоторые узлы являются кратными, то разделенная

разность выражается через значения функции $f(t)$ в простых узлах и ее производные в кратных узлах.

Свойство 3. Если $f(t) = k_g g(t) + k_h h(t)$, то

$$f[t_1, t_2, \dots, t_{m+1}] = k_g g[t_1, t_2, \dots, t_{m+1}] + k_h h[t_1, t_2, \dots, t_{m+1}].$$

Это равенство следует из единственности интерполяционного полинома.

Свойство 4. Если $f(t) = g(t)h(t)$, то

$$f[t_0, t_1, \dots, t_m] = \sum_{k=0}^m g[t_0, t_1, \dots, t_k] \cdot h[t_k, t_{k+1}, \dots, t_m]. \quad (2.7.7)$$

Для доказательства этой формулы аппроксимируем каждую из функций $g(t)$ и $h(t)$ полиномом (2.7.5) и перемножим их. Коэффициент при наивысшей степени t^m результирующего полинома будет равен правой части (2.7.7), что и требовалось доказать. Равенство (2.7.7) называется **формулой Лейбница**.

Свойство 5. Если для функции $f(t)$ построен полином $p_{m-1}(t)$ степени $m-1$ на последовательности узлов t_0, t_1, \dots, t_{m-1} , то

$$f(t) = p_{m-1}(t) + f[t_0, t_1, \dots, t_{m-1}, t] (t-t_0)(t-t_1)\dots(t-t_{m-1}).$$

Это равенство получим, рассматривая t как дополнительный узел в последовательности узлов и равенство как необходимое условие в дополнительном узле для полинома степени m .

Рассмотрим **усеченную степенную функцию со смещенным началом**:

$$\sigma_m(z) = (z-t)_+^m \equiv (\max(0, z-t))^m. \quad (2.7.8)$$

Графики функций $\sigma_m(z)$ приведены на рис. 2.7.1. Чем выше степень функции $\sigma_m(z)$, тем она более гладкая в точке $z = t$.

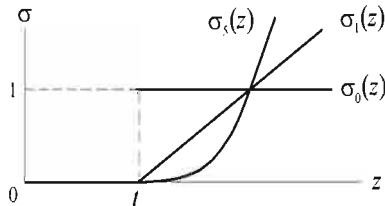


Рис. 2.7.1

Параметр t определяет смещение начала усеченной степенной функции. Усеченная степенная функция со смещенным началом равна нулю при $z \leq t$ и равна $(z-t)^m$ в противном случае. Функция $\sigma_m(z)$ кусочно-монотонная непрерывная при $m > 0$, при $m = 0$ она имеет скачок в точке

$z = t$, равный единице. Функция $\sigma_m(z)$ имеет непрерывные производные до $(m - 1)$ -го порядка включительно. Производная m -го порядка имеет разрыв в точке $z = t$. Функция (2.7.8) примечательна тем, что ее производные являются аналогичными функциями, степень которых на единицу меньше.

Разделенные разности $\sigma_m[t_0, t_1, \dots, t_{m+1}]$ усеченной степенной функции со смещенным началом обладают свойствами, которые позволяют на их основе строить рациональные кривые. Разделенная разность вычисляется при фиксированном параметре t , но является функцией параметра t . Если все $t_i < t$, $i = 0, 1, \dots, m + 1$, то все $\sigma_m(t_i) = 0$, и разделенная разность $\sigma_m[t_0, t_1, \dots, t_{m+1}]$ равна нулю. Если все $t_i > t$, $i = 0, 1, \dots, m + 1$, то разделенная разность $\sigma_m[t_0, t_1, \dots, t_{m+1}]$ равна нулю, так как все узлы находятся на участке, где функция представляет собой полином степени, меньшей $(m + 1)$. Только тогда, когда параметр t смещения начала лежит внутри отрезка, образованного последовательностью узлов t_0, t_1, \dots, t_{m+1} ($\min(t_0, t_1, \dots, t_{m+1}) \leq t \leq \max(t_0, t_1, \dots, t_{m+1})$), разделенная разность $\sigma_m[t_0, t_1, \dots, t_{m+1}]$ усеченной степенной функции m -й степени отлична от нуля.

2.8. В-сплайны

B-сплайны представляют собой скалярные функции, обладающие набором полезных свойств. *B*-сплайны строятся на основе разделенных разностей и используются для построения кривых. Основу теории *B*-сплайнов заложили Фергюсон (J.C. Ferguson) и Шенберг (I.J. Schoenberg). Гордон (W.J. Gordon) и Розенфельд (R.F. Riesenfeld) установили связь между кривыми Безье и *B*-сплайнами и показали, что *B*-сплайны являются обобщением функций Бернштейна. На основе *B*-сплайнов строятся рациональные кривые, которые в частных случаях описывают конические сечения, кривые Безье и другие кривые.

B-сплайн для последовательности $m + 2$ узлов есть разделенная разность $(m + 1)$ -го порядка усеченной степенной функции со смещенным началом m -й степени, умноженная на разность максимального и минимального значений узлов последовательности.

B-сплайн, построенный на последовательности узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$, определяется равенством

$$N_i^m(t) = (t_{\max} - t_{\min}) \sigma_m[t_i, t_{i+1}, \dots, t_{i+m+1}], \quad (2.8.1)$$

где $t_{\max} = \max(t_i, t_{i+1}, \dots, t_{i+m+1})$, $t_{\min} = \min(t_i, t_{i+1}, \dots, t_{i+m+1})$ — максимальное и минимальное значения узлов последовательности. Индексы в обозначении *B*-сплайна призваны охарактеризовать последовательность узлов, на которой он построен, и в различной литературе могут иметь различное значение. В одних источниках *B*-сплайну приписывается порядок, рав-

ной степени усеченной степенной функции, в других источниках B -сплайну приписывается порядок, равный порядку разделенной разности. Кроме того, B -сплайн привязывается к одному из узлов последовательности. Мы будем считать порядок B -сплайна равным степени функции, по которой он построен, и привяжем его к первому узлу. На рис. 2.8.1 приведены B -сплайны нулевого, первого и второго порядков.

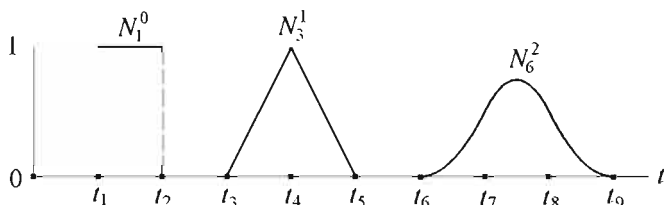


Рис. 2.8.1

На рис. 2.8.2 приведены B -сплайн третьего порядка $N_1^3(t) = (t_5 - t_1)\sigma_3[t_1, t_2, t_3, t_4, t_5]$ и B -сплайн пятого порядка $N_5^5(t) = (t_9 - t_3)\sigma_5[t_3, t_4, t_5, t_6, t_7, t_8, t_9]$, построенные на равноотстоящих узлах.

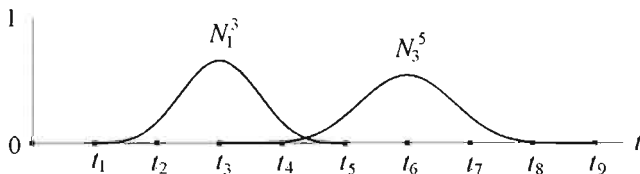


Рис. 2.8.2

Если последовательность узлов, на которой построен B -сплайн, содержит кратные узлы, то производные B -сплайна теряют непрерывность. На рис. 2.8.3 приведены B -сплайны третьего порядка с кратными узлами $N_1^3(t) = (t_2 - t_1)\sigma_3[t_1, t_1, t_1, t_2]$, $N_3^3(t) = (t_5 - t_3)\sigma_3[t_3, t_4, t_4, t_5]$, $N_6^3(t) = (t_8 - t_6)\sigma_3[t_6, t_6, t_6, t_7, t_8]$.

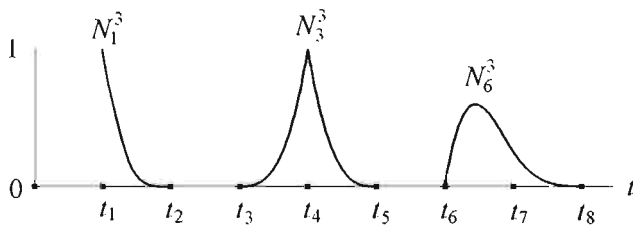


Рис. 2.8.3

В формуле (2.8.1) индекс m указывает на порядок B -сплайна, а индекс i указывает на первый узел последовательности. Расположение индекса i перед индексом m указывает на то, что i -й узел является первым узлом последовательности. В дальнейшем мы определим B -сплайн $N_{i+m+1}^m(t)$, имеющий индекс последнего узла последовательности, на что будет указывать расположение индекса $i + m + 1$ после индекса m . B -сплайны $N_i^m(t)$ и $N_{i+m+1}^m(t)$, построенные на одной и той же последовательности узлов, равны.

Для вычисления B -сплайнов вводятся ненормированные B -сплайны, которые представляют собой разделенные разности усеченной степенной функции.

Ненормированный B -сплайн для последовательности $m + 2$ узлов есть разделенная разность $(m + 1)$ -го порядка усеченной степенной функции со смещенным началом m -й степени.

Ненормированный B -сплайн, построенный на последовательности узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$, определяется равенством

$$M_i^m(t) = \sigma_m[t_i, t_{i+1}, \dots, t_{i+m+1}]. \quad (2.8.2)$$

В обозначении ненормированного B -сплайна в формуле (2.8.2) индекс i указывает на первый узел последовательности, а индекс m указывает на порядок B -сплайна.

B -сплайн и ненормированный B -сплайн связаны между собой коэффициентом, и с точностью до множителя их определения совпадают. Вычисление B -сплайнов производится с помощью ненормированных B -сплайнов.

Пусть имеется последовательность $m + 2$ узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$, хотя бы два из которых имеют отличные друг от друга значения. Положим, что $t_i \neq t_{i+m+1}$. Используем формулу Лейбница (2.7.7) для вычисления B -сплайна, представив функцию $\sigma_m(z) = (z-t)_+^m$ в виде произведения $\sigma_m(z) = g(z) \cdot \sigma_{m-1}(z)$, где $\sigma_{m-1}(z) = (z-t)_+^{m-1}$ и $g(z) = (z-t)$. Тогда разделенная разность

$$\begin{aligned} & \sigma_m[t_i, t_{i+1}, \dots, t_{i+m+1}] = \\ & = g[t_i] \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m+1}] + g[t_i, t_{i+1}] \sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] + 0 + \dots + 0 = \\ & = (t_i - t) \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m+1}] + \sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}]. \end{aligned} \quad (2.8.3)$$

В (2.8.3) использовалось равенство нулю разделенных разностей выше первого порядка для функции $g(z)$. Подставим в (2.8.3) вместо $\sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m+1}]$ определение разделенной разности (2.7.2)

$$\sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m+1}] = \frac{\sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] - \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m}]}{t_{i+m+1} - t_i}$$

и получим

$$\begin{aligned}
& \sigma_m[t_i, t_{i+1}, \dots, t_{i+m+1}] = \\
& = \frac{t_i - t}{t_{i+m+1} - t_i} (\sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] - \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m}]) + \\
& \quad + \frac{t_{i+m+1} - t_i}{t_{i+m+1} - t_i} \sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] = \\
& = \frac{t_{i+m+1} - t}{t_{i+m+1} - t_i} \sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] + \frac{t - t_i}{t_{i+m+1} - t_i} \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m}]. \quad (2.8.4)
\end{aligned}$$

Перепишем равенство (2.8.4) с использованием определения ненормированного B -сплайна (2.8.2) и получим формулу

$$M_i^m(t) = \frac{t_{i+m+1} - t}{t_{i+m+1} - t_i} M_{i+1}^{m-1}(t) + \frac{t - t_i}{t_{i+m+1} - t_i} M_i^{m-1}(t). \quad (2.8.5)$$

Рекуррентное соотношение (2.8.5) было получено независимо Мэнсфилдом, Коксом (M.G. Cox) и Де Буром (C. De Boor) и называется **формулой Кокса–Де Бура**. Это соотношение занимает центральное место в теории B -сплайнов. Оно побуждает забыть о разделенных разностях и определить B -сплайн m -го порядка для последовательности $m + 2$ узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$ как функцию, вычисляемую по рекуррентному соотношению (2.8.5) при начальных значениях

$$M_i^0(t) = \begin{cases} \frac{1}{(t_{i+1} - t_i)}, & \text{если } \min(t_i, t_{i+1}) \leq t < \max(t_i, t_{i+1}), \\ 0 & , \text{ в остальных случаях.} \end{cases} \quad (2.8.6)$$

Равенство (2.8.6) выражает ненормированный B -сплайн нулевого порядка через разделенную разность первого порядка для функции $\sigma_0(z) = (z - t)_+^0$.

Отвлечемся в формуле (2.8.5) от индексации и рассмотрим ее с использованием барицентрических координат. Пусть дана произвольная последовательность $m + 2$ узлов, хотя бы два из которых имеют отличные друг от друга значения. Вспомним, что разделенная разность не зависит от порядка следования узлов в списке аргументов. Из последовательности $m + 2$ узлов выберем два произвольных несовпадающих узла t_a и t_b . На базе этих двух узлов введем барицентрические координаты

$$a(t) = \frac{t_b - t}{t_b - t_a}, \quad b(t) = \frac{t_a - t}{t_a - t_b}.$$

Введем следующие обозначения. Последовательность $m + 2$ узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$ обозначим через T . Последовательность $m + 1$ узлов, полученную из T удалением узла t_a , обозначим через $T \setminus a$. Последовательность

$m + 1$ узлов, полученную из T удалением узла t_b , обозначим через $T \setminus b$. Разделенную разность $(m + 1)$ -го порядка усеченной степенной функции $(z - t)_+$ на последовательности узлов T обозначим через $M^T(t)$. Разделенную разность m -го порядка усеченной степенной функции $(z - t)_+$ на последовательности узлов $T \setminus a$ обозначим через $M^{T \setminus a}(t)$. Разделенную разность m -го порядка усеченной степенной функции $(z - t)_+$ на последовательности узлов $T \setminus b$ обозначим через $M^{T \setminus b}(t)$. Тогда формула Кокса—Де Бура (2.8.5) для вычисления ненормированного B -сплайна на последовательности узлов T примет вид

$$M^T(t) = a(t) M^{T \setminus a}(t) + b(t) M^{T \setminus b}(t). \quad (2.8.7)$$

Так как разделенная разность является симметричной функцией своих аргументов, то последнее рекуррентное соотношение не зависит от выбора узлов t_a и t_b . Рекуррентное соотношение (2.8.7) начинается с вычисления ненормированных B -сплайнов нулевого порядка (2.8.6).

Для наглядного представления соотношения (2.8.5) будем считать, что узлы последовательности расположены в порядке возрастания их значений: $t_i \leq t_{i+1} \leq \dots \leq t_{i+m+1}$, тогда один из двух B -сплайнов $(m - 1)$ -го порядка будет примыкать к началу вычисляемого B -сплайна, а другой — к концу вычисляемого B -сплайна, рис. 2.8.4. Каждый из двух B -сплайнов $(m - 1)$ -го порядка умножается на коэффициент, пропорциональный расстоянию параметра t от крайнего узла заданной последовательности, к которому B -сплайн $(m - 1)$ -го порядка примыкает. Сумма этих двух коэффициентов равна единице.

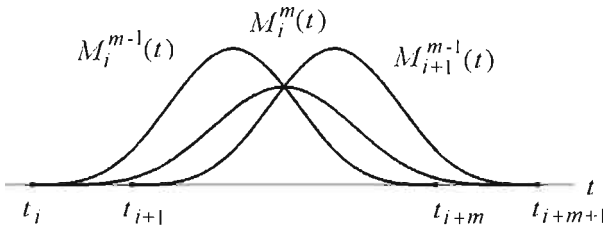


Рис. 2.8.4

С помощью B -сплайнов мы будем строить кривые на базе точек, в которых индекс узла привязки B -сплайна будет равен индексу точки. Поэтому для построения кривых используются неубывающие последовательности узлов, так как узлы связаны с параметром кривой. Кроме того, совокупность B -сплайнов, построенная на неубывающих последовательностях узлов, при заданном параметре t обладает определенным свойством. Таким образом, в практических задачах B -сплайны строятся на узлах, являющихся частью общей последовательности узлов, расположенных подряд.

Свойство 1. Пусть B -сплайн $N_i^m(t)$ построен на неубывающей последовательности $m + 2$ узлов $t_j, t_{j+1}, \dots, t_{j+m+1}$. Тогда B -сплайн $N_i^m(t)$ отличен от нуля только в области $t_j \leq t \leq t_{j+m+1}$ и $N_i^m(t) = 0$ при $t < t_j$ и при $t > t_{j+m+1}$. Действительно, как только $t < t_j$, то между точками t_i и t_{i+m-1} усеченная функция примет вид $\sigma_m(z) = (z - t)^m$ и в силу того, что $(z - t)^m$ является полиномом степени меньшей $m + 1$, любая разделенная разность $(m + 1)$ -го порядка, построенная по этому полиному, равна нулю. Как только $t > t_{j+m+1}$, то разделенная разность равна нулю в силу того, что между узлами t_i и t_{i+m+1} $\sigma_m(z) = 0$. Только при $t_j \leq t \leq t_{j+m+1}$ разделенная разность $\sigma_m[t_j, t_{j+1}, \dots, t_{j+m+1}]$ отлична от нуля. Таким образом, B -сплайны являются локальными функциями, принимающими ненулевые значения на отрезке, содержащем узлы построения B -сплайна.

Свойство 2. $N_i^m(t) \geq 0$, то есть B -сплайны при любых значениях параметра принимают неотрицательные значения. Это следует из предыдущего свойства и рекуррентного соотношения (2.8.5) при неубывающем расположении узлов.

Свойство 3. Пусть задана бесконечная (или достаточно длинная) неубывающая последовательность узлов t_i и на каждом $(m + 2)$ -х подряд расположенных узлах построен B -сплайн m -го порядка. Покажем, что для любого $t_j \leq t < t_{j+1}$ отличны от нуля только $m + 1$ B -сплайнов а именно, $N_{i-m}^m(t), N_{i+1-m}^m(t), \dots, N_{i-1}^m(t), N_i^m(t)$, и их сумма равна единице. Действительно, все B -сплайны $N_j^m(t)$, где $j < i - m$, равны нулю, так как построены на узлах, где $\sigma_m(z) = 0$; все B -сплайны $N_j^m(t)$, где $j \geq i + 1$, равны нулю, так как построены на узлах, где $\sigma_m(z) = (z - t)^m$ и являются полиномом степени, меньшей $m + 1$. Таким образом, сумма всех B -сплайнов при $t_j \leq t < t_{j+1}$ зависит от конечного числа B -сплайнов, а именно:

$$\sum_j N_j^m(t) = \sum_{j=i-m}^i N_j^m(t). \quad (2.8.8)$$

Если воспользоваться определением разделенной разности (2.7.2), то B -сплайн (2.8.1) можно записать в виде

$$N_j^m(t) = \sigma_m[t_{j+1}, t_{j+2}, \dots, t_{j+m+1}] - \sigma_m[t_j, t_{j+1}, \dots, t_{j+m}]. \quad (2.8.9)$$

Используя равенства (2.8.8) и (2.8.9), найдем сумму всех B -сплайнов при $t_j \leq t < t_{j+1}$

$$\begin{aligned} \sum_{j=i-m}^i N_j^m(t) &= \sum_{j=i-m}^i (\sigma_m[t_{j+1}, t_{j+2}, \dots, t_{j+m+1}] - \sigma_m[t_j, t_{j+1}, \dots, t_{j+m}]) = \\ &= \sum_{j=i-m}^i \sigma_m[t_{j+1}, t_{j+2}, \dots, t_{j+m+1}] - \sum_{j=i-m}^i \sigma_m[t_j, t_{j+1}, \dots, t_{j+m}] = \\ &= \sigma_m[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] - \sigma_m[t_{i-m}, t_{i+1-m}, \dots, t_i] = 1 - 0 = 1. \end{aligned} \quad (2.8.10)$$

В (2.8.10) использовалось то, что $\sigma_m[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] = 1$ как коэффициент при наивысшей степени аргумента полинома $(z - t)^m$, которым опи-

сывается усеченная степенная функция на открытом интервале $t_{i+1} < z < t_{i+m+1}$. При этом на интервале $t_{i-m} < z < t_i$ усеченная степенная функция равна нулю, следовательно, ее раздельная разность $\sigma_m[t_{i-m}, t_{i+1-m}, \dots, t_i] = 0$. Так как участок $t_i \leq t < t_{i+1}$ был выбран произвольно, то аналогичное равенство выполняется для любого другого участка. Таким образом, при любом t для суммы всех B -сплайнов справедливо равенство

$$\sum_j N_j^m(t) = 1. \quad (2.8.11)$$

Соотношение (2.8.11) B -сплайнов аналогично соотношению (2.4.3) для базиса Бернштейна. Таким образом, при любом фиксированном параметре t совокупность всех ненулевых функций $N_i^m(t)$ представляет собой разложение единицы.

Свойство 4. Для неубывающей последовательности узлов $t_j, t_{j+1}, \dots, t_{i+m+1}$ B -сплайн вычисляется с помощью рекуррентного соотношения

$$N_i^m(t) = \frac{t_{i+m+1} - t}{t_{i+m+1} - t_{i+1}} N_{i+1}^{m-1}(t) + \frac{t - t_i}{t_{i+m} - t_i} N_i^{m-1}(t), \quad (2.8.12)$$

начиная с B -сплайнов нулевого порядка

$$N_i^0(t) = \begin{cases} 1, & \text{если } t_j \leq t < t_{i+1}, \\ 0, & \text{в противном случае.} \end{cases} \quad (2.8.13)$$

Соотношение (2.8.12) следует из рекуррентной формулы Кокса–Де Бура (2.8.5) при подстановке в нее соотношений между B -сплайнами и соответствующими ненормированными B -сплайнами

$$N_i^m(t) = (t_{i+m+1} - t_i) M_i^m(t). \quad (2.8.14)$$

Свойство 5. Производная B -сплайна m -го порядка выражается через B -сплайны $(m - 1)$ -го порядка. Для доказательства этого утверждения продифференцируем B -сплайн, построенный на неубывающей последовательности узлов $t_j, t_{j+1}, \dots, t_{i+m+1}$ и представленный в виде (2.8.9)

$$\frac{dN_i^m(t)}{dt} = \frac{d\sigma_m[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}]}{dt} - \frac{d\sigma_m[t_j, t_{i+1}, \dots, t_{i+m}]}{dt}.$$

Подставив в последнее выражение производные по t усеченной степенной функции

$$\frac{d\sigma_m}{dt} = -m(z - t)^{m-1} = -m\sigma_{m-1},$$

получим

$$\begin{aligned}
\frac{dN_i^m(t)}{dt} &= -m \sigma_{m-1}[t_{i+1}, t_{i+2}, \dots, t_{i+m+1}] + m \sigma_{m-1}[t_i, t_{i+1}, \dots, t_{i+m}] = \\
&= -m M_{i+1}^{m-1}(t) + m M_i^{m-1}(t) = \\
&= -\frac{m}{t_{i+m+1} - t_{i+1}} N_{i+1}^{m-1}(t) + \frac{m}{t_{i+m} - t_i} N_i^{m-1}(t). \tag{2.8.15}
\end{aligned}$$

Аналогично получим, что производная ненормированного B -сплайна m -го порядка выражается через ненормированные B -сплайны $(m - 1)$ -го порядка формулой

$$\frac{dM_i^m(t)}{dt} = m \frac{M_i^{m-1}(t) - M_{i+1}^{m-1}(t)}{t_{i+m+1} - t_i}.$$

Свойство 6. Пусть B -сплайн построен на неубывающей последовательности узлов $t_i, t_{i+1}, \dots, t_{i+m+1}$. Из соотношения (2.8.12) и равенства (2.8.15) следует, что B -сплайн m -го порядка является m раз дифференцируемой функцией. При отсутствии кратных узлов B -сплайн и его производные до $(m - 1)$ -го порядка включительно являются непрерывными функциями и при $t = t_i$ и $t = t_{i+m+1}$ равны нулю.

Свойство 7. Площадь под любым ненормированным B -сплайном удовлетворяет равенству

$$\int_{-\infty}^{+\infty} M_i^m(t) dt = \int_{t_{\min}}^{t_{\max}} M_i^m(t) dt = \frac{1}{m+1}, \tag{2.8.16}$$

где t_{\max} и t_{\min} — максимальное и минимальное значения узлов последовательности, на которой вычислен B -сплайн. Равенство (2.8.16) получим, проинтегрировав усеченную степенную функцию (2.7.8) в указанных пределах

$$\int_{t_{\min}}^{t_{\max}} (z-t)^m dt = \frac{1}{m+1} (z-t_{\min})^{m+1} + -\frac{1}{m+1} (z-t_{\max})^{m+1} +$$

и вычислив разделенную разность $(m + 1)$ -го порядка для полученной функции на последовательности узлов B -сплайна. Второе слагаемое правой части на области $t_{\min} \leq z \leq t_{\max}$ равно нулю, а первое слагаемое равно полиному $(m + 1)$ -й степени, разделенная разность $(m + 1)$ -го порядка которого в соответствии с соотношением (2.7.6) равна коэффициенту при z^{m+1} , то есть равна $1 / (m + 1)$, что доказывает равенство (2.8.16).

Из (2.8.16) следует равенство

$$\int_{-\infty}^{+\infty} N_i^m(t) dt = (t_{\max} - t_{\min}) \int_{t_{\min}}^{t_{\max}} M_i^m(t) dt = \frac{t_{\max} - t_{\min}}{m+1}$$

для вычисления площади под B -сплайном.

2.9. В-кривые

Основу теории В-кривых (*B-spline curves*) заложили Кунс (S.A. Coons) и Де Бур (C. De Boor). Развил теорию В-кривых и применил ее для проектирования Версприлл (K.J. Versprille).

Построим неоднородную (non-uniform) рациональную (rational) кривую на основе В-сплайнов $N_j^m(t)$ по контрольным точкам \mathbf{p}_j , обладающим весами $w_j, j = 0, 1, \dots, n$ ($n \geq m$). Радиус-вектор кривой вычислим по формуле

$$\mathbf{r}(t) = \frac{\sum_{j=0}^n N_j^m(t) w_j \mathbf{p}_j}{\sum_{j=0}^n N_j^m(t) w_j}. \quad (2.9.1)$$

Каждый из В-сплайнов m -го порядка $N_j^m(t)$ построен на неубывающей последовательности из $m + 2$ узлов $t_j, t_{j+1}, \dots, t_{j+m+1}$. Кривая (2.9.1) в общем случае является неоднородной и рациональной. Неоднородность кривой говорит о том, что ее узлы отстоят друг от друга на неравных расстояниях, а рациональность кривой говорит о наличии весов у контрольных точек, что влечет за собой присутствие знаменателя в формуле (2.9.1). Кривую (2.9.1) называют **NURBS кривой** или **В-кривой**. Аббревиатура NURBS получена из первых букв словосочетания Non-Uniform Rational B-Spline.

Когда все контрольные точки кривой (2.9.1) имеют равные веса, то в силу равенства (2.8.11) радиус-вектор В-кривой вычислим по формуле

$$\mathbf{r}(t) = \sum_{j=0}^n N_j^m(t) \mathbf{p}_j. \quad (2.9.2)$$

Такой же вид формула (2.9.1) имеет в терминах однородных координат (1.4.8).

Для построения совокупности $n + 1$ В-сплайнов m -го порядка незамкнутой В-кривой требуется $n + m + 2$ узлов. Последовательность узлов $t_0, t_1, \dots, t_{m+n+2}$, на которой строится В-кривая, называют *узловым вектором*. Форма В-кривой зависит не только от расположения контрольных точек, но и от значений узлов. Для того чтобы продемонстрировать зависимость формы кривой от значений узлов, построим две В-кривые одного и того же порядка по одним и тем же контрольным точкам, но с разными узловыми векторами. На рис. 2.9.1 приведены две В-кривые и ломаная линия, построенные по одним и тем же контрольным точкам. Сплошной линией на рис. 2.9.1 показана В-кривая третьего порядка с узловым вектором 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Штриховой линией на рис. 2.9.1 показана В-кривая третьего порядка с узловым вектором 0, 1, 11, 12, 22, 23, 33, 34, 44, 45, 55. В качестве узловых векторов взяты две различные неубывающие последовательности из одиннадцати чисел.

Тонкой линией на рис. 2.9.1 показана ломаная линия, построенная по контрольным точкам B -кривых.

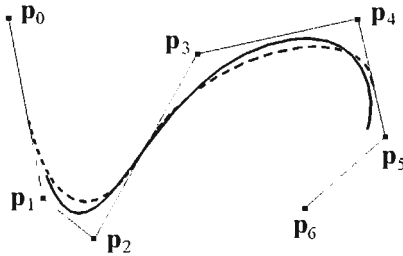


Рис. 2.9.1

На рис. 2.9.2 приведена B -кривая 3-го порядка, построенная по тем же контрольным точкам, что и B -кривая на рис. 2.9.1, но имеющая узловый вектор $\hat{0}, \hat{0}, \hat{0}, \hat{1}, \hat{2}, \hat{3}, \hat{4}, \hat{4}, \hat{4}, \hat{4}$. \hat{B} -кривая, приведенная на рис. 2.9.2, имеет кратные узлы в начале и в конце узлового вектора и поэтому проходит через свои крайние контрольные точки. Чтобы незамкнутая кривая проходила через первую и последнюю контрольные точки, первый B -сплайн должен иметь кратными первые $m + 1$ узлы из $m + 2$ узлов, на которых он строится, а последний B -сплайн должен иметь кратными последние $m + 1$ узлы из $m + 2$ узлов, на которых он строится. То есть для незамкнутой B -кривой, проходящей через крайние точки, требуется выполнение равенств $t_0 = t_1 = \dots = t_m$ и $t_{n+1} = t_{n+2} = \dots = t_{n+m+1}$. Параметр B -кривой изменяется от значения узла $t_{\min} = t_m$ до значения узла $t_{\max} = t_{n+1}$.

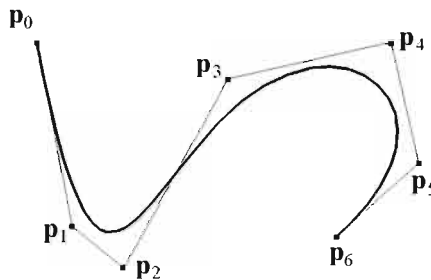


Рис. 2.9.2

Для построения совокупности $n + 1$ B -сплайнов m -го порядка циклически замкнутой B -кривой требуется $n + 2m + 2$ узлов (узловой вектор должен быть расширен на m узлов). Последовательность узлов должна отражать замкнутость: первые $2m$ узлов должны идти через интервалы, повторяющие интервалы, через которые идут последние $2m$ узлов. То

есть для замкнутой B -кривой требуется выполнение равенств $t_1 - t_0 = t_{n+2} - t_{n+1}$, $t_2 - t_1 = t_{n+3} - t_{n+2}$, ..., $t_{2m+1} - t_{2m} = t_{n+2m+2} - t_{n+2m+1}$. На рис. 2.9.3 приведена циклически замкнутая B -кривая 3-го порядка, имеющая узловой вектор $-3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11$, а также ломаная линия, построенная по восьми контрольным точкам B -кривой.

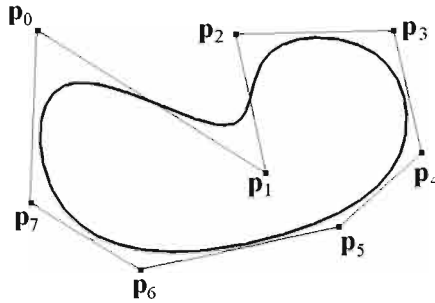


Рис. 2.9.3

Параметр циклически замкнутой B -кривой изменяется от значения узла $t_{\min} = t_m$ до значения узла $t_{\max} = t_{n+m+1}$. Число узлов B -кривых всегда больше числа контрольных точек, поэтому узловый вектор называют расширенным.

B -кривые, построенные на узловых векторах, не содержащих кратных узлов или содержащих кратные узлы только в начале и конце узлового вектора (для незамкнутых кривых), имеют непрерывные производные до $(m - 1)$ -го порядка включительно на всей области определения. Если узловый вектор содержит кратные узлы, кроме краевых, то непрерывность соответствующих производных кривой будет нарушена.

В соответствии с формулой Кокса–Де Бура (2.8.5) схема вычисления одного B -сплайна $N_j^m(t)$ формально выглядит следующим образом (порядок B -сплайна возрастает сверху вниз, а номера узлов возрастают слева направо):

$$\begin{array}{r}
 N_j^0(t) \quad N_{j+1}^0(t) \dots N_{j+m-2}^0(t) \quad N_{j+m-1}^0(t) N_{j+m}^0(t) \\
 N_j^1(t) \quad N_{j+1}^1(t) \quad \dots \quad N_{j+m-2}^1(t) \quad N_{j+m-1}^1(t) \\
 N_j^2(t) \quad N_{j+1}^2(t) \quad \dots \quad N_{j+m-2}^2(t) \\
 \dots \quad \dots \quad \dots \\
 N_j^{m-1}(t) \quad N_{j+1}^{m-1}(t) \\
 N_j^m(t)
 \end{array}$$

Но при любом параметре t среди всех B -сплайнов нулевого порядка (2.8.13) только один отличен от нуля, поэтому для вычисления радиуса-вектора B -кривой используется следующая схема. По значению параметра t из условия $t_i \leq t < t_{i+1}$ определяется номер i отличного от нуля

B -сплайна нулевого порядка и вычисляется значение ненормированного B -сплайна:

$$M_i^0(t) = \frac{1}{t_{i+1} - t_i}. \quad (2.9.3)$$

Далее, используя рекуррентное соотношение Кокса–Де Бура

$$M_i^m(t) = \frac{t_{i+m+1} - t}{t_{i+m+1} - t_i} M_{i+1}^{m-1}(t) + \frac{t - t_i}{t_{i+1} - t} M_i^{m-1}(t), \quad (2.9.4)$$

последовательно вычисляются все отличные от нуля при данном параметре t ненормированные B -сплайны до m -го порядка включительно (порядок B -сплайна возрастает сверху вниз, а номера узлов возрастают слева направо):

$$\begin{array}{ccccccc} & & & & & & M_i^0(t) \\ & & & & & & M_{i-1}^1(t) M_i^1(t) \\ & & & & & M_{i-2}^2(t) & M_{i-1}^2(t) M_i^2(t) \\ & & & \dots & & \dots & \dots \\ & & M_{i-m+1}^{m-1}(t) & \dots & M_{i-2}^{m-1}(t) & M_{i-1}^{m-1}(t) & M_i^{m-1}(t) \\ M_{i-m}^m(t) & M_{i-m+1}^m(t) & \dots & & M_{i-2}^m(t) & M_{i-1}^m(t) & M_i^m(t) \end{array}$$

Данная треугольная таблица вычисляется построчно, поскольку из (2.9.3) известен элемент первой строки, а каждый элемент следующей строки можно построить по двум соседним элементам в предыдущей строке с помощью (2.9.4). При вычислении крайних элементов каждой строки используется тот факт, что один из элементов в предыдущей строке, отсутствующий в таблице, равен нулю. Далее B -сплайны нормируются

$$N_j^m(t) = (t_{j+m+1} - t_j) M_j^m(t), j = i-m, i-m+1, \dots, i$$

и подставляются в формулу (2.9.1), которая для $t_i \leq t < t_{i+1}$ примет вид

$$\mathbf{r}(t) = \frac{\sum_{j=i-m}^i N_j^m(t) w_j \mathbf{p}_j}{\sum_{j=i-m}^i N_j^m(t) w_j}.$$

С помощью соотношений (1.58) может проводиться одновременное вычисление B -сплайнов и их производных, что используется в итерационных процессах пересечения кривых.

Пусть для значения параметра t найден номер i узла t_i из условия $t_i \leq t < t_{i+1}$. При данном параметре t отличными от нуля будут B -сплайны: $N_{i-m}^m(t), N_{i-m+1}^m(t), \dots, N_{i-1}^m(t), N_i^m(t)$. Продифференцируем числитель правой части равенства (2.9.1) с учетом (2.8.15) и получим

$$\begin{aligned}
\frac{d(w\mathbf{r})}{dt} &\equiv \frac{d}{dt} \left(\sum_{j=i-m}^i N_j^m w_j \mathbf{p}_j \right) = \sum_{j=i-m}^i \frac{dN_j^m}{dt} w_j \mathbf{p}_j = \\
&= m \sum_{j=i-m}^i N_j^{m-1} \frac{w_j \mathbf{p}_j}{t_{j+m} - t_j} - m \sum_{j=i-m}^i N_{j+1}^{m-1} \frac{w_j \mathbf{p}_j}{t_{i+m+1} - t_{j+1}} = \\
&= m N_{i-m}^{m-1} \frac{w_{i-m} \mathbf{p}_{i-m}}{t_i - t_{i-m}} + m \sum_{j=i-m+1}^i N_j^{m-1} \frac{w_j \mathbf{p}_j}{t_{j+m} - t_j} - \\
&- m N_{i+1}^{m-1} \frac{w_i \mathbf{p}_i}{t_{i+m+1} - t_{i+1}} - m \sum_{j=i-m+1}^i N_j^{m-1} \frac{w_{j-1} \mathbf{p}_{j-1}}{t_{j+m} - t_j} = \\
&= m \sum_{j=i-m+1}^i N_j^{m-1} \frac{w_j \mathbf{p}_j - w_{j-1} \mathbf{p}_{j-1}}{t_{j+m} - t_j} = \sum_{j=i-m+1}^i N_j^{m-1} \mathbf{p}_j^{(1)}, \quad (2.9.5)
\end{aligned}$$

где введено обозначение

$$\mathbf{p}_j^{(1)} = m \frac{w_j \mathbf{p}_j - w_{j-1} \mathbf{p}_{j-1}}{t_{j+m} - t_j}, \quad j = 1, 2, \dots, n.$$

В равенстве (2.9.5) использовались свойства $N_{i-m}^{m-1}(t) = 0$ и $N_{i+1}^{m-1}(t) = 0$ при $t_i \leq t < t_{i+1}$. Равенство (2.9.5) показывает, что первая производная функции $\sum_{j=i-m}^i N_j^m w_j \mathbf{p}_j$ представляет собой аналогичную функцию

$$\sum_{j=i-m+1}^i N_j^{m-1} \mathbf{p}_j^{(1)}, \text{ порядок и число слагаемых которой на единицу меньше.}$$

Продолжая дифференцирование, можно найти производную требуемого порядка для числителя выражения (2.9.1)

$$\frac{d^k(w\mathbf{r})}{dt^k} \equiv \frac{d^k}{dt^k} \left(\sum_{j=0}^n N_j^m w_j \mathbf{p}_j \right) = \sum_{j=k}^n N_j^{m-k} \mathbf{p}_j^{(k)}, \quad (2.9.6)$$

$$\text{где } \mathbf{p}_j^{(k)} = (m+1-k) \frac{\mathbf{p}_j^{(k-1)} - \mathbf{p}_{j-1}^{(k-1)}}{t_{j+m+1-k} - t_j}, \quad k = 1, 2, \dots, m, j = k, k+1, \dots, n, \mathbf{p}_j^{(0)} = w_j \mathbf{p}_j.$$

Для производных знаменателя (2.9.1) получим выражение, аналогичное (2.9.6):

$$\frac{d^k w}{dt^k} \equiv \frac{d^k}{dt^k} \left(\sum_{j=0}^n N_j^m w_j \right) = \sum_{j=k}^n N_j^{m-k} w_j^{(k)}, \quad (2.9.7)$$

$$\text{где } w_j^{(k)} = (m+1-k) \frac{w_j^{(k-1)} - w_{j-1}^{(k-1)}}{t_{j+m+1-k} - t_j}, \quad k = 1, 2, \dots, m, j = k, k+1, \dots, n, w_j^{(0)} = w_j.$$

Радиус-вектор B -кривой (2.9.1) вычисляется как частное от деления двух функций параметра кривой t , поэтому при вычислении производной B -кривой правую часть (2.9.1) следует рассматривать как сложную функцию, первая производная которой равна

$$\frac{d\mathbf{r}}{dt} = \frac{d}{dt} \left(\frac{w\mathbf{r}}{w} \right) = \frac{1}{w} \frac{d(w\mathbf{r})}{dt} - \frac{w\mathbf{r}}{w^2} \frac{dw}{dt}, \quad (2.9.8)$$

где $w\mathbf{r} = \sum_{j=0}^n N_j^m(t) w_j \mathbf{p}_j$ – числитель правой части равенства (2.9.1);

$w = \sum_{i=0}^n N_i^m(t) w_i$ – знаменатель правой части равенства (2.9.1);

$\frac{d(w\mathbf{r})}{dt} = m \sum_{j=1}^n N_j^{m-1}(t) \frac{w_j \mathbf{p}_j - w_{j-1} \mathbf{p}_{j-1}}{t_{j+m} - t_j}$ – производная числителя;

$\frac{dw}{dt} = m \sum_{j=1}^n N_j^{m-1}(t) \frac{w_j - w_{j-1}}{t_{j+m} - t_j}$ – производная знаменателя.

Аналогично вычисляются производные B -кривой более высокого порядка:

$$\frac{d^{(k)}\mathbf{r}}{dt^{(k)}} = \frac{d^{(k)}}{dt^{(k)}} \left(\frac{w\mathbf{r}(t)}{w(t)} \right).$$

Следует заметить, что радиус-вектор точки и ее вес ($w\mathbf{r}$ или $w\mathbf{p}$) в формулах B -кривой выступают как единое целое.

2.10. Алгоритм Де Бура

Преобразуем числитель выражения (2.9.1), используя формулу (2.8.12), следующим образом:

$$\begin{aligned} w\mathbf{r}(t) &\equiv \sum_{j=0}^n N_j^m w_j \mathbf{p}_j = \sum_{j=n-m}^i N_j^m w_j \mathbf{p}_j = \\ &\sum_{j=n-m}^i \left(\frac{t_{j+m+1} - t}{t_{j+m+1} - t_{j+1}} N_{j+1}^{m-1} + \frac{t - t_j}{t_{j+m} - t_j} N_j^{m-1} \right) w_j \mathbf{p}_j = \\ &= \frac{t_{i+m+1} - t}{t_{i+m+1} - t_{i+1}} N_{i+1}^{m-1} w_i \mathbf{p}_i + \sum_{j=n-m+1}^i \frac{t_{j+m} - t}{t_{j+m} - t_j} N_j^{m-1} w_{j-1} \mathbf{p}_{j-1} + \\ &+ \frac{t - t_{i-m}}{t_i - t_{i-m}} N_{i-m}^{m-1}(t) w_{i-m} \mathbf{p}_{i-m} + \sum_{j=n-m+1}^i \frac{t - t_j}{t_{j+m} - t_j} N_j^{m-1} w_j \mathbf{p}_j = \\ &= \sum_{i=n-m+1}^i N_j^{m-1} \left(\frac{t_{j+m} - t}{t_{j+m} - t_j} w_{j-1} \mathbf{p}_{j-1} + \frac{t - t_j}{t_{j+m} - t_j} w_j \mathbf{p}_j \right) = \\ &= \sum_{j=n-m+1}^i N_j^{m-1} \mathbf{r}_j^{(1)}, \end{aligned} \quad (2.10.1)$$

где введено обозначение

$$\mathbf{r}_j^{(1)} = \frac{t_{j+m} - t}{t_{j+m} - t_j} w_{j-1} \mathbf{p}_{j-1} + \frac{t - t_j}{t_{j+m} - t_j} w_j \mathbf{p}_j.$$

В равенстве (2.10.1) использовалось то обстоятельство, что при $t_i \leq t < t_{i+1}$ отличны от нуля только $N_{i-m}^m(t)$, $N_{i-m+1}^m(t)$, ..., $N_{i-1}^m(t)$, $N_i^m(t)$, а также равенства $N_{i-m}^{m-1}(t) = 0$ и $N_{i+1}^{m-1}(t) = 0$. Мы свели сумму $m + 1$ слагаемых к сумме m слагаемых и понизили на единицу порядок B -сплайнов в этой сумме. Продолжим упрощение числителя выражения (2.9.1) аналогичным образом и получим

$$\begin{aligned} w\mathbf{r}(t) &\equiv \sum_{j=0}^n N_j^m w_j \mathbf{p}_j = \sum_{j=n}^l N_j^m w_j \mathbf{p}_j = \sum_{j=i-n+1}^l N_j^{m-1} \mathbf{r}_j^{(1)} = \\ &= \sum_{j=i-m+2}^l N_j^{m-2} \left(\frac{t_{j+m-1} - t}{t_{j+m-1} - t_j} \mathbf{r}_{j-1}^{(1)} + \frac{t - t_j}{t_{j+m-1} - t_j} \mathbf{r}_j^{(1)} \right) = \\ &= \sum_{j=i-m+2}^l N_j^{m-2} \mathbf{r}_j^{(2)} = \dots = \sum_{j=i-m+m}^l N_j^0(t) \mathbf{r}_j^{(m)} = \mathbf{r}_i^{(m)}, \end{aligned} \quad (2.10.2)$$

где введены обозначения $\mathbf{r}_j^{(0)} = w_j \mathbf{p}_j$, $\mathbf{r}_j^{(k)} = \frac{t_{j+m+1-k} - t}{t_{j+m+1-k} - t_j} \mathbf{r}_{j-1}^{(k-1)} + \frac{t - t_j}{t_{j+m+1-k} - t_j} \mathbf{r}_j^{(k-1)}$, при изменении k от 1 до m . Для знаменателя выражения (2.9.1) получим аналогичное выражение

$$\begin{aligned} w(t) &\equiv \sum_{j=0}^n N_j^m w_j = \sum_{j=i-n}^l N_j^m w_j = \sum_{j=i-n+1}^l N_j^{m-1} w_j^{(1)} = \\ &= \sum_{j=i-m+2}^l N_j^{m-2} w_j^{(2)} = \dots = \sum_{j=i-m+m}^l N_j^0(t) w_j^{(m)} = w_i^{(m)}, \end{aligned} \quad (2.10.3)$$

где введены обозначения $w_j^{(0)} = w_j$, $w_j^{(k)} = \frac{t_{j+m+1-k} - t}{t_{j+m+1-k} - t_j} w_{j-1}^{(k-1)} + \frac{t - t_j}{t_{j+m+1-k} - t_j} w_j^{(k-1)}$, при изменении k от 1 до m .

Мы пришли к выводу, что положение точки B -кривой (2.9.1) для заданного параметра $t_i \leq t < t_{i+1}$ может быть определено по формуле

$$\mathbf{r}(t) = \frac{w\mathbf{r}(t)}{w(t)} = \frac{\mathbf{r}_i^{(m)}}{w_i^{(m)}} \quad (2.10.4)$$

с помощью рекуррентных соотношений

$$\mathbf{r}_j^{(k)} = \frac{t_{j+m+1-k} - t}{t_{j+m+1-k} - t_j} \mathbf{r}_{j-1}^{(k-1)} + \frac{t - t_j}{t_{j+m+1-k} - t_j} \mathbf{r}_j^{(k-1)},$$

$$w_j^{(k)} = \frac{t_{j+m+1-k} - t_j}{t_{j+m+1-k} - t_j} w_{j-1}^{(k-1)} + \frac{t - t_j}{t_{j+m+1-k} - t_j} w_j^{(k-1)},$$

начиная со значений $\mathbf{r}_j^{(0)} = w_j \mathbf{p}_j$, $w_j^{(0)} = w_j$. Эти соотношения являются обобщением алгоритма Де Кастелье и называются **алгоритмом Де Бура**.

$$\begin{array}{ccccccc}
 \mathbf{r}_{i-m}^{(0)} & \mathbf{r}_{i-m+1}^{(0)} & \mathbf{r}_{i-m+2}^{(0)} & \dots & \mathbf{r}_{i-2}^{(0)} & \mathbf{r}_{i-1}^{(0)} & \mathbf{r}_i^{(0)} \\
 & \mathbf{r}_{i-m+1}^{(1)} & \mathbf{r}_{i-m+2}^{(1)} & \dots & \mathbf{r}_{i-2}^{(1)} & \mathbf{r}_{i-1}^{(1)} & \mathbf{r}_i^{(1)} \\
 & & \mathbf{r}_{i-m+2}^{(2)} & \dots & \mathbf{r}_{i-2}^{(2)} & \mathbf{r}_{i-1}^{(2)} & \mathbf{r}_i^{(2)} \\
 & & & \dots & & & \\
 & & & & & \mathbf{r}_{i-2}^{(m-2)} & \mathbf{r}_{i-1}^{(m-2)} & \mathbf{r}_i^{(m-2)} \\
 & & & & & & \mathbf{r}_{i-1}^{(m-1)} & \mathbf{r}_i^{(m-1)} \\
 & & & & & & & \mathbf{r}_i^{(m)}
 \end{array}$$

Рис. 2.10.1

Алгоритм вычисления радиуса-вектора точки B -кривой для параметра $t_i \leq t < t_{i+1}$ иллюстрирует рис. 2.10.1.

2.11. Вставка точки и узла B -кривой

Число узлов и контрольных точек B -кривой можно увеличить, сохранив неизменной ее форму и параметрическую длину. Вставим дополнительный узел v в последовательность узлов $t_0, t_1, \dots, t_{m+m+2}$ кривой (2.9.2).

Пусть значение v удовлетворяет условиям $t_{\min} < v < t_{\max}$ и не совпадает ни с одним из узлов B -кривой m -го порядка. Найдем номер i узла t_i из условия $t_i \leq v < t_{i+1}$ и построим последовательность узлов $v_0 = t_0, v_1 = t_1, \dots, v_i = t_i, v_{i+1} = v, v_{i+2} = t_{i+1}, \dots, v_{n+m+3} = t_{n+m+2}$. Новая кривая должна иметь на единицу больше узлов и контрольных точек. Радиус-вектор новой кривой опишем формулой

$$\mathbf{r}(t) = \sum_{j=0}^{n+1} \underline{N}_j^m(t) \mathbf{q}_j.$$

B -сплайны новой кривой обозначим подчеркнутыми буквами. В силу значений узлов будут выполняться равенства

$$\begin{aligned}
 \underline{N}_j^m(t) &= N_j^m(t), & j &= 0, 1, \dots, i-m-1, \\
 \underline{N}_j^m(t) &= \underline{N}_{j+1}^m(t), & j &= i+1, i+2, \dots, n,
 \end{aligned}$$

$$N_j^m(t) = \frac{v_{j+m+2} - v}{v_{j+m+2} - v_{j+1}} \underline{N}_{j+1}^m(t) + \frac{v - v_j}{v_{j+m+1} - v_j} \underline{N}_j^m(t), j = i-m, i-m+1, \dots, i.$$

Для ненормированных B -сплайнов последнее равенство имеет вид

$$M_j^m(t) = \frac{v_{j+m+2} - v}{v_{j+m+2} - v_j} \underline{M}_{j+1}^m(t) + \frac{v - v_j}{v_{j+m+2} - v_j} \underline{M}_j^m(t). \quad (2.11.1)$$

$$j = i-m, i-m+1, \dots, i.$$

Равенство (2.11.1) докажем методом индукции. Непосредственной подстановкой убеждаемся, что равенство (2.11.1) для ненормированных B -сплайнов нулевого порядка выполняется

$$\frac{1}{t_{j+1} - t_j} = \frac{v_{i+2} - v}{v_{i+2} - v_i} \frac{1}{v_{i+2} - v_{i+1}}.$$

Предположим, что равенство (2.11.1) справедливо для ненормированных B -сплайнов $(m-1)$ -го порядка и убедимся, что из этого предположения следует его верность для ненормированных B -сплайнов m -го порядка. Для этого подставим в (2.8.5) равенства (2.11.1) для $M_{j+1}^{m-1}(t)$ и $M_j^{m-1}(t)$ и получим

$$\begin{aligned} M_j^m(t) &= \frac{t_{j+m+1} - t}{t_{j+m+1} - t_j} \left(\frac{v_{j+m+2} - v}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+2}^{m-1}(t) + \frac{v - v_{j+1}}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+1}^{m-1}(t) \right) + \\ &+ \frac{t - t_j}{t_{j+m+1} - t_j} \left(\frac{v_{j+m+1} - v}{v_{j+m+1} - v_j} \underline{M}_{j+1}^{m-1}(t) + \frac{v - v_j}{v_{j+m+1} - v_j} \underline{M}_j^{m-1}(t) \right) = \\ &= \frac{t_{j+m+1} - t}{t_{j+m+1} - t_j} \frac{v_{j+m+2} - v}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+2}^{m-1}(t) + \frac{t - t_j}{t_{j+m+1} - t_j} \frac{v_{j+m+1} - v}{v_{j+m+1} - v_j} \underline{M}_{j+1}^{m-1}(t) + \\ &+ \frac{t_{j+m+1} - t}{t_{j+m+1} - t_j} \frac{v - v_{j+1}}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+1}^{m-1}(t) + \frac{t - t_j}{t_{j+m+1} - t_j} \frac{v - v_j}{v_{j+m+1} - v_j} \underline{M}_j^{m-1}(t). \end{aligned}$$

Во втором и третьем слагаемых выполним замену, используя равенство

$$\frac{(t - t_j)(v_{j+m+1} - v)}{v_{j+m+1} - v_j} + \frac{(t_{j+m+1} - t)(v - v_{j+1})}{v_{j+m+2} - v_{j+1}} = \frac{(v_{j+m+2} - v)(t - v_{j+1})}{v_{j+m+2} - v_{j+1}} + \frac{(v - v_j)(v_{j+m+1} - t)}{v_{j+m+1} - v_j},$$

которое можно проверить непосредственными вычислениями. После замены получим

$$M_j^m(t) = \frac{v_{j+m+2} - v}{v_{j+m+2} - v_j} \frac{v_{j+m+2} - t}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+2}^{m-1}(t) + \frac{v_{j+m+2} - v}{v_{j+m+2} - v_j} \frac{t - v_{j+1}}{v_{j+m+2} - v_{j+1}} \underline{M}_{j+1}^{m-1}(t) +$$

$$\begin{aligned}
& + \frac{v - v_j}{v_{j+m+2} - v_j} \frac{v_{j+m+1} - t}{v_{j+m+1} - v_j} \underline{M}_{j+1}^{m-1}(t) + \frac{v - v_j}{v_{j+m+2} - v_j} \frac{t - v_j}{v_{j+m+1} - v_j} \underline{M}_j^{m-1}(t) = \\
& = \frac{v_{j+m+2} - v}{v_{j+m+2} - v_j} \underline{M}_{j+1}^m(t) + \frac{v - v_j}{v_{j+m+2} - v_j} \underline{M}_j^m(t),
\end{aligned}$$

что и требовалось доказать.

Для новой кривой должно выполняться равенство

$$\sum_{j=0}^n N_j^m(t) \mathbf{p}_j = \sum_{j=0}^{n+1} \underline{N}_j^m(t) \mathbf{q}_j.$$

В силу независимости B -сплайнов получим равенства

$$\begin{aligned}
\mathbf{p}_j &= \mathbf{q}_j, j = 0, 1, \dots, i-m-1, \\
\mathbf{p}_j &= \mathbf{q}_{j+1}, j = i+1, i+2, \dots, n, \\
\sum_{j=i-m}^i N_j^m(t) \mathbf{p}_j &= \sum_{j=i-m}^{i+1} \underline{N}_j^m(t) \mathbf{q}_j.
\end{aligned}$$

Подставим в последнее равенство выражение B -сплайнов кривой (2.9.2) через B -сплайны новой кривой

$$\begin{aligned}
& \left(\frac{v_{i+2} - v}{v_{i+2} - v_{i-m+1}} \underline{N}_{i-m+1}^m(t) + \frac{v - v_{i-m}}{v_{i+1} - v_{i-m}} \underline{N}_{i-m}^m(t) \right) \mathbf{p}_{i-m} + \\
& + \left(\frac{v_{i+3} - v}{v_{i+3} - v_{i-m+2}} \underline{N}_{i-m+2}^m(t) + \frac{v - v_{i-m+1}}{v_{i+2} - v_{i-m+1}} \underline{N}_{i-m+1}^m(t) \right) \mathbf{p}_{i-m+1} + \\
& \quad \dots \\
& + \left(\frac{v_{i+m+2} - v}{v_{i+m+2} - v_{i+1}} \underline{N}_{i+1}^m(t) + \frac{v - v_i}{v_{i+m+1} - v_i} \underline{N}_i^m(t) \right) \mathbf{p}_i = \\
& = \underline{N}_{i-m}^m(t) \mathbf{q}_{i-m} + \underline{N}_{i-m+1}^m(t) \mathbf{q}_{i-m+1} + \dots + \underline{N}_i^m(t) \mathbf{q}_i + \underline{N}_{i+1}^m(t) \mathbf{q}_{i+1}.
\end{aligned}$$

После перегруппировки слагаемых получим

$$\begin{aligned}
& \underline{N}_{i-m}^m(t) (\mathbf{q}_{i-m} - \mathbf{p}_{i-m}) + \\
& \underline{N}_{i-m+1}^m(t) \left(\mathbf{q}_{i-m+1} - \frac{v - t_{i-m+1}}{t_{i+1} - t_{i-m+1}} \mathbf{p}_{i-m+1} - \frac{t_{i+1} - v}{t_{i+1} - t_{i-m+1}} \mathbf{p}_{i-m} \right) + \\
& \quad \dots \\
& + \underline{N}_i^m(t) \left(\mathbf{q}_i - \frac{v - t_i}{t_{i+m} - t_i} \mathbf{p}_i - \frac{t_{i+m} - v}{t_{i+m} - t_i} \mathbf{p}_{i-1} \right) + \underline{N}_{i+1}^m(t) (\mathbf{q}_{i+1} - \mathbf{p}_i) = 0.
\end{aligned}$$

Введем обозначение $\alpha_j = \frac{v-t_j}{t_{j,m}-t_j}$ для $j=i-m+1, i-m+2, \dots, i$ и получим равенства, выражающие контрольные точки новой кривой через контрольные точки \mathbf{p}_j

$$\begin{aligned} \mathbf{q}_j &= \mathbf{p}_j, & j &= 0, 1, \dots, i-m, \\ \mathbf{q}_j &= \alpha_j \mathbf{p}_j + (1-\alpha_j) \mathbf{p}_{j-1}, & j &= i-m+1, i-m+2, \dots, i, \\ \mathbf{q}_{j+1} &= \mathbf{p}_j, & j &= i, i+2, \dots, n. \end{aligned}$$

Для вставки узла $v, t_i \leq v < t_{i+1}$ в кривую (2.9.1) в полученных равенствах контрольные точки должны быть умножены на свои веса. Веса контрольных точек новой кривой выражаются через веса контрольных точек исходной кривой аналогичными формулами:

$$\begin{aligned} \underline{w}_j &= w_j, & j &= 0, 1, \dots, i-m, \\ \underline{w}_j &= \alpha_j w_j + (1-\alpha_j) w_{j-1}, & j &= i-m+1, i-m+2, \dots, i, \\ \underline{w}_{j+1} &= w_j, & j &= i, i+2, \dots, n. \end{aligned}$$

Вставляемый новый узел в последовательность узлов B -кривой может совпадать с имеющимся узлом, например, $v = t_i$. Таким образом, новый узел может быть вставлен несколько раз, при условии, что его кратность не превысит порядок B -кривой.

2.12. Примеры B -кривых

В качестве примеров построим B -кривые по $n + 1$ контрольным точкам с целочисленными значениями узлов. Параметризация с равноотстоящими значениями узлов называется равномерной.

Пусть первые $m + 1$ узлов незамкнутой B -кривой имеют значения, равные нулю: $t_0 = t_1 = \dots = t_m = 0$; следующие $n - m$ узлов принимают целочисленные значения от 1 до $n - m$: $t_{m+i} = i, i = 1, 2, \dots, n - m$; оставшиеся $m + 1$ узлов принимают значение $n - m + 1$: $t_{n+1} = t_{n+2} = \dots = t_{n+m+1} = n - m + 1$. Параметр незамкнутой B -кривой изменяется в пределах от $t_m = 0$ до $t_{n+1} = n - m + 1$. На рис. 2.12.1 приведен набор B -сплайнов третьего порядка, построенных на расширенном множестве узлов 0, 0, 0, 0, 1, 2, 3, 4, 5, 6, 6, 6, 6 для девяти контрольных точек.

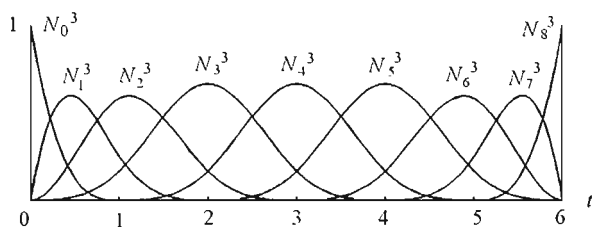


Рис. 2.12.1

Построенная по этим точкам B -кривая и ее контрольная ломаная приведены на рис. 2.12.2. Как и кривая Безье, B -кривая не проходит через свои контрольные точки за исключением крайних точек. Параметр t кривой, приведенной на рис. 2.12.2, принимает значения на отрезке $0 \leq t \leq 6$.

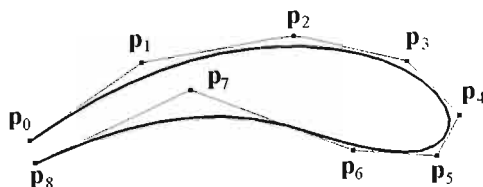


Рис. 2.12.2

Пусть узлы циклически замкнутой B -кривой принимают значения $t_i = i - m$, $i = 0, 1, 2, \dots, n + 2m + 1$. Параметр замкнутой B -кривой изменяется в пределах от $t_m = 0$ до $t_{n+m+1} = n + 1$. На рис. 2.12.3 приведен набор B -сплайнов третьего порядка, построенных на расширенном множестве узлов $-3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ для шести контрольных точек. Все B -сплайны для замкнутой кривой похожи друг на друга.

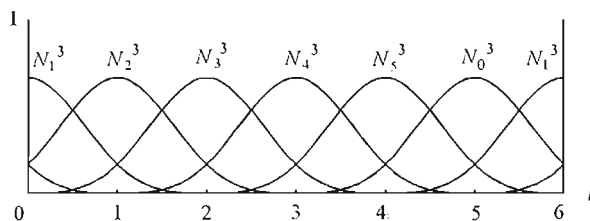


Рис. 2.12.3

Построенная по этим точкам циклически замкнутая B -кривая и ее контрольная ломаная приведены на рис. 2.12.4. Параметр кривой, приведенной на рис. 2.12.4, принимает значения на отрезке $0 \leq t \leq 6$.

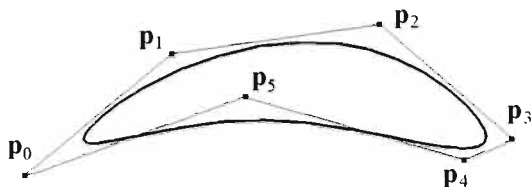


Рис. 2.12.4

На рис. 2.12.5 по одним и тем же восьми контрольным точкам построены B -кривые первого, третьего, пятого и седьмого порядков. На рис. 2.12.5 мы наблюдаем, что чем выше порядок кривой, тем она более гладкая. B -кривая первого порядка совпадает с контрольной ломаной.

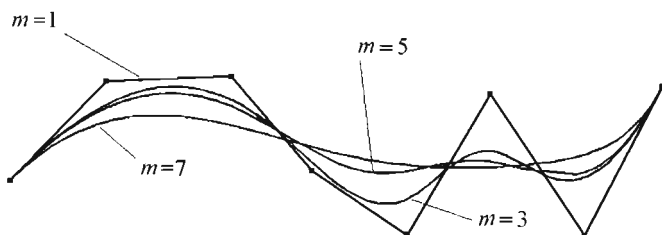


Рис. 2.12.5

На рис. 2.12.6 показано влияние веса контрольной точки p_4 на форму кривой пятого порядка (вес остальных контрольных точек равен единице). Чем больше вес вершины, тем ближе к ней проходит B -кривая.

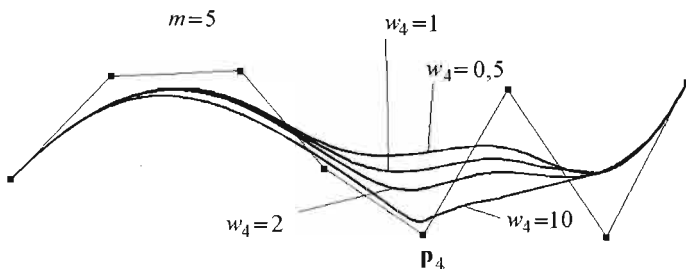


Рис. 2.12.6

В общем случае вес вершины может быть нулевым и отрицательным. На практике чаще других используются B -кривые второго и третьего порядков.

2.13. B -кривые и кривые Безье

Покажем, что кривые Безье являются частными случаями B -кривых. Для этого изменим индексацию B -сплайнов. B -сплайну мы приписывали индекс, равный индексу первого узла. С равным правом B -сплайну можно приписать индекс, равный индексу последнего узла.

Введем определение B -сплайна, в котором он имеет индекс последнего узла последовательности, на которой он построен. Построим B -кривые с помощью таких B -сплайнов и выведем для них алгоритм Де Бура.

Наряду с (2.8.1) определим B -сплайн m -го порядка, построенный на неубывающей последовательности узлов $t_{i-m-1}, t_{i-m}, \dots, t_i$, равенством

$$N_i^m(t) = (t_i - t_{i-m-1}) \sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_i]. \quad (2.13.1)$$

Кроме того, наряду с (2.8.2) определим ненормированный B -сплайн m -го порядка, построенный на неубывающей последовательности узлов $t_{i-m-1}, t_{i-m}, \dots, t_i$, равенством

$$M_i^m(t) = \sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_i]. \quad (2.13.2)$$

Ранее было отмечено, что B -сплайн нижним индексом может быть связан с любым узлом последовательности, на которой он построен. В формулах (2.13.1) и (2.13.2) индекс i B -сплайнов указывает на последний узел последовательности и расположен после индекса m , характеризующего порядок B -сплайна. Из определений следует, что $N_{i-m-1}^m(t) = N_i^m(t)$ и $M_{i-m-1}^m(t) = M_i^m(t)$.

Большинство свойств B -сплайнов при определении (2.13.1) сохраняются, так как разделенная разность не зависит от порядка следования узлов, на которой она построена. Некоторые изменения появятся в рекуррентных соотношениях, привязанных к неубывающей последовательности узлов $t_{i-m-1}, t_{i-m}, \dots, t_i$, которая используется в B -кривых.

Уменьшим в равенстве (2.8.4) все индексы на $m + 1$ и получим

$$\begin{aligned} & \sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_i] = \\ & = \frac{t_i - t}{t_i - t_{i-m-1}} \sigma_{m-1}[t_{i-m}, t_{i-m+1}, \dots, t_i] + \frac{t - t_{i-m-1}}{t_i - t_{i-m-1}} \sigma_{m-1}[t_{i-m-1}, t_{i-m}, \dots, t_{i-1}]. \end{aligned}$$

Запишем последнее равенство с использованием определения (2.13.2) и получим формулу Кокса–Де Бура для ненормированных B -сплайнов, связанных с последним узлом последовательностей, на которых они построены:

$$M_{i,j}^m(t) = \frac{t_i - t}{t_i - t_{i-m-1}} M_{i,j}^{m-1}(t) + \frac{t - t_{i-m-1}}{t_i - t_{i-m-1}} M_{i-1,j}^{m-1}(t). \quad (2.13.3)$$

Рекуррентное соотношение (2.13.3) начинается с ненормированных B -сплайнов нулевого порядка:

$$M_{i,j}^0(t) = \begin{cases} \frac{1}{t_j - t_{j-1}}, & \text{если } t_{j-1} \leq t < t_j, \\ 0, & \text{в остальных случаях.} \end{cases} \quad (2.13.4)$$

Используя равенство $N_{i,j}^m(t) = (t_i - t_{i-m-1})M_{i,j}^m(t)$, из соотношения (2.13.3) получим формулу Кокса–Де Бура для B -сплайнов, связанных с последним узлом неубывающих последовательностей, на которых они построены:

$$N_{i,j}^m(t) = \frac{t_i - t}{t_i - t_{i-m}} N_{i,j}^{m-1}(t) + \frac{t - t_{i-m-1}}{t_{i-1} - t_{i-m-1}} N_{i-1,j}^{m-1}(t). \quad (2.13.5)$$

Рекуррентное соотношение (2.13.5) начинается с B -сплайнов нулевого порядка:

$$N_{i,j}^0(t) = \begin{cases} 1, & \text{если } t_{j-1} \leq t < t_j, \\ 0, & \text{в противном случае.} \end{cases} \quad (2.13.6)$$

Используя определение разделенной разности (2.7.2), получим

$$\begin{aligned} N_{i,j}^m(t) &= (t_{i-m-1} - t_i) \sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_i] = \\ &= \sigma_m[t_{i-m}, t_{i-m+1}, \dots, t_i] - \sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_{i-1}]. \end{aligned} \quad (2.13.7)$$

С помощью (2.13.7) вычислим производную B -сплайна $N_{i,j}^m(t)$:

$$\frac{dN_{i,j}^m(t)}{dt} = \frac{d\sigma_m[t_{i-m}, t_{i-m+1}, \dots, t_i]}{dt} - \frac{d\sigma_m[t_{i-m-1}, t_{i-m}, \dots, t_{i-1}]}{dt}.$$

Подставив в последнее выражение производные по t усеченной степенной функции

$$\frac{d\sigma_m}{dt} = -m(z-t)^{m-1}, \quad z = t_{i-m-1},$$

получим

$$\begin{aligned}
\frac{dN_{i-1}^m(t)}{dt} &= -m\sigma_{m-1}[t_{i-m}, t_{i-m+1}, \dots, t_i] + m\sigma_{m-1}[t_{i-m-1}, t_{i-m}, \dots, t_{i-1}] = \\
&= -mM_{i-1}^{m-1}(t) + mM_{i-1}^{m-1}(t) = \\
&= -\frac{m}{t_i - t_{i-m}}N_{i-1}^{m-1}(t) + \frac{m}{t_{i-1} - t_{i-m-1}}N_{i-1}^{m-1}(t). \tag{2.13.8}
\end{aligned}$$

Аналогично получим формулу для производной ненормированного B -сплайна $M_i^m(t)$:

$$\frac{dM_i^m(t)}{dt} = m \frac{M_{i-1}^{m-1}(t) - M_i^{m-1}(t)}{t_i - t_{i-m-1}}.$$

Пусть задана бесконечная (или достаточно длинная) неубывающая последовательность узлов t_i и на каждых $(m + 2)$ -х подряд расположенных узлах построен B -сплайн m -го порядка. Для определения (2.13.1) при любом $t_{i-1} \leq t < t_i$ отличны от нуля только $m + 1$ B -сплайнов, а именно $N_i^m(t), N_{i+1}^m(t), \dots, N_{i+m}^m(t)$, и их сумма равна единице. Для доказательства этого свойства нужно в формуле (2.8.10) использовать (2.13.7) вместо (2.8.9).

B -сплайны $N_{i-m}^m(t)$ и $N_i^m(t)$, построенные на одной и той же последовательности узлов, совпадают, поэтому B -кривая (2.9.1) совпадет с кривой

$$\mathbf{r}(t) = \frac{\sum_{j=0}^n N_j^m(t) w_j \mathbf{p}_j}{\sum_{j=0}^n N_j^m(t) w_j}, \tag{2.13.9}$$

при уменьшении на $m + 1$ значений индексов всех узлов кривой (2.9.1). Кривую (2.13.9) отличают от кривой (2.9.1) только рекуррентные соотношения, с помощью которых вычисляются радиус-вектор кривой и его производные.

Производная k -го порядка для числителя правой части (2.13.9) равна

$$\frac{d^k(w\mathbf{r})}{dt^k} \equiv \frac{d^k}{dt^k} \left(\sum_{j=0}^n N_j^m w_j \mathbf{p}_j \right) = \sum_{j=0}^{n-k} N_j^{m-k} \mathbf{p}_j^{(k)}, \tag{2.13.10}$$

где $\mathbf{p}_j^{(k)} = (m + 1 - k) \frac{\mathbf{p}_{j+1}^{(k-1)} - \mathbf{p}_j^{(k-1)}}{t_j - t_{j-1-k}}$, $k = 1, 2, \dots, m, j = 0, 1, \dots, n - k$, $\mathbf{p}_j^{(0)} = w_j \mathbf{p}_j$.

Производная k -го порядка для знаменателя правой части (2.13.9) равна

$$\frac{d^k w}{dt^k} \equiv \frac{d^k}{dt^k} \left(\sum_{j=0}^n N_j^m w_j \right) = \sum_{j=0}^{n-k} N_j^{m-k} w_j^{(k)}, \tag{2.13.11}$$

где $w_j^{(k)} = (m+1-k) \frac{w_{j+1}^{(k-1)} - w_j^{(k-1)}}{t_j - t_{j-m+1+k}}$, $k = 1, 2, \dots, m, j = 0, 1, \dots, n-k$, $w_j^{(0)} = w_j$.

Доказательства равенств (2.13.10) и (2.13.11) аналогичны доказательствам равенств (2.9.6) и (2.9.7).

Преобразуем числитель выражения (2.13.9), используя формулу (2.13.5), следующим образом:

$$\begin{aligned} w\mathbf{r}(t) &\equiv \sum_{j=0}^n N_j^m, w_j \mathbf{p}_j = \sum_{j=0}^{i+m} N_j^m, w_j \mathbf{p}_j = \\ &= \sum_{j=0}^{i+m-1} N_j^{m-1} \left(\frac{t_j - t}{t_j - t_{j-m}} w_j \mathbf{p}_j + \frac{t - t_{j-m}}{t_j - t_{j-m}} w_{j+1} \mathbf{p}_{j+1} \right) = \sum_{j=0}^{i+m-1} N_j^{m-1} \mathbf{r}_j^{(1)} = \\ &= \sum_{j=0}^{i+m-2} N_j^{m-2} \left(\frac{t_j - t}{t_j - t_{j-m+1}} \mathbf{r}_j^{(1)} + \frac{t - t_{j-m+1}}{t_j - t_{j-m+1}} \mathbf{r}_{j+1}^{(1)} \right) = \\ &= \sum_{j=0}^{i+m-2} N_j^{m-2} \mathbf{r}_j^{(2)} = \dots = \sum_{j=0}^{i+m-m} N_j^0(t) \mathbf{r}_j^{(m)} = \mathbf{r}^{(m)}, \end{aligned}$$

где введены обозначения $\mathbf{r}_j^{(0)} = w_j \mathbf{p}_j$, $\mathbf{r}_j^{(k)} = \frac{t_j - t}{t_j - t_{j-m+1+k}} \mathbf{r}_j^{(k-1)} + \frac{t - t_{j-m+1+k}}{t_j - t_{j-m+1+k}} \mathbf{r}_{j+1}^{(k-1)}$,

при изменении k от 1 до m . Для знаменателя выражения (2.13.9) получим аналогичное выражение:

$$w(t) = \sum_{j=0}^{i+m} N_j^m w_j = \sum_{j=0}^{i+m-1} N_j^{m-1} w_j^{(1)} = \sum_{j=0}^{i+m-2} N_j^{m-2} w_j^{(2)} = \dots = w_j^{(m)},$$

где введены обозначения $w_j^{(0)} = w_j$, $w_j^{(k)} = \frac{t_j - t}{t_j - t_{j-m+1+k}} w_j^{(k-1)} + \frac{t - t_{j-m+1+k}}{t_j - t_{j-m+1+k}} w_{j+1}^{(k-1)}$,

при изменении k от 1 до m .

Мы получили алгоритм Де Бура для вычисления радиуса-вектора точки B -кривой (2.13.9) при определении B -сплайнов (2.13.1) для параметра $t_{i-1} \leq t < t_i$

$$\mathbf{r}(t) = \frac{w\mathbf{r}(t)}{w(t)} = \frac{\mathbf{r}^{(m)}}{w_j^{(m)}},$$

где $\mathbf{r}_j^{(m)}$ и $w_j^{(m)}$ вычисляются с помощью рекуррентных соотношений:

$$\begin{aligned} \mathbf{r}_j^{(k)} &= \frac{t_j - t}{t_j - t_{j-m+1+k}} \mathbf{r}_j^{(k-1)} + \frac{t - t_{j-m+1+k}}{t_j - t_{j-m+1+k}} \mathbf{r}_{j+1}^{(k-1)}, \\ w_j^{(k)} &= \frac{t_j - t}{t_j - t_{j-m+1+k}} w_j^{(k-1)} + \frac{t - t_{j-m+1+k}}{t_j - t_{j-m+1+k}} w_{j+1}^{(k-1)}, \end{aligned}$$

начиная со значений $\mathbf{r}_j^{(0)} = w_j \mathbf{p}_j$, $w_j^{(0)} = w_j$.

B -кривые (2.9.1) и (2.13.9) эквивалентны, и вся разница заключается в предпочтении использования определения B -сплайнов (2.8.1) или определения B -сплайнов (2.13.1).

Рассмотрим частный случай B -кривой, B -сплайны n -го порядка $N_i^n(t)$, $i = 0, 1, \dots, n$ которой построены на последовательности узлов $t_{-1-n} = t_{-n} = \dots = t_{-1} = 0$, $t_0 = t_1 = \dots = t_n = 1$. Рекуррентное соотношение (2.13.3) для такой последовательности узлов примет вид

$$M_i^n(t) = (1-t) M_{i-1}^{n-1}(t) + t M_{i-1}^{n-1}(t).$$

Так как каждый B -сплайн отличен от нуля на отрезке $0 \leq t \leq 1$, то B -сплайны будут связаны соотношением

$$N_i^n(t) = (1-t) N_{i-1}^{n-1}(t) + t N_{i-1}^{n-1}(t),$$

которое начинается с B -сплайна $N_0^n = 1$. Сравнив полученное рекуррентное соотношение с (2.4.4), придем к выводу, что $N_i^n(t) = B_i^n(t)$ на последовательности узлов $t_{-1-n} = t_{-n} = \dots = t_{-1} = 0$, $t_0 = t_1 = \dots = t_n = 1$. Из равенства частного случая B -сплайнов $N_i^n(t) = B_i^n(t)$ и функций Бернштейна $B_i^n(t)$ следует равенство частного случая B -кривой (2.13.9) и рациональной кривой Безье (2.6.2). Так как $N_{i-n-1}^n(t) = N_i^n(t)$, то на последовательности узлов $t_0 = t_1 = \dots = t_n = 0$, $t_{n+1} = t_{n+2} = \dots = t_{2n+1} = 1$ B -сплайны $N_i^n(t)$ будут равны функциям Бернштейна $B_i^n(t)$ и, следовательно, частный случай B -кривой (2.9.1) совпадет с рациональной кривой Безье (2.6.2). Таким образом, мы видим, что B -кривая является обобщением кривой Безье.

2.14. Частные случаи B -кривых

На последовательности узлов $t_0 = t_1 = 0$, $t_2 = t_3 = 1$ построим два B -сплайна первого порядка:

$$N_0^1(t) = \frac{t_2 - t}{t_2 - t_1} N_1^0(t) = 1 - t, \quad N_1^1(t) = \frac{t - t_1}{t_2 - t_1} N_1^0(t) = t.$$

B -кривая (2.9.2) на базе этих B -сплайнов представляет собой отрезок прямой

$$r(t) = N_0^1(t) p_0 + N_1^1(t) p_1 = (1-t) p_0 + t p_1,$$

соединяющий точки p_0 и p_1 .

Если построить B -кривую (2.9.2) по $n + 1$ вершинам на последовательности узлов $t_0 = t_1 < t_2 < \dots < t_{n+1} = t_{n+2}$ на базе B -сплайнов первого порядка, то она будет представлять собой ломаную линию. Область изменения параметра B -кривой в этом случае равна $t_1 \leq t \leq t_{n+2}$. B -сплайны первого порядка определяются равенством

$$N_i^1(t) = \begin{cases} \frac{t-t_i}{t_{i+1}-t_i}, & \text{если } t_i \leq t < t_{i+1}, \\ \frac{t_{i+2}-t}{t_{i+2}-t_{i+1}}, & \text{если } t_{i+1} \leq t < t_{i+2}, \\ 0, & \text{если } t < t_i \text{ или } t \geq t_{i+2}. \end{cases}$$

Мы видим, что $N_i^1(t)$ представляют собой кусочно-линейные функции, принимающие значения $N_i^1(t_{i+1}) = 1$ и $N_j^1(t_{i+1}) = 0, j \neq i$. B -сплайны первого порядка для незамкнутой ломаной, построенной на шести контрольных точках, приведены на рис. 2.14.1.

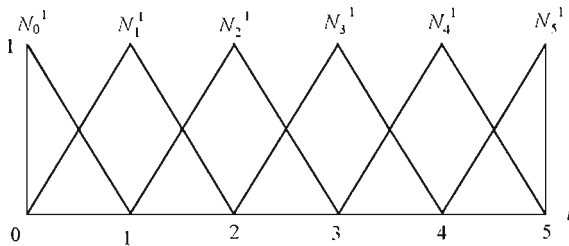


Рис. 2.14.1

Замкнутую ломаную линию как частный случай B -кривой (2.9.2) можно построить на последовательности узлов $t_0 = -1, t_1 = 0, t_2 = 1, \dots, t_i = i-1, \dots, t_{n+2} = n+1, t_{n+3} = n+2$. Параметр B -кривой в этом случае изменяется в пределах $t_1 \leq t \leq t_{n+2}$. B -сплайны первого порядка для циклически замкнутой ломаной, построенной на четырех контрольных точках, приведены на рис. 2.14.2.

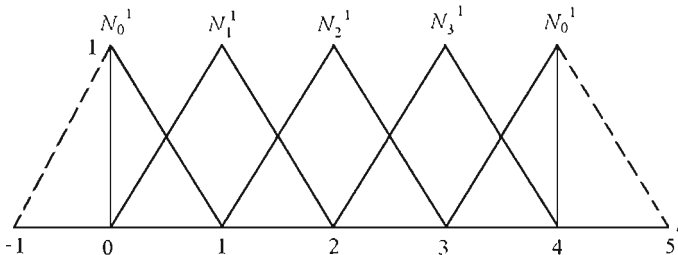


Рис. 2.14.2

$$\begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\
 a_{00} & a_{01} & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\
 0 & a_{21} & a_{22} & a_{23} & 0 & \dots & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & a_{32} & a_{33} & a_{34} & \dots & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & a_{43} & a_{44} & \dots & 0 & 0 & 0 & 0 & 0 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
 0 & 0 & 0 & 0 & 0 & \dots & a_{n-1,n-1} & a_{n-1,n} & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & \dots & a_{nn-1} & a_{nn} & a_{nn+1} & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & a_{n+1,n+1} & a_{n+1,n+2} & 0 \\
 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 1
 \end{bmatrix}
 \begin{bmatrix}
 p_0 \\
 p_1 \\
 p_2 \\
 p_3 \\
 p_4 \\
 \dots \\
 p_{n-1} \\
 p_n \\
 p_{n+1} \\
 p_{n+2}
 \end{bmatrix}
 =
 \begin{bmatrix}
 a_0 \\
 a_1 \\
 a_2 \\
 a_3 \\
 \dots \\
 a_{n-2} \\
 a_{n-1} \\
 a_n \\
 a_{n+1}
 \end{bmatrix}$$

На рис. 2.14.3 приведена B -кривая третьего порядка, совпадающая с кубическим сплайном, построенным по шести точкам ($n = 5$).

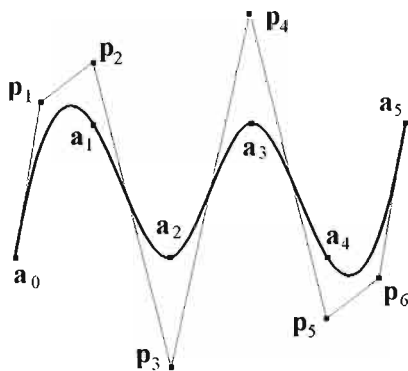


Рис. 2.14.3

На рис. 2.14.4 приведен набор B -сплайнов B -кривой третьего порядка, совпадающей с кубическим сплайном и показанной на рис. 2.14.3.

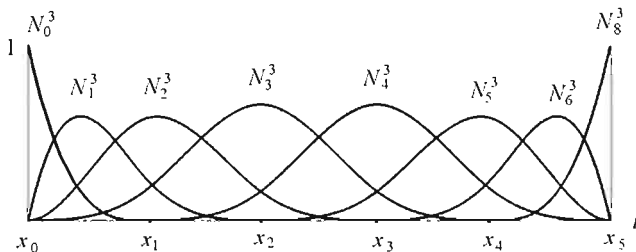


Рис. 2.14.4

Для циклически замкнутого кубического сплайна контрольные точки $\mathbf{p}_j, j = 0, 1, \dots, n$ найдем из условий $\mathbf{r}(x_i) = \mathbf{a}_i, i = 0, 1, \dots, n$. Число контрольных точек B -кривой будет равно числу заданных точек сплайна.

Рассмотрим **NURBS представление** незамкнутого составного **сплайна Эрмита** третьего порядка (2.3.10), при значениях параметра x_i проходящего через точки \mathbf{a}_i и имеющего в этих точках производные \mathbf{q}_i , где $i = 0, 1, 2, \dots, n$ – номера точек. Сплайн Эрмита представляет собой составную кривую третьей степени с разрывными производными второго и третьего порядков в точках x_i , поэтому для построения B -кривой будем использовать B -сплайны третьего порядка $N_j^3(t)$ с кратными узлами. Для совпадения начальной и конечной точек сплайна и B -кривой начальные и конечные четыре узла последовательности должны иметь кратность, равную четырем, остальные узлы последовательности должны иметь кратность, равную двум. Построим следующую последовательность узлов: $t_0 = t_1 = t_2 = t_3 = x_0, t_4 = t_5 = x_1, t_6 = t_7 = x_2, t_8 = t_9 = x_3, \dots, t_{2i+2} = t_{2i+3} = x_i, \dots, t_{2n} = t_{2n+1} = x_{n-1}, t_{2n+2} = t_{2n+3} = t_{2n+4} = t_{2n+5} = x_n$. Всего последовательность содержит $2(n+1) + 4$ узла. B -кривая будет содержать $2n + 2$ контрольных точек. Вес каждой контрольной точки положим равным единице. Таким образом, B -кривая, совпадающая со сплайном Эрмита, будет иметь вид

$$\mathbf{r}(t) = \sum_{j=0}^{2n+1} N_j^3(t) \mathbf{p}_j .$$

B -кривая, так же как и сплайн Эрмита, будет иметь область определения $x_0 \leq t \leq x_n$.

Начальная и конечная контрольные точки B -кривой равны соответственно начальной и конечной точкам сплайна: $\mathbf{p}_0 = \mathbf{a}_0, \mathbf{p}_{2n+1} = \mathbf{a}_n$. Соседние с ними контрольные точки B -кривой построим следующим образом:

$$\mathbf{p}_1 = \mathbf{a}_0 + \frac{x_1 - x_0}{3} \mathbf{q}_0, \quad \mathbf{p}_{2n} = \mathbf{a}_n - \frac{x_n - x_{n-1}}{3} \mathbf{q}_n .$$

Далее на основе каждой «внутренней» контрольной точки $\mathbf{a}_i, i = 1, 2, \dots, n-1$ сплайна Эрмита построим пару контрольных точек B -кривой следующим образом:

$$\mathbf{p}_{2i} = \mathbf{a}_i - \frac{x_i - x_{i-1}}{3} \mathbf{q}_i, \quad \mathbf{p}_{2i+1} = \mathbf{a}_i + \frac{x_{i+1} - x_i}{3} \mathbf{q}_i . \quad (2.14.3)$$

Используя формулы (2.9.3) и (2.9.4), определим значения B -сплайнов при значениях параметра $t = x_i, i = 0, 1, \dots, n$. Для параметра $t = x_0$ отличен от нуля только один B -сплайн третьего порядка $N_0^3(x_0) = 1$ и один B -сплайн второго порядка $N_1^2(x_0) = 1$. Для параметра $t = x_n$ отличен от нуля также только один B -сплайн третьего порядка $N_{2n+1}^3(x_n) = 1$ и один B -сплайн второго порядка $N_{2n+1}^2(x_n) = 1$. Из (2.14.1) следует, что для каж-

дого значения параметра $t = t_j = t_{j-1}$, $j = 5, 6, \dots, 2n + 1$ отличны от нуля только два B -сплайна третьего порядка:

$$N_{j-3}^3(t_j) = \frac{t_{j+1} - t_j}{t_{j+1} - t_{j-2}} \frac{t_{j+1} - t_j}{t_{j+1} - t_{j-1}}, \quad N_{j-2}^3(t_j) = \frac{t_{j+1} - t_j}{t_{j+1} - t_{j-2}} \frac{t_j - t_{j-2}}{t_{j+1} - t_{j-1}}.$$

Или после подстановки $t_{2i+2} = t_{2i+3} = x_i$, $i = 1, 2, \dots, n - 1$:

$$N_{2i-3}^3(x_i) = \frac{x_{i+1} - x_i}{x_{i+1} - x_{i-1}}, \quad N_{2i+1}^3(x_i) = \frac{x_i - x_{i-1}}{x_{i+1} - x_{i-1}}.$$

Из (2.14.2) следует, что для каждого значения параметра $t = t_j = t_{j-1}$, $j = 5, 6, \dots, 2n + 1$ отличен от нуля только один B -сплайн второго порядка:

$$N_{j-2}^2(t_j) = \frac{t_{j+1} - t_j}{t_{j+1} - t_{j-1}}.$$

После подстановки $t_{2i+2} = t_{2i+3} = x_i$, $i = 1, 2, \dots, n - 1$: $N_{2i+1}^2(x_i) = 1$. Непосредственной подстановкой проверим, что при значениях параметра $t = x_i$, $i = 0, 1, \dots, n$ B -кривая имеет радиус-вектор \mathbf{a}_i и производные \mathbf{q}_i . Действительно,

$$\begin{aligned} \mathbf{r}(x_0) &= N_0^3(x_0) \mathbf{p}_0 = \mathbf{a}_0, \\ \mathbf{r}(x_i) &= N_{2i}^3(x_i) \mathbf{p}_{2i} + N_{2i+1}^3(x_i) \mathbf{p}_{2i+1} = \mathbf{a}_i, \\ \mathbf{r}(x_n) &= N_{2n+1}^3(x_n) \mathbf{p}_{2n+1} = \mathbf{a}_n, \end{aligned}$$

и в соответствии с (2.9.5) и (2.14.3)

$$\begin{aligned} \frac{d\mathbf{r}}{dt}(x_0) &= 3N_1^2(x_0) \frac{\mathbf{p}_1 - \mathbf{p}_0}{t_4 - t_1} = \mathbf{q}_0, \\ \frac{d\mathbf{r}}{dt}(x_i) &= 3N_{2i+1}^2(x_i) \frac{\mathbf{p}_{2i+1} - \mathbf{p}_{2i}}{t_{2i+4} - t_{2i+1}} = \mathbf{q}_i, \\ \frac{d\mathbf{r}}{dt}(x_n) &= 3N_{2n+1}^2(x_n) \frac{\mathbf{p}_{2n+1} - \mathbf{p}_{2n}}{t_{2n+4} - t_{2n+1}} = \mathbf{q}_n. \end{aligned}$$

На рис. 2.14.5 приведена B -кривая третьего порядка, совпадающая с составным сплайном Эрмита, построенным по четырем точкам ($n=3$).

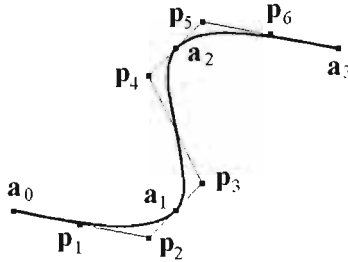


Рис. 2.14.5

На рис. 2.14.6 приведен набор B -сплайнов B -кривой третьего порядка, совпадающей с составным сплайном Эрмита и показанной на рис. 2.14.5.

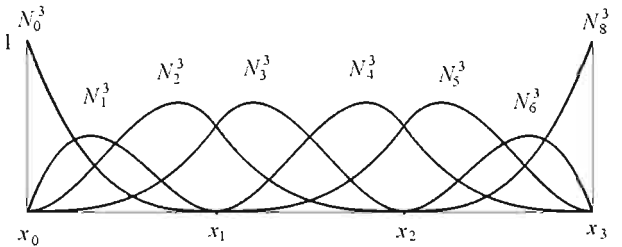


Рис. 2.14.6

Для циклически замкнутого сплайна Эрмита последовательность узлов B -кривой будет содержать только двукратные узлы. Контрольные точки найдем по формулам (2.14.3). Число контрольных точек B -кривой будет равно удвоенному числу заданных точек сплайна.

Рассмотрим **NURBS представление кривой второго порядка**. Выше мы показали, что кривые второго порядка могут быть представлены в виде рациональных кривых Безье. B -кривые являются обобщением рациональных кривых Безье, поэтому некоторая часть любой кривой второго порядка может быть представлена в виде B -кривой второго порядка, построенной по трем точкам на последовательности узлов $t_0 = t_1 = t_2 < t_3 = t_4 = t_5$.

Циклически замкнутую кривую второго порядка полностью можно представить в виде B -кривой, имеющей кратные узлы. На рис. 2.14.7 приведена эллиптическая B -кривая второго порядка и ее контрольная ломаная.

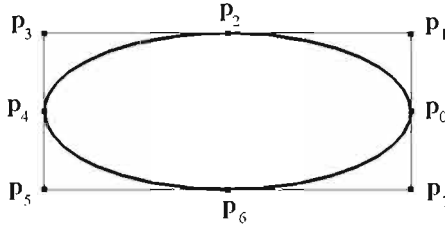


Рис. 2.14.7

Она построена по точкам $p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7$. Ее контрольная ломаная представляет собой прямоугольник, описанный около эллиптической кривой. Веса контрольных точек равны: $w_0 = w_2 = w_4 = w_6 = 1$, $w_1 = w_3 = w_5 = w_7 = \cos(\pi/4)$. Последовательность узлов имеет вид $t_0 = -1$, $t_1 = t_2 = 0$, $t_3 = t_4 = 1$, $t_5 = t_6 = 2$, $t_7 = t_8 = 3$, $t_9 = t_{10} = 4$, $t_{11} = t_{12} = 5$. Если мы хотим, чтобы параметрическая длина кривой, показанной на рис. 2.14.7, была равна 2π , то вместо указанной выше последовательности узлов можно использовать последовательность $t_0 = -\pi/2$, $t_1 = t_2 = 0$, $t_3 = t_4 = \pi/2$, $t_5 = t_6 = \pi$, $t_7 = t_8 = 3\pi/2$, $t_9 = t_{10} = 2\pi$, $t_{11} = t_{12} = 2\pi + \pi/2$. B -кривая может иметь изломы в точках с кратными узлами.

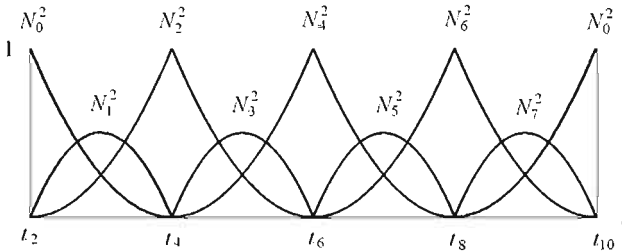


Рис. 2.14.8

На рис. 2.14.8 приведен набор B -сплайнов B -кривой второго порядка, совпадающей с эллипсом и показанной на рис. 2.14.7.

2.15. Кривые, построенные на базе кривых

Кривые могут быть построены на базе других кривых. Мы рассмотрим продолженную, усеченную, эквидистантную, репараметризованную, ссылочную кривую и кривую перехода. Кривую, на основе которой строится новая кривая, будем называть *базовой кривой*. В качестве базо-

вой кривой не должна использоваться кривая того же типа, что и создаваемая кривая.

Продолженная кривая может быть построена на базе любой кривой путем изменения области определения параметра кривой. Пусть требуется продолжить кривую $\mathbf{r}_b(t)$, $t_{\min} \leq t \leq t_{\max}$ путем расширения области определения параметров до $a + t_{\min} \leq t \leq t_{\max} + b$. При $a < 0$ и $b > 0$ кривая продлевается за свои пределы, при $a > 0$ и $b < 0$ кривая усекается.

Если кривая не является периодической, а ее параметр вышел за границу области определения, то кривую можно продолжить в соответствии с законом вычисления радиуса-вектора кривой. К таким кривым относятся аналитические кривые, а также кривые, построенные на основе кривых с известным законом вычисления радиуса-вектора. Например, радиус-вектор цилиндрической спирали за пределами области определения продолжает изменяться по закону цилиндрической спирали.

В общем случае при отсутствии закона поведения неперiodической кривой за пределами ее области определения продолжим кривую по касательной, которую она имела на соответствующем конце. Радиус-вектор продолженной кривой вычислим по формуле

$$\mathbf{r}(t) = \begin{cases} \mathbf{r}_b(t_{\min}) + (t - t_{\min}) \left. \frac{d\mathbf{r}_b}{dt} \right|_{t_{\min}}, & \text{если } t < t_{\min}, \\ \mathbf{r}_b(t), & \text{если } t_{\min} \leq t \leq t_{\max}, \\ \mathbf{r}_b(t_{\max}) + (t - t_{\max}) \left. \frac{d\mathbf{r}_b}{dt} \right|_{t_{\max}}, & \text{если } t_{\max} > t. \end{cases} \quad (2.15.1)$$

Если базовая кривая является периодической, то при выходе параметра за границу области определения выполним его циклический пересчет. В этом случае радиус-вектор продолженной кривой вычислим по формуле

$$\mathbf{r}(t) = \mathbf{r}_b(t - p \cdot \Phi((t - t_{\min})/p)), \quad (2.15.2)$$

где p – период базовой кривой; $\Phi(w)$ – функция, вычисляющая наибольшее целое число, не превосходящее w .

В качестве базовой кривой для продолженной кривой не должна использоваться другая продолженная кривая, а должна использоваться базовая кривая последней с соответствующим пересчетом параметров a и b .

По формуле (2.15.1) или (2.15.2) может вычисляться радиус-вектор не только продолженной, а любой кривой при выходе ее параметра за область определения.

Усеченная кривая представляет собой некоторую часть любой другой кривой с измененным на противоположное направлением или с сохра-

ненным направлением. Пусть параметр базовой кривой t изменяется в пределах $t_{\min} \leq t \leq t_{\max}$. Усеченную кривую определим как часть базовой кривой, начинающейся при параметре t_1 и оканчивающейся при параметре t_2 . Направление усеченной кривой может совпадать с направлением базовой кривой или быть ему противоположным, например, при $t_2 < t_1$. Если кривая циклически замкнута, то движение от точки t_1 к точке t_2 можно выполнить двумя способами: в положительном направлении базовой кривой и в противоположном направлении. Чтобы преодолеть эту неоднозначность для замкнутых кривых, вводится параметр sign , характеризующий совпадение ее направления с направлением базовой кривой и принимающий значения $+1$ или -1 . Параметру базовой кривой t_1 соответствует параметр усеченной кривой $w_{\min} = 0$, параметру базовой кривой t_2 соответствует параметр усеченной кривой $w_{\max} = s$, где s — параметрическое расстояние между t_2 и t_1 с учетом замкнутости кривой. Если кривая не замкнута, то $s = |t_2 - t_1|$. Радиус-вектор усеченной кривой описывается формулой

$$\begin{aligned} \mathbf{r}(w) &= \mathbf{r}_b(t_1 + w \text{sign}), \\ 0 &\leq w \leq s, \end{aligned} \quad (2.15.3)$$

где $\mathbf{r}_b(t)$ — базовая кривая.

В качестве базовой кривой для усеченной кривой не должна использоваться другая усеченная кривая, а должна использоваться базовая кривая последней с соответствующим пересчетом параметров усечения.

Усеченная кривая может применяться для изменения направления кривой. В таком случае $t_1 = t_{\max}$, $t_2 = t_{\min}$ и $\text{sign} = -1$.

Усеченная кривая может применяться для изменения положения начальной точки циклически замкнутой кривой. Для этого базовая кривая должна быть циклически замкнутой и $t_1 = t_2$. В этом случае усеченная кривая так же будет циклически замкнутой.

Репараметризованная кривая может быть построена на базе любой кривой путем изменения значений ее предельных параметров. Пусть требуется, чтобы кривая $\mathbf{r}_b(t)$, $t_{\min} \leq t \leq t_{\max}$, имела область определения параметра $w_{\min} \leq w \leq w_{\max}$. В этом случае построим репараметризованную кривую

$$\mathbf{r}(w) = \mathbf{r}_b(t(w)), \quad w_{\min} \leq w \leq w_{\max}, \quad (2.15.4)$$

$$\text{где } t(w) = t_{\min} \frac{w_{\max} - w}{w_{\max} - w_{\min}} + t_{\max} \frac{w - w_{\min}}{w_{\max} - w_{\min}}.$$

В качестве базовой кривой для репараметризованной кривой не должна использоваться другая репараметризованная кривая, а должна использоваться базовая кривая последней.

Репараметризованная кривая совпадает с базовой кривой, но имеет другую область определения параметра. Кривая с измененной длиной параметра применяется для согласования областей изменения параметра двух кривых.

Эквидистантная кривая описывается векторной функцией

$$\mathbf{r}(t) = \mathbf{r}_b(t) + \mathbf{a} \times \mathbf{t}_b(t), \quad (2.15.5)$$

где $\mathbf{r}_b(t)$ – базовая кривая, $\mathbf{t}_b(t) = \frac{\dot{\mathbf{r}}_b}{\sqrt{\dot{\mathbf{r}}_b \cdot \dot{\mathbf{r}}_b}}$ – единичный касательный вектор

к базовой кривой в данной точке; \mathbf{a} – заданный вектор. Область изменения параметра эквидистантной кривой может совпадать с областью изменения параметра базовой кривой, а может быть расширена в соответствии с правилами построения продолженной кривой.

Эквидистантная кривая оправдывает свое название, если $\mathbf{r}_b(t)$ – плоская кривая, а вектор \mathbf{a} ортогонален плоскости базовой кривой. В этом случае второе слагаемое в правой части (2.15.5) есть вектор, который лежит в плоскости базовой кривой, ортогонален ей и имеет длину вектора \mathbf{a} . В результате получим кривую, каждая точка которой отстоит по нормали от соответствующей точки базовой кривой на длину вектора \mathbf{a} . Пример эквидистантной кривой приведен на рис. 2.15.1.

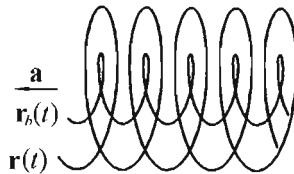


Рис. 2.15.1

В качестве базовой кривой для эквидистантной кривой не должна использоваться другая эквидистантная кривая, а должна использоваться базовая кривая последней с соответствующим пересчетом вектора эквидистанты. Например, если нужно построить эквидистантную кривую на базе эквидистантной кривой $\mathbf{r}^*(t) = \mathbf{r}_b(t) + \mathbf{a}' \times \mathbf{t}_b(t)$, то в качестве базовой кривой должна использоваться кривая $\mathbf{r}_b(t)$, а вектор эквидистанты должен быть равен $\mathbf{a} + \mathbf{a}'$.

Ссылочная кривая представляет собой кривую, каждая точка которой получена путем некоторого преобразования соответствующей точки базовой кривой. Область изменения параметра ссылочной кривой совпадает с областью изменения параметра базовой кривой. Радиус-вектор ссылочной кривой описывается формулой

$$\mathbf{r}(t) = \mathbf{M} \cdot \mathbf{r}_b(t), \quad (2.15.6)$$

где $\mathbf{r}_b(t)$ – базовая кривая, \mathbf{M} – расширенная матрица преобразования (1.4.2). При работе с расширенной матрицей мы предполагаем, что радиусы-векторы используются в виде (1.4.3), а свободные векторы – в виде (1.4.4).

В качестве базовой кривой для ссылочной кривой не должна использоваться другая ссылочная кривая, а должна использоваться базовая кривая последней с соответствующим пересчетом матрицы преобразования. Например, если требуется построить ссылочную кривую на базе ссылочной кривой $\mathbf{r}^*(t) = \mathbf{M}^* \mathbf{r}_b(t)$, то в качестве базовой кривой должна использоваться кривая $\mathbf{r}_b(t)$, а матрицу преобразования получим как произведение матриц $\mathbf{M}^* \mathbf{M}$.

Кривая перехода плавно соединяет края двух кривых. Пусть, например, требуется плавно сопрячь конец кривой $\mathbf{a}(t)$ с началом кривой $\mathbf{b}(t)$. Радиус-вектор кривой перехода описывается формулой

$$\begin{aligned} \mathbf{r}(u, w) = & \mathbf{a}(t_{\max}) (1 - 3w^2 + 2w^3) + \mathbf{b}(t_{\min}) (3w^2 - 2w^3) + \\ & + \mathbf{a}'(t_{\max}) (w - 2w^2 + w^3) k_a + \mathbf{b}'(t_{\min}) (-w^2 + w^3) k_b, \\ & w \in [0, 1], \end{aligned}$$

где $\mathbf{a}'(t_{\max}) = \left. \frac{d\mathbf{a}(t)}{dt} \right|_{t=t_{\max}}$ – производная радиуса-вектора первой сопрягаемой кривой;

$\mathbf{b}'(t_{\min}) = \left. \frac{d\mathbf{b}(t)}{dt} \right|_{t=t_{\min}}$ – производная радиуса-вектора второй

сопрягаемой кривой; k_a и k_b – коэффициенты, нормирующие производные $\mathbf{a}'(t_{\max})$ и $\mathbf{b}'(t_{\min})$ соответственно. Коэффициенты k_a и k_b получим путем деления расстояния между сопрягаемыми точками кривых на длину производных $\mathbf{a}'(t_{\max})$ и $\mathbf{b}'(t_{\min})$.

2.16. Составная кривая

Составная кривая базируется на других кривых и представляет собой наиболее общий вид кривой. Кривые, образующие составную кривую, будем называть **сегментами**. При построении составной кривой должны быть выполнены условия: начало каждого последующего сегмента должно совпадать с концом предыдущего сегмента. Замкнутую составную кривую будем называть **контуром**. Для контура начало первого сегмента должно совпадать с концом последнего сегмента. Если сегменты составной кривой стыкуются не гладко, то составная кривая будет иметь изломы. В общем случае в местах стыковки сегментов производные со-

ставной кривой терпят разрыв по длине и направлению. Составная кривая приведена на рис. 2.16.1.

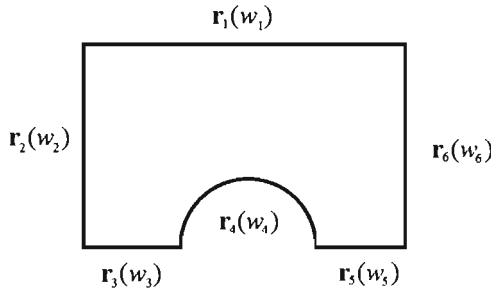


Рис. 2.16.1

Пусть составная кривая содержит n сегментов:

$$\mathbf{r}_i(w_i), w_{\min} \leq w_i \leq w_{\max}, i = 1, 2, \dots, n.$$

Начальное значение параметра t составной кривой положим равным нулю. Параметрическую длину составной кривой положим равной сумме параметрических длин составляющих ее сегментов:

$$t_{\min} = 0, t_{\max} = \sum_{i=1}^n (w_{i,\max} - w_{i,\min}).$$

При вычислении радиуса-вектора составной кривой сначала необходимо определить сегмент, которому соответствует значение параметра составной кривой. Далее нужно определить соответствующее значение собственного параметра этого сегмента и, наконец, вычислить радиус-вектор сегмента, который и будет радиусом-вектором составной кривой. Пусть для параметра t составной кривой найден номер сегмента k , для которого выполняется соотношение

$$\sum_{i=1}^{k-1} (w_{i,\max} - w_{i,\min}) \leq t < \sum_{i=1}^k (w_{i,\max} - w_{i,\min}).$$

Тогда в соответствии со сказанным радиус-вектор составной кривой определится равенством

$$\mathbf{r}(t) = \mathbf{r}_k \left(w_{k,\min} + t - \sum_{i=1}^{k-1} (w_{i,\max} - w_{i,\min}) \right) = \mathbf{r}_k(w_k), \quad (2.16.1)$$

$$0 \leq t \leq t_{\max},$$

где параметр k -го сегмента равен

$$w_k = w_{kmin} + t - \sum_{i=1}^{k-1} (w_{imax} - w_{imin}). \quad (2.16.2)$$

Как можно заметить, подход к вычислению радиуса-вектора составной кривой аналогичен подходу к вычислению радиуса-вектора ломаной линии и сплайна Эрмита.

В качестве сегментов составной кривой не должны использоваться другие составные кривые, но это не означает, что нельзя построить составную кривую из набора составных кривых. Если составную кривую нужно построить на основе других составных кривых, то последние должны рассматриваться как совокупность составляющих их кривых, а не как единые кривые.

Упражнения

1. Вычислите максимальную и минимальную кривизну эллипса.
2. Опишите векторной функцией кардиоиду.
3. Выразите первую производную кривой Безье на краях через радиусы-векторы крайних точек.
4. Выразите параболу в виде B -кривой.

Глава 3. Поверхности

Поверхности являются основным элементом описания формы моделируемых объектов. В общем случае поверхность имеет сложную границу, которая описывается с помощью двумерных кривых в пространстве параметров поверхности. Такая граница может быть построена для любой поверхности с областью определения простой формы, например, прямоугольной. В данной главе рассмотрены методы построения поверхностей. Поверхности могут быть построены с помощью аналитических функций, на основе кривых, по набору точек и на базе других поверхностей.

3.1. Поверхность

Поверхностью будем называть векторную функцию

$$\mathbf{r}(u, v) = \begin{bmatrix} r_1(u, v) \\ r_2(u, v) \\ r_3(u, v) \end{bmatrix} \quad (3.1.1)$$

двух скалярных параметров u и v , принимающих значения на двумерной связной области Ω . Будем считать, что координаты $r_1(u, v)$, $r_2(u, v)$, $r_3(u, v)$ точки поверхности $\mathbf{r}(u, v)$ являются однозначными непрерывными функциями параметров u и v . Такое описание поверхности называется *параметрическим*. Описание поверхности уравнением, которому удовлетворяют координаты ее точек, не является инвариантным относительно преобразования координат, то есть при переходе в другую систему координат изменяется уравнение поверхности.

Поверхность представляет собой непрерывное отображение двумерной связной области Ω в трехмерное пространство. Область Ω будем описывать в двумерной декартовой системе координат. В частном случае область Ω представляет собой прямоугольник, и параметры поверхности принимают значения в пределах $u_{\min} \leq u \leq u_{\max}$, $v_{\min} \leq v \leq v_{\max}$. В общем случае область Ω описывается двумерными кривыми.

Поверхность будем называть *периодической по первому параметру*, если существует $p_u > 0$ такое, что $\mathbf{r}(u \pm kp_u, v) = \mathbf{r}(u, v)$, где k – целое число. Поверхность будем называть *периодической по второму параметру*, если существует $p_v > 0$ такое, что $\mathbf{r}(u, v \pm kp_v) = \mathbf{r}(u, v)$, где k – целое число. Для устранения неоднозначности область определения периодической поверхности должна лежать в пределах одного периода для каждого параметра. Периодическую поверхность с прямоугольной областью опреде-

ления параметров Ω будем называть *циклически замкнутой по первому параметру*, если $p_u = u_{\max} - u_{\min}$, и, соответственно, *циклически замкнутой по второму параметру*, если $p_v = v_{\max} - v_{\min}$.

Введем обозначения

$$\mathbf{r}_1 = \frac{\partial \mathbf{r}}{\partial u}, \quad \mathbf{r}_2 = \frac{\partial \mathbf{r}}{\partial v}, \quad \mathbf{r}_{11} = \frac{\partial^2 \mathbf{r}}{\partial u^2}, \quad \mathbf{r}_{22} = \frac{\partial^2 \mathbf{r}}{\partial v^2}, \quad \mathbf{r}_{12} = \frac{\partial^2 \mathbf{r}}{\partial u \partial v} = \mathbf{r}_{21} = \frac{\partial^2 \mathbf{r}}{\partial v \partial u} \quad (3.1.2)$$

для частных производных поверхности по ее параметрам. Векторы частных производных обозначим численными нижними индексами, соответствующими номеру параметра, по которому выполнено дифференцирование. Далее будем предполагать, что координатные функции $r_1(u, v)$, $r_2(u, v)$, $r_3(u, v)$ имеют непрерывные производные по каждому параметру любого порядка, который нам потребуется.

Точку поверхности будем называть *обыкновенной*, если в этой точке не обращаются в нуль длины частных производных поверхности по обоим параметрам и частные производные поверхности не параллельны. В противном случае точку поверхности будем называть *особой*.

В обыкновенной точке по векторам \mathbf{r}_1 и \mathbf{r}_2 можно построить плоскость. Плоскость, проходящая через точку поверхности и параллельная векторам \mathbf{r}_1 и \mathbf{r}_2 в этой точке, называется *касательной плоскостью* поверхности.

Если зафиксировать один из параметров, а другой изменять в некоторых пределах, то мы получим кривую, которая лежит на поверхности. Такие кривые называют *координатными линиями поверхности*. Будем называть *u-линиями* кривые, вдоль которых меняется только параметр u , а *v-линиями* будем называть кривые, вдоль которых меняется только параметр v . Производные \mathbf{r}_1 и \mathbf{r}_2 поверхности представляют собой векторы, касательные к соответствующим координатным линиям. Произвольную кривую на поверхности можно построить, если ввести зависимость параметров поверхности u и v от некоторого общего для них параметра, например, $u = u(t)$, $v = v(t)$.

Кривая на поверхности описывается зависимостью

$$\mathbf{r}(t) = \begin{bmatrix} r_1(u(t), v(t)) \\ r_2(u(t), v(t)) \\ r_3(u(t), v(t)) \end{bmatrix}. \quad (3.1.3)$$

Производная кривой на поверхности

$$\frac{d\mathbf{r}}{dt} = \mathbf{r}_1 \frac{du}{dt} + \mathbf{r}_2 \frac{dv}{dt}$$

лежит в касательной плоскости, построенной в рассматриваемой точке. Параметры поверхности u и v являются координатами двумерной точки $\mathbf{r}[u \ v]^T$ в некоторой выбранной двумерной декартовой системе координат. Таким образом, кривая на поверхности представляет собой совокупность поверхности $\mathbf{r}(u, v)$ и двумерной кривой (2.1.5).

Двумерные кривые могут быть построены с помощью аналитических функций, по набору точек, на базе других двумерных кривых. Двумерные кривые будем строить методами построения трехмерных кривых с той разницей, что вместо трехмерных объектов будем использовать двумерные точки, векторы и базовые кривые. Большинство методов построения трехмерных кривых, описанных в предыдущей главе, применимы для построения двумерных кривых.

Метрические свойства поверхности выражаются через метрические свойства кривых на ней. Исследуем метрические свойства поверхности в малой окрестности некоторой ее точки, определяемой параметрами u и v . Сместимся из рассматриваемой точки по некоторой кривой на поверхности в бесконечно близкую ей точку, определяемую параметрами $u + du$, $v + dv$ и вычислим длину дуги. С точностью до слагаемых, линейно зависящих от бесконечно малых величин, длина дуги равна

$$ds = |\mathbf{r}_1 du + \mathbf{r}_2 dv|.$$

Квадрат длины бесконечно малой дуги равен

$$ds^2 = \mathbf{r}_1' \mathbf{r}_1 du^2 + 2\mathbf{r}_1' \mathbf{r}_2 du dv + \mathbf{r}_2' \mathbf{r}_2 dv^2.$$

Введем обозначения

$$g_{11} = \mathbf{r}_1' \mathbf{r}_1, g_{12} = g_{21} = \mathbf{r}_1' \mathbf{r}_2 = \mathbf{r}_2' \mathbf{r}_1, g_{22} = \mathbf{r}_2' \mathbf{r}_2. \quad (3.1.4)$$

Тогда квадрат длины бесконечно малой дуги кривой на поверхности будет определяться формулой

$$ds^2 = g_{11} du^2 + 2g_{12} dudv + g_{22} dv^2.$$

Выражение в правой части является квадратичной формой дифференциалов du и dv и называется *первой квадратичной формой поверхности*. Величины $g_{11}(u, v)$, $g_{12}(u, v)$, $g_{21}(u, v)$, $g_{22}(u, v)$ определяют метрические свойства поверхности и называются коэффициентами первой квадратичной формы поверхности.

С помощью первой квадратичной формы можно вычислить **длину дуги кривой** на поверхности. Пусть задан участок кривой на поверхности $u = u(t)$, $v = v(t)$, $t_1 \leq t \leq t_2$. В пределе сумма бесконечно малых длин дуг ds даст длину соответствующего участка кривой:

$$s = \int_{t_1}^{t_2} \sqrt{g_{11} \left(\frac{du}{dt} \right)^2 + 2g_{12} \frac{du}{dt} \frac{dv}{dt} + g_{22} \left(\frac{dv}{dt} \right)^2} dt.$$

Первая квадратичная форма поверхности позволяет вычислять **углы между кривыми на поверхности**. Пусть имеются две кривые на поверхности, проходящие через общую точку, определяемую параметрами u и v . Обозначим через du и dv дифференциалы параметров поверхности, соответствующие бесконечно малому смещению вдоль первой кривой на поверхности, а через δu и δv – дифференциалы параметров поверхности, соответствующие бесконечно малому смещению вдоль второй кривой на поверхности. Эти бесконечно малые смещения определяются векторами

$$d\mathbf{r} = \mathbf{r}_1 du + \mathbf{r}_2 dv, \quad \delta\mathbf{r} = \mathbf{r}_1 \delta u + \mathbf{r}_2 \delta v.$$

Найдем косинус угла φ между кривыми на поверхности как скалярное произведение касательных к ним векторов, деленное на произведение длин этих векторов:

$$\begin{aligned} \cos\varphi &= \frac{d\mathbf{r} \cdot \delta\mathbf{r}}{|d\mathbf{r}| |\delta\mathbf{r}|} = \\ &= \frac{g_{11} du \delta u + g_{12} du \delta v + g_{21} dv \delta u + g_{22} dv \delta v}{\sqrt{(g_{11} du du + 2g_{12} du dv + g_{22} dv dv)(g_{11} \delta u \delta u + 2g_{12} \delta u \delta v + g_{22} \delta v \delta v)}}. \end{aligned}$$

Это выражение позволяет найти угол между координатной u -линией и v -линией в рассматриваемой точке поверхности, если мы положим в ней $du \neq 0, dv = 0, \delta u = 0, \delta v \neq 0$:

$$\cos\varphi = \frac{g_{12}}{\sqrt{g_{11} g_{22}}}.$$

Если $g_{12} = 0$, то координатные линии в рассматриваемой точке ортогональны.

Первая квадратичная форма используется и для вычисления **площади поверхности**. Рассмотрим в точке, определяемой параметрами u и v , бесконечно малый криволинейный четырехугольник на поверхности со сторонами du и dv . Площадь четырехугольника в первом приближении равна

$$dS = |\mathbf{r}_1 du \times \mathbf{r}_2 dv| = |\mathbf{r}_1 \times \mathbf{r}_2| du dv.$$

Представим квадрат длины вектора $\mathbf{r}_1 \times \mathbf{r}_2$ следующим образом:

$$\begin{aligned} |\mathbf{r}_1 \times \mathbf{r}_2|^2 &= |\mathbf{r}_1|^2 |\mathbf{r}_2|^2 \sin^2 \varphi = |\mathbf{r}_1|^2 |\mathbf{r}_2|^2 (1 - \cos^2 \varphi) = \\ &= g_{11} g_{22} \left(1 - \frac{g_{12}^2}{g_{11} g_{22}} \right) = g_{11} g_{22} - g_{12}^2. \end{aligned}$$

Таким образом, площадь бесконечно малого криволинейного четырехугольника в первом приближении определится формулой

$$dS = \sqrt{g_{11}g_{22} - g_{12}^2} dudv .$$

Тогда площади поверхности, параметры которой принимают значения на двумерной связной области Ω , вычислим с помощью интеграла

$$S = \iint_{\Omega} \sqrt{g_{11}g_{22} - g_{12}^2} dudv .$$

Геометрические свойства поверхности, которые можно установить с помощью первой квадратичной формы, называются *внутренней геометрией поверхности*.

В обыкновенной точке **нормаль** поверхности определяется через векторное произведение $\mathbf{r}_1 \times \mathbf{r}_2$. Так как векторы \mathbf{r}_1 и \mathbf{r}_2 лежат в касательной к поверхности плоскости, то их векторное произведение ортогонально касательной плоскости. Поделив вектор $\mathbf{r}_1 \times \mathbf{r}_2$ на его длину, получим формулу для единичной нормали поверхности в рассматриваемой точке:

$$\mathbf{m} = \frac{\mathbf{r}_1 \times \mathbf{r}_2}{\sqrt{g_{11}g_{22} - g_{12}^2}} .$$

Рассмотрим некоторую кривую на поверхности, определяемую функциями $u = u(t)$, $v = v(t)$. Вычислим приращение радиуса-вектора кривой на поверхности, которое он получит при бесконечно малом приращении ее параметра dt с точностью до второго порядка малости относительно dt :

$$\Delta \mathbf{r} = \frac{d\mathbf{r}}{dt} dt + \frac{1}{2} \frac{d^2\mathbf{r}}{dt^2} dt^2 .$$

Найдем проекцию $\Delta \mathbf{r}$ на нормаль поверхности в рассматриваемой точке. Первая производная кривой ортогональна нормали поверхности, поэтому основную роль в проекции $\Delta \mathbf{r}$ на нормаль будет играть вторая производная кривой:

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{r}_{11} \left(\frac{du}{dt} \right)^2 + 2\mathbf{r}_{12} \frac{du}{dt} \frac{dv}{dt} + \mathbf{r}_{22} \left(\frac{dv}{dt} \right)^2 + \mathbf{r}_1 \frac{d^2u}{dt^2} + \mathbf{r}_2 \frac{d^2v}{dt^2} .$$

Проекция второй производной кривой на нормаль поверхности характеризует искривление поверхности, именно поверхности, а не кривой. Действительно, если кривая построена на плоскости, то как бы она искривлена ни была, проекция вектора $\Delta \mathbf{r}$ на нормаль плоскости будет равна нулю. Умножив скалярно $\Delta \mathbf{r}$ на нормаль поверхности, получим

$$\Delta \mathbf{r} \cdot \mathbf{m} = \frac{1}{2} \frac{d^2\mathbf{r}}{dt^2} \cdot \mathbf{m} dt^2 = \frac{1}{2} \left(\mathbf{r}_{11} \cdot \mathbf{m} \left(\frac{du}{dt} \right)^2 + 2\mathbf{r}_{12} \cdot \mathbf{m} \frac{du}{dt} \frac{dv}{dt} + \mathbf{r}_{22} \cdot \mathbf{m} \left(\frac{dv}{dt} \right)^2 \right) dt^2 .$$

Введем обозначения

$$b_{11} = \mathbf{r}_{11} \cdot \mathbf{m}, \quad b_{12} = b_{21} = \mathbf{r}_{12} \cdot \mathbf{m} = \mathbf{r}_{21} \cdot \mathbf{m}, \quad b_{22} = \mathbf{r}_{22} \cdot \mathbf{m}. \quad (3.1.5)$$

Тогда для главной части отклонения кривой на поверхности от касательной плоскости получим значение

$$\Delta \mathbf{r} \cdot \mathbf{m} = \frac{1}{2} (b_{11} du^2 + 2b_{12} dudv + b_{22} dv^2).$$

Выражение в скобках правой части данного равенства является квадратичной формой дифференциалов du и dv . Оно называется *второй квадратичной формой поверхности*. Величины $b_{11}(u, v)$, $b_{12}(u, v)$, $b_{21}(u, v)$, $b_{22}(u, v)$ определяют искривленность поверхности и называются коэффициентами второй квадратичной формы поверхности.

Коэффициенты второй квадратичной формы можно выразить несколько иначе. Используем тот факт, что вектор нормали \mathbf{m} всегда ортогонален векторам \mathbf{r}_1 и \mathbf{r}_2 . Дифференцируя равенства $\mathbf{m} \cdot \mathbf{r}_1 = 0$ и $\mathbf{m} \cdot \mathbf{r}_2 = 0$ по u и v , получим

$$b_{11} = -\mathbf{m}_1 \cdot \mathbf{r}_1, \quad b_{12} = -\mathbf{m}_2 \cdot \mathbf{r}_1, \quad b_{21} = -\mathbf{m}_1 \cdot \mathbf{r}_2, \quad b_{22} = -\mathbf{m}_2 \cdot \mathbf{r}_2,$$

где $\mathbf{m}_1 = \frac{\partial \mathbf{m}}{\partial u}$ и $\mathbf{m}_2 = \frac{\partial \mathbf{m}}{\partial v}$ — частные производные нормали поверхности по параметрам поверхности. Перемножив скалярно дифференциалы $d\mathbf{r} = \mathbf{r}_1 du + \mathbf{r}_2 dv$, $d\mathbf{m} = \mathbf{m}_1 du + \mathbf{m}_2 dv$, придем к равенству

$$-d\mathbf{r} \cdot d\mathbf{m} = b_{11} du^2 + 2b_{12} du dv + b_{22} dv^2. \quad (3.1.6)$$

Вектор \mathbf{m} имеет единичную длину. Дифференцируя по параметрам поверхности равенство $\mathbf{m} \cdot \mathbf{m} = 1$, получим

$$\mathbf{m}_1 \cdot \mathbf{m} = 0, \quad \mathbf{m}_2 \cdot \mathbf{m} = 0.$$

Из приведенных равенств следует, что производные нормали по параметрам поверхности лежат в касательной плоскости поверхности, то есть \mathbf{m}_1 и \mathbf{m}_2 можно представить в виде разложений по векторам \mathbf{r}_1 и \mathbf{r}_2 .
Деривационные формулы Вейнгартена

$$\mathbf{m}_1 = -\frac{b_{11}g_{22} - b_{12}g_{12}}{g} \mathbf{r}_1 - \frac{b_{12}g_{11} - b_{11}g_{21}}{g} \mathbf{r}_2,$$

$$\mathbf{m}_2 = -\frac{b_{21}g_{22} - b_{22}g_{12}}{g} \mathbf{r}_1 - \frac{b_{22}g_{11} - b_{21}g_{21}}{g} \mathbf{r}_2$$

выражают частные производные нормали поверхности через производные поверхности и коэффициенты первой и второй квадратичных форм. Коэффициенты деривационных формул Вейнгартена можно определить

из системы уравнений, полученной путем умножения формальных равенств $\mathbf{m}_1 = a_{11}\mathbf{r}_1 + a_{12}\mathbf{r}_2$ и $\mathbf{m}_2 = a_{21}\mathbf{r}_1 + a_{22}\mathbf{r}_2$ скалярно на \mathbf{r}_1 и \mathbf{r}_2 .

Установим зависимость кривизны кривой на поверхности от ориентации ее касательного вектора на соприкасающейся плоскости. Рассмотрим равенство

$$\mathbf{m} \cdot \frac{d^2\mathbf{r}}{dt^2} = b_{11} \left(\frac{du}{dt} \right)^2 + 2b_{12} \frac{du}{dt} \frac{dv}{dt} + b_{22} \left(\frac{dv}{dt} \right)^2.$$

Вторая производная кривой на поверхности равна

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{d^2s}{dt^2} \mathbf{t} + \left(\frac{ds}{dt} \right)^2 k \mathbf{n},$$

где \mathbf{t} – касательный вектор кривой на поверхности; \mathbf{n} – главная нормаль кривой на поверхности; k – кривизна кривой на поверхности; s – длина дуги кривой на поверхности. Касательный вектор \mathbf{t} кривой лежит в касательной плоскости поверхности и ортогонален нормали поверхности \mathbf{m} , поэтому справедливо равенство

$$ds^2 k \mathbf{n} \cdot \mathbf{m} = b_{11} du^2 + 2b_{12} dudv + b_{22} dv^2.$$

Подставив квадрат дифференциала длины дуги кривой, выраженный с помощью первой квадратичной формы, получим

$$k \mathbf{n} \cdot \mathbf{m} = \frac{b_{11} du^2 + 2b_{12} dudv + b_{22} dv^2}{g_{11} du^2 + 2g_{12} dudv + g_{22} dv^2}. \quad (3.1.7)$$

Геометрический смысл последнего соотношения будет более понятен, если мы рассмотрим *нормальное сечение поверхности* – кривую пересечения поверхности и плоскости, проходящей через нормаль поверхности и касающейся кривой на поверхности.

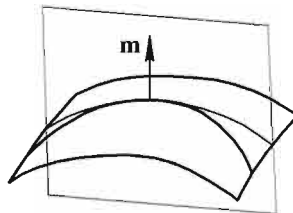


Рис. 3.1.1

На рис. 3.1.1 показана некоторая кривая на поверхности и соответствующее ей нормальное сечение поверхности, проходящее через рассмат-

риваемую точку кривой. Правая часть равенства (3.1.7) зависит только от положения рассматриваемой точки и направления кривой на поверхности, определяемого отношением du к dv . Следовательно, любая кривая на поверхности, проходящая через рассматриваемую точку и имеющая общую касательную с рассматриваемым нормальным сечением, будет иметь одно и то же значение $k\mathbf{n}\cdot\mathbf{m}$, несмотря на то, что у нее другая кривизна. Нормальное сечение лежит как на поверхности, так и на секущей плоскости, поэтому его главная нормаль также лежит в секущей плоскости и, следовательно, для нормального сечения $|\mathbf{n}\cdot\mathbf{m}| = 1$. Поэтому из всех кривых на поверхности, проходящих через рассматриваемую точку и имеющих в ней общую касательную, нормальное сечение имеет наименьшую кривизну, равную

$$\mu = \frac{b_{11}du^2 + 2b_{12}dudv + b_{22}dv^2}{g_{11}du^2 + 2g_{12}dudv + g_{22}dv^2}. \quad (3.1.8)$$

Кривизна нормального сечения называется *нормальной кривизной* поверхности в заданной точке в заданном направлении.

Угол между нормалью поверхности и главной нормалью кривой на поверхности равен углу между плоскостью нормального сечения и соприкасающейся плоскостью кривой. Если известна кривизна нормального сечения, то можно определить кривизну кривой на поверхности, касательной к этому нормальному сечению, при условии, что известен угол между нормалью поверхности и главной нормалью кривой.

Теорема Менье констатирует описанный факт и гласит: *радиус кривизны $\rho = 1/k$ в заданной точке кривой на поверхности равен произведению радиуса кривизны $\rho_n = 1/\mu$ соответствующего нормального сечения в этой точке на косинус угла между нормалью поверхности и главной нормалью кривой:*

$$\frac{1}{k} = \frac{1}{\mu} \mathbf{n} \cdot \mathbf{m}.$$

Если нормальное сечение касательно к координатной u -линии, то $dv = 0$ и

$$\mu_u = \frac{b_{11}}{g_{11}},$$

где через μ_u обозначена нормальная кривизна поверхности в u -направлении. Аналогично нормальная кривизна поверхности μ_v в v -направлении равна

$$\mu_v = \frac{b_{22}}{g_{22}}.$$

В заданной точке поверхности можно построить множество нормальных сечений, которые отличаются направлением, определяемым отношением du к dv . Направление нормального сечения, кривизна которого равна нулю, называется *асимптотическим направлением* в рассматриваемой точке. В каждой точке поверхности существует не более двух асимптотических направлений, если не считать те случаи, когда в точке все коэффициенты второй квадратичной формы равны нулю.

Мы рассмотрели проекцию вектора кривизны $k\mathbf{n}$ кривой на поверхность на нормаль поверхности \mathbf{m} . Теперь рассмотрим оставшуюся часть вектора кривизны — его проекцию на касательную плоскость, равную

$$\mathbf{h} = k\mathbf{n} - \mathbf{m}(k\mathbf{n} \cdot \mathbf{m}).$$

Длина вектора \mathbf{h} называется *геодезической кривизной* кривой на поверхности. Геодезическая кривизна нормального сечения равна нулю. Если построить ортогональную проекцию на касательную плоскость кривой на поверхности, то кривизна этой проекции в рассматриваемой точке будет равна длине вектора \mathbf{h} . Нормальная кривизна является характеристикой поверхности, а геодезическая кривизна является характеристикой кривой на ней.

Главные кривизны поверхности определяют направления, характеризующиеся взаимным поведением векторов $d\mathbf{r}$ и $d\mathbf{m}$. Из равенств (3.1.6) и (3.1.8) следует, что для векторов $d\mathbf{r}$ и $d\mathbf{m}$ справедливо равенство

$$\mu(d\mathbf{r} \cdot d\mathbf{r}) = -(d\mathbf{r} \cdot d\mathbf{m}). \quad (3.1.9)$$

В заданной точке поверхности кривизна нормального сечения зависит от направлений векторов $d\mathbf{r}$ и $d\mathbf{m}$ относительно друг друга. Заданная целью найти такое направление движения по поверхности, при котором векторы $d\mathbf{r}$ и $d\mathbf{m}$ были бы коллинеарными, то есть для которого выполняется равенство $d\mathbf{m} = -\lambda d\mathbf{r}$. Из равенств (3.1.8) и (3.1.9) следует, что для коллинеарности векторов $d\mathbf{r}$ и $d\mathbf{m}$ должно выполняться равенство

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \cdot \begin{bmatrix} du \\ dv \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \cdot \begin{bmatrix} du \\ dv \end{bmatrix}. \quad (3.1.10)$$

Итак, для определения искомого направления мы пришли к системе линейных алгебраических уравнений для du и dv :

$$\left(\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} - \frac{1}{\lambda} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right) \cdot \begin{bmatrix} du \\ dv \end{bmatrix} = 0.$$

Данная система является однородной и имеет ненулевое решение, если определитель ее матрицы равен нулю:

$$(g_{11}g_{22} - g_{12}^2)\lambda^2 - (b_{11}g_{22} + b_{22}g_{11} - 2b_{12}g_{12})\lambda + (b_{11}b_{22} - b_{12}^2) = 0. \quad (3.1.11)$$

Раскрыв определитель, приходим к квадратному уравнению относительно λ , откуда в общем случае найдем два корня: λ_1 и λ_2 . Подставив каждый из корней в любое из уравнений (3.1.10), получим два направления на поверхности, определяемые отношениями $du_1 : dv_1$ и $du_2 : dv_2$.

Направления движения на поверхности, для которых векторы $d\mathbf{r}$ и $d\mathbf{m}$ коллинеарны, называются *главными направлениями поверхности*. Нормальные сечения в данной точке поверхности, касательные к которым идут по главным направлениям, называются *главными сечениями*, а их кривизны называются *главными кривизнами* в данной точке поверхности. Из (3.1.11) легко получить сумму и произведение главных кривизн:

$$\begin{aligned} \lambda_1 + \lambda_2 &= \frac{b_{11}g_{22} + b_{22}g_{11} - 2b_{12}g_{12}}{g}, \\ \lambda_1\lambda_2 &= \frac{b_{11}b_{22} - b_{12}b_{21}}{g}. \end{aligned} \quad (3.1.12)$$

Полусумма главных кривизн называется *средней кривизной поверхности* в данной точке, а произведение главных кривизн называется *гауссовой кривизной поверхности* в данной точке.

Обозначим через \mathbf{t}_1 и \mathbf{t}_2 касательные векторы нормальных сечений вдоль главных направлений. Покажем, что главные направления поверхности ортогональны друг другу. Выразим главные направления через производные радиуса-вектора:

$$\begin{aligned} \mathbf{t}_1 &= \mathbf{r}_1 du_1 + \mathbf{r}_2 dv_1, \\ \mathbf{t}_2 &= \mathbf{r}_1 du_2 + \mathbf{r}_2 dv_2. \end{aligned}$$

Их скалярное произведение равно

$$\mathbf{t}_1 \cdot \mathbf{t}_2 = [du_1 \ dv_1] \cdot \mathbf{G} \cdot [du_2 \ dv_2]^T, \quad (3.1.13)$$

где $\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$.

Покажем, что оно в общем случае равно нулю. Для этого систему двух уравнений (3.1.10) запишем для первого главного направления, первое из этих уравнений умножим на du_2 , второе уравнение умножим на dv_2 и сложим с первым, в результате получим равенство

$$[du_2 \ dv_2] \cdot (\mathbf{B} - \lambda_1 \mathbf{G}) \cdot [du_1 \ dv_1]^T = 0,$$

где $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$.

Аналогично получим второе равенство, поменяв местами главные направления:

$$[du_1 \, dv_1] \cdot (\mathbf{B} - \lambda_2 \mathbf{G}) \cdot [du_2 \, dv_2]^T = 0.$$

Вычтем последние два равенства одно из другого и получим равенство

$$[du_1 \, dv_1] \cdot ((\lambda_2 - \lambda_1) \mathbf{G}) \cdot [du_2 \, dv_2]^T = 0,$$

из которого следует, что при $\lambda_1 \neq \lambda_2$ главные направления ортогональны. Если главные кривизны поверхности равны, то за главные могут быть выбраны любые два ортогональных направления (такую ситуацию мы имеем на сфере и плоскости). Точка, в которой $\lambda_2 = \lambda_1$, называется точкой закружения.

Так как главные направления в общем случае ортогональны, то производные радиуса-вектора поверхности и ее нормали в любом направлении можно разложить по единичным векторам \mathbf{t}_1 и \mathbf{t}_2 главных направлений:

$$\begin{aligned} \frac{d\mathbf{r}}{ds} &= \mathbf{t}_1 \cos\varphi + \mathbf{t}_2 \sin\varphi, \\ \frac{d\mathbf{m}}{ds} &= -\lambda_1 \mathbf{t}_1 \cos\varphi - \lambda_2 \mathbf{t}_2 \sin\varphi, \end{aligned}$$

где угол φ отсчитывается в касательной плоскости от первого главного направления ко второму. Кривизна нормального сечения в произвольно выбранном направлении с учетом последних равенств определится равенством

$$\mu = -\frac{d\mathbf{r}}{ds} \cdot \frac{d\mathbf{m}}{ds} = \lambda_1 \cos^2\varphi + \lambda_2 \sin^2\varphi. \quad (3.1.14)$$

Формула (3.1.14) называется *формулой Эйлера*. Она выражает кривизну произвольного нормального сечения в точке через главные кривизны и угол между нормальным сечением и первым главным направлением. Из этого равенства мы видим, что главные кривизны поверхности λ_1 и λ_2 являются максимальной кривизной и минимальной кривизной соответственно.

Гауссова кривизна поверхности (3.1.12) может быть использована для определения поведения поверхности в некоторой ее точке. Так как знаменатель в (3.1.12) больше нуля, то знак гауссовой кривизны зависит от знака числителя, то есть от знака определителя матрицы \mathbf{B} . Если $|\mathbf{B}| > 0$, то точка называется эллиптической. Поведение поверхности в эллиптической точке приведено на рис. 3.1.2. Если $|\mathbf{B}| < 0$, то точка называется гиперболической. Поведение поверхности в гиперболической точке показано на рис. 3.1.3.

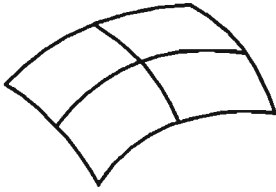


Рис. 3.1.2

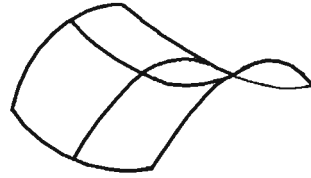


Рис. 3.1.3

При движении от рассматриваемой точки в любом направлении поверхность изгибается или в сторону нормали, или в противоположную сторону в зависимости от знаков λ_1 и λ_2 . Согласно (3.1.14), существуют такие нормальные сечения, для которых выполняется равенство

$$\mu_1 \cos^2 \varphi + \mu_2 \sin^2 \varphi = 0.$$

Касательные к нормальным сечениям под углами

$$\varphi = \pm \operatorname{arctg} \left(\sqrt{\frac{-\mu_1}{\mu_2}} \right)$$

расположены в касательной плоскости симметрично относительно главных направлений и определяют асимптотические направления в точке. Если в рассматриваемой точке $|\mathbf{B}| = 0$, то такая точка называется параболической. Поведение поверхности в параболической точке приведено на рис. 3.1.4.

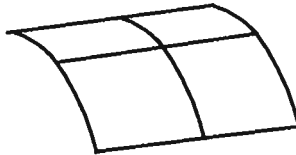


Рис. 3.1.4

В случае $\lambda_1 = \lambda_2 = 0$ каждое из направлений является асимптотическим. В противном случае главным направлением является асимптотическое направление, для которого кривизна равна нулю. Соответствующее нормальное сечение в такой точке имеет точку распрямления.

Кривая на поверхности называется *линией кривизны*, если касательная в каждой точке к ней параллельна одному из главных направлений в этой точке поверхности. Линиями кривизны часто являются координат-

ные линии. Пусть u -линии и v -линии являются линиями кривизны. В этом случае в каждой точке поверхности выполняются равенства

$$g_{12} = g_{21} = b_{12} = b_{21} = 0 \quad (3.1.15)$$

в силу ортогональности главных направлений. Справедливо и обратное утверждение: если в каждой точке поверхности выполняются равенства (3.1.15), то координатные линии являются линиями кривизны.

3.2. Аналитические поверхности

Аналитическими поверхностями будем называть поверхности, координаты которых в некоторой локальной системе координат можно описать с помощью аналитических функций, не используя точки, векторы, кривые и другие поверхности.

Для аналитических поверхностей используются локальные системы координат, в которых поверхности имеют канонический вид. Построим локальную декартову систему координат с началом в точке \mathbf{p} и базисными векторами $\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z$. Поверхность, координаты которой в локальной системе координат равны соответственно $x(u, v), y(u, v), z(u, v)$, будет описываться векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + x(u, v) \mathbf{i}_x + y(u, v) \mathbf{i}_y + z(u, v) \mathbf{i}_z. \quad (3.2.1)$$

Пусть p_i – координаты начала \mathbf{p} локальной системы координат, x_i – компоненты базисного вектора \mathbf{i}_x , y_i – компоненты базисного вектора \mathbf{i}_y , z_i – компоненты базисного вектора \mathbf{i}_z , $i = 1, 2, 3$. Тогда аналитическая поверхность будет представлять собой функцию

$$\begin{bmatrix} r_1(u, v) \\ r_2(u, v) \\ r_3(u, v) \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \cdot \begin{bmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{bmatrix}.$$

Координаты точки поверхности (3.2.1) равны

$$r_i(u, v) = p_i + x(u, v) x_i + y(u, v) y_i + z(u, v) z_i.$$

При изменении положения или ориентации подобным образом описанной аналитической поверхности изменяются координаты начала местной системы координат и ее базисные векторы, а аналитические функции поверхности остаются неизменными, сохраняя канонический вид. Рассмотрим примеры аналитических поверхностей.

Коническую поверхность опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + (r + hv \operatorname{tg} \gamma)(\cos u \mathbf{i}_x + \sin u \mathbf{i}_y) + hv \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [v_{\min}, v_{\max}],$$

где r — радиус одного из оснований конуса, h — длина конуса, γ — угол между образующей и осью конуса. Начало локальной системы координат мы расположили в центре одного из оснований конуса, базисный вектор \mathbf{i}_z направили вдоль оси поверхности. Коническая поверхность является периодической по первому параметру и усечена по второму параметру. Круговой конус показан на рис. 3.2.1.

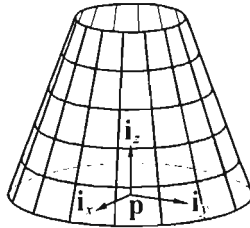


Рис. 3.2.1

Если угол γ положить равным нулю, то получим **цилиндрическую поверхность**.

Поверхность тора опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + (R + r \cos v) \cos u \mathbf{i}_x + (R + r \cos v) \sin u \mathbf{i}_y + r \sin v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [v_{\min}, v_{\max}],$$

где r — радиус трубки тора; R — радиус оси трубки. Начало локальной системы координат мы расположили в центре тора, базисный вектор \mathbf{i}_z направили вдоль оси тора, рис. 3.2.2. Скалярные функции $x(u, v) = (R + r \cos v) \cdot \cos u$, $y(u, v) = (R + r \cos v) \cdot \sin u$, $z(u, v) = r \sin v$ тора связаны уравнением

$$\left(\sqrt{x^2 + y^2} - R \right)^2 + z^2 = r^2.$$

При $r \leq R$ поверхность тора циклически замкнута по обоим параметрам и имеет форму бублика, $v_{\min} = -\pi$, $v_{\max} = \pi$. Если $r > R$, то для предотвращения перехода через места самопересечения поверхности ограничим область изменения второго параметра. Положим $v_{\min} = -\pi + v_0$, $v_{\max} = \pi - v_0$, где $v_0 = \arccos(R/r)$, если $-|r| < R < r$. При $0 < R < r$ поверхность тора имеет форму яблока, рис. 3.2.3. При $R = 0$ поверхность тора превратится в **сферу** радиуса r . При $-|r| < R < 0$ поверхность тора имеет форму лимона.

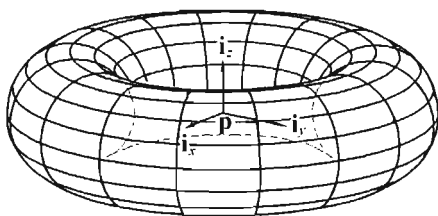


Рис. 3.2.2

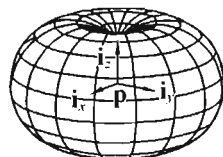


Рис. 3.2.3

Если радиус оси трубки не превосходит радиуса трубки, то поверхность тора имеет особые точки при $v = \pm(\pi - \arccos(R/r))$, так как в этих точках обращается в ноль производная поверхности по первому параметру.

Эллипсоид опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + a \cos v \cos u \mathbf{i}_x + b \cos v \sin u \mathbf{i}_y + c \sin v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [-\pi/2, \pi/2],$$

где a, b, c – полуоси эллипсоида. Эллипсоид, как и сфера, является периодической по первому параметру поверхностью. Скалярные функции $x(u, v) = a \cdot \cos v \cdot \cos u$, $y(u, v) = b \cdot \cos v \cdot \sin u$, $z(u, v) = c \cdot \sin v$ эллипсоида связаны уравнением

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1.$$

Однополостный гиперboloид опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + a \operatorname{ch} v \cos u \mathbf{i}_x + b \operatorname{ch} v \sin u \mathbf{i}_y + c \operatorname{sh} v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [v_{\min}, v_{\max}],$$

где a и b – действительные полуоси гиперboloида; c – мнимая полуось гиперboloида. Действительные полуоси равны расстояниям от точки \mathbf{p} до точек пересечения поверхности с осями \mathbf{i}_x и \mathbf{i}_y локальной системы координат. Однополостный гиперboloид является периодической по первому параметру поверхностью. Скалярные функции $x(u, v) = a \cdot \operatorname{ch} v \cdot \cos u$, $y(u, v) = b \cdot \operatorname{ch} v \cdot \sin u$, $z(u, v) = c \cdot \operatorname{sh} v$ гиперboloида связаны уравнением

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 1.$$

Двуполостный гиперboloид опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + a \operatorname{sh} v \cos u \mathbf{i}_x + b \operatorname{sh} v \sin u \mathbf{i}_y + c \operatorname{ch} v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [0, v_{\max}],$$

где a и b – мнимые полуоси гиперболоида; c – действительная полуось гиперболоида. Действительная полуось равна расстоянию от точки \mathbf{p} до точек пересечения поверхности с осью \mathbf{i}_z локальной системы координат. Двуполостный гиперболоид является периодической по первому параметру поверхностью. Скалярные функции $x(u, v) = a \cdot \operatorname{sh} v \cdot \cos u$, $y(u, v) = b \cdot \operatorname{sh} v \cdot \sin u$, $z(u, v) = c \cdot \operatorname{ch} v$ гиперболоида связаны уравнением

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = -1.$$

Эллиптический параболоид опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + a v \cos u \mathbf{i}_x + b v \sin u \mathbf{i}_y + v^2 \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [0, v_{\max}],$$

где a и b – полуоси параболоида. Эллиптический параболоид является периодической по первому параметру поверхностью. Скалярные функции $x(u, v) = a \cdot v \cdot \cos u$, $y(u, v) = b \cdot v \cdot \sin u$, $z(u, v) = v^2$ параболоида связаны уравнением

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = z.$$

Гиперболический параболоид опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + a u \mathbf{i}_x + b v \mathbf{i}_y + (u^2 - v^2) \mathbf{i}_z, \\ u \in [u_{\min}, u_{\max}], v \in [v_{\min}, v_{\max}],$$

где a и b – полуоси параболоида. Скалярные функции $x(u, v) = a \cdot u$, $y(u, v) = b \cdot v$, $z(u, v) = u^2 - v^2$ гиперболического параболоида связаны уравнением

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = z.$$

Гиперболический параболоид не является периодической поверхностью.

3.3. Поверхности, полученные движением кривой

Поверхность можно получить путем движения кривой по заданной траектории. Траекторией может служить произвольная кривая. Кривая, движением которой получается поверхность, называется *образующей*; кривая, служащая траекторией, называется *направляющей*. Пусть образующая описывается кривой $\mathbf{c}(u)$, $u_{\min} \leq u \leq u_{\max}$, а направляющей является кривая $\mathbf{g}(v)$, $v_{\min} \leq v \leq v_{\max}$. Пусть первый параметр поверхности движения

совпадает с параметром образующей кривой $\mathbf{c}(u)$, а второй параметр поверхности совпадает с параметром направляющей кривой $\mathbf{g}(v)$. Область изменения параметров поверхностей движения представляет собой прямоугольник. Если образующая кривая $\mathbf{c}(u)$ циклически замкнута, то поверхность движения также будет циклически замкнутой по первому параметру.

При движении образующей кривой вдоль направляющей кривой ориентация образующей может оставаться неизменной в пространстве или оставаться неизменной относительно направляющей. В первом случае образующая выполняет плоскопараллельное движение. Назовем такую поверхность *поверхностью сдвига*. Во втором случае среди всевозможных направляющих выделим две наиболее простые — отрезок прямой и дугу окружности. Если направляющей служит отрезок прямой, то поверхность назовем *поверхностью выдавливания*. Если направляющей служит дуга окружности или вся окружность, то поверхность назовем *поверхностью вращения*. В остальных случаях поверхность движения будем называть *кинематической поверхностью*.

Назовем кривую *монотонной*, если скалярное произведение касательных векторов в любых двух точках кривой всегда больше нуля. Монотонная кривая не может быть замкнутой. К таким кривым относятся отрезок, спираль и некоторые другие кривые.

Поверхность сдвига, полученную движением кривой $\mathbf{c}(u)$, $u_{\min} \leq u \leq u_{\max}$ вдоль монотонной направляющей кривой $\mathbf{g}(v)$, опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{g}(v) + \mathbf{c}(u) - \mathbf{g}(v_{\min}). \quad (3.3.1)$$

Радиус-вектор поверхности сдвига строится как сумма двух векторов: вектора точки на направляющей кривой $\mathbf{g}(v)$ и вектора положения точки образующей относительно начальной точки направляющей $\mathbf{c}(u) - \mathbf{g}(v_{\min})$. Поверхность сдвига приведена на рис. 3.3.1.

Поверхность выдавливания, полученную движением кривой $\mathbf{c}(u)$, $u_{\min} \leq u \leq u_{\max}$ вдоль вектора \mathbf{d} , опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{c}(u) + v\mathbf{d}. \quad (3.3.2)$$

Поверхность выдавливания приведена на рис. 3.3.2.

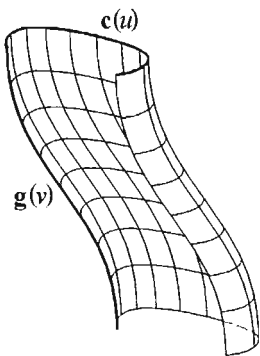


Рис. 3.3.1

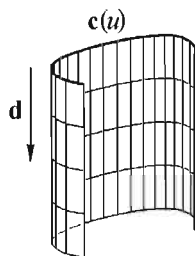


Рис. 3.3.2

Поверхность вращения построим, используя локальную декартову прямоугольную систему координат, связанную с осью вращения. Ее начало \mathbf{p} расположим на оси вращения, вдоль оси вращения направим третий базисный вектор \mathbf{i}_3 локальной системы координат. Первый базисный вектор \mathbf{i}_1 локальной системы координат направим в сторону образующей кривой. Для этого построим вектор $\mathbf{d} = \mathbf{c}(u_1) - \mathbf{p}$ и вычтем из \mathbf{d} его же составляющую, параллельную оси вращения: $\mathbf{d}_1 = \mathbf{d} - (\mathbf{i}_3 \cdot \mathbf{d})\mathbf{i}_3$. Точкой $\mathbf{c}(u_1)$ может служить произвольная точка образующей кривой, не лежащая на оси вращения. Базисный вектор \mathbf{i}_1 направим вдоль \mathbf{d}_1 :

$$\mathbf{i}_1 = \frac{\mathbf{d}_1}{|\mathbf{d}_1|}.$$

Базисный вектор \mathbf{i}_2 локальной системы координат направим параллельно вектору $\mathbf{i}_2 = \mathbf{i}_3 \times \mathbf{i}_1$. Для вычисления радиуса-вектора точки поверхности вращения построим матрицу

$$\mathbf{A} = [\mathbf{i}_1 \quad \mathbf{i}_2 \quad \mathbf{i}_3],$$

столбцы которой составлены из компонент базисных векторов локальной системы координат. Матрица \mathbf{A} является матрицей преобразования координат радиуса-вектора точки из локальной системы координат в глобальную систему координат.

Поверхность, полученную вращением кривой $\mathbf{c}(u)$, $u_{\min} \leq u \leq u_{\max}$ на угол α вокруг оси, заданной вектором \mathbf{i}_3 и точкой \mathbf{p} , опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + \mathbf{M}(v) \cdot (\mathbf{c}(u) - \mathbf{p}), \quad u \in [u_{\min}, u_{\max}], \quad v \in [0, \alpha], \quad (3.3.3)$$

$$\text{где } \mathbf{M}(v) = \mathbf{A} \cdot \begin{bmatrix} \cos v & -\sin v & 0 \\ \sin v & \cos v & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{A}^{-1} \text{ — матрица вращения. Матрица } \mathbf{M}(v)$$

переводит вектор $\mathbf{c}(u) - \mathbf{p}$ в локальную систему координат, поворачивает его в ней на угол v вокруг оси вращения и возвращает повернутый вектор обратно в глобальную систему координат. Если угол вращения $\alpha = 2\pi$, то поверхность (3.3.3) циклически замкнута по второму параметру. Поверхность вращения приведена на рис. 3.3.3.

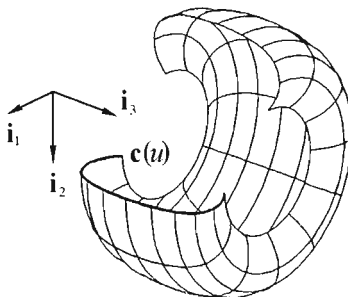


Рис. 3.3.3

Кинематическую поверхность построим, используя подвижную декартову прямоугольную систему координат, связанную с направляющей кривой $\mathbf{g}(v)$. Ее начало \mathbf{p}_0 расположим в текущей точке направляющей кривой. Первый базисный вектор $\mathbf{i}_1 = \mathbf{g}' / |\mathbf{g}'|$ подвижной системы координат направим по касательной к направляющей кривой, второй базисный вектор \mathbf{i}_2 построим ортогонально первому, а третий базисный вектор $\mathbf{i}_3 = \mathbf{i}_1 \times \mathbf{i}_2$ пусть образует с первым и вторым базисными векторами тройку векторов. Базисный вектор \mathbf{i}_2 направим параллельно вектору

$$\mathbf{d}_2(v) = \mathbf{d} - (\mathbf{i}_1 \cdot \mathbf{d}) \mathbf{i}_1, \quad (3.3.4)$$

где \mathbf{d} — некоторый вектор, который не параллелен касательной к направляющей $\mathbf{g}(v)$ ни в одной ее точке. Вектор $\mathbf{d}_2(v)$ представляет собой составляющую вектора \mathbf{d} , перпендикулярную вектору \mathbf{i}_1 . Вектор $\mathbf{d}_2(v)$ никогда не равен нулю, так как мы предполагаем, что вектор \mathbf{d} никогда не параллелен вектору \mathbf{i}_1 . Будем считать, что для направляющей кривой существует такой вектор \mathbf{d} , который никогда не параллелен касательной к кривой. Подвижная система координат является функцией параметра v направляющей кривой

$$\mathbf{p}_0(v) = \mathbf{g}(v), \quad (3.3.5)$$

$$\mathbf{i}_1(v) = \frac{\mathbf{g}'}{|\mathbf{g}'|}, \quad \mathbf{i}_2(v) = \frac{\mathbf{d}_2}{|\mathbf{d}_2|}, \quad \mathbf{i}_3(v) = \mathbf{i}_1 \times \mathbf{i}_2,$$

где \mathbf{d}_2 определяется равенством (3.3.4), а $\mathbf{g}' = \frac{d\mathbf{g}}{dv}$.

Для вычисления радиуса-вектора точки кинематической поверхности построим матрицу

$$\mathbf{A}(v) = [\mathbf{i}_1(v) \quad \mathbf{i}_2(v) \quad \mathbf{i}_3(v)], \quad (3.3.6)$$

столбцы которой составлены из компонент базисных векторов подвижной системы координат. Матрица $\mathbf{A}(v)$ является матрицей преобразования координат радиуса-вектора точки из подвижной системы координат в глобальную систему координат и зависит от параметра направляющей кривой.

Запомним положение образующей кривой $\mathbf{c}(u)$ в подвижном касательном базисе в начале направляющей и будем сохранять его при движении вдоль направляющей. Радиус-вектор точки образующей кривой в подвижной системе координат при $v = v_{\min}$ равен

$$\mathbf{x}(u, v_{\min}) = \mathbf{A}^{-1}(v_{\min}) \cdot (\mathbf{c}(u) - \mathbf{g}(v_{\min})),$$

где $\mathbf{A}^{-1}(v_{\min})$ – матрица преобразования координат радиуса-вектора точки из глобальной системы координат в подвижную систему координат, вычисленную для начальной точки направляющей кривой. При движении вдоль направляющей кривой подвижный касательный базис меняет свое положение и ориентацию в пространстве и увлекает за собой жестко связанную с ним образующую кривую. Вектор $\mathbf{x}(u, v_{\min})$ выражает положение точки образующей относительно точки на направляющей кривой в подвижном касательном базисе, которое сохраняется для произвольного параметра v . Переходя из подвижной системы координат в глобальную систему координат при текущем параметре v , получим радиус-вектор точки на кинематической поверхности:

$$\mathbf{r}(u, v) = \mathbf{g}(v) + \mathbf{A}(v) \cdot \mathbf{x}(u, v_{\min}).$$

Таким образом, радиус-вектор кинематической поверхности опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{g}(v) + \mathbf{M}(v) \cdot (\mathbf{c}(u) - \mathbf{g}(v_{\min})), \quad (3.3.7)$$

где $\mathbf{M}(v)$ – матрица поворота текущего подвижного базиса относительно его начального положения. Эта матрица вычисляется по формуле

$$\mathbf{M}(v) = \mathbf{A}(v) \cdot \mathbf{A}^{-1}(v_{\min}). \quad (3.3.8)$$

В начале направляющей кривой матрица $\mathbf{M}(v_{\min})$ равна единичной матрице.

Для вычисления производных радиуса-вектора кинематической поверхности нужно вычислить производные матрицы $\mathbf{M}(v)$, что в свою очередь сводится к вычислению производных базисных векторов $\mathbf{i}_1(v)$, $\mathbf{i}_2(v)$, $\mathbf{i}_3(v)$, например:

$$\mathbf{i}_1 = \frac{\mathbf{g}'}{|\mathbf{g}'|}, \quad \frac{d\mathbf{i}_1}{dv} = \frac{\mathbf{g}''}{|\mathbf{g}'|} - \left(\mathbf{i}_1 \cdot \frac{\mathbf{g}''}{|\mathbf{g}'|} \right) \mathbf{i}_1,$$

$$\frac{d^2\mathbf{i}_1}{d^2v} = \frac{\mathbf{g}'''}{|\mathbf{g}'|} - 2 \left(\mathbf{i}_1 \cdot \frac{\mathbf{g}''}{|\mathbf{g}'|} \right) \frac{d\mathbf{i}_1}{dv} - \left(\mathbf{i}_1 \cdot \frac{\mathbf{g}'''}{|\mathbf{g}'|} + \frac{d\mathbf{i}_1}{dv} \cdot \frac{\mathbf{g}''}{|\mathbf{g}'|} \right) \mathbf{i}_1. \quad (3.3.9)$$

Из (3.3.9) следует, что для вычисления частных производных второго порядка радиуса-вектора кинематической поверхности требуются производные радиуса-вектора направляющей кривой третьего порядка.

Замкнутость кинематической поверхности по параметру u совпадает с замкнутостью образующей кривой $\mathbf{c}(v)$, а замкнутость поверхности по параметру v совпадает с замкнутостью направляющей кривой $\mathbf{g}(v)$. Кинематическая поверхность приведена на рис. 3.3.4.

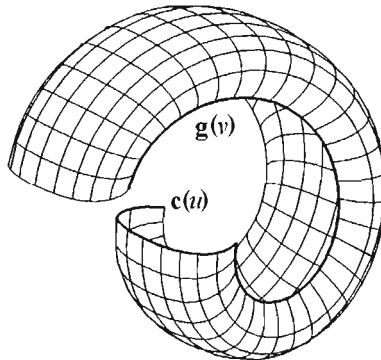


Рис. 3.3.4

При неудачном сочетании образующей и направляющей кривых поверхность сдвига и кинематическая поверхность могут пересекать сами себя.

В общем случае большинство поверхностей можно представить в виде поверхности движения, если в процессе движения образующей кривой и направляющей кривой дать возможность изменяться. На рис. 3.3.5 приведена поверхность выдавливания, у которой плоская образующая кривая в процессе движения уменьшается в три раза вдоль меньшего попе-

речного размера и поворачивается на 45° . Уменьшение меньшего поперечного размера образующей кривой линейно зависит от увеличения второго параметра поверхности. Поворот образующей кривой является синусоидальной функцией второго параметра поверхности. На рис. 3.3.6 приведена поверхность вращения, у которой расстояние неплоской образующей кривой от оси в процессе вращения изменяется по закону укороченной эпициклоиды.

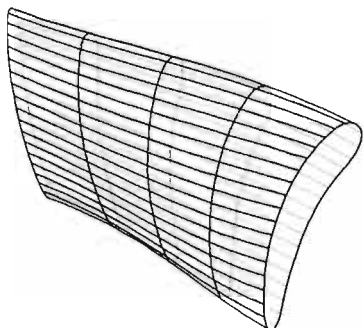


Рис. 3.3.5

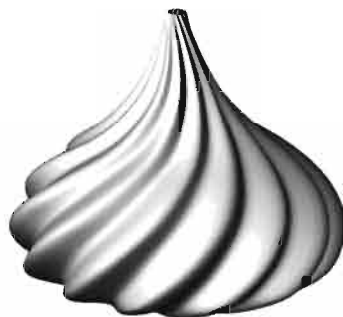


Рис. 3.3.6

Более сложные поверхности можно получить, если в процессе движения изменение образующей кривой поставить в зависимость от второго параметра поверхности, а изменение направляющей кривой поставить в зависимость от первого параметра поверхности.

3.4. Поверхности, построенные на семействе кривых

Рассмотрим построение поверхностей на основе совокупности не пересекающих друг друга кривых $c_i(u)$, $i = 0, 1, \dots, n$. Путем изменения параметризации (2.15.4) приведем все кривые к одной параметрической длине, так чтобы область изменения параметра каждой кривой лежала в пределах от u_{\min} до u_{\max} . Совокупность не пересекающих друг друга кривых, имеющих одинаковую параметрическую длину, назовем *семейством кривых*. Пусть первый параметр поверхности $r(u, v)$ совпадает с общим параметром кривых. Для второго параметра поверхности выберем возрастающую последовательность значений v_i и поставим ей в соответствие семейство кривых $c_i(u)$, $i = 0, 1, \dots, n$. Если все кривые семейства циклически замкнуты, то создаваемая поверхность будет циклически замкнута по первому параметру.

Линейчатая поверхность строится на семействе из двух кривых и представляет собой геометрическое место отрезков прямых, соединяю-

ших соответствующие точки кривых семейства. Пусть заданы две кривые $\mathbf{c}_0(u)$ и $\mathbf{c}_1(u)$. Линейчатую поверхность опишем векторной функцией

$$\mathbf{r}(u, v) = (1-v) \mathbf{c}_0(u) + v \mathbf{c}_1(u). \quad (3.4.1)$$

На рис. 3.4.1 приведена линейчатая поверхность.

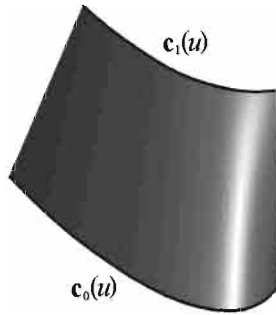


Рис. 3.4.1

Секторная поверхность является частным случаем линейчатой поверхности. Она получается из линейчатой поверхности (3.4.1), если одну из кривых, например $\mathbf{c}_1(u)$, заменить точкой \mathbf{p} . Радиус-вектор секторной поверхности опишем формулой

$$\mathbf{r}(u, v) = (1-v) \mathbf{c}_0(u) + v \mathbf{p}; \quad v \leq 1.$$

Если кривая $\mathbf{c}_0(u)$ циклически замкнута, то секторная поверхность также циклически замкнута по первому параметру. Точки поверхности при $v = 1$ являются особыми, так как для них частная производная по первому параметру радиуса-вектора поверхности равна нулю.

Построим на семействе кривых $\mathbf{c}_i(t)$, $i = 0, 1, \dots, n$ поверхность аналогично построению по семейству точек кривой Лагранжа (2.3.2). Выберем возрастающую последовательность значений второго параметра v_i , при которых поверхность $\mathbf{r}(u, v)$ должна совпасть с кривыми $\mathbf{c}_i(u)$. Заметим, что форма поверхности зависит от значений v_i , $i = 0, 1, \dots, n$. **Поверхность Лагранжа**, построенную по семейству кривых $\mathbf{c}_i(u)$, $i = 0, 1, 2, \dots, n$, опишем векторной функцией

$$\mathbf{r}(u, v) = \sum_{i=0}^n L_i^n(v) \mathbf{c}_i(u), \quad (3.4.2)$$

где $L_i^n(v)$ – коэффициенты Лагранжа (2.3.3). Каждая v -линия $\mathbf{r}(\text{const}, v)$ такой поверхности представляет собой кривую Лагранжа.

Пусть кроме кривых $\mathbf{c}_i(u)$ нам известны частные производные поверхности по второму параметру $\mathbf{r}_i(u, v) \left. \frac{\partial \mathbf{r}(u, v)}{\partial v} \right|_{v=v_i}$ на координатных линиях вдоль выбранных значений второго параметра $v = v_i, i = 0, 1, 2, \dots, n$. **Поверхность Эрмита**, построенную по семейству кривых, опишем векторной функцией

$$\begin{aligned} \mathbf{r}(u, v) &= \\ &= \mathbf{c}_i(u)(1-3w^2+2w^3) + \mathbf{c}_{i+1}(u)(3w^2-2w^3) + \\ &+ \mathbf{r}_i(u, v_i)(v_{i+1}-v_i)(w-2w^2+w^3) + \mathbf{r}_i(u, v_{i+1})(v_{i+1}-v_i)(-w^2+w^3) = \\ &= [\alpha_0(w) \quad \alpha_1(w) \quad \beta_0(w) \quad \beta_1(w)] \cdot \begin{bmatrix} \mathbf{c}_i(u) \\ \mathbf{c}_{i+1}(u) \\ \mathbf{r}_i(u, v_i)(v_{i+1}-v_i) \\ \mathbf{r}_i(u, v_{i+1})(v_{i+1}-v_i) \end{bmatrix}, \end{aligned} \quad (3.4.3)$$

где $w = w(v) = \frac{v - v_i}{v_{i+1} - v_i}$ – местный параметр поверхности, а индекс i найден из условия $v_i \leq v \leq v_{i+1}$. Данная поверхность построена аналогично построению составного сплайна Эрмита (2.3.10) и является кубической по v -направлению. При изменении параметра w от 0 до 1 параметр поверхности изменяется от v_i до v_{i+1} .

Если производные $\mathbf{r}_i(u, v_i)$ семейства неизвестны, то они могут быть вычислены по соседним кривым аналогично тому, как вычисляются производные в контрольных точках составного сплайна Эрмита, например, для внутренних кривых ($i = 1, 2, \dots, n-1$):

$$\begin{aligned} \mathbf{r}_i(u, v_i) &= \mathbf{c}_{i-1}(u) \frac{v_i - v_{i+1}}{(v_{i-1} - v_i)(v_{i-1} - v_{i+1})} + \\ &+ \mathbf{c}_i(u) \frac{2v_i - v_{i-1} - v_{i+1}}{(v_i - v_{i-1})(v_i - v_{i+1})} + \mathbf{c}_{i+1}(u) \frac{v_i - v_{i-1}}{(v_{i+1} - v_{i-1})(v_{i+1} - v_i)}, \end{aligned}$$

для крайних кривых ($i = 0, i = n$):

$$\begin{aligned} \mathbf{r}_0(u, v_0) &= \frac{3\mathbf{c}_1(u) - \mathbf{c}_0(u)}{2} \frac{1}{v_1 - v_0} - \frac{1}{2} \mathbf{r}_1(u, v_1), \\ \mathbf{r}_n(u, v_n) &= \frac{3\mathbf{c}_n(u) - \mathbf{c}_{n-1}(u)}{2} \frac{1}{v_n - v_{n-1}} - \frac{1}{2} \mathbf{r}_n(u, v_{n-1}). \end{aligned}$$

Формулы вычисления $\mathbf{r}_i(u, v_0)$ и $\mathbf{r}_i(u, v_n)$ обеспечивают равенство нулю вторых производных на крайних кривых поверхности в перпендикулярных к краю поверхности направлениях. Если все кривые семейства яв-

ляются циклически замкнутыми, то поверхность (3.4.2) будет циклически замкнута по первому параметру. Поверхность может быть циклически замкнута и по второму параметру. Поверхность Эрмита, построенная на семействе кривых, приведена на рис. 3.4.2.

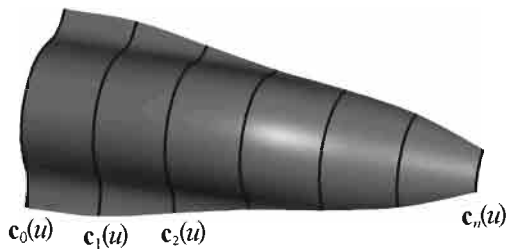


Рис. 3.4.2

В частном случае поверхность (3.4.3) может служить переходным мостиком от одной поверхности к другой. Сопрягаемые поверхности должны иметь прямоугольные области определения параметров. Пусть, например, требуется плавно сопрячь край поверхности $\mathbf{a}(u, v_{\max})$ с краем поверхности $\mathbf{b}(u, v_{\min})$. Поверхность перехода опишем векторной функцией

$$\begin{aligned} \mathbf{r}(u, w) = & \mathbf{a}(u, v_{\max}) (1 - 3w^2 + 2w^3) + & (3.4.4) \\ & + \mathbf{b}(u, v_{\min}) (3w^2 - 2w^3) + \\ & + \mathbf{a}_v(u, v_{\max})(w - 2w^2 + w^3) k_a + \\ & + \mathbf{b}_v(u, v_{\min})(-w^2 + w^3) k_b, \\ & u \in [u_{\min}, u_{\max}], w \in [0, 1], \end{aligned}$$

где $\mathbf{a}_v(u, v_{\max}) = \left. \frac{\partial \mathbf{a}(u, v)}{\partial v} \right|_{v=v_{\max}}$ – частная производная по фиксированному

параметру радиуса-вектора первой сопрягаемой поверхности;

$\mathbf{b}_v(u, v_{\min}) = \left. \frac{\partial \mathbf{b}(u, v)}{\partial v} \right|_{v=v_{\min}}$ – частная производная по фиксированному пара-

метру радиуса-вектора второй сопрягаемой поверхности, k_a и k_b – коэффициенты, нормирующие производные $\mathbf{a}_v(u, v_{\max})$ и $\mathbf{b}_v(u, v_{\min})$ соответственно. Коэффициенты k_a и k_b получим путем деления среднего расстояния между крайними сопрягаемыми поверхностями на среднюю длину частных производных $\mathbf{a}_v(u, v_{\max})$ и $\mathbf{b}_v(u, v_{\min})$. Поверхность перехода приведена на рис. 3.4.3.

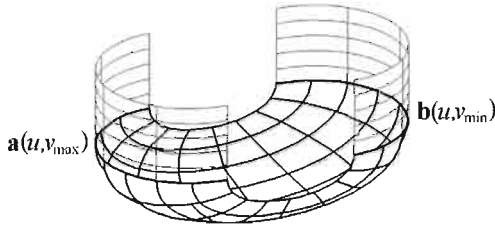


Рис. 3.4.3

На семействе кривых построим поверхность аналогично построению по семейству точек кривой Безье. Для этого в (2.6.2) вместо контрольных точек подставим кривые семейства. **Поверхность Безье, построенную на семействе кривых $\mathbf{c}_i(u)$, $i = 0, 1, \dots, n$,** опишем векторной функцией

$$\mathbf{r}(u, v) = \frac{\sum_{i=0}^n B_i^n(v) w_i(u) \mathbf{c}_i(u)}{\sum_{i=0}^n B_i^n(v) w_i(u)}, \quad (3.4.5)$$

где $B_i^n(v)$ – функции Бернштейна (2.4.2). В общем случае каждая кривая $\mathbf{c}_i(u)$ семейства имеет свой вес $w_i(u)$, представляющий собой скалярную функцию параметра соответствующей кривой. Каждая линия поверхности при $u = \text{const}$ является рациональной кривой Безье. В частном случае, когда все кривые семейства имеют не зависящие от параметра одинаковые веса, радиус-вектор поверхности Безье на семействе кривых вычислим по формуле

$$\mathbf{r}(u, v) = \sum_{i=0}^n B_i^n(v) \mathbf{c}_i(u). \quad (3.4.6)$$

В отличие от других поверхностей, построенных на семействе кривых, поверхность (3.4.5) не проходит по кривым семейства.

На семействе кривых построим поверхность аналогично построению B -кривой по семейству точек. Для этого в (2.9.1) вместо контрольных точек подставим кривые семейства. **B -поверхность, построенную на семействе кривых $\mathbf{c}_i(u)$, $i = 0, 1, \dots, n$,** опишем векторной функцией

$$\mathbf{r}(u, v) = \frac{\sum_{j=0}^n N_j^m(v) w_j(u) \mathbf{c}_j(u)}{\sum_{j=0}^n N_j^m(v) w_j(u)}, \quad (3.4.7)$$

где $N_j^m(v)$ – B -сплайны (2.8.1) m -порядка; $w_j(u)$ – функции веса кривой $\mathbf{c}_j(u)$ семейства. Каждый B -сплайн $N_j^m(v)$ построен на неубывающей последовательности из $m + 2$ узлов $v_j, v_{j+1}, \dots, v_{j+m+1}$.

Для построения совокупности $n + 1$ B -сплайнов m -го порядка незамкнутой по второму параметру B -поверхности требуется $n + m + 2$ узлов $v_j, j = 0, 1, \dots, n + m + 1$. Для построения совокупности $n + 1$ B -сплайнов m -го порядка циклически замкнутой по второму параметру B -поверхности требуется $n + 2m + 2$ узлов $v_j, j = 0, 1, \dots, n + 2m + 1$. Последовательность узлов второго параметра поверхности строится аналогично последовательности узлов B -кривой (2.9.1). Второй параметр B -поверхности изменяется от значения узла $v_{\min} = v_m$ до значения узла $v_{\max} = v_{n+m+1}$.

Каждая линия $u = \text{const}$ на поверхности (3.4.7) является B -кривой (2.9.1). На рис. 3.4.4 приведена B -поверхность и семейство кривых, по которым построена поверхность.

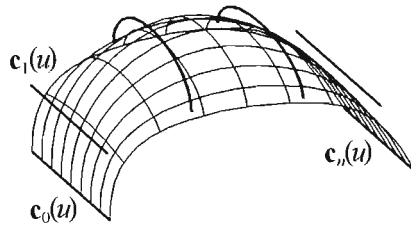


Рис. 3.4.4

В частном случае, когда все кривые семейства имеют не зависящие от параметра одинаковые веса, радиус-вектор B -поверхности на семействе кривых вычислим по формуле

$$\mathbf{r}(u, v) = \sum_{j=0}^n N_j^m(v) \mathbf{c}_j(u). \quad (3.4.8)$$

B -поверхность, построенная на семействе кривых, не проходит по кривым семейства.

Пусть дано семейство кривых $\mathbf{c}_i(u), i = 0, 1, \dots, n$, направляющая кривая $\mathbf{g}(v)$ и возрастающая последовательность значений параметра направляющей v_j , поставленная в соответствие семейству кривых. Построим поверхность $\mathbf{r}(u, v)$, проходящую по семейству кривых, учитывающую поведение направляющей кривой. Пусть при значениях второго параметра, располагающихся между v_j и v_{j+1} , поверхность представляет собой взвешенную сумму кинематических поверхностей, построенных на направляющей $\mathbf{g}(v)$ с образующими $\mathbf{c}_j(u)$ и $\mathbf{c}_{j+1}(u)$. В соответствии с (3.3.7) радиус-вектор кинематической поверхности, построенной на образующей $\mathbf{c}_j(u)$, вычислим по формуле

$$\mathbf{r}_j(u, v) = \mathbf{g}(v) + \mathbf{M}_j(v) \cdot (\mathbf{c}_j(u) - \mathbf{g}(v_j)).$$

В соответствии с (3.3.8) матрицу $M_i(v)$ поворота текущего подвижного базиса относительно его начального положения вычислим по формуле

$$M_i(v) = A(v) \cdot A^{-1}(v_i).$$

Поверхность, построенную на семействе кривых $c_i(u)$, $i = 0, 1, \dots, n$ и направляющей $g(v)$, опишем векторной функцией

$$r(u, v) = r_i(u, v) (1 - 3w^2 + 2w^3) + r_{i+1}(u, v) (3w^2 - 2w^3), \quad (3.4.9)$$

где индекс i найден из условия $v_i \leq v \leq v_{i+1}$; $w = \frac{v - v_i}{v_{i+1} - v_i}$ — местный параметр.

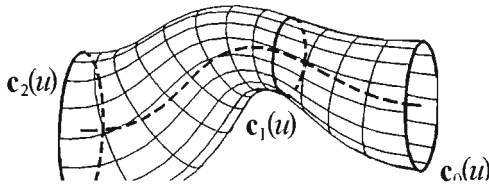


Рис. 3.4.5

На рис. 3.4.5 приведена поверхность на семействе из трех кривых и направляющей.

3.5. Поверхности, построенные на сетке кривых

Пусть имеется два семейства кривых: семейство кривых $c_i(u)$, $i = 0, 1, 2, \dots, n$, $u_{\min} \leq u \leq u_{\max}$ и семейство кривых $b_j(v)$, $j = 0, 1, 2, \dots, m$, $v_{\min} \leq v \leq v_{\max}$, рис. 3.5.1. Пусть при $u = u_j$, $j = 0, 1, 2, \dots, m$ каждая кривая $c_i(u)$ первого семейства пересекает каждую кривую $b_j(v)$ второго семейства при $v = v_i$, $i = 0, 1, 2, \dots, n$. Пусть параметры точек пересечения кривых образуют возрастающие последовательности: $u_j < u_{j+1}$ и $v_i < v_{i+1}$. Два семейства кривых, удовлетворяющие перечисленным требованиям, назовем *сеткой кривых*.

Простейшей сеткой кривых являются две пары отрезков, построенных на четырех общих точках p_{00} , p_{10} , p_{01} и p_{11} . Пусть первое семейство образуют отрезки $c_0(u) = p_{00}(1 - u) + p_{10}u$ и $c_1(u) = p_{01}(1 - u) + p_{11}u$, а второе семейство образуют отрезки $b_0(v) = p_{00}(1 - v) + p_{10}v$ и $b_1(v) = p_{01}(1 - v) + p_{11}v$. **Билинейная поверхность**, построенная на данной сетке кривых, определится векторной функцией

$$r(u, v) = (1 - u)(1 - v) p_{00} + u(1 - v) p_{10} + (1 - u)v p_{01} + uv p_{11}. \quad (3.5.1)$$

Билинейная поверхность приведена на рис 3.5.2.

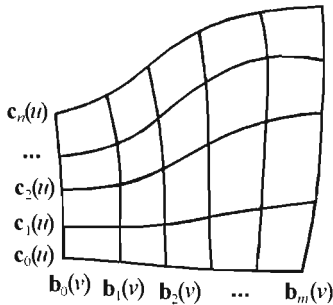


Рис. 3.5.1

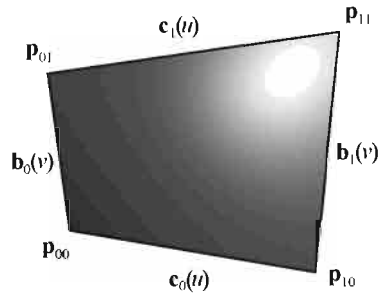


Рис. 3.5.2

Поверхность (3.5.1) является частным случаем линейчатой поверхности (3.4.1), построенной на двух отрезках $c_0(u)$ и $c_1(u)$. Билинейная поверхность является линейчатой по обоим параметрам.

Для построения билинейной поверхности достаточно знать радиус-векторы точек p_{00} , p_{10} , p_{01} и p_{11} . Если точки p_{00} и p_{01} или p_{10} и p_{11} совпадают, то мы получим *треугольную* поверхность с особой точкой.

Построим линейчатую поверхность (3.4.1) с приведением кривых $c_0(u)$ и $c_1(u)$ к параметрической длине $0 \leq u \leq 1$. К векторной функции этой линейчатой поверхности прибавим и вычтем из нее векторную функцию билинейной поверхности (3.5.1), где p_{00} и p_{10} – начальная и конечная точки кривой $c_0(u)$, а p_{01} и p_{11} – начальная и конечная точки кривой $c_1(u)$ линейчатой поверхности. В результате этих действий векторная функция не изменится, но примет вид

$$\begin{aligned} \mathbf{r}(u, v) = & (1 - v) \mathbf{c}_0(u) + v \mathbf{c}_1(u) + \\ & + (1 - u) (\mathbf{p}_{00}(1 - v) + \mathbf{p}_{01}v) + u (\mathbf{p}_{10}(1 - v) + \mathbf{p}_{11}v) - \\ & - (1 - u)(1 - v) \mathbf{p}_{00} - u(1 - v) \mathbf{p}_{10} - (1 - u)v \mathbf{p}_{01} - uv \mathbf{p}_{11}. \end{aligned} \quad (3.5.2)$$

В (3.5.2) введем обозначения для отрезков, соединяющих концы кривых $c_0(u)$ и $c_1(u)$: $\mathbf{b}_0(v) = \mathbf{p}_{00}(1 - v) + \mathbf{p}_{01}v$ и $\mathbf{b}_1(v) = \mathbf{p}_{10}(1 - v) + \mathbf{p}_{11}v$. Если в качестве $\mathbf{b}_0(v)$ и $\mathbf{b}_1(v)$ использовать произвольные кривые, начинающиеся и оканчивающиеся в тех же точках, то получим **поверхность Кунса**. Радиус-вектор поверхности Кунса, построенной по четырем кривым $c_0(u)$, $c_1(u)$, $\mathbf{b}_0(v)$, $\mathbf{b}_1(v)$, описывается функцией

$$\begin{aligned} \mathbf{r}(u, v) = & (1 - v) \mathbf{c}_0(u) + v \mathbf{c}_1(u) + (1 - u) \mathbf{b}_0(v) + u \mathbf{b}_1(v) - \\ & - (1 - u)(1 - v) \mathbf{p}_{00} - u(1 - v) \mathbf{p}_{10} - (1 - u)v \mathbf{p}_{01} - uv \mathbf{p}_{11}. \end{aligned} \quad (3.5.3)$$

Поверхность Кунса приведена на рис. 3.5.3.

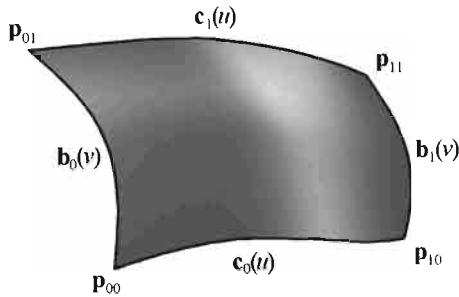


Рис. 3.5.3

В качестве примера построим поверхность Кунса по четырем дугам одного эллипса:

$$\begin{aligned} \mathbf{r}(u, v) = & \mathbf{p} + \\ & + a \left(1 - u - v - (1 - v) \cos \frac{\pi u}{2} + v \sin \frac{\pi u}{2} - (1 - u) \cos \frac{\pi v}{2} + u \sin \frac{\pi v}{2} \right) \mathbf{i}_1 + \\ & + b \left(u - v - (1 - v) \sin \frac{\pi u}{2} + v \cos \frac{\pi u}{2} + (1 - u) \sin \frac{\pi v}{2} - u \cos \frac{\pi v}{2} \right) \mathbf{i}_2, \\ & u \in [0, 1], v \in [0, 1], \end{aligned}$$

где \mathbf{p} – радиус-вектор центра; a и b – полуоси эллипса; $\mathbf{i}_1, \mathbf{i}_2$ – базисные векторы, лежащие в плоскости эллипса, рис. 3.5.4.

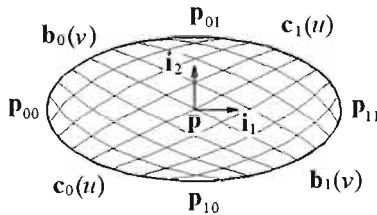


Рис. 3.5.4

В точках стыковки дуг эллипса частные производные радиуса-вектора поверхности по параметрам u и v параллельны.

Введем обозначения:

$$\begin{aligned} s(u, 0) = c_0(u), s(u, 1) = c_1(u), s(0, v) = b_0(v), s(1, v) = b_1(v), \\ s(0, 0) = p_{00}, s(1, 0) = p_{10}, s(0, 1) = p_{01}, s(1, 1) = p_{11}, \\ \mu_0(w) = 1 - w, \mu_1(t) = w. \end{aligned}$$

В результате векторная функция (3.5.3) примет вид

$$\begin{aligned} \mathbf{r}(u, v) = & [\mu_0(u) \quad \mu_1(u)] \cdot \begin{bmatrix} \mathbf{s}(0, v) \\ \mathbf{s}(1, v) \end{bmatrix} + [\mu_0(v) \quad \mu_1(v)] \cdot \begin{bmatrix} \mathbf{s}(u, 0) \\ \mathbf{s}(u, 1) \end{bmatrix} - \\ & - [\mu_0(u) \quad \mu_1(u)] \cdot \begin{bmatrix} \mathbf{s}(0, 0) & \mathbf{s}(0, 1) \\ \mathbf{s}(1, 0) & \mathbf{s}(1, 1) \end{bmatrix} \cdot \begin{bmatrix} \mu_0(v) \\ \mu_1(v) \end{bmatrix}. \end{aligned} \quad (3.5.4)$$

Поверхность (3.5.4) построена на четырех кривых, коэффициенты при которых являются линейными функциями. Такое линейное объединение четырех кривых называется *линейной поверхностью Кунса*. Функции $\mu_0(w)$ и $\mu_1(w)$ называются *функциями смещения*. В общем случае кривые поверхности Кунса могут не пересекаться, а точки \mathbf{p}_{00} , \mathbf{p}_{10} , \mathbf{p}_{01} и \mathbf{p}_{11} могут быть произвольными точками.

Из поверхностей (3.5.4) можно сконструировать составную поверхность, стыкуя их по граничным кривым, то есть делая граничные кривые общими. Для того чтобы на стыке поверхностей производная в перпендикулярном направлении к границе не претерпевала излом, добавим к описанию поверхности производные в перпендикулярном к границе направлении. То есть добавим к описанию поверхности ее частные производные вдоль граничных кривых $s_u(i, v)$, $s_v(u, j)$ и частные производные в углах поверхности $s_u(i, j)$, $s_v(i, j)$, $s_{uv}(i, j)$, где $i = 0, 1$, $j = 0, 1$. Линейные функции смещения заменим кубическими функциями смещения:

$$\begin{aligned} \alpha_0(w) &= 1 - 3w^2 + 2w^3, \\ \alpha_1(w) &= 3w^2 - 2w^3, \\ \beta_0(w) &= w - 2w^2 + w^3, \\ \beta_1(w) &= -w^2 + w^3. \end{aligned}$$

Радиус-вектор построенной таким образом поверхности описывается векторной функцией

$$\begin{aligned} \mathbf{r}(u, v) = & \quad \quad \quad (3.5.5) \\ = & [\alpha_0(u) \quad \alpha_1(u) \quad \beta_0(u) \quad \beta_1(u)] \cdot \begin{bmatrix} \mathbf{s}(0, v) \\ \mathbf{s}(1, v) \\ s_u(0, v) \\ s_u(1, v) \end{bmatrix} + [\alpha_0(v) \quad \alpha_1(v) \quad \beta_0(v) \quad \beta_1(v)] \cdot \begin{bmatrix} \mathbf{s}(u, 0) \\ \mathbf{s}(u, 1) \\ s_v(u, 0) \\ s_v(u, 1) \end{bmatrix} - \\ & - [\alpha_0(u) \quad \alpha_1(u) \quad \beta_0(u) \quad \beta_1(u)] \cdot \begin{bmatrix} \mathbf{s}(0, 0) & \mathbf{s}(0, 1) & s_u(0, 0) & s_u(0, 1) \\ \mathbf{s}(1, 0) & \mathbf{s}(1, 1) & s_u(1, 0) & s_u(1, 1) \\ s_{uv}(0, 0) & s_{uv}(0, 1) & s_{uv}(0, 0) & s_{uv}(0, 1) \\ s_{uv}(1, 0) & s_{uv}(1, 1) & s_{uv}(1, 0) & s_{uv}(1, 1) \end{bmatrix} \cdot \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix}. \end{aligned}$$

Поверхность (3.5.5) называется *кубической поверхностью Кунса*.

Форрест (A.R. Forrest) предложил обобщить поверхности (3.5.4) и (3.5.5) и использовать вместо линейных или кубических функций смещения некоторые *обобщенные функции смещения* $\alpha_i^k(u)$. Индекс k указывает на то, что обобщенная функция смещения умножается на векторную функцию, представляющую собой производную k -го порядка от граничной кривой. Производной нулевого порядка функции будем называть саму функцию. В терминах обобщенных функций смещения $\alpha_i(u) = \alpha_i^0(u)$, $\beta_i(u) = \alpha_i^1(u)$, $i = 0, 1$. Обобщенные функции смещения должны удовлетворять равенствам

$$\left. \frac{d^n \alpha_i^k}{du^n} \right|_{u=u_j} = \delta_i^j \delta_k^n = \begin{cases} 1, & \text{если } i = j \text{ и } k = n \\ 0, & \text{если } i \neq j \text{ или } k \neq n \end{cases}$$

Обобщенную поверхность Кунса опишем векторной функцией

$$\begin{aligned} \mathbf{r}(u, v) = & \sum_{i=0}^1 \sum_{k=0}^n \alpha_i^k(u) \mathbf{s}_n^{(k)}(i, v) + \sum_{j=0}^1 \sum_{l=0}^m \alpha_j^l(v) \mathbf{s}_m^{(l)}(u, j) - \\ & - \sum_{i=0}^1 \sum_{j=0}^1 \sum_{k=0}^n \sum_{l=0}^m \alpha_i^k(u) \alpha_j^l(v) \mathbf{s}_{mn}^{(kl)}(i, j), \\ & u \in [0, 1], v \in [0, 1], \end{aligned}$$

где используются заданные значения производных на краях поверхности

$$\begin{aligned} \mathbf{s}_n^{(k)}(i, v) &= \left. \frac{\partial^k \mathbf{r}(u, v)}{\partial u^k} \right|_{u=i}, \quad i = 0, 1, \\ \mathbf{s}_m^{(l)}(u, j) &= \left. \frac{\partial^l \mathbf{r}(u, v)}{\partial v^l} \right|_{v=j}, \quad j = 0, 1 \end{aligned}$$

и заданные значения производных в угловых точках параметрической области поверхности

$$\mathbf{s}_{mn}^{(kl)}(i, j) = \left. \frac{\partial^{k+l} \mathbf{r}(u, v)}{\partial u^k \partial v^l} \right|_{\substack{u=i \\ v=j}}, \quad i, j = 0, 1.$$

В качестве обобщенных функций смещения могут быть использованы полиномы.

Поверхность Гордона, построенная на сетке кривых, образованной семейством $\mathbf{c}_i(u)$, $i = 0, 1, 2, \dots, n$, $u_0 \leq u \leq u_n$ и семейством $\mathbf{b}_j(v)$, $j = 0, 1, 2, \dots, m$, $v_0 \leq v \leq v_m$, описывается векторной функцией

$$\mathbf{r}(u, v) = \sum_{i=0}^n L_i^n(v) \mathbf{c}_i(u) + \sum_{j=0}^m L_j^m(u) \mathbf{b}_j(v) - \sum_{j=0}^m \sum_{i=0}^n L_j^m(u) L_i^n(v) \mathbf{p}_j, \quad (3.5.6)$$

где $L_i^n(v)$ и $L_j^m(u)$ – коэффициенты Лагранжа (2.3.3); \mathbf{p}_j – точки пересечения кривых $\mathbf{p}_j = \mathbf{c}_i(u_j) = \mathbf{b}_j(v)$. Поверхность Гордона приведена на рис. 3.5.5.

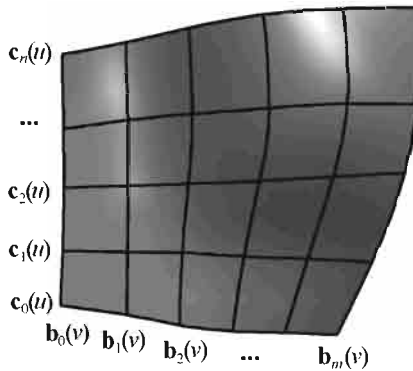


Рис. 3.5.5

Если все кривые какого-либо семейства являются циклически замкнутыми, то поверхность (3.5.6) будет циклически замкнута в соответствующем направлении.

Вместо функций $L_i^n(v)$ и $L_j^m(u)$ могут использоваться и другие сплайны. В качестве примера опишем поверхность типа (3.5.6), используя принцип построения поверхности Эрмита (3.4.3). Для параметров u и v поверхности найдем индекс j из условия $u_{j+1} \leq u \leq u_j$ и индекс i из условия $v_{i+1} \leq v \leq v_i$. Составную поверхность Кунса, находящуюся между кривыми $\mathbf{c}_i(u)$, $\mathbf{c}_{i+1}(u)$, $\mathbf{b}_j(v)$, $\mathbf{b}_{j+1}(v)$, опишем векторной функцией

$$\mathbf{r}(u, v) = F_i^T \cdot \begin{bmatrix} \mathbf{c}_i(u) \\ \mathbf{c}_{i+1}(u) \\ \mathbf{s}_v(u, v_j) \\ \mathbf{s}_v(u, v_{i+1}) \end{bmatrix} + F_j^T \cdot \begin{bmatrix} \mathbf{b}_j(v) \\ \mathbf{b}_{j+1}(v) \\ \mathbf{s}_u(u_j, v) \\ \mathbf{s}_u(u_{j+1}, v) \end{bmatrix} -$$

$$- F_j^T \cdot \begin{bmatrix} \mathbf{p}_i & \mathbf{p}_{i+1} & \mathbf{s}_v(u_j, v_j) & \mathbf{s}_v(u_j, v_{i+1}) \\ \mathbf{p}_{j+1} & \mathbf{p}_{j+2} & \mathbf{s}_v(u_{j+1}, v_j) & \mathbf{s}_v(u_{j+1}, v_{i+1}) \\ \mathbf{s}_u(u_j, v_j) & \mathbf{s}_u(u_j, v_{i+1}) & \mathbf{s}_u(u_j, v_j) & \mathbf{s}_u(u_j, v_{i+1}) \\ \mathbf{s}_u(u_{j+1}, v_j) & \mathbf{s}_u(u_{j+1}, v_{i+1}) & \mathbf{s}_u(u_{j+1}, v_j) & \mathbf{s}_u(u_{j+1}, v_{i+1}) \end{bmatrix} \cdot F_i, \quad (3.5.7)$$

$$\text{где } F_i = \begin{bmatrix} \alpha_0(w) \\ \alpha_1(w) \\ (v_{i+1} - v_i)\beta_0(w) \\ (v_{i+1} - v_i)\beta_1(w) \end{bmatrix}, \quad F_j = \begin{bmatrix} \alpha_0(t) \\ \alpha_1(t) \\ (u_{j+1} - u_j)\beta_0(t) \\ (u_{j+1} - u_j)\beta_1(t) \end{bmatrix} - \text{матрицы-столбцы; } t = \frac{u - u_j}{u_{j+1} - u_j},$$

$w = \frac{v - v_i}{v_{i+1} - v_i}$ – местные параметры поверхности, изменяющиеся от 0 до 1;

$\mathbf{p}_{ji} = \mathbf{c}(u_j) = \mathbf{b}_j(v_i)$ – точки пересечения кривых; $\mathbf{s}_u(u, v)$ – частная производная радиуса-вектора поверхности по первому параметру; $\mathbf{s}_v(u, v)$ – частная производная радиуса-вектора поверхности по второму параметру; $\mathbf{s}_{uv}(u, v)$ – смешанная производная радиуса-вектора поверхности по первому и второму параметрам. Производные поверхности $\mathbf{s}_v(u, v)$, $\mathbf{s}_u(u, v)$ и $\mathbf{s}_{uv}(u, v)$ на кривых сетки вычислим по соседним кривым аналогично тому, как по соседним точкам вычисляются производные составного сплайна Эрмита (2.3.10) в опорных точках:

$$\begin{aligned} \mathbf{s}_u(u_j, v) &= \\ &= \mathbf{b}_{j-1}(v) \frac{u_j - u_{j+1}}{(u_{j-1} - u_j)(u_{j-1} - u_{j+1})} + \mathbf{b}_j(v) \frac{2u_j - u_{j-1} - u_{j+1}}{(u_j - u_{j-1})(u_j - u_{j+1})} + \mathbf{b}_{j+1}(v) \frac{u_j - u_{j-1}}{(u_{j+1} - u_{j-1})(u_{j+1} - u_j)}, \\ \mathbf{s}_v(u, v_i) &= \\ &= \mathbf{c}_{i-1}(u) \frac{v_i - v_{i+1}}{(v_{i-1} - v_i)(v_{i-1} - v_{i+1})} + \mathbf{c}_i(u) \frac{2v_i - v_{i-1} - v_{i+1}}{(v_i - v_{i-1})(v_i - v_{i+1})} + \mathbf{c}_{i+1}(u) \frac{v_i - v_{i-1}}{(v_{i+1} - v_{i-1})(v_{i+1} - v_i)} \end{aligned}$$

для внутренних кривых ($j = 1, 2, \dots, m-1$, $i = 1, 2, \dots, n-1$) и

$$\begin{aligned} \mathbf{s}_u(u_0, v) &= \frac{3\mathbf{b}_1(v) - \mathbf{b}_0(v)}{2} \frac{1}{u_1 - u_0} - \frac{1}{2} \mathbf{s}_u(u_1, v), \quad \mathbf{s}_v(u_m, v) = \frac{3\mathbf{b}_m(v) - \mathbf{b}_{m-1}(v)}{2} \frac{1}{u_m - u_{m-1}} - \frac{1}{2} \mathbf{s}_v(u_{m-1}, v), \\ \mathbf{s}_v(u, v_0) &= \frac{3\mathbf{c}_1(u) - \mathbf{c}_0(u)}{2} \frac{1}{v_1 - v_0} - \frac{1}{2} \mathbf{s}_v(u, v_1), \\ \mathbf{s}_v(u, v_n) &= \frac{3\mathbf{c}_n(u) - \mathbf{c}_{n-1}(u)}{2} \frac{1}{v_n - v_{n-1}} - \frac{1}{2} \mathbf{s}_v(u, v_{n-1}) \end{aligned}$$

для крайних кривых. По частным производным $\mathbf{s}_v(u, v)$ и $\mathbf{s}_u(u, v)$ вычислим смешанные производные:

$$\mathbf{s}_{uv}(u_j, v_i) = \frac{1}{2} \left(\left. \frac{\partial \mathbf{s}_u(u, v)}{\partial v} \right|_{u=u_j} + \left. \frac{\partial \mathbf{s}_v(u, v)}{\partial u} \right|_{u=u_j, v=v_i} \right).$$

Предположим, что в формуле (3.5.7) кривые $\mathbf{c}(u)$ и $\mathbf{b}_j(v)$ сетки представляют собой кубические сплайны Эрмита (2.3.10), построенные по набору точек \mathbf{p}_{ji} , и определяются зависимостями

$$\mathbf{c}_j(u) = F_j^T \cdot \begin{bmatrix} \mathbf{p}_j \\ \mathbf{p}_{j+1} \\ \mathbf{s}_u(u_j, v_j) \\ \mathbf{s}_u(u_{j+1}, v_j) \end{bmatrix}, \quad \mathbf{b}_j(v) = F_j^T \cdot \begin{bmatrix} \mathbf{p}_j \\ \mathbf{p}_{j+1} \\ \mathbf{s}_v(u_j, v_j) \\ \mathbf{s}_v(u_j, v_{j+1}) \end{bmatrix}.$$

После подстановки этих зависимостей в (3.5.7) получим три одинаковых матричных члена, два из них со знаком плюс и один со знаком минус. В результате формула (3.5.7) примет вид

$$\mathbf{r}(u, v) = F_j^T \cdot \begin{bmatrix} \mathbf{p}_j & \mathbf{p}_{j+1} & \mathbf{s}_v(u_j, v_j) & \mathbf{s}_v(u_j, v_{j+1}) \\ \mathbf{p}_{j+1} & \mathbf{p}_{j+1} & \mathbf{s}_v(u_{j+1}, v_j) & \mathbf{s}_v(u_{j+1}, v_{j+1}) \\ \mathbf{s}_u(u_j, v_j) & \mathbf{s}_u(u_j, v_{j+1}) & \mathbf{s}_u(u_j, v_j) & \mathbf{s}_u(u_j, v_{j+1}) \\ \mathbf{s}_u(u_{j+1}, v_j) & \mathbf{s}_u(u_{j+1}, v_{j+1}) & \mathbf{s}_u(u_{j+1}, v_j) & \mathbf{s}_u(u_{j+1}, v_{j+1}) \end{bmatrix} \cdot F_j, \quad (3.5.8)$$

где использованы обозначения

$$\begin{aligned} \mathbf{s}_u(u_j, v_j) &= \\ &= \mathbf{p}_{j+1} \frac{u_j - u_{j+1}}{(u_{j+1} - u_j)(u_{j+1} - u_{j+1})} + \mathbf{p}_j \frac{2u_j - u_{j+1} - u_{j+1}}{(u_j - u_{j+1})(u_j - u_{j+1})} + \mathbf{p}_{j+1} \frac{u_j - u_{j+1}}{(u_{j+1} - u_{j+1})(u_{j+1} - u_j)}, \\ \mathbf{s}_v(u_j, v_j) &= \\ &= \mathbf{p}_{j+1} \frac{v_j - v_{j+1}}{(v_{j+1} - v_j)(v_{j+1} - v_{j+1})} + \mathbf{p}_j \frac{2v_j - v_{j+1} - v_{j+1}}{(v_j - v_{j+1})(v_j - v_{j+1})} + \mathbf{p}_{j+1} \frac{v_j - v_{j+1}}{(v_{j+1} - v_{j+1})(v_{j+1} - v_j)}, \\ \mathbf{s}_{u,v}(u_j, v_j) &= \begin{bmatrix} \frac{u_j - u_{j+1}}{(u_{j+1} - u_j)(u_{j+1} - u_{j+1})} \\ \frac{2u_j - u_{j+1} - u_{j+1}}{(u_j - u_{j+1})(u_j - u_{j+1})} \\ \frac{u_j - u_{j+1}}{(u_{j+1} - u_{j+1})(u_{j+1} - u_j)} \end{bmatrix}^T \cdot \begin{bmatrix} \mathbf{p}_{j+1} & \mathbf{p}_{j+1} & \mathbf{p}_{j+1} \\ \mathbf{p}_{j+1} & \mathbf{p}_j & \mathbf{p}_{j+1} \\ \mathbf{p}_{j+1} & \mathbf{p}_{j+1} & \mathbf{p}_{j+1} \end{bmatrix} \cdot \begin{bmatrix} \frac{v_j - v_{j+1}}{(v_{j+1} - v_j)(v_{j+1} - v_{j+1})} \\ \frac{2v_j - v_{j+1} - v_{j+1}}{(v_j - v_{j+1})(v_j - v_{j+1})} \\ \frac{v_j - v_{j+1}}{(v_{j+1} - v_{j+1})(v_{j+1} - v_j)} \end{bmatrix}. \end{aligned}$$

Поверхность (3.5.8) полностью определяется точками \mathbf{p}_j .

3.6. Поверхность на сетке точек

Построим **бикубическую поверхность** путем модернизации билинейной поверхности (3.5.1), построенной на четырех точках \mathbf{p}_{00} , \mathbf{p}_{10} , \mathbf{p}_{01} и \mathbf{p}_{11} . Пусть поверхность является кубической функцией как первого параметра, так и второго параметра. Пусть область определения параметров поверхности изменяется в пределах $0 \leq u \leq 1$ и $0 \leq v \leq 1$. Потребуем, чтобы в

углах радиус-вектор поверхности совпадал с точками $\mathbf{p}_{00}, \mathbf{p}_{10}, \mathbf{p}_{01}, \mathbf{p}_{11}$, чтобы в углах частные производные радиуса-вектора по параметру u принимали значения $\mathbf{u}_{00}, \mathbf{u}_{10}, \mathbf{u}_{01}, \mathbf{u}_{11}$, чтобы в углах частные производные радиуса-вектора по параметру v принимали значения $\mathbf{v}_{00}, \mathbf{v}_{10}, \mathbf{v}_{01}, \mathbf{v}_{11}$, чтобы в углах вторые смешанные производные радиуса-вектора по параметрам u и v принимали значения $\mathbf{t}_{00}, \mathbf{t}_{10}, \mathbf{t}_{01}, \mathbf{t}_{11}$. Таким образом, для бикубической поверхности $\mathbf{r}(u, v)$ должны выполняться условия:

$$\begin{aligned} \mathbf{r}(0,0) &= \mathbf{p}_{00}, \mathbf{r}(1,0) = \mathbf{p}_{10}, \mathbf{r}(0,1) = \mathbf{p}_{01}, \mathbf{r}(1,1) = \mathbf{p}_{11}, \\ \mathbf{r}_u(0,0) &= \mathbf{u}_{00}, \mathbf{r}_u(1,0) = \mathbf{u}_{10}, \mathbf{r}_u(0,1) = \mathbf{u}_{01}, \mathbf{r}_u(1,1) = \mathbf{u}_{11}, \\ \mathbf{r}_v(0,0) &= \mathbf{v}_{00}, \mathbf{r}_v(1,0) = \mathbf{v}_{10}, \mathbf{r}_v(0,1) = \mathbf{v}_{01}, \mathbf{r}_v(1,1) = \mathbf{v}_{11}, \\ \mathbf{r}_{uv}(0,0) &= \mathbf{t}_{00}, \mathbf{r}_{uv}(1,0) = \mathbf{t}_{10}, \mathbf{r}_{uv}(0,1) = \mathbf{t}_{01}, \mathbf{r}_{uv}(1,1) = \mathbf{t}_{11}. \end{aligned}$$

Заданным условиям удовлетворяет поверхность

$$\mathbf{r}(u, v) = \begin{bmatrix} \mathbf{p}_{00} & \mathbf{p}_{01} & \mathbf{v}_{00} & \mathbf{v}_{01} \\ \mathbf{p}_{10} & \mathbf{p}_{11} & \mathbf{v}_{10} & \mathbf{v}_{11} \\ \mathbf{u}_{00} & \mathbf{u}_{01} & \mathbf{t}_{00} & \mathbf{t}_{01} \\ \mathbf{u}_{10} & \mathbf{u}_{11} & \mathbf{t}_{10} & \mathbf{t}_{11} \end{bmatrix} \cdot \begin{bmatrix} \alpha_0(v) \\ \alpha_1(v) \\ \beta_0(v) \\ \beta_1(v) \end{bmatrix}, \quad (3.6.1)$$

где $\alpha_0(w) = 1 - 3w^2 + 2w^3$, $\alpha_1(w) = 3w^2 - 2w^3$, $\beta_0(w) = w - 2w^2 + w^3$; $\beta_1(w) = -w^2 + w^3$.

Построим бикубическую поверхность, состоящую из нескольких стыкующихся по краям и углам поверхностей (3.6.1). Пусть имеется совокупность контрольных точек \mathbf{p}_{ij} , $i = 0, 1, \dots, n$, $j = 0, 1, \dots, m$, условно расположенных в виде $n + 1$ строк и $m + 1$ столбцов по $n + 1$ точек в каждом столбце и по $m + 1$ точек в каждой строке. Индексы точки означают, что точка \mathbf{p}_{ij} расположена в i -м ряду j -й по счету. Нумеровать ряды будем от 0 до n , а точки в них будем нумеровать от 0 до m . По расположению строк контрольных точек построим совокупность узлов v_i , $i = 0, 1, \dots, n$ так, чтобы изменение узловых значений было бы пропорционально расстоянию между соседними рядами контрольных точек:

$$\frac{v_{i+1} - v_i}{v_i - v_{i-1}} \approx \frac{\sum_{j=0}^{m-1} |\mathbf{p}_{i+1,j} - \mathbf{p}_{i,j}|}{\sum_{j=1}^m |\mathbf{p}_{i,j} - \mathbf{p}_{i-1,j}|}.$$

По расположению столбцов контрольных точек построим совокупность узлов u_j , $j = 0, 1, \dots, m$, чтобы изменение узловых значений было пропорционально расстоянию между соседними столбцами контрольных точек:

$$\frac{u_{j+1} - u_j}{u_j - u_{j-1}} \approx \frac{\sum_{i=0}^{n-1} |\mathbf{p}_{j+1} - \mathbf{p}_i|}{\sum_{i=1}^j |\mathbf{p}_j - \mathbf{p}_{i-1}|}.$$

Потребуем, чтобы при значениях параметров $u = u_j$ и $v = v_i$ поверхность $\mathbf{r}(u, v)$ проходила через точки \mathbf{p}_j , то есть $\mathbf{r}(u_j, v_i) = \mathbf{p}_j$. Для построения бикубической поверхности требуется определить ее частные производные в контрольных точках $\mathbf{u}_j, \mathbf{v}_j, \mathbf{t}_j$.

Частные производные \mathbf{u}_j и \mathbf{v}_j радиуса-вектора поверхности в контрольных точках вычислим по аналогии с вычислением производных в контрольных точках сплайна Эрмита. Для этого построим векторный полином Лагранжа (2.3.4) второй степени, который при значениях параметра v_{i-1}, v_i, v_{i+1} проходит по трем соседним точкам $\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1}$, вычислим производную этого полинома при v_i и возьмем ее в качестве частной производной по второму параметру радиуса-вектора поверхности в точке $v_i, i = 1, 2, \dots, n-1$:

$$\begin{aligned} \mathbf{v}_i &\equiv \frac{\partial \mathbf{r}(u_j, v_i)}{\partial v} = \\ &= \mathbf{p}_{i-1} \frac{v_i - v_{i+1}}{(v_{i-1} - v_i)(v_{i-1} - v_{i+1})} + \mathbf{p}_i \frac{2v_i - v_{i-1} - v_{i+1}}{(v_i - v_{i-1})(v_i - v_{i+1})} + \mathbf{p}_{i+1} \frac{v_i - v_{i-1}}{(v_{i+1} - v_{i-1})(v_{i+1} - v_i)}. \end{aligned}$$

Аналогично вычислим частные производные по первому параметру радиуса-вектора поверхности в точках $u_j, j = 1, 2, \dots, m-1$:

$$\begin{aligned} \mathbf{u}_j &\equiv \frac{\partial \mathbf{r}(u_j, v_i)}{\partial u} = \\ &= \mathbf{p}_{u-1} \frac{u_j - u_{j+1}}{(u_{j-1} - u_j)(u_{j-1} - u_{j+1})} + \mathbf{p}_j \frac{2u_j - u_{j-1} - u_{j+1}}{(u_j - u_{j-1})(u_j - u_{j+1})} + \mathbf{p}_{u+1} \frac{u_j - u_{j-1}}{(u_{j+1} - u_{j-1})(u_{j+1} - u_j)}. \end{aligned}$$

Производные в крайних контрольных точках поверхности получим из условия обращения в нуль второй производной радиуса-вектора поверхности по параметру v

$$\frac{\partial \mathbf{r}(u_j, v_0)}{\partial v} = \frac{3 \mathbf{p}_{1j} - \mathbf{p}_{0j}}{2 v_1 - v_0} - \frac{1}{2} \frac{\partial \mathbf{r}(u_j, v_1)}{\partial v}, \quad \frac{\partial \mathbf{r}(u_j, v_n)}{\partial v} = \frac{3 \mathbf{p}_j - \mathbf{p}_{n-1j}}{2 v_n - v_{n-1}} - \frac{1}{2} \frac{\partial \mathbf{r}(u_j, v_{n-1})}{\partial v}$$

и по параметру u

$$\begin{aligned} \frac{\partial \mathbf{r}(u_0, v_i)}{\partial u} &= \frac{3 \mathbf{p}_{i1} - \mathbf{p}_{i0}}{2 u_1 - u_0} - \frac{1}{2} \frac{\partial \mathbf{r}(u_1, v_i)}{\partial u}, \\ \frac{\partial \mathbf{r}(u_m, v_i)}{\partial u} &= \frac{3 \mathbf{p}_{im} - \mathbf{p}_{im-1}}{2 u_m - u_{m-1}} - \frac{1}{2} \frac{\partial \mathbf{r}(u_{m-1}, v_i)}{\partial u} \end{aligned}$$

в крайних контрольных точках.

Частные производные t_{ij} радиуса-вектора поверхности в контрольных точках вычислим по конечно разностным формулам:

$$t_{ij} \equiv \frac{\partial^2 \mathbf{r}(u_j, v_i)}{\partial u \partial v} = \frac{\mathbf{p}_{i+1, j+1} - \mathbf{p}_{i-1, j+1} - \mathbf{p}_{i+1, j-1} + \mathbf{p}_{i-1, j-1}}{(u_{i+1} - u_{i-1})(v_{j+1} - v_{j-1})}.$$

Если при вычислении производных t_{ij} значение индекса $i - 1$ или $j - 1$ меньше нуля, то положим его равным нулю, если значение индекса $i + 1$ больше $n + 1$, то положим его равным $n + 1$, если значение индекса $j + 1$ больше $m + 1$, то положим его равным $m + 1$.

Для вычисления радиуса-вектора поверхности, построенной на сетке точек, выполним следующие действия. По заданным параметрам u и v найдем индексы i и j угловой контрольной точки \mathbf{p}_{ij} рассматриваемого участка поверхности из условий $u_j \leq u \leq u_{j+1}$ и $v_i \leq v \leq v_{i+1}$ и вычислим радиус-вектор поверхности по формуле

$$\begin{aligned} \mathbf{r}(u, v) = \\ = [\alpha_0(x) \quad \alpha_1(x) \quad \beta_0(y) \quad \beta_1(y)] \cdot \begin{bmatrix} \mathbf{p}_{ij} & \mathbf{p}_{i+1j} & \mathbf{v}_{ij} & \mathbf{v}_{i+1j} \\ \mathbf{p}_{i+1j} & \mathbf{p}_{i+1j+1} & \mathbf{v}_{i+1j} & \mathbf{v}_{i+1j+1} \\ \mathbf{u}_{ij} & \mathbf{u}_{i+1j} & \mathbf{t}_{ij} & \mathbf{t}_{i+1j} \\ \mathbf{u}_{i+1j} & \mathbf{u}_{i+1j+1} & \mathbf{t}_{i+1j} & \mathbf{t}_{i+1j+1} \end{bmatrix} \cdot \begin{bmatrix} \alpha_0(y) \\ \alpha_1(y) \\ \beta_0(y) \\ \beta_1(y) \end{bmatrix}, \end{aligned} \quad (3.6.2)$$

где $x = \frac{u - u_j}{u_{j+1} - u_j}$; $y = \frac{v - v_i}{v_{i+1} - v_i}$.

Следует заметить, что поверхность (3.6.2) может иметь складки и самопересечения в местах, где расстояние между соседними контрольными точками не пропорционально разности значений соответствующих параметров и где частные производные t_{ij} принимают большие значения.

3.7. Поверхность Безье

Пусть имеется совокупность точек \mathbf{p}_{ij} , $i = 0, 1, \dots, n, j = 0, 1, \dots, m$, условно расположенных в узлах прямоугольной сетки, то есть условно расположенных в виде $n + 1$ рядов по $m + 1$ точек в каждом ряду. Индексы точки означают, что точка \mathbf{p}_{ij} расположена j -ой по счету в i -ом ряду. Нумеровать ряды будем от 0 до n , а точки в них будем нумеровать от 0 до m . **Поверхность Безье**, построенная на совокупности точек \mathbf{p}_{ij} , описывается векторной функцией

$$\mathbf{r}(u, v) = \sum_{i=0}^n \sum_{j=0}^m B_i^n(v) B_j^m(u) \mathbf{p}_{ij}, \quad u \in [0, 1], \quad v \in [0, 1], \quad (3.7.1)$$

где $B_i^n(v) = \frac{n!}{(n-i)!i!} v^i (1-v)^{n-i}$ и $B_j^m(u) = \frac{m!}{(m-j)!j!} u^j (1-u)^{m-j}$ — функции базиса Бернштейна. В соответствии с (2.4.3) выполняется равенство

$$\sum_{i=0}^n \sum_{j=0}^m B_i^n(v) B_j^m(u) = 1. \quad (3.7.2)$$

Точки \mathbf{p}_{ij} называются *контрольными точками*. Сечения поверхности (3.7.1) по линиям $u = \text{const}$ или $v = \text{const}$ представляют собой кривые Безье степени m и n соответственно. Будем говорить, что по первому параметру поверхность (3.7.1) имеет степень m , а по второму параметру поверхность имеет степень n . Область изменения параметров поверхности представляет собой квадрат со стороной, равной единице.

Если представить, что на каждой из четырех соседних точек $\mathbf{p}_{ij}, \mathbf{p}_{j+1}, \mathbf{p}_{i+1}, \mathbf{p}_{i+1+j+1}$ построена билинейная поверхность, то мы получим многогранник, который будем называть *контрольным многогранником* или *характеристическим многогранником*. Контрольный многогранник дает общее представление о поверхности Безье. Поверхность Безье вместе со своим контрольным многогранником приведена на рис. 3.7.1. Так как из коэффициентов Бернштейна только нулевой и последний принимают максимально возможные значения, а остальные при этом равны нулю, то поверхность Безье проходит только через угловые контрольные точки сетки: $\mathbf{p}_{00}, \mathbf{p}_{0m}, \mathbf{p}_{n0}, \mathbf{p}_{nm}$.

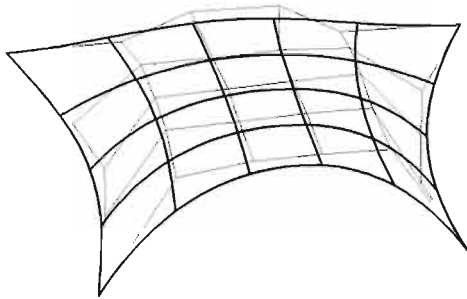


Рис. 3.7.1

Припишем каждой контрольной точке \mathbf{p}_{ij} вес w_{ij} соответственно. **Рациональная поверхность Безье**, построенная на совокупности точек \mathbf{p}_{ij} , описывается векторной функцией

$$\mathbf{r}(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m B_i^n(v) B_j^m(u) w_{ij} \mathbf{p}_{ij}}{\sum_{i=0}^n \sum_{j=0}^m B_i^n(v) B_j^m(u) w_{ij}}. \quad (3.7.3)$$

В терминах однородных координат радиус-вектор рациональной поверхности Безье описывается равенством, по форме совпадающим с равенством (3.7.1), в которой \mathbf{p}_j – расширенные векторы (1.4.8) контрольных точек поверхности. У рациональной поверхности играют роль не абсолютные значения весов точек, а их относительные значения. Чем больше вес контрольной точки по отношению к весам других контрольных точек, тем ближе к этой точке подходит поверхность. Если веса всех контрольных точек равны, то рациональная поверхность Безье примет вид (3.7.1).

Радиус-вектор рациональной поверхности вычисляется как частное от деления двух функций параметров u и v , поэтому при вычислении производных рациональной поверхности правую часть (3.7.3) следует рассматривать как сложную функцию. Если условно обозначим радиус-вектор рациональной поверхности как $\mathbf{r} = \frac{w\mathbf{r}}{w}$, то частные производные радиуса-вектора рациональной поверхности определяются формулами

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{1}{w} \frac{\partial (w\mathbf{r})}{\partial u} - \frac{w\mathbf{r}}{w^2} \frac{\partial w}{\partial u}, \quad \frac{\partial \mathbf{r}}{\partial v} = \frac{1}{w} \frac{\partial (w\mathbf{r})}{\partial v} - \frac{w\mathbf{r}}{w^2} \frac{\partial w}{\partial v}.$$

Производные более высокого порядка вычисляются аналогично.

Рациональной поверхностью Безье можно описать некоторую часть поверхности второго порядка, а также часть поверхности тора. Для этого проще всего использовать рациональную поверхность Безье второй степени, построенную по девяти контрольным точкам:

$$\mathbf{r}(u, v) = \frac{\begin{bmatrix} (1-v)^2 & 2v(1-v) & v^2 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{p}_{00} & w_u \mathbf{p}_{01} & \mathbf{p}_{02} \\ w_v \mathbf{p}_{10} & w_v w_u \mathbf{p}_{11} & w_v \mathbf{p}_{12} \\ \mathbf{p}_{20} & w_v \mathbf{p}_{21} & \mathbf{p}_{22} \end{bmatrix} \cdot \begin{bmatrix} (1-u)^2 \\ 2u(1-u) \\ u^2 \end{bmatrix}}{\left((1-u)^2 + 2u(1-u)w_u + u^2 \right) \cdot \left((1-v)^2 + 2v(1-v)w_v + v^2 \right)}. \quad (3.7.4)$$

Веса угловых точек равны единице, а для остальных точек введены веса w_u и w_v , имеющие индексы параметрических направлений поверхности. Тип поверхности, получаемой по формуле (3.7.4), зависит от положения контрольных точек и от весов w_u и w_v . Если сечения поверхности (3.7.4) по линиям $u = \text{const}$ или $v = \text{const}$ представляют собой дуги окружности с углами раствора α_u и α_v соответственно, то $w_u = \cos(\alpha_u / 2)$ и $w_v = \cos(\alpha_v / 2)$.

Составная поверхность Безье позволяет преодолеть жесткую связь степени гладкости поверхности с количеством контрольных точек, по которым она построена. Составную поверхность Безье получим в результате стыковки отдельных поверхностей Безье, имеющих вдоль стыкуемых краев общие контрольные точки и одинаковые степени. Для того

чтобы отдельные поверхности Безье гладко стыковались друг с другом, нужно чтобы отрезки контрольных многогранников, прилегающие к общим точкам стыкуемых поверхностей, лежали в одной плоскости.

3.8. B -поверхность

Построим неоднородную рациональную поверхность на основе B -сплайнов так же, как мы построили поверхность Безье на основе функций Бернштейна. Пусть имеется совокупность точек \mathbf{p}_{ij} , $i = 0, 1, \dots, n$, $j = 0, 1, \dots, m$, условно расположенных в узлах прямоугольной сетки в виде $n + 1$ рядов по $m + 1$ точек в каждом ряду. Неоднородную рациональную поверхность на основе B -сплайнов k -го порядка по первому параметру и l -го порядка по второму параметру опишем векторной функцией:

$$\mathbf{r}(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_{ij} \mathbf{p}_{ij}}{\sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_{ij}}, \quad (3.8.1)$$

где $N_i^l(v)$ и $N_j^k(u)$ – B -сплайны (2.8.1). Поверхность (3.8.1) называют **NURBS поверхностью** (Non-Uniform Rational B -Spline surface) или **B -поверхностью**.

Каждый из B -сплайнов k -го порядка $N_j^k(u)$ построен на неубывающей последовательности из $k + 2$ узлов $u_j, u_{j+1}, \dots, u_{j+k+1}$, всего для построения совокупности $m + 1$ B -сплайнов k -го порядка требуется $m + k + 2$ узлов u_j в случае незамкнутой поверхности и $m + 2k + 2$ узлов в случае циклически замкнутой поверхности. Каждый из B -сплайнов l -го порядка $N_i^l(v)$ построен на неубывающей последовательности из $l + 2$ узлов $v_i, v_{i+1}, \dots, v_{i+l+1}$, всего для построения совокупности $n + 1$ B -сплайнов l -го порядка требуется $n + l + 2$ узлов v_i в случае незамкнутой поверхности и $n + 2l + 2$ узлов в случае циклически замкнутой поверхности. Как следует из (3.8.1), B -поверхность имеет две неубывающие последовательности узлов, одну – для первого параметра, а другую – для второго параметра.

Сечения поверхности (3.8.1) по линиям $u = \text{const}$ или $v = \text{const}$ представляют собой B -кривые степени k и l , соответственно. Будем говорить, что по первому параметру поверхность (3.8.1) имеет порядок k , а по второму параметру поверхность имеет порядок l . Область изменения параметров B -поверхности представляет собой прямоугольник: $u_k \leq u \leq u_{m+k+1}$, $v_l \leq v \leq v_{n+l+1}$. Таким образом, $u_{\min} = u_k$, $u_{\max} = u_{m+k+1}$, $v_{\min} = v_l$, $v_{\max} = v_{n+l+1}$.

Пусть нумерация узлов начинается с нуля. Для незамкнутой по первому параметру B -поверхности, как правило, используются B -сплайны $N_j^k(u)$, построенные на неубывающей последовательности узлов, в которой первые и последние $k + 1$ узлов совпадают: $u_0 = u_1 = \dots = u_k$ и

$u_{m+1} = u_{m+2} = \dots = u_{m+k+1}$. Для незамкнутой по второму параметру B -поверхности, как правило, используются B -сплайны $N'_i(v)$, построенные на неубывающей последовательности узлов, в которой первые и последние $l + 1$ узлов совпадают: $v_0 = v_1 = \dots = v_l$ и $v_{n+1} = v_{n+2} = \dots = v_{n+l+1}$. Для циклически замкнутой B -поверхности по первому (второму) параметру соответствующая последовательность узлов должна отражать замкнутость: первые $2k$ ($2l$) узлов должны идти через интервалы, повторяющие интервалы, через которые идут последние $2k$ ($2l$) узлов.

Когда все контрольные точки B -поверхности имеют равные веса, то в силу равенства (2.8.11) формула (3.8.1) для вычисления радиуса-вектора поверхности на основе B -сплайнов примет вид

$$\mathbf{r}(u, v) = \sum_{i=0}^n \sum_{j=0}^m N'_i(v) N_j^k(u) \mathbf{p}_j. \quad (3.8.2)$$

Если использовать расширенные векторы, то равенство (3.8.1) примет вид, совпадающий по форме с равенством (3.8.2), в которой \mathbf{p}_j – расширенные векторы (1.4.8) контрольных точек поверхности (3.8.1).

Если на каждых четырех соседних контрольных точках построить билинейные поверхности, то мы получим некоторый *контрольный многогранник* поверхности. Если построить B -поверхность на базе B -сплайнов первого порядка, то она совпадает со своим контрольным многогранником.

Для вычисления радиуса-вектора B -поверхности используем формулы (2.9.3) и (2.9.4). По значению параметра u из условия $u_j \leq u < u_{j+1}$ определим номер j отличного от нуля ненормированного B -сплайна нулевого порядка:

$$M_j^0(u) = \frac{1}{u_{j+1} - u_j}.$$

Далее, используя рекуррентное соотношение Кокса–Де Бура

$$M_r^k(u) = \frac{u_{r+k+1} - u}{u_{r+k+1} - u_r} M_{r+1}^{k-1}(u) + \frac{u - u_r}{u_{r+k+1} - u_r} M_r^{k-1}(u),$$

последовательно вычислим все отличные от нуля при данном параметре u B -сплайны до k -го порядка включительно: $M_{j-k}^k(u), M_{j-k+1}^k(u), \dots, M_j^k(u)$. Далее B -сплайны k -го порядка нормируем:

$$N_r^k(u) = (u_{r+k+1} - u_r) M_r^k(u), \quad r = j-k, j-k+1, \dots, j.$$

Аналогично по значению параметра v из условия $v_i \leq v < v_{i+1}$ определим номер i отличного от нуля ненормированного B -сплайна нулевого порядка:

$$M_i^0(v) = \frac{1}{v_{i+1} - v_i}.$$

Используя рекуррентное соотношение Кокса–Де Бура

$$M_q^l(v) = \frac{v_{q+l+1} - v}{v_{q+l+1} - v_q} M_{q+l}^{l-1}(v) + \frac{v - v_q}{v_{q+l+1} - v_q} M_q^{l-1}(v),$$

последовательно вычислим все отличные от нуля при данном параметре v B -сплайны до l -го порядка включительно: $M_{i-l}^l(v), M_{i-l+1}^l(v), \dots, M_i^l(v)$, затем нормируем их:

$$N_q^l(v) = (v_{q+l+1} - v_q) M_q^l(v), \quad q = i-l, i-l+1, \dots, i.$$

Так как в общем случае только часть B -сплайнов отлична от нуля, радиус-вектор точки B -поверхности (3.8.1) для параметров $u_j \leq u < u_{j+1}$ и $v_l \leq v < v_{l+1}$ вычислим по формуле

$$\mathbf{r}(u, v) = \frac{\sum_{q=i-l}^i \sum_{r=j-k}^j N_q^l(v) N_r^k(u) w_{qr} \mathbf{p}_{qr}}{\sum_{q=i-l}^i \sum_{r=j-k}^j N_q^l(v) N_r^k(u) w_{qr}}.$$

Одновременно с вычислением B -сплайнов может проводиться вычисление их производных с помощью соотношений (2.8.15).

Частные производные поверхности вычислим дифференцированием числителя и знаменателя правой части равенства (3.8.1) с учетом (2.8.15) и (2.8.5):

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{\partial}{\partial u} \left(\frac{w \mathbf{r}}{w} \right) = \frac{1}{w} \frac{\partial (w \mathbf{r})}{\partial u} - \frac{w \mathbf{r}}{w^2} \frac{\partial w}{\partial u},$$

$$\frac{\partial \mathbf{r}}{\partial v} = \frac{\partial}{\partial v} \left(\frac{w \mathbf{r}}{w} \right) = \frac{1}{w} \frac{\partial (w \mathbf{r})}{\partial v} - \frac{w \mathbf{r}}{w^2} \frac{\partial w}{\partial v},$$

где $w \mathbf{r} = \sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_{ij} \mathbf{p}_{ij}$ – числитель правой части (3.8.1);

$w = \sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_{ij}$ – знаменатель правой части (3.8.1);

$\frac{\partial (w \mathbf{r})}{\partial u} = k \sum_{i=0}^n \sum_{j=1}^m N_i^l(v) N_j^{k-1}(u) \frac{w_{ij} \mathbf{p}_{ij} - w_{i,j-1} \mathbf{p}_{i,j-1}}{u_{j,k} - u_j}$ – производная по u числителя;

$\frac{\partial w}{\partial u} = k \sum_{i=0}^n \sum_{j=1}^m N_i^l(v) N_j^{k-1}(u) \frac{w_{ij} - w_{i,j-1}}{u_{j,k} - u_j}$ – производная по u знаменателя;

$\frac{\partial(w\mathbf{r})}{\partial v} = l \sum_{i=1}^n \sum_{j=0}^m N_i^{l-1}(v) N_j^k(u) \frac{w_j \mathbf{p}_j - w_{i-1} \mathbf{p}_{i-1}}{v_{i+1} - v_i}$ – производная по v числителя;

$\frac{\partial w}{\partial v} = l \sum_{i=1}^n \sum_{j=0}^m N_i^{l-1}(v) N_j^k(u) \frac{w_j - w_{i-1}}{v_{i+1} - v_i}$ – производная по v знаменателя.

Пусть для значений параметров u и v найдены номера j и i узлов u_j и v_i из условий $u_j \leq u < u_{j+1}$ и $v_i \leq v < v_{i+1}$. При данных параметрах отличными от нуля будут B -сплайны $N_{j-k}^k(u), \dots, N_{j-1}^k(u), N_j^k(u)$ и $N_{i-1}^l(v), \dots, N_{i-1}^l(v), N_i^l(v)$, а также $N_{j-k+1}^{k-1}(u), \dots, N_{j-1}^{k-1}(u), N_j^{k-1}(u)$ и $N_{i-1}^{l-1}(v), \dots, N_{i-1}^{l-1}(v), N_i^{l-1}(v)$, поэтому суммирование в приведенных выше формулах выполняется по индексам, которым соответствуют отличные от нуля B -сплайны.

Аналогично вычисляются производные радиуса-вектора B -поверхности более высокого порядка:

$$\frac{\partial^{(l+q)} \mathbf{r}}{\partial u^{(l)} \partial v^{(q)}} = \frac{\partial^{(l+q)}}{\partial u^{(l)} \partial v^{(q)}} \left(\frac{w\mathbf{r}(u, v)}{w(u, v)} \right).$$

Если нумерацию узлов u_j начать не с нуля, а с $-k - 1$ (уменьшить на $k + 1$ значения индексов всех узлов u_j), нумерацию узлов v_i также начать не с нуля, а с $-l - 1$ (уменьшить на $l + 1$ значения индексов всех узлов v_i), то поверхность (3.8.1) совпадет с поверхностью

$$\mathbf{r}(u, v) = \frac{\sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_j \mathbf{p}_j}{\sum_{i=0}^n \sum_{j=0}^m N_i^l(v) N_j^k(u) w_j}, \quad (3.8.3)$$

где $N_i^l(v)$ и $N_j^k(u)$ – B -сплайны определения (2.13.1). B -поверхность (3.8.3) использует определение B -сплайна, в котором он имеет индекс последнего узла последовательности, на которой B -сплайн построен. Поверхность (3.8.3) отличают от поверхности (3.8.1) только рекуррентные соотношения, с помощью которых вычисляются радиус-вектор поверхности и его производные. Вместо приведенных выше формул используются соотношения (2.13.3) – (2.13.8).

В частном случае, когда $k = m$, $l = n$ и B -сплайны $N_j^k(u)$ построены на последовательности узлов $u_{-1-k} = u_{-k} = \dots = u_{-1} = 0$, $u_0 = u_1 = \dots = u_k = 1$, а B -сплайны $N_i^l(v)$ построены на последовательности узлов $v_{-1-l} = v_{-l} = \dots = v_{-1} = 0$, $v_0 = v_1 = \dots = v_l = 1$, поверхность (3.8.3) совпадает с рациональной поверхностью Безье (3.7.3).

В качестве примера построим B -поверхность с целочисленными значениями узлов.

Пусть первые $k + 1$ узлов u -последовательности незамкнутой B -поверхности имеют значения, равные нулю: $u_0 = u_1 = \dots = u_k = 0$; следующие $m - k$ узлов принимают целочисленные значения от единицы до

$m - k$: $u_{k+j} = j, j = 1, 2, \dots, m - k$; оставшиеся $k + 1$ узлов принимают значение $m - k + 1$: $u_{m+1} = u_{m+2} = \dots = u_{m+k+1} = m - k + 1$. Пусть первые $l + 1$ узлов v -последовательности имеют значения, равные нулю: $v_0 = v_1 = \dots = v_l = 0$; следующие $n - l$ узлов принимают целочисленные значения от единицы до $n - l$: $v_{l+i} = i, i = 1, 2, \dots, n - l$; оставшиеся $l + 1$ узлов принимают значение $n - l + 1$: $v_{n+1} = v_{n+2} = \dots = v_{n+l+1} = n - l + 1$.

На рис. 3.8.1 приведена B -поверхность третьего порядка по обоим параметрическим направлениям, построенная на сетке, состоящей из шестидесяти трех контрольных точек. Веса всех контрольных точек одинаковые.

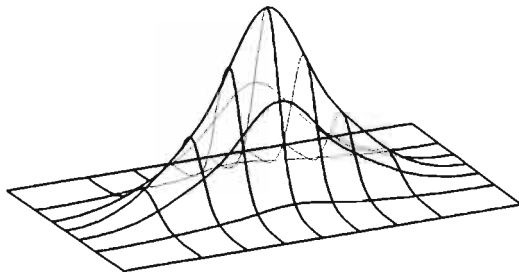


Рис. 3.8.1

Контрольный многогранник поверхности, приведенной на рис. 3.8.1, показан на рис. 3.8.2.

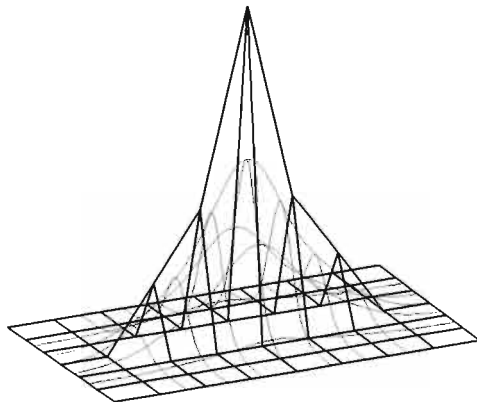


Рис. 3.8.2

Для циклически замкнутой B -поверхности по первому (второму) параметру можно использовать равномерные последовательности узлов с единичным шагом, например, $u_j = j - k, j = 0, 1, \dots, m + 2k + 1$ ($v_i = i - l, i = 0, 1, \dots, n + 2l + 1$).

Если $n = l, m = k$, узлы u -последовательности имеют значения $u_0 = u_1 = \dots = u_m = 0, u_{m+1} = u_{m+2} = \dots = u_{2m+1} = 1$, узлы v -последовательности имеют значения $v_0 = v_1 = \dots = v_n = 0, v_{n+1} = v_{n+2} = \dots = v_{2n+1} = 1$, то B -поверхность (3.8.1) совпадает с рациональной поверхностью Безье (3.7.3). Это следует из того, что функции Бернштейна являются частными случаями B -сплайнов. Область изменения параметров B -поверхности в этом случае представляет собой квадрат со стороной, равной единице.

В отличие от поверхности Безье порядок B -поверхности не связан жестко с количеством контрольных точек и предоставляет возможность строить поверхности невысокого порядка на большом числе контрольных точек, что придает поверхности большую гибкость.

3.9. T -поверхности

B -поверхность строится на совокупности контрольных точек \mathbf{p}_i , условно расположенных в узлах прямоугольной сетки. Если требуется построить поверхность, имеющую локальную особенность, то необходимо локальное сгущение контрольных точек в соответствующем месте. Это сгущение повлечет увеличение контрольных точек во всех рядах и колонках сетки, проходящих через локальную особенность. Если поверхность имеет несколько локальных особенностей, то нам придется построить неоправданно сложную поверхность. Устранить описанный недостаток позволяет некоторая модификация B -поверхности.

Рассмотрим поверхность на основе B -сплайнов, которая описывается векторной функцией

$$\mathbf{r}(u, v) = \frac{\sum_{i=0}^n N_i^k(u) N_i^k(v) w_i \mathbf{p}_i}{\sum_{i=0}^n N_i^k(u) N_i^k(v) w_i}, \quad (3.9.1)$$

где $N_i^k(u)$ и $N_i^k(v)$ – B -сплайны k -го порядка; $\mathbf{p}_i, i = 0, 1, \dots, n$, – контрольные точки, w_i – их веса. Поверхность (3.9.1) может иметь уникальные B -сплайны $N_i^k(u)$ и $N_i^k(v)$ для каждой контрольной точки, так как могут быть построены на уникальных последовательностях узлов. Поверхность (3.9.1) называют **T -поверхностью**.

Для описания последовательностей узлов B -сплайнов в области определения параметров поверхности (3.9.1) построим *сетку*, состоящую из ребер и вершин. На рис. 3.9.1–3.9.3 приведены фрагменты сетки в

области параметров T -поверхности. Каждая ячейка сетки должна иметь прямоугольную форму. Стороны ячейки могут быть составлены из нескольких ребер, но длины противоположных сторон каждой ячейки должны быть равны. Так, для длин ребер ячейки, расположенной в центре рис. 3.9.1, должны выполняться равенства, $u_1 + u_2 = u_3 + u_4 + u_5$, $v_1 = v_2 + v_3$.

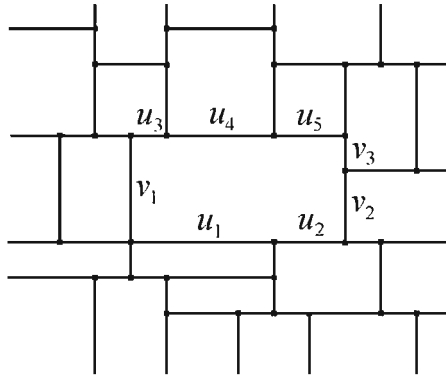


Рис. 3.9.1

С вершинами сетки ассоциируем контрольные точки поверхности.

Когда сетка T -поверхности построена, то последовательность узлов для B -сплайнов $N_i^k(u)$ и $N_i^k(v)$ i -й контрольной точки найдем следующим образом. Через i -й узел сетки проведем две прямые вдоль параметрических направлений – горизонтальную и вертикальную. Для последовательности узлов B -сплайна $N_i^k(u)$ возьмем $k + 2$ точки пересечения горизонтальной прямой с вертикальными ребрами сетки. Для последовательности узлов B -сплайна $N_i^k(v)$ возьмем $k + 2$ точки пересечения вертикальной прямой с горизонтальными ребрами сетки. Пусть $k / 2$ узла последовательности лежат до i -й вершины сетки, а остальные $k + 2 - (k / 2)$ узла последовательности лежат начиная с i -й вершины сетки. На рис. 3.9.2 серыми штриховыми линиями показаны области ненулевых значений B -сплайнов пятого порядка $N_i^5(u)$ и $N_i^5(v)$ i -й контрольной точки T -поверхности.

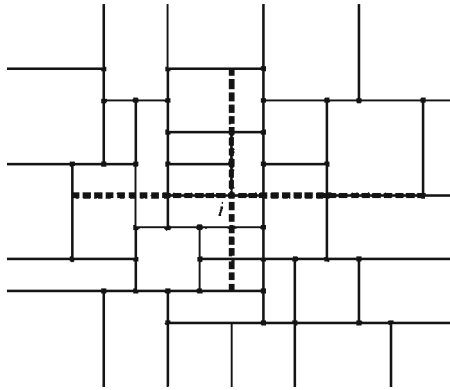


Рис. 3.9.2

Если T -поверхность циклически не замкнута и с какой-то стороны от i -й вершины сетки отсутствует необходимое количество пересечений, то крайний узел будем делать кратным. Для контрольной точки, лежащей на краю сетки, ассоциированная вершина сетки будет входить в последовательность узлов несколько раз. На рис. 3.9.3 кратность j -го узла в последовательности узлов B -сплайна пятого порядка $N_j^s(u)$ равна четырем, а кратность k -го узла в последовательности узлов B -сплайна пятого порядка $N_k^s(v)$ равна трем.

Циклически не замкнутая T -поверхность не проходит через крайние контрольные точки.

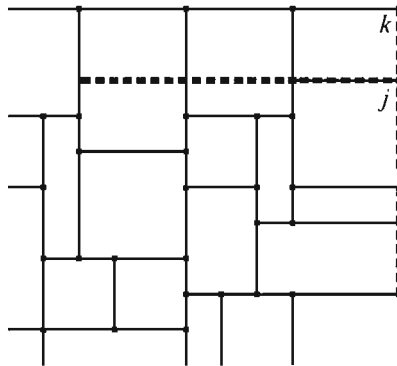


Рис. 3.9.3

Сетка в области параметров может быть построена и для B -поверхности. Она представляет собой регулярную прямоугольную решетку. Как правило, T -поверхность строят путем модификации B -поверхности. Модификация может выполняться или путем вставки новых контрольных точек в B -поверхность, или упрощением B -поверхности путем удаления контрольных точек и изменением положения оставшихся контрольных точек. На рис. 3.9.4 приведена B -поверхность третьего порядка по обоим координатным направлениям и ее контрольные точки.

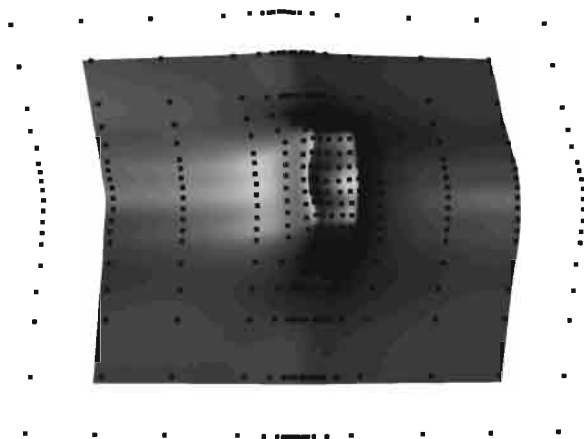


Рис. 3.9.4

B -сплайны крайних точек поверхности построены на последовательностях из пяти узлов, крайние три из которых совпадают, поэтому поверхность не проходит через крайние контрольные точки. В средней части поверхности требуется локальное сгущение контрольных точек, которое влечет увеличение контрольных точек во всех средних рядах и колонках сетки. Сетка в области определения параметров рассматриваемой B -поверхности представляет собой регулярную прямоугольную решетку, каждой вершине которой соответствует контрольная точка. Сетку T -поверхности получим из сетки B -поверхности путем удаления некоторых ребер и вершин. На рис. 3.9.5 приведена T -поверхность, полученная упрощением B -поверхности, изображенной на рис. 3.9.4.

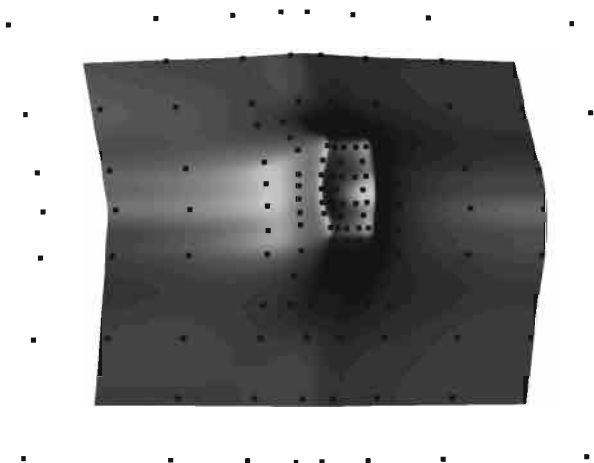


Рис. 3.9.5

При упрощении B -поверхности ее форма может меняться. Для минимизации изменения формы поверхности используются различные методы. Пусть требуется построить T -поверхность, аппроксимирующую некоторую поверхность, радиус-вектор которой известен и описывается векторной функцией $\mathbf{b}(u, v)$. Положение контрольных точек и их веса поверхности (3.9.1) найдем из условия минимума функции:

$$\Psi(\mathbf{p}_0, \dots, \mathbf{p}_n, w_0, \dots, w_n) = \frac{1}{2} \iint \left(\frac{\sum_{i=0}^n N_i^k(u) N_i^k(v) w_i \mathbf{p}_i}{\sum_{i=0}^n N_i^k(u) N_i^k(v) w_i} - \mathbf{b}(u, v) \right)^2 dudv.$$

Интегрирование можно заменить суммированием значений интегрируемой функции при некоторых выбранных значениях параметров u и v . Условие минимума функции даст систему уравнений, из решения которой получим координаты контрольных точек и их веса T -поверхности.

3.10. Поверхности треугольной формы

Рассмотрим поверхности, которые имеют треугольную форму и треугольную область определения параметров. Для треугольных областей в двумерном пространстве удобно использовать барицентрические координаты.

Пусть в двумерном пространстве параметров u и v поверхности выбрана декартова прямоугольная система координат. Пусть в этой систе-

ме заданы три точки $p_a = [u_a \ v_a]^T$, $p_b = [u_b \ v_b]^T$ и $p_c = [u_c \ v_c]^T$, не лежащие на одной прямой, рис. 3.10.1.

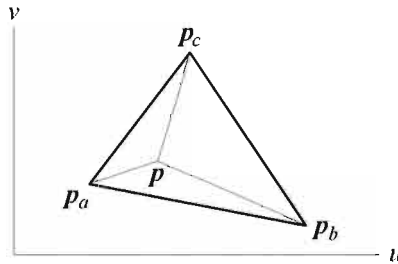


Рис. 3.10.1

Положение любой другой точки $p = [u \ v]^T$ можно описать с помощью точек p_a, p_b, p_c равенством

$$p = a p_a + b p_b + c p_c, \quad (3.10.1)$$

где коэффициенты a, b, c определены с точностью до множителя. Для полной определенности потребуем, чтобы их сумма была равна единице:

$$a + b + c = 1. \quad (3.10.2)$$

Значения коэффициентов a, b, c , соответствующие декартовым координатам u и v , мы найдем из системы уравнений

$$\begin{aligned} a u_a + b u_b + c u_c &= u, \\ a v_a + b v_b + c v_c &= v, \\ a + b + c &= 1. \end{aligned} \quad (3.10.3)$$

Коэффициенты a, b, c определяются равенствами

$$a = \frac{1}{\Delta_{abc}} \begin{vmatrix} u & u_b & u_c \\ v & v_b & v_c \\ 1 & 1 & 1 \end{vmatrix}, \quad b = \frac{1}{\Delta_{abc}} \begin{vmatrix} u_a & u & u_c \\ v_a & v & v_c \\ 1 & 1 & 1 \end{vmatrix}, \quad c = \frac{1}{\Delta_{abc}} \begin{vmatrix} u_a & u_b & u \\ v_a & v_b & v \\ 1 & 1 & 1 \end{vmatrix}, \quad (3.10.4)$$

где определитель системы уравнений (3.10.3)

$$\Delta_{abc} = \begin{vmatrix} u_a & u_b & u_c \\ v_a & v_b & v_c \\ 1 & 1 & 1 \end{vmatrix} = u_a v_b + u_b v_c + u_c v_a - v_a u_b - v_b u_c - v_c u_a \quad (3.10.5)$$

равен удвоенной площади треугольника $p_a p_b p_c$: $\Delta_{abc} = 2S_{abc}$.

Во многих случаях для описания положения точки в двумерном пространстве удобно перейти от координат u и v к трем *барицентрическим*

координатам, которыми являются коэффициенты a, b, c . Бариецентрические координаты произвольной точки $\mathbf{p} = [u \ v]^T$ равны отношению площади треугольника, полученного из треугольника $\mathbf{p}_a\mathbf{p}_b\mathbf{p}_c$ заменой соответствующей его вершины точкой \mathbf{p} , к площади треугольника $\mathbf{p}_a\mathbf{p}_b\mathbf{p}_c$. Бариецентрические координаты a, b, c являются функциями точки \mathbf{p} .

Отметим свойства бариецентрических координат. Они подчинены равенству (3.10.2). Если точка \mathbf{p} лежит внутри треугольника $\mathbf{p}_a\mathbf{p}_b\mathbf{p}_c$, то ее бариецентрические координаты неотрицательные. Если точка \mathbf{p} совпадает с одной из вершин треугольника $\mathbf{p}_a\mathbf{p}_b\mathbf{p}_c$, то соответствующая бариецентрическая координата точки \mathbf{p} равна единице, а две другие ее бариецентрические координаты равны нулю. Если точка \mathbf{p} принадлежит стороне треугольника $\mathbf{p}_a\mathbf{p}_b\mathbf{p}_c$ или лежит на ее продолжении, то одна бариецентрическая координата точки \mathbf{p} равна нулю. Если точка \mathbf{p} лежит вне треугольника ABC , то, по крайней мере, одна бариецентрическая координата точки \mathbf{p} отрицательна.

Любой двумерный вектор $\mathbf{t} = [u \ v]^T$ можно описать с помощью точек $\mathbf{p}_a, \mathbf{p}_b, \mathbf{p}_c$ равенством

$$\mathbf{t} = \alpha \mathbf{p}_a + \beta \mathbf{p}_b + \gamma \mathbf{p}_c. \quad (3.10.6)$$

Коэффициенты α, β, γ называют бариецентрическими компонентами вектора. Двумерный вектор \mathbf{t} можно представить в виде разности радиус-векторов двух точек, например, $\mathbf{p}_1 = a_1\mathbf{p}_a + b_1\mathbf{p}_b + c_1\mathbf{p}_c$ и $\mathbf{p}_0 = a_0\mathbf{p}_a + b_0\mathbf{p}_b + c_0\mathbf{p}_c$. Так как $a_1 + b_1 + c_1 = 1$ и $a_0 + b_0 + c_0 = 1$, то бариецентрические компоненты α, β, γ двумерного вектора связаны равенством

$$\alpha + \beta + \gamma = 0. \quad (3.10.7)$$

Значения бариецентрических компонент α, β, γ , соответствующие декартовым компонентам u и v вектора \mathbf{t} , мы найдем из системы уравнений

$$\begin{aligned} \alpha u_a + \beta u_b + \gamma u_c &= u, \\ \alpha v_a + \beta v_b + \gamma v_c &= v, \\ \alpha + \beta + \gamma &= 0. \end{aligned} \quad (3.10.8)$$

Компоненты α, β, γ определяются равенствами

$$\alpha = \frac{1}{\Delta_{abc}} \begin{vmatrix} u & u_b & u_c \\ v & v_b & v_c \\ 0 & 1 & 1 \end{vmatrix}, \quad \beta = \frac{1}{\Delta_{abc}} \begin{vmatrix} u_a & u & u_c \\ v_a & v & v_c \\ 1 & 0 & 1 \end{vmatrix}, \quad \gamma = \frac{1}{\Delta_{abc}} \begin{vmatrix} u_a & u_b & u \\ v_a & v_b & v \\ 1 & 1 & 0 \end{vmatrix}, \quad (3.10.9)$$

где определитель Δ_{abc} вычисляется по формуле (3.10.5).

Треугольная поверхность строится по трем точкам. Пусть область определения параметров поверхности ограничена треугольником с вершинами в точках $\mathbf{p}_a, \mathbf{p}_b, \mathbf{p}_c$, рис. 3.10.1. Введем бариецентрические координаты a, b, c (3.10.4). Радиус-вектор треугольной поверхности, построенной по

трем точкам $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, опишем векторной функцией трех параметров a, b, c :

$$\mathbf{r}(a, b, c) = a\mathbf{p}_1 + b\mathbf{p}_2 + c\mathbf{p}_3, \quad a + b + c = 1. \quad (3.10.10)$$

Пусть даны три кривые $\mathbf{c}_1(t_1), \mathbf{c}_2(t_2), \mathbf{c}_3(t_3)$, попарно пересекающиеся в крайних точках $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, как показано на рис. 3.10.2. Построим поверхность внутри треугольника, образованного кривыми. Путем изменения параметризации преобразуем кривые так, чтобы точкам \mathbf{p}_3 и \mathbf{p}_2 на кривой $\mathbf{c}_1(t_1)$ соответствовали параметры 0 и 1, точкам \mathbf{p}_1 и \mathbf{p}_3 на кривой $\mathbf{c}_2(t_2)$ соответствовали параметры 0 и 1, точкам \mathbf{p}_2 и \mathbf{p}_1 на кривой $\mathbf{c}_3(t_3)$ соответствовали параметры 0 и 1. **Поверхность на трех кривых**, построенную по данным трем кривым, опишем векторной функцией трех барицентрических координат a, b, c :

$$\begin{aligned} \mathbf{r}(a, b, c) = & a(\mathbf{c}_3(1-b) + \mathbf{c}_2(c) - \mathbf{p}_1) + \\ & + b(\mathbf{c}_1(1-c) + \mathbf{c}_3(a) - \mathbf{p}_2) + \\ & + c(\mathbf{c}_2(1-a) + \mathbf{c}_1(b) - \mathbf{p}_3), \end{aligned} \quad (3.10.11)$$

$$a \in [0, 1], b \in [0, 1], c \in [0, 1], a + b + c = 1.$$

Края поверхности (3.10.11) совпадают с кривыми, по которым она построена. При $c = 0, b = 1 - a$ и вектор

$$\mathbf{r}(a, 1-a, 0) = a(\mathbf{c}_3(a) + \mathbf{c}_2(0) - \mathbf{p}_1) + (1-a)(\mathbf{c}_1(1) + \mathbf{c}_3(a) - \mathbf{p}_2) = \mathbf{c}_3(a)$$

описывает кривую $\mathbf{c}_3(t_3)$. При $a = 0, c = 1 - b$ и вектор

$$\mathbf{r}(0, b, 1-b) = b(\mathbf{c}_1(b) + \mathbf{c}_3(0) - \mathbf{p}_2) + (1-b)(\mathbf{c}_2(1) + \mathbf{c}_1(b) - \mathbf{p}_3) = \mathbf{c}_1(b)$$

описывает кривую $\mathbf{c}_1(t_1)$. При $b = 0, a = 1 - c$ и вектор

$$\mathbf{r}(1-c, 0, c) = c(\mathbf{c}_2(c) + \mathbf{c}_1(0) - \mathbf{p}_3) + (1-c)(\mathbf{c}_3(1) + \mathbf{c}_2(c) - \mathbf{p}_1) = \mathbf{c}_2(c)$$

описывает кривую $\mathbf{c}_2(t_2)$. Правая часть (3.10.11) симметрична относительно циклической перестановки индексов. Поверхность (3.10.11) является треугольным аналогом поверхности Кунса (3.5.3).

На рис. 3.10.3 приведена треугольная поверхность, построенная на трех одинаковых дугах окружностей. Плоскости дуг окружностей ортогональны друг другу. Дуги пересекаются в точках $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$. Каждая дуга содержит четверть окружности.

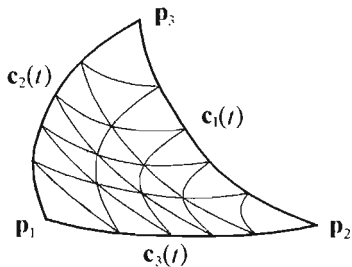


Рис. 3.10.2

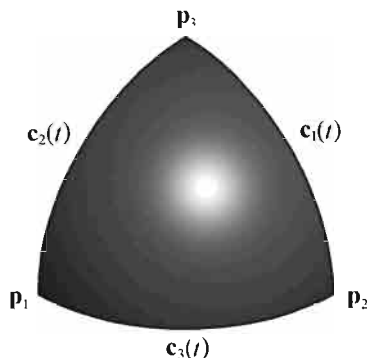
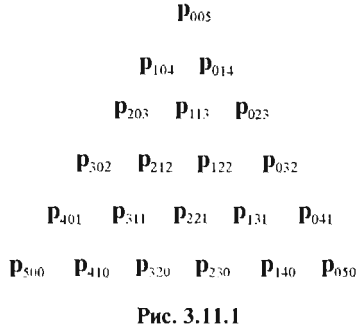


Рис. 3.10.3

Фактически радиусы-векторы поверхностей (3.10.10) и (3.10.11) зависят от двух параметров, так как три параметра a, b, c связаны равенством (3.10.2). Можно переписать зависимости (3.10.10) и (3.10.11) как функции двух параметров u и v , например: $a(u, v) = u, b(u, v) = v, c(u, v) = 1 - u - v$, где u и v принимают значения на треугольной области. Можно переписать зависимости (3.10.10) и (3.10.11) как функции тех же двух параметров u и v , принимающих значения на обычной квадратной области $0 \leq u \leq 1, 0 \leq v \leq 1$, например: $a = u, b = v(1 - u), c = 1 - u - v + uv$ или $a = u(1 - v), b = v, c = 1 - u - v + uv$.

3.11. Треугольные поверхности Безье

Пусть имеется совокупность контрольных точек p_{ijk} , условно расположенных в узлах треугольной сетки. На рис. 3.11.1 показана треугольная сетка, вдоль каждой стороны которой расположены шесть контрольных точек. Точки расположены в виде треугольника, причем вдоль каждой стороны треугольника расположено одинаковое число точек. Пусть это число равно $n + 1$. В приведенном на рис. 3.11.1. примере $n = 5$. Общее число контрольных точек равно $(n + 1)(n + 2) / 2$. Каждый из трех индексов точки соответствует своей вершине треугольника. Точка, лежащая в вершине треугольника, имеет значение индекса, соответствующего этой вершине, равное n (максимальное значение), а значение других индексов равно нулю. Точки, принадлежащие ряду вдоль одной из трех сторон треугольника, имеют нулевой индекс, соответствующий противоположной вершине. Чем ближе ряд, в котором лежат точки, к вершине, тем большее значение имеет индекс этой вершины.



Индексы точки p_{ijk} означают целочисленное расстояние от нее до соответствующей стороны треугольника, измеренное количеством отделяющих рядов. Сумма индексов контрольной точки p_{ijk} имеет постоянное значение

$$i + j + k = n. \quad (3.11.1)$$

Треугольную поверхность Безье опишем с помощью барицентрических координат a, b, c . Будем считать, что поверхность Безье имеет ту же область определения параметров, что и поверхность (3.10.10). **Треугольная поверхность Безье**, построенная по контрольным точкам p_{ijk} , описывается векторной функцией

$$\mathbf{r}(a, b, c) = \sum_{i, j, k=0, i+j+k=n} B_{ijk}^n(a, b, c) \mathbf{p}_{ijk} = \sum_{i, j, k=0, i+j+k=n} \frac{n!}{i!j!k!} a^i b^j c^k \mathbf{p}_{ijk}, \quad (3.11.2)$$

$$a \in [0, 1], b \in [0, 1], c \in [0, 1], a + b + c = 1.$$

Суммирование в (3.11.2) выполняется по всем контрольным точкам, а *трехиндексные функции Бернштейна* имеют вид

$$B_{ijk}^n(a, b, c) = \frac{n!}{i!j!k!} a^i b^j c^k, \quad (3.11.3)$$

где должно выполняться равенство (3.11.1). Треугольная поверхность Безье приведена на рис. 3.11.2.

Крайние линии поверхности (3.11.2) описываются векторами $\mathbf{r}(a, 1-a, 0)$, $\mathbf{r}(0, b, 1-b)$, $\mathbf{r}(1-c, 0, c)$ и представляют собой кривые Безье (2.4.1). Покажем это на примере крайней кривой $\mathbf{r}(a, 1-a, 0)$. Для нее $c = 0, k = 0$ и функции (3.11.3) примут вид

$$B_{i00}^n(a) = \frac{n!}{i!j!} a^i b^j = \frac{n!}{i!(n-i)!} a^i (1-a)^{n-i}, \quad (3.11.4)$$

что совпадает с (2.4.2).

Покажем, что трехиндексные функции Бернштейна удовлетворяют равенству

$$\sum_{i,j,k=0,j+i+k=n}^n B_{ijk}^n(a,b,c) = 1. \quad (3.11.5)$$

Суммирование (3.11.5) выполним по строчкам третьего индекса $k = 0, 1, \dots, n$. В k -й строке третьего индекса содержится $m = n - k$ контрольных точек:

$$\begin{aligned} \sum_{i,j,k=0,j+i+k=n}^n B_{ijk}^n(a,b,c) &= \sum_{i,j,k=0,j+i+k=n}^n \frac{n!}{i!j!k!} a^i b^j c^k = \\ &= \sum_{k=0}^n \frac{n!}{k!(n-k)!} c^k \sum_{i=0}^{n-k} \frac{(n-k)!}{i!(n-k-i)!} a^i b^{n-k-i} = \\ &= \sum_{k=0}^n C_n^k c^k \sum_{i=0}^{n-k} C_{n-k}^i a^i b^{n-k-i} = \sum_{k=0}^n C_n^k c^k (a+b)^{n-k} = (a+b+c)^n = 1, \end{aligned}$$

что и требовалось доказать.

Если представить, что на каждых трех соседних точках $\mathbf{P}_{ijk}, \mathbf{P}_{i-1,j,k+1}, \mathbf{P}_{i-1,j+1,k}$ построена треугольная поверхность, то мы получим *контрольный многогранник*, который дает общее представление о поверхности. Контрольный многогранник треугольной поверхности Безье, приведенной на рис. 3.11.2, показан на рис. 3.11.3.

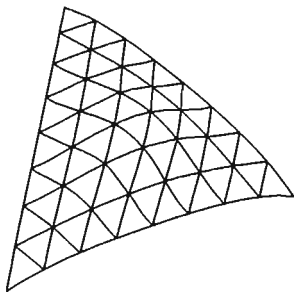


Рис. 3.11.2

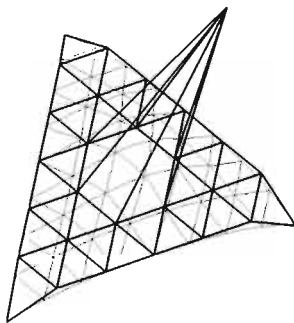


Рис. 3.11.3

Треугольная поверхность Безье не проходит через свои контрольные точки за исключением трех угловых точек

Трехиндексные функции Бернштейна удовлетворяют рекуррентному соотношению

$$B_{ijk}^n = aB_{i-1,j,k}^{n-1} + bB_{i,j-1,k}^{n-1} + cB_{i,j,k-1}^{n-1}. \quad (3.11.6)$$

Это соотношение доказывается непосредственной подстановкой:

$$\begin{aligned}
 & aB_{i-1,j,k}^{n-1}(a,b,c) + bB_{i,j-1,k}^{n-1}(a,b,c) + cB_{i,j,k-1}^{n-1}(a,b,c) = \\
 & = a \frac{(n-1)!}{(i-1)!j!k!} a^{i-1}b^j c^k + b \frac{(n-1)!}{i!(j-1)!k!} a^i b^{j-1} c^k + c \frac{(n-1)!}{i!j!(k-1)!} a^i b^j c^{k-1} = \\
 & = \frac{i}{n} \frac{n!}{i!j!k!} a^i b^j c^k + \frac{j}{n} \frac{n!}{i!j!k!} a^i b^j c^k + \frac{k}{n} \frac{n!}{i!j!k!} a^i b^j c^k = \\
 & = \frac{i+j+k}{n} \frac{n!}{i!j!k!} a^i b^j c^k = B_{ijk}^n(a,b,c),
 \end{aligned}$$

что и требовалось доказать. Начнем вычисление с функции

$$B_{000}^0(a,b,c) = 1.$$

Далее получим $B_{100}^1(a,b,c) = a$, $B_{010}^1(a,b,c) = b$, $B_{001}^1(a,b,c) = c$. Функция Бернштейна, один из индексов которой равен отрицательному числу, считается равной нулю. Подставим рекуррентное соотношение (3.11.6) в выражение (3.11.2), выделив крайние точки и учитывая, что $B_{n00}^n = a^n$, $B_{0n0}^n = b^n$, $B_{00n}^n = c^n$, $B_{0jk}^n = b^j c^k$, $B_{i0k}^n = a^i c^k$, $B_{ij0}^n = a^i b^j$, получим

$$\begin{aligned}
 \mathbf{r}(a,b,c) &= \sum B_{ijk}^n \mathbf{p}_{ijk} = \sum (aB_{i-1,j,k}^{n-1} + bB_{i,j-1,k}^{n-1} + cB_{i,j,k-1}^{n-1}) \mathbf{p}_{ijk} = \\
 & a \sum_{i,j,k=0, i+j+k=n-1}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i+1,j,k} + b \sum_{i,j,k=0, i+j+k=n-1}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i,j+1,k} + c \sum_{i,j,k=0, i+j+k=n-1}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i,j,k+1} = \\
 & = a \mathbf{r}_{i+1,j,k}^{(n-1)} + b \mathbf{r}_{i,j+1,k}^{(n-1)} + c \mathbf{r}_{i,j,k+1}^{(n-1)},
 \end{aligned}$$

где введены обозначения

$$\begin{aligned}
 \mathbf{r}_{i+1,j,k}^{(n-1)} &= \sum_{\substack{i,j,k=0, \\ i+j+k=n-1}}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i+1,j,k}, \quad \mathbf{r}_{i,j+1,k}^{(n-1)} = \sum_{\substack{i,j,k=0, \\ i+j+k=n-1}}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i,j+1,k}, \\
 \mathbf{r}_{i,j,k+1}^{(n-1)} &= \sum_{\substack{i,j,k=0, \\ i+j+k=n-1}}^{n-1} B_{ijk}^{n-1} \mathbf{p}_{i,j,k+1}.
 \end{aligned}$$

Выполнив описанные действия для $\mathbf{r}_{i+1,j,k}^{(n-1)}$, $\mathbf{r}_{i,j+1,k}^{(n-1)}$ и $\mathbf{r}_{i,j,k+1}^{(n-1)}$, в конце придем к равенству

$$\mathbf{r}_{ijk}^{(0)} = B_{000}^0(a,b,c) \mathbf{p}_{ijk} = \mathbf{p}_{ijk}.$$

Обозначим $\mathbf{r}(a,b,c)$ через $\mathbf{r}_{000}^{(n)}$. В результате получим рекуррентное соотношение для вычисления точки треугольной поверхности Бэзье:

$$\mathbf{r}_{ijk}^{(m)} = a \mathbf{r}_{i+1,j,k}^{(m-1)} + b \mathbf{r}_{i,j+1,k}^{(m-1)} + c \mathbf{r}_{i,j,k+1}^{(m-1)}, \quad (3.11.7)$$

где $i+j+k = n-m$ ($i,j,k \geq 0$), $\mathbf{r}_{ijk}^{(0)} = \mathbf{p}_{ijk}$, $\mathbf{r}_{000}^{(n)} = \mathbf{r}(a,b,c)$. Алгоритм, описываемый соотношением (3.11.7), называется *алгоритмом Де Кастелье*. Точки $\mathbf{r}_{ijk}^{(m)}$, получаемые в процессе вычисления, называются точками Де

Кастелье. Алгоритм Де Кастелье позволяет вычислить любую точку треугольной поверхности Безье по контрольным точкам, ничего не зная о функции Бернштейна. Коэффициенты a , b и c при координатах точек $\Gamma_{i+jk}^{(m-1)}$, $\Gamma_{ij+1k}^{(m-1)}$ и $\Gamma_{ijk+1}^{(m-1)}$ являются барицентрическими координатами. Треугольную поверхность Безье можно определить как поверхность, точки которой определяются рекуррентным соотношением (3.11.7) по совокупности контрольных точек.

От треугольной поверхности Безье n -й степени можно перейти к треугольной поверхности $(n + 1)$ -й степени аналогично тому, как это было сделано для кривых Безье. Для этого умножим правую часть равенства (3.11.2) на сумму $a + b + c$, равную 1, и получим

$$\begin{aligned}
 \mathbf{r}(a,b,c) &= \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n B_{ijk}^n \mathbf{p}_{ijk} = \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{n!}{i!j!k!} a^i b^j c^k \mathbf{p}_{ijk} (a+b+c) = \\
 &= \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{n!}{i!j!k!} a^{i+1} b^j c^k \mathbf{p}_{ijk} + \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{n!}{i!j!k!} a b^{j+1} c^k \mathbf{p}_{ijk} + \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{n!}{i!j!k!} a b^j c^{k+1} \mathbf{p}_{ijk} = \\
 &= \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{(n+1)!}{(i+1)!j!k!} a^{i+1} b^j c^k \mathbf{p}_{ijk} \frac{i+1}{n+1} + \\
 &+ \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{(n+1)!}{i!(j+1)!k!} a b^{j+1} c^k \mathbf{p}_{ijk} \frac{j+1}{n+1} + \\
 &+ \sum_{\substack{i,j,k=0 \\ i+j+k=n}}^n \frac{(n+1)!}{i!j!(k+1)!} a b^j c^{k+1} \mathbf{p}_{ijk} \frac{k+1}{n+1} = \\
 &= \sum_{\substack{i=0, j,k=0 \\ i+j+k=n+1}}^{n+1} \frac{(n+1)!}{i!j!k!} a^i b^j c^k \frac{i \mathbf{p}_{i-1, jk}}{n+1} + \\
 &+ \sum_{\substack{j=0, i,k=0 \\ i+j+k=n+1}}^{n+1} \frac{(n+1)!}{i!j!k!} a^i b^j c^k \frac{j \mathbf{p}_{i, j-1, k}}{n+1} + \\
 &+ \sum_{\substack{i,j=0, k=1 \\ i+j+k=n+1}}^{n+1} \frac{(n+1)!}{i!j!k!} a^i b^j c^k \frac{k \mathbf{p}_{ij, k-1}}{n+1} = \\
 &= \sum_{\substack{i,j,k=0 \\ i+j+k=n+1}}^{n+1} B_{ijk}^{n+1} \mathbf{p}_{ijk}^*, \tag{3.11.8}
 \end{aligned}$$

где $\mathbf{p}_{ijk}^* = \frac{i \mathbf{p}_{i-1, jk} + j \mathbf{p}_{i, j-1, k} + k \mathbf{p}_{ij, k-1}}{n+1}$ – новые контрольные точки. Новых контрольных точек на один ряд вдоль каждой из трех сторон треугольника стало больше. Точки в углах треугольной сетки остались прежние: $\mathbf{p}_{n+100}^* = \mathbf{p}_{n00}$, $\mathbf{p}_{0n+10}^* = \mathbf{p}_{0n0}$, $\mathbf{p}_{00n+1}^* = \mathbf{p}_{00n}$. Выражение (3.11.8) описывает ту же поверхность, что и выражение (3.11.2), только через другие контрольные точки.

Пусть треугольная поверхность Безье построена на совокупности контрольных точек \mathbf{p}_{ijk} , образующих треугольную сетку, показанную на рис. 3.11.1. Припишем каждой точке \mathbf{p}_{ijk} вес w_{ijk} . **Треугольная рациональная поверхность Безье** описывается векторной функцией

$$\mathbf{r}(a,b,c) = \frac{\sum_{i,j,k=0, i+j+k=n}^n B_{ijk}^n(a,b,c) w_{ijk} \mathbf{p}_{ijk}}{\sum_{i,j,k=0, i+j+k=n}^n B_{ijk}^n(a,b,c) w_{ijk}}, \quad (3.11.9)$$

$$a \in [0, 1], b \in [0, 1], c \in [0, 1], a + b + c = 1,$$

где суммирование выполняется по всем контрольным точкам. В силу свойства (3.11.5), поверхность Безье (3.11.2) является частным случаем рациональной поверхности (3.11.9) при одинаковых весах точек. Чем больше вес точки, тем ближе к ней проходит поверхность. Напомним, что роль играет не абсолютное значение веса каждой точки, а отношения весов точек друг к другу. Если веса всех характерных точек одинаковы, то мы приходим к обычной поверхности Безье (3.11.2).

В терминах однородных координат расширенный радиус-вектор (1.4.8) треугольной поверхности Безье определяется равенством

$$\mathbf{r}(a,b,c) = \sum_{i,j,k=0, i+j+k=n}^n B_{ijk}^n(a,b,c) w_{ijk} \mathbf{p}_{ijk},$$

по форме совпадающей с равенством (3.11.2), где \mathbf{p}_{ijk} – расширенные радиусы-векторы (1.4.8) контрольных точек поверхности.

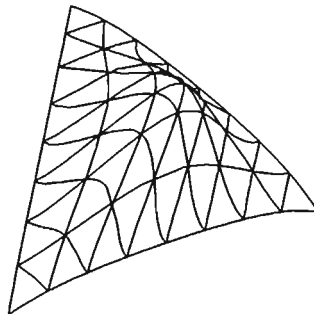


Рис. 3.11.4

На рис. 3.11.4 приведена треугольная поверхность Безье, аналогичная приведенной на рис. 3.11.2 поверхности, с весом средней точки, равным четырем.

3.12. Однородные разделенные разности

Однородные разделенные разности используются для построения сплайновых поверхностей. На их основе строятся симплексные сплайны, представляющие собой аналоги B -сплайнов.

Пусть даны двумерные точки $\mathbf{p}_1 = [u_1 \ v_1]^T$, $\mathbf{p}_2 = [u_2 \ v_2]^T$ и $\mathbf{p}_3 = [u_3 \ v_3]^T$. В однородных координатах каждая из них получит по третьей координате, которые обозначим через w_1 , w_2 , w_3 соответственно. Рассмотрим функцию

$$d(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = w_1 u_2 v_3 + w_2 u_3 v_1 + w_3 u_1 v_2 - w_1 v_2 u_3 - w_2 v_3 u_1 - w_3 v_1 u_2. \quad (3.12.1)$$

Будем считать, что у всех двумерных точек присутствует третья координата — вес точки, которая равна единице, если не оговорено противное. Как можно заметить, если среди аргументов функции (3.12.1) две точки совпадают или все три точки лежат на одной прямой, то она равна нулю. Функция $d(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ не меняется при циклической перестановке точек. При перестановке двух соседних точек местами функция $d(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ меняет свой знак на противоположный. То есть она обладает свойствами

$$d(\mathbf{p}, \mathbf{p}, \mathbf{p}) = d(\mathbf{p}, \mathbf{p}, \mathbf{q}) = d(\mathbf{p}, \mathbf{q}, \mathbf{p}) = d(\mathbf{q}, \mathbf{p}, \mathbf{p}) = 0, \quad (3.12.2)$$

$$\begin{aligned} d(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) &= d(\mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_1) = d(\mathbf{p}_3, \mathbf{p}_1, \mathbf{p}_2) = \\ &= -d(\mathbf{p}_2, \mathbf{p}_1, \mathbf{p}_3) = -d(\mathbf{p}_1, \mathbf{p}_3, \mathbf{p}_2) = -d(\mathbf{p}_3, \mathbf{p}_2, \mathbf{p}_1). \end{aligned} \quad (3.12.3)$$

Функции (3.12.1) линейна по каждому из аргументов, то есть для точки, заданной выражением (3.10.1), справедливо равенство

$$d(\mathbf{r}, \mathbf{q}, \lambda_1 \mathbf{p}_1 + \lambda_2 \mathbf{p}_2 + \lambda_3 \mathbf{p}_3) = \lambda_1 d(\mathbf{r}, \mathbf{q}, \mathbf{p}_1) + \lambda_2 d(\mathbf{r}, \mathbf{q}, \mathbf{p}_2) + \lambda_3 d(\mathbf{r}, \mathbf{q}, \mathbf{p}_3). \quad (3.12.4)$$

Пусть задано множество двумерных точек $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m$, $m \geq 3$ и функция $f(\mathbf{r}, \mathbf{q}, \dots)$, аргументами которой являются двумерные точки. Пусть все точки множества различные и ни одна тройка точек множества не лежит на одной прямой. О таком множестве будем говорить, что оно образует *нормальную последовательность* точек. Рассмотрим выражение

$$\frac{f(\mathbf{p}_i, \mathbf{p}_j, \dots)}{\prod_{\substack{n=1, \\ n \neq i, \\ n \neq j}}^m d(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_n)}, \quad i \neq j. \quad (3.12.5)$$

Для нормальной последовательности ни один из сомножителей $d(\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_n)$ в (3.12.5) не равен нулю. Выражение (3.12.5) соответствует паре точек \mathbf{p}_i и

p_j . В знаменателе (3.12.5) находится $m - 2$ сомножителей. Если функция f обладает свойством

$$f(\mathbf{r}, \mathbf{q}, \dots) = (-1)^{m-2} f(\mathbf{q}, \mathbf{r}, \dots), \quad (3.12.6)$$

то выражение (3.12.5) не меняет знак при перестановке точек p_i и p_j местами.

Однородная разделенная разность для функции $f(\mathbf{r}, \mathbf{q}, \dots)$, обладающей свойством (3.12.6), на последовательности точек p_1, p_2, \dots, p_m , $m \geq 3$ определяется формулой

$$f[p_1, p_2, \dots, p_m] = \sum_{i=1}^m \sum_{j=i+1}^m \frac{f(p_i, p_j, \dots)}{\prod_{\substack{\alpha=1, \\ \alpha \neq i, \\ \alpha \neq j}}^m d(p_i, p_j, p_\alpha)}. \quad (3.12.7)$$

Обозначим через P последовательность точек p_1, p_2, \dots, p_m : $P = \{p_1, p_2, \dots, p_m\}$, а разделенную разность $f[p_1, p_2, \dots, p_m]$ для функции f будем обозначать также через $f[P]$. В литературе для разделенной разности используются также обозначения $[p_1, p_2, \dots, p_m]f$ и $[P]f$.

Однородные разделенные разности обладают следующими свойствами.

Свойство 1. В силу независимости значения каждого слагаемого в правой части (3.12.7) от порядка следования точек p_i и p_j разделенная разность является симметричной функцией своих аргументов, то есть значение $f[P]$ не зависит от порядка, в котором следуют точки p_1, p_2, \dots, p_m в списке параметров:

$$f[P] = f[p_1, p_2, \dots, p_m] = f[p_2, p_1, \dots, p_m] = \dots = f[p_m, p_1, \dots, p_1]. \quad (3.12.8)$$

Свойство 2. Если $f(\mathbf{r}, \mathbf{q}, \dots) = k_g g(\mathbf{r}, \mathbf{q}, \dots) + k_h h(\mathbf{r}, \mathbf{q}, \dots)$, то

$$f[P] = k_g g[P] + k_h h[P], \quad (3.12.9)$$

так как каждое выражение (3.12.5) будет состоять из двух слагаемых.

Свойство 3. Пусть точка p принадлежит последовательности P . Тогда разделенная разность функции $f(\mathbf{r}, \mathbf{q}, \dots) = d(\mathbf{r}, \mathbf{q}, p)g(\mathbf{r}, \mathbf{q}, \dots)$ определится формулой

$$f[P] = g[P \setminus p], \quad (3.12.10)$$

где $g[P \setminus p]$ есть разделенная разность для функции $g(\mathbf{r}, \mathbf{q}, \dots)$ на последовательности точек $P \setminus p$, полученной из последовательности P путем исключения из нее точки p . Действительно, слагаемые

$$\frac{d(p_i, p_j, p)g(p_i, p_j, \dots)}{\prod_{\substack{\alpha=1, \\ \alpha \neq i, \\ \alpha \neq j}}^m d(p_i, p_j, p_\alpha)}, \quad i \neq j,$$

в которых точка p равна p_i или p_j , обращаются в ноль в силу (3.12.2). В остальных слагаемых произойдет сокращение одинаковых отличных от нуля сомножителей $d(p_i, p_j, p)$. Тем самым разделенная разность будет вычислена для функции $g(r, q, \dots)$, а последовательность точек, на которой она построена, не будет содержать точку p .

Свойство 4. Это свойство является следствием предыдущего свойства. Пусть точка p представлена выражением (3.10.1) через некоторые три точки последовательности p_1, p_2, \dots, p_m :

$$p = \sum_i \lambda_i p_i, \lambda_i \geq 0, p_i \in P. \quad (3.12.11)$$

Тогда разделенная разность функции $f(r, q, \dots) = d(r, q, p)g(r, q, \dots)$ определяется формулой

$$f[P] = \sum_i \lambda_i g[P \setminus p_i], \quad (3.12.12)$$

где $g[P \setminus p_i]$ есть разделенная разность для функции $g(r, q, \dots)$ на последовательности точек $P \setminus p_i$, полученной из последовательности P путем исключения из нее точки p_i .

Пусть в двумерном пространстве заданы некоторые непересекающиеся области. Каждая двумерная точка не может принадлежать более чем одной области. Для ответа на вопрос, какой из двумерных областей принадлежит точка, лежащая на линии стыковки двух областей или совпадающая с точкой стыковки нескольких областей, вводится понятие *полуоткрытой выпуклой области*.

Через $[V]$ обозначим выпуклую область, которую покрывает некоторый выпуклый многоугольник V , а через $[V]$ обозначим полуоткрытую выпуклую область многоугольника. Не все граничные точки выпуклой области $[V]$ принадлежат полуоткрытой выпуклой области $[V]$. Точка $p = [u \ v]^T$ принадлежит полуоткрытой выпуклой области $[V]$, если существуют такие $\epsilon > 0$ и $\eta > 0$, при которых точки $p = [u \ v]^T$, $p_1 = [u + \epsilon \ v]^T$ и $p_2 = [u + \epsilon \ v + \eta]^T$ принадлежат $[V]$.

На рис. 3.12.1 приведена полуоткрытая выпуклая область.

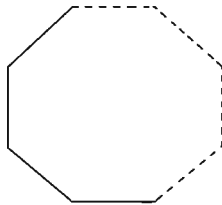


Рис. 3.12.1

Для области, показанной на рис. 3.12.1, внутренние точки многоугольника и точки сплошных линий принадлежат $[U]$, а точки пунктирных линий и точки их стыковки со сплошными линиями не принадлежат $[U]$.

Для полуоткрытой выпуклой области $[U]$ и точки p определим функцию

$$\delta_{[U]}(p) = \begin{cases} 1, & \text{если } p \in [U] \\ 0, & \text{если } p \notin [U] \end{cases}, \quad (3.12.13)$$

характеризующую принадлежность точки p полуоткрытой выпуклой области.

Рассмотрим степенную функцию двумерных точек:

$$f_n(r, q, p) = (d(r, q, p))^n \cdot d(r, q, x) \cdot \frac{\delta_{[r, q, x]}(p)}{|d(r, q, x)|}, \quad (3.12.14)$$

где x – произвольная двумерная точка, $\delta_{[r, q, x]}(p)$ – функция (3.12.13), характеризующая принадлежность точки p полуоткрытой выпуклой области $[r, q, x]$, построенной на треугольнике с вершинами в точках r , q и x .

На последовательности точек $P = \{p_1, p_2, \dots, p_m\}$ вычислим однородную разделенную разность для степенной функции $f_{m-3}(r, q, p)$:

$$f_{m-3}[P] = \frac{\sum_{i=1}^m \sum_{j=i+1}^m (d(p_i, p_j, p))^{m-3} \cdot d(p_i, p_j, x) \delta_{[p_i, p_j, x]}(p)}{\prod_{\substack{n=1, \\ n \neq i, \\ n \neq j}}^m d(p_i, p_j, p_n)} |d(p_i, p_j, x)|}. \quad (3.12.15)$$

Припишем данной разделенной разности порядок, равный степени функции (3.12.14), и введем для нее обозначения:

$$f_{m-3}[P] = M(p | p_1, p_2, \dots, p_m) \equiv M^p(p).$$

Однородная разделенная разность (3.12.15) степенной функции двумерных точек, кроме перечисленных выше свойств, обладает следующими свойствами.

Свойство 5. При $m = 3$ разделенная разность (3.12.15) принимает значение

$$M(p | p_1, p_2, p_3) = \begin{cases} \frac{1}{|d(p_1, p_2, p_3)|}, & \text{если } p \in [p_1, p_2, p_3] \\ 0, & \text{если } p \notin [p_1, p_2, p_3] \end{cases}. \quad (3.12.16)$$

Докажем это утверждение на последовательности трех точек p_1, p_2, p_3 :

$$M(p | p_1, p_2, p_3) = \frac{d(p_1, p_2, x)}{d(p_1, p_2, p_3)} \frac{\delta_{(p, p_2, x)}(p)}{|d(p, p_2, x)|} + \frac{d(p_2, p_3, x)}{d(p_2, p_3, p_1)} \frac{\delta_{(p, p_3, x)}(p)}{|d(p_2, p_3, x)|} + \frac{d(p_3, p_1, x)}{d(p_3, p_1, p_2)} \frac{\delta_{(p, p_1, x)}(p)}{|d(p_3, p_1, x)|}.$$

Выразив точку x через точки последовательности $x = \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3$ и используя равенство (3.12.4), получим

$$M(p | p_1, p_2, p_3) = \lambda_3 \frac{\delta_{(p, p_2, x)}(p)}{|\lambda_3 d(p_1, p_2, p_3)|} + \lambda_1 \frac{\delta_{(p, p_3, x)}(p)}{|\lambda_1 d(p_2, p_3, p_1)|} + \lambda_2 \frac{\delta_{(p, p_1, x)}(p)}{|\lambda_2 d(p_3, p_1, p_2)|}. \quad (3.12.17)$$

Если точка x находится внутри или на границе треугольника с вершинами в точках p_1, p_2, p_3 , то все $\lambda_i \geq 0, i = 1, 2, 3$, из слагаемых правой части (3.12.17) не более одного будет отлично от нуля (следует из определения полукрытой выпуклой области), и мы приходим к формуле (3.12.16). Если точка x находится вне треугольника с вершинами в точках p_1, p_2, p_3 , то барицентрические координаты $\lambda_i, i = 1, 2, 3$ будут иметь разные знаки, из слагаемых правой части (3.12.17) могут быть отличны от нуля два, одно или ни одного слагаемых. Одно слагаемое будет отлично от нуля, если точка p находится внутри треугольника с вершинами в точках p_1, p_2, p_3 , соответствующая $\lambda_i > 0$, и мы также приходим к формуле (3.12.16). Точка p может находиться вне треугольника с вершинами в точках p_1, p_2, p_3 , но внутри двух из треугольников с вершинами в точках x, p_2, p_3 , с вершинами в точках p_1, x, p_3 и с вершинами в точках p_1, p_2, x . В этом случае два слагаемых правой части (3.12.17) будут отличны от нуля. Если точка p (как и точка x) находится вне треугольника с вершинами в точках p_1, p_2, p_3 , соответствующие им λ_i будут иметь разные знаки и разнесенная разность будет равна нулю. Если точка p находится вне упомянутых треугольников, то все слагаемые правой части (3.12.17) будут равны нулю. Таким образом, во всех случаях мы приходим к формуле (3.12.16). Разнесенная разность нулевого порядка является разрывной константной функцией.

Свойство 6. При $m > 3$ степенную функцию (3.12.14) представим в виде $f_k(r, q, p) = d(r, q, p) f_{k-1}(r, q, p)$, а точку p с помощью (3.12.11) выразим через некоторые три точки последовательности p_1, p_2, \dots, p_m . Тогда свойство (3.12.12) для функции (3.12.15) принимает вид

$$M(p | p_1, p_2, \dots, p_m) = \sum_i \lambda_i M(p | p_1, p_2, \dots, p_m \setminus p_i), \quad (3.12.18)$$

где $M(p | p_1, p_2, \dots, p_m \setminus p_i)$ есть разделенная разность для функции $f_{m-4}(r, q, p)$ на последовательности точек, полученной из исходной последовательности путем исключения из нее точки p_i . Разложение (3.12.18) не зависит от выбора точек, через которые выражается точка p с помощью формулы (3.12.11). При равенстве аргумента p разделенной разности (3.12.15) одной из точек последовательности P из (3.12.18) следует, что

$$M(p | p_1, p_2, \dots, p_m) = M(p | p_1, p_2, \dots, p_m \setminus p), p \in P. \quad (3.12.19)$$

Применяя формулу (3.12.18) многократно к ее слагаемым правой части, мы приходим к тому, что разделенная разность будет представлена суммой ненулевых разделенных разностей (3.12.16), умноженных на положительные значения барицентрических координат. Из этого следуют следующие свойства.

Свойство 7. Из свойств (3.12.16) и (3.12.18) следует, что разделенная разность (3.12.15) не зависит от точки x .

Свойство 8. Разделенная разность (3.12.15) является функцией точки p .

Свойство 9. Для последовательности точек p_1, p_2, \dots, p_m можно построить выпуклый многоугольник V , содержащий все точки последовательности, вершинами которого являются также точки последовательности (часть точек последовательности может лежать внутри многоугольника V). Разделенная разность (3.12.15) отлична от нуля на полуоткрытой выпуклой области $[V]$, построенной для заданной последовательности p_1, p_2, \dots, p_m . Другими словами, однородные разделенные разности для степенной функции являются локальными функциями. Их носителем является полуоткрытая выпуклая область, очерченная заданной последовательностью точек.

Свойство 10. Разделенная разность (3.12.15) принимает неотрицательные значения.

Свойство 11. При $m > 3$ разделенная разность (3.12.15) на границе выпуклой области $[V]$, построенной для последовательности p_1, p_2, \dots, p_m , равна нулю. Покажем это. Если точка p совпадает с одной из вершин выпуклой области $[V]$, то в соответствии с (3.12.19) разделенная разность будет рассчитана на выпуклой области, не содержащей точку p , и будет равна нулю. Пусть p принадлежит отрезку $p_i p_j$, являющемуся граничным для области $[V]$. Выразим точку p через точки граничного отрезка и еще одну не совпадающую с ними произвольную точку p_k последовательности $p = \lambda p_i + \lambda p_j + \lambda p_k$. Так как точка p лежит на отрезке $p_i p_j$, то $\lambda_k = 0$. Тогда в соответствии с (3.12.18)

$$M(p | p_1, p_2, \dots, p_m) = \lambda_i M(p | p_1, p_2, \dots, p_m \setminus p_i) + \lambda_j M(p | p_1, p_2, \dots, p_m \setminus p_j).$$

Но разделенные разности $M(p | p_1, p_2, \dots, p_m \setminus p_i) = 0$ и $M(p | p_1, p_2, \dots, p_m \setminus p_j) = 0$ в силу того, что точка p не принадлежит ни полуоткрытой области $[p_1, p_2, \dots, p_m \setminus p_i)$, ни полуоткрытой области $[p_1, p_2, \dots, p_m \setminus p_j)$. Действительно, если из выпуклой области $[V]$ удалить точку p_i или p_j , являющуюся вершиной многоугольника V , то вновь полученная область не будет содержать интервал pp_j .

Рассмотрим **производную по направлению** для функции двумерных точек. Пусть дан двумерный единичный вектор $t = [u, v, w]^T$, определяющий некоторое направление, и функция $f(p, \dots)$, одним из аргументов которой является двумерная точка $p = [u_p, v_p, w_p]^T$. Выражение

$$\frac{\partial f}{\partial u_p} u + \frac{\partial f}{\partial v_p} v + \frac{\partial f}{\partial w_p} w = \frac{\partial f}{\partial t}$$

называется производной функции $f(p, \dots)$ по направлению вектора t . Производная функции по направлению t представляет собой скалярное произведение градиента функции $f(p, \dots)$ и вектора t . Будем считать, что у двумерных векторов присутствует третья координата – вес вектора. Если выразить двумерный вектор через три барицентрические координаты относительно трех точек с помощью выражения (3.10.6), то при единичных весах точек в силу равенства (3.10.7) получим, что вес двумерного вектора равен нулю. Тогда производная функции $d(r, q, p)$ по направлению t по аргументу p равна

$$\frac{\partial(d(r, q, p))}{\partial t} = d(r, q, t). \quad (3.12.20)$$

Аналогичным образом определяются производные функции $d(r, q, p)$ по направлению t по другим аргументам.

Пусть двумерный единичный вектор t представлен через некоторые три точки в виде:

$$t = \sum_i \alpha_i p_i, \quad p_i \in P, \quad (3.12.21)$$

где α_i , – барицентрические координаты (3.10.6) вектора t . Тогда производная функции $d(r, q, p)$ по последнему аргументу по направлению t равна

$$\frac{\partial}{\partial t} d(r, q, p) = \sum_i \alpha_i d(r, q, p_i). \quad (3.12.22)$$

Производная по направлению степенной функции двумерных точек (3.12.14) является функцией, аналогичной самой степенной функции.

Действительно,

$$\begin{aligned} \frac{\partial (f_n(r, q, p))}{\partial t} &= n(d(r, q, p))^{n-1} d(r, q, x) \cdot \frac{\delta_{[r, q, x]}(p)}{|d(r, q, x)|} \cdot \frac{\partial (d(r, q, p))}{\partial t} = \\ &= n(d(r, q, p))^{n-1} d(r, q, x) \cdot \frac{\delta_{[r, q, x]}(p)}{|d(r, q, x)|} \cdot \sum_i \alpha_i d(r, q, p_i). \end{aligned}$$

Найдем производную однородной разделенной разности $M(p | p_1, p_2, \dots, p_m)$ по аргументу p по направлению t . Аналогично (3.12.21) представим вектор t через произвольные три точки последовательности p_1, p_2, \dots, p_m и получим

$$\begin{aligned} &\frac{\partial}{\partial t} M(p | p_1, p_2, \dots, p_m) = \\ &= (m-3) \sum_{i=1}^m \sum_{j=i+1}^m \frac{(d(p_i, p_j, p))^{m-4} \cdot d(p_i, p_j, x) \delta_{[p_i, p_j, x]}(p)}{\prod_{\substack{n=1, \\ n \neq i, \\ n \neq j}}^m d(p_i, p_j, p_n)} \frac{\delta_{[p_i, p_j, x]}(p)}{|d(p_i, p_j, x)|} \frac{\partial d(p_i, p_j, p)}{\partial t} = \\ &= (m-3) \sum_k \alpha_k M(p | p_1, p_2, \dots, p_m \setminus p_k), \end{aligned} \quad (3.12.23)$$

где $M(p | p_1, p_2, \dots, p_m \setminus p_k)$ есть разделенная разность для функции $f_{m-4}(r, q, p)$ на последовательности точек, полученной из исходной последовательности путем исключения из нее точки p_k . Действительно, после подстановки (3.12.22) слагаемые

$$\alpha_k (m-3) \frac{(d(p_i, p_j, p))^{m-4} \cdot d(p_i, p_j, x) \delta_{[p_i, p_j, x]}(p)}{\prod_{\substack{n=1, \\ n \neq i, \\ n \neq j}}^m d(p_i, p_j, p_n)} \frac{\delta_{[p_i, p_j, x]}(p)}{|d(p_i, p_j, x)|} d(p_i, p_j, p_k),$$

$$i \neq j, p_k \in P$$

в которых точка p_k равна p_i или p_j , обращаются в ноль в силу (3.12.2). В остальных слагаемых произойдет сокращение одинаковых отличных от нуля сомножителей $d(p_i, p_j, p_k)$.

Формула (3.12.23) показывает, что производная однородной разделенной разности выражается через разделенные разности меньшего порядка. Формула (3.12.23) не зависит от выбора точек, через которые выражается вектор t .

3.13. Симплексные сплайны

На основе однородных разделенных разностей определим симплексные сплайны. Соотношение (3.12.18) побуждает забыть об однородных разделенных разностях степенной функции двумерных точек (3.12.15) и определить симплексный сплайн для последовательности двумерных точек как функцию, вычисляемую по рекуррентному соотношению (3.12.18) при начальных значениях (3.12.16). При этом мы распространим рекуррентное соотношение и на последовательности с совпадающими точками. Совпадающие точки будем называть *кратными*. Разделенная разность (3.12.16) при совпадении двух или всех трех точек, на которых она построена, равна нулю в силу вырождения ее полуоткрытой выпуклой области.

Симплексные сплайны являются обобщением трехиндексных функций Бернштейна и аналогами B -сплайнов. На основе симплексных сплайнов строятся поверхности.

Симплексный сплайн n -го порядка, построенный на последовательности точек $P = \{p_1, p_2, \dots, p_m\}$, $m = n + 3$, хотя бы три из которых не лежат на одной прямой, определим как функцию, вычисляемую с помощью рекуррентного соотношения

$$M^p(p) = aM^{p_a}(p) + bM^{p_b}(p) + cM^{p_c}(p), \quad (3.13.1)$$

при начальных значениях

$$M^r(p) = \begin{cases} \frac{1}{|d(p, p_j, p_k)|}, & \text{если } p \in [p_j, p_j, p_k] \\ 0, & \text{если } p \notin [p_j, p_j, p_k] \end{cases}, \quad (3.13.2)$$

где a, b, c – барицентрические координаты (3.10.4) точки p относительно точек p_a, p_b, p_c ; $M^{p_a}(p)$, $M^{p_b}(p)$, $M^{p_c}(p)$ – симплексные сплайны $(n - 1)$ -го порядка, построенные на последовательностях точек, полученных из последовательности P путем исключения из нее соответственно точки p_a, p_b, p_c ; $M^r(p)$ – симплексный сплайн нулевого порядка, построенный по трем точкам p_j, p_j, p_k ; функция $d(p, p_j, p_k)$ вычисляется по формуле (3.12.1).

Точки p_a, p_b, p_c последовательности точек P не должны лежать на одной прямой или совпадать. Такая тройка точек в последовательности P по условию всегда существует. Точки p_1, p_2, \dots, p_m последовательности называют *узлами*. Вычисление значения симплексного сплайна n -го порядка начинается с вычисления значений симплексных сплайнов нулевого порядка на треугольных областях. Пусть вершинами некоторого треугольника T являются точки p_i, p_j, p_k . Значение симплексного сплайна нулевого порядка на треугольной области T не равно нулю, если

область T не вырождена и аргумент p принадлежит полуоткрытой выпуклой области $[p_i, p_j, p_k)$.

Для последовательности $P = \{p_1, p_2, \dots, p_m\}$, $m \geq 3$ построим выпуклый многоугольник V с вершинами в точках последовательности, так чтобы все точки заданной последовательности принадлежали $[V]$. В общем случае часть точек заданной последовательности будет лежать в вершинах многоугольника V , часть — на его сторонах и часть — внутри многоугольника V . Полуоткрытую выпуклую область $[V)$ многоугольника V называют *носителем* симплексного сплайна $M^p(p)$ или его областью определения.

Обратим внимание на аналогию формулы Кокса—Де Бура (2.8.7) и формулы (3.13.1). Обе формулы описывают рекуррентное соотношение для сплайна $(m - \lambda)$ -го порядка, построенного на последовательности m узлов, где λ — число барицентрических координат. В обоих случаях вычисления начинаются со сплайнов нулевого порядка для λ несовпадающих узлов.

Симплексный сплайн (3.13.1) совпадает с однородной разделенной разностью степенной функции на нормальной последовательности точек P . При наличии в последовательности кратных узлов симплексный сплайн сохраняет алгоритм вычисления (3.12.18). Симплексный сплайн наследует свойства однородной разделенной разности степенной функции.

Свойство 1. Симплексный сплайн нулевого порядка является разрывной константной функцией положения точки p . Внутри своей треугольной области определения T симплексный сплайн нулевого порядка является неотрицательной функцией. Если вершины треугольника T совпадают или лежат на одной прямой, то симплексный сплайн нулевого порядка $M^T(p)$ равен нулю.

Свойство 2. Симплексный сплайн $M^p(p)$ принимает неотрицательные значения. Он отличен от нуля внутри своей полуоткрытой выпуклой области $[V)$ и равен нулю вне $[V)$.

Свойство 3. Результат формулы (3.13.1) не зависит от выбранной тройки точек p_a, p_b, p_c .

Свойство 4. Симплексный сплайн (3.13.1) является функцией точки p .

Свойство 5. Симплексный сплайн (3.13.1) ненулевого порядка на границе выпуклой области $[V)$, построенной для нормальной последовательности p_1, p_2, \dots, p_m , равен нулю.

Свойство 6. В общем случае симплексный сплайн n -го порядка является кусочно-полиномиальной функцией степени n , имеющей непрерывные производные до $(n - 1)$ -го порядка включительно. В используемых нами обозначениях порядок симплексного сплайна совпадает с наивысшей степенью полиномов, описывающих его.

Для симплексного сплайна можно вычислить производную в направлении вектора на плоскости параметров. Пусть дан единичный двумерный вектор $t = [u \ v \ 0]^T$. По формулам (3.10.9) этот вектор можно описать с помощью барицентрических компонент α, β, γ относительно точек p_a, p_b, p_c равенством

$$t = \alpha p_a + \beta p_b + \gamma p_c.$$

В соответствии с производной однородной разделенной разности (3.12.23) производная симплексного сплайна (3.13.1) по направлению t вычисляется по формуле

$$\frac{\partial M^r(p)}{\partial t} = n (\alpha M^{r_{\alpha}}(p) + \beta M^{r_{\beta}}(p) + \gamma M^{r_{\gamma}}(p)). \quad (3.13.3)$$

Производные симплексных сплайнов n -го порядка определяются через симплексные сплайны $(n - 1)$ -го порядка. В качестве точек p_a, p_b, p_c могут быть взяты любые три не совпадающие и не лежащие на одной прямой узла дифференцируемого симплексного сплайна.

Произвольную точку p мы можно выразить через линейную комбинацию всех точек последовательности P . То есть выражение (3.12.11) мы можем использовать в виде

$$p = \sum_{i=1}^m \lambda_i p_i, \quad p_i \in P. \quad (3.13.4)$$

Тогда формула (3.13.1) для вычисления симплексного сплайна примет симметричный относительно узлов вид:

$$M^r(p) = \sum_{i=1}^m \lambda_i M^{r_i}(p). \quad (3.13.5)$$

Аналогично можно представить с помощью линейной комбинации всех узлов единичный вектор t . Тогда симметричный вид примет и формула (3.13.3) для производной симплексного сплайна по направлению t :

$$\frac{\partial M^r(p)}{\partial t} = (m-3) \sum_{i=1}^m \alpha_i M^{r_i}(p). \quad (3.13.6)$$

Трехиндексные функции Бернштейна (3.11.3) являются частными случаями симплексных сплайнов. Функция Бернштейна

$$B_{ijk}^n(a, b, c) = \frac{n!}{i!j!k!} a^i b^j c^k$$

с точностью до множителя равна симплексному сплайну n -го порядка, построенному на последовательности $(n + 3)$ узлов, из которых $(i + 1)$ узлов совпадают с точкой p_a , $(j + 1)$ узлов совпадают с точкой p_b и $(k + 1)$

узлов совпадают с точкой p_c . Пусть точки p_a, p_b, p_c являются вершинами треугольника T . На треугольной области T можно построить $C_{n+2} = \frac{(n+2)!}{2!n!} = \frac{(n+1)(n+2)}{2}$ симплексных сплайнов n -го порядка. Эти симплексные сплайны и соответствующие им функции Бернштейна связаны равенством

$$M_{ijk}(\mathbf{p}) = \frac{1}{\Delta_{abc}} \frac{n!}{i!j!k!} a^i b^j c^k, \quad (3.13.7)$$

где $\mathbf{p} = ap_a + bp_b + cp_c$; Δ_{abc} – удвоенная площадь треугольника T . Симплексный сплайн $M_{ijk}(\mathbf{p})$ построен на последовательности узлов: p_a – кратности $(i+1)$, p_b – кратности $(j+1)$, p_c – кратности $(k+1)$.

На рис. 3.13.1 приведен симплексный сплайн нулевого порядка, построенный на треугольной области (значение симплексного сплайна откладывается в направлении, перпендикулярном плоскости параметров сплайна).

На рис. 3.13.2 приведен симплексный сплайн первого порядка, построенный на прямоугольной области.

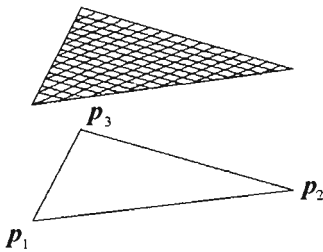


Рис. 3.13.1

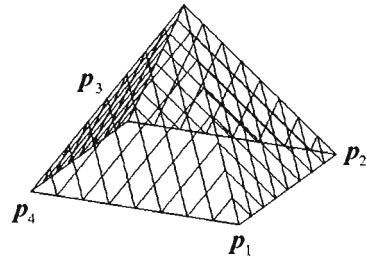


Рис. 3.13.2

На рис. 3.13.3 приведен симплексный сплайн первого порядка, построенный на треугольной области. На рис. 3.13.4 приведен симплексный сплайн второго порядка, построенный на пяти несовпадающих узлах.

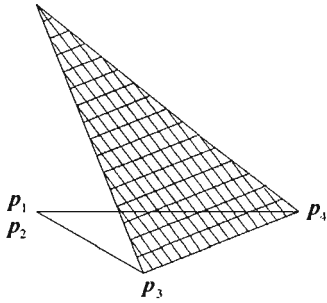


Рис. 3.13.3

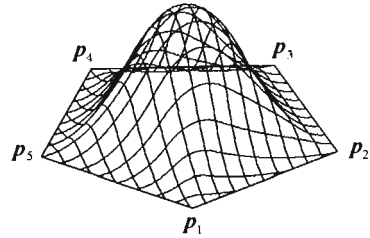


Рис. 3.13.4

Последовательность точек симплексного сплайна может содержать кратные узлы. В кратных узлах высшие производные симплексного сплайна теряют гладкость. На рис. 3.13.5 приведен симплексный сплайн третьего порядка, построенный на шести узлах, образующих треугольную область. В каждой вершине треугольной области находится по два совпадающих узла.

Некоторые узлы симплексного сплайна могут лежать внутри его выпуклого многоугольника. На рис. 3.13.6 приведен симплексный сплайн третьего порядка, построенный на шести узлах, три из которых лежат в вершинах треугольной области, а остальные три узла совпадают и лежат в центре этой треугольной области. Производные в центре такого сплайна теряют непрерывность.

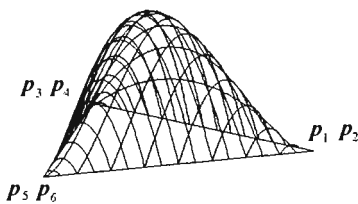


Рис. 3.13.5

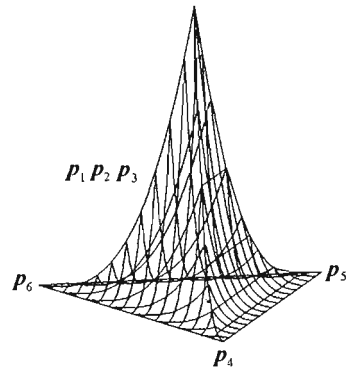
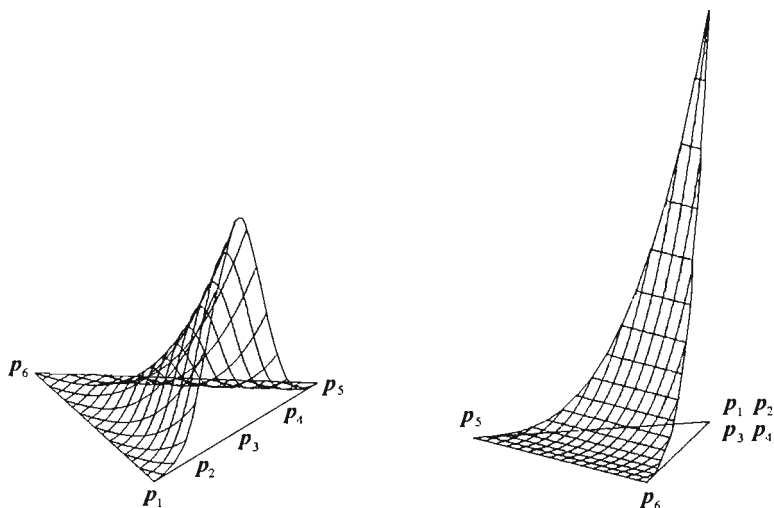


Рис. 3.13.6

Некоторые узлы симплексного сплайна могут лежать на границе его выпуклого многоугольника. На рис. 3.13.7 приведен симплексный

сплайн третьего порядка, построенный на шести узлах, пять из которых лежат вдоль стороны треугольника. На рис. 3.13.8 приведен симплексный сплайн третьего порядка, построенный на шести узлах, четыре из которых лежат в одной из вершин треугольника.



В местах концентрации узлов значение симплексного сплайна заметно больше, чем в остальных местах. Если все узлы симплексного сплайна расположены в вершинах его выпуклого многоугольника, то чем больше порядок сплайна, тем меньше значение сплайна в центральной его части. На рис. 3.13.9 приведен симплексный сплайн, узлы которого лежат в вершинах правильного восьмиугольника.

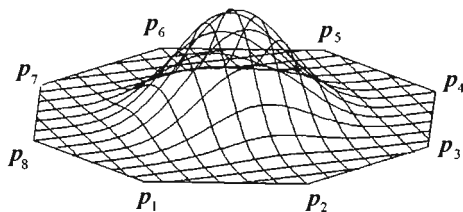


Рис. 3.13.9

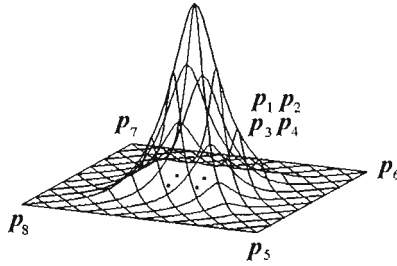


Рис. 3.13.10

На рис. 3.13.10 приведен симплексный сплайн пятого порядка, построенный на четырехугольной области, с четырьмя узлами внутри четырехугольной области.

3.14. S-поверхности

Поверхности, построенные с использованием симплексных сплайнов, будем называть **S-поверхностями**. Они строятся в виде взвешенной суммы некоторых контрольных точек. Каждая контрольная точка входит в сумму с коэффициентом, который является функцией координат двумерной точки на области определения параметров поверхности. Каждая такая функция отлична от нуля на некоторой ограниченной двумерной области.

Рассмотрим S-поверхность, использующую симплексные сплайны третьего порядка. Для каждой контрольной точки построим симплексный сплайн третьего порядка на шести двумерных точках, расположенных в вершинах правильного шестиугольника, рис. 3.14.1.

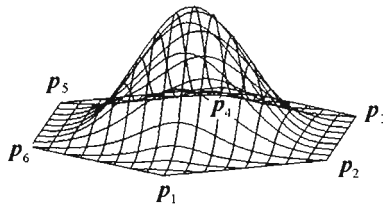


Рис. 3.14.1

Центры шестиугольников расположим в виде регулярной решетки, в которой любые три центра областей определения соседних симплексных

сплайнов располагаются в вершинах правильного треугольника. Центры областей определения симплексных сплайнов и граница области определения одного из них приведены на рис. 3.14.2. Шестиугольные области соседних симплексных сплайнов частично перекрываются.

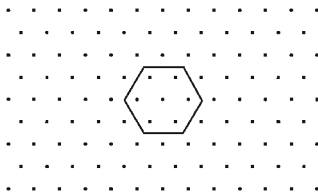


Рис. 3.14.2

S -поверхность опишем векторной функцией

$$\mathbf{r}(\mathbf{p}) = \frac{\sum_i M^p_i(\mathbf{p}) w_i \mathbf{r}_i}{\sum_i M^p_i(\mathbf{p}) w_i}, \quad \mathbf{p} = [u \ v]^T \in \Omega. \quad (3.14.1)$$

Радиус-вектор поверхности (3.14.1) является функцией двух параметров u и v . Производную векторной функции (3.14.1) по первому параметру вычислим по формуле

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{\sum_i \frac{\partial M^p_i(\mathbf{p})}{\partial u} w_i \mathbf{r}_i}{\sum_i M^p_i(\mathbf{p}) w_i} - \frac{\sum_i M^p_i(\mathbf{p}) w_i \mathbf{r}_i}{\left(\sum_i M^p_i(\mathbf{p}) w_i\right)^2} \sum_i \frac{\partial M^p_i(\mathbf{p})}{\partial u} w_i,$$

где $\mathbf{t} = [1 \ 0]^T$. Производную векторной функции (3.14.1) по второму параметру вычислим по аналогичной формуле при $\mathbf{t} = [0 \ 1]^T$. Фрагмент S -поверхности (3.14.1) и ее контрольные точки приведены на рис. 3.14.3.

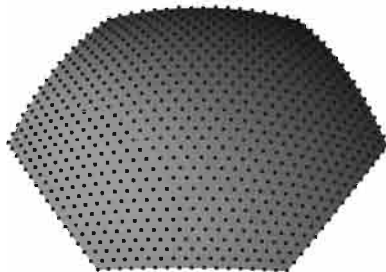


Рис. 3.14.3

Один из способов построения поверхностей на базе симплексных сплайнов предложили Дамен (W. Dahmen), Микели (С. Micchelli) и Зайдель (Н. Seidel). Поверхности, построенные этим способом, называют **DMS-сплайнами**. Областью определения DMS-сплайна служит правильная триангуляция некоторой двумерной области.

3.15. Поверхности, построенные на базе поверхностей

Поверхности могут быть построены на базе других поверхностей. Мы рассмотрим продолженную, эквидистантную и ссылочную поверхности. Поверхность, на основе которой строится новая поверхность, будем называть *базовой поверхностью*. В качестве базовой поверхности не должна использоваться поверхность того же типа, что и создаваемая поверхность.

Продолженная поверхность может быть построена на базе любой поверхности путем изменения области определения ее параметров. Пусть требуется продолжить поверхность $\mathbf{b}(u, v)$, $u_{\min} \leq u \leq u_{\max}$, $v_{\min} \leq v \leq v_{\max}$ путем расширения области определения параметров до $a_u + u_{\min} \leq u \leq u_{\max} + b_u$, $a_v + v_{\min} \leq v \leq v_{\max} + b_v$. Если $a_u \leq 0$, $a_v \leq 0$, $b_u \geq 0$, $b_v \geq 0$, то поверхность будет продолжена за свои пределы. Если $a_u > 0$, $a_v > 0$, $b_u < 0$, $b_v < 0$, то поверхность будет усечена.

Если поверхность не является периодической, а ее параметр вышел за границу области определения, то поверхность можно продолжить в соответствии с законом вычисления радиуса-вектора поверхности. К таким поверхностям относятся аналитические поверхности, а также поверхности, построенные на основе кривых или поверхностей с известным законом вычисления радиуса-вектора.

В общем случае при отсутствии закона поведения непериодической поверхности за пределами ее области определения продолжим поверхность по касательной на ближайшей границе и вычислим по продолженной поверхности необходимые геометрические характеристики. Радиус-вектор продолженной поверхности вычислим по формулам

$$\mathbf{r}(u, v) = \mathbf{b}(u, v), \quad (3.15.1)$$

если $u_{\min} \leq u \leq u_{\max}$, $v_{\min} \leq v \leq v_{\max}$;

$$\mathbf{r}(u, v) = \mathbf{b}(u_0, v) + (u - u_0) \left. \frac{\partial \mathbf{b}}{\partial u} \right|_{u_0, v},$$

если $v_{\min} \leq v \leq v_{\max}$, а параметр u вышел за граничное значение u_0 ;

$$\mathbf{r}(u, v) = \mathbf{b}(u, v_0) + (v - v_0) \left. \frac{\partial \mathbf{b}}{\partial v} \right|_{u, v_0},$$

если $u_{\min} \leq u \leq u_{\max}$, а параметр v вышел за граничное значение v_0 ;

$$\mathbf{r}(u, v) = \mathbf{b}(u_0, v_0) + (u - u_0) \left. \frac{\partial \mathbf{b}}{\partial u} \right|_{u_0, v_0} + (v - v_0) \left. \frac{\partial \mathbf{b}}{\partial v} \right|_{u_0, v_0} + (u - u_0)(v - v_0) \left. \frac{\partial^2 \mathbf{b}}{\partial u \partial v} \right|_{u_0, v_0},$$

если параметры u и v вышли за граничные значения u_0 и v_0 соответственно. Граничным значением u_0 мы называем u_{\min} или u_{\max} , граничным значением v_0 мы называем v_{\min} или v_{\max} . На рис. 3.15.1 приведены продолженная поверхность и ее базовая поверхность (тонкими линиями).

Если поверхность периодическая по какому-либо из параметров, то при выходе этого параметра за границу области определения выполним его циклический пересчет. Если поверхность $\mathbf{b}(u, v)$ периодическая по первому параметру, то радиус-вектор продолженной поверхности вычислим по формуле

$$\mathbf{r}(u, v) = \mathbf{b}(u - p_u \cdot \Phi((u - u_{\min})/p_u), v), \quad (3.15.2)$$

где p_u – период первого параметра базовой поверхности; $\Phi(w)$ – функция, вычисляющая наибольшее целое число, не превосходящее w . Если поверхность $\mathbf{b}(u, v)$ периодическая по второму параметру, то радиус-вектор продолженной поверхности вычислим по формуле

$$\mathbf{r}(u, v) = \mathbf{b}(u, v - p_v \cdot \Phi((v - v_{\min})/p_v)), \quad (3.15.3)$$

где p_v – период второго параметра базовой поверхности; $\Phi(w)$ – функция, вычисляющая наибольшее целое число, не превосходящее w .

В качестве базовой поверхности для продолженной поверхности не должна использоваться другая продолженная поверхность, а должна использоваться базовая поверхность последней с соответствующим пересчетом параметров a_u, a_v, b_u, b_v .

Формулы (3.15.1) – (3.15.3) могут быть использованы для вычисления радиуса-вектора поверхности и его производных при выходе параметров поверхности за область их определения.

Эквидистантная поверхность описывается векторной функцией

$$\mathbf{r}(u, v) = \mathbf{b}(u, v) + h \mathbf{m}(u, v), \quad (3.15.4)$$

$$u, v \in \Omega,$$

где $\mathbf{b}(u, v)$ – базовая поверхность; h – расстояние от базовой поверхности; $\mathbf{m}(u, v)$ – нормаль базовой поверхности; Ω – область определения параметров базовой поверхности. Нормаль $\mathbf{m}(u, v)$ имеет единичную длину и направление вектора $\mathbf{b}_u \times \mathbf{b}_v$. Область изменения параметров эквиди-

стантной поверхности Ω может совпадать с областью изменения параметров базовой поверхности, а может быть расширена в соответствии с правилами построения продолженной поверхности. На рис. 3.15.2 приведена эквидистантная поверхность и ее базовая поверхность (тонкими линиями).

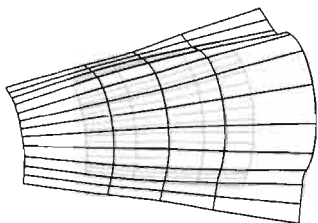


Рис. 3.15.1

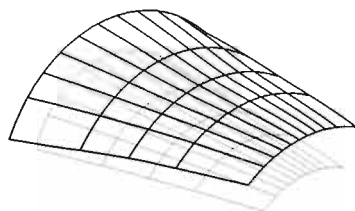


Рис. 3.15.2

В качестве базовой поверхности для эквидистантной поверхности не должна использоваться другая эквидистантная поверхность, а должна использоваться базовая поверхность последней с соответствующим пересчетом длины эквидистанты. То есть если нужно построить эквидистантную поверхность на расстоянии h_2 от другой эквидистантной поверхности $\mathbf{r}_1(u, v) = \mathbf{b}(u, v) + h_1 \mathbf{m}(t)$, то в качестве базовой поверхности должна быть использована поверхность $\mathbf{b}(u, v)$, а длина эквидистанты должна быть равна $h = h_2 + h_1$.

Ссылочная поверхность представляет собой поверхность, каждая точка которой получена путем некоторого преобразования соответствующей точки базовой поверхности. Ссылочная поверхность описывается векторной функцией

$$\mathbf{r}(u, v) = \mathbf{M} \cdot \mathbf{b}(u, v), \quad u, v \in \Omega, \quad (3.15.5)$$

где $\mathbf{b}(u, v)$, $u, v \in \Omega$ – базовая поверхность; Ω – область определения параметров базовой поверхности; \mathbf{M} – расширенная матрица преобразования (1.4.2) базовой поверхности. При работе с расширенной матрицей мы предполагаем, что радиусы-векторы используются в виде (1.4.3), а свободные векторы используются в виде (1.4.4). Область изменения параметров ссылочной поверхности совпадает с областью изменения параметров базовой поверхности. Ссылочная поверхность может использоваться для работы в глобальной декартовой системе координат с поверхностью $\mathbf{b}(u, v)$, описанной в некоторой локальной декартовой системе координат.

В качестве базовой поверхности для ссылочной поверхности не должна использоваться другая ссылочная поверхность, а должна использоваться базовая поверхность последней с соответствующим пересчетом

матрицы преобразования. Например, если требуется построить ссылочную поверхность на базе ссылочной поверхности $\mathbf{r}(u, v) = \mathbf{M}\mathbf{b}(u, v)$, то в качестве базовой поверхности должна использоваться поверхность $\mathbf{b}(u, v)$, а матрицу преобразования получим как произведение матриц $\mathbf{M}\mathbf{M}$.

3.16. Деформируемая поверхность

Для коррекции формы поверхности на ее базе может быть построена деформируемая поверхность. Исходную поверхность, на основе которой строится деформируемая поверхность, будем называть базовой поверхностью.

Деформируемая поверхность представляет собой эквидистантную поверхность с переменным расстоянием от базовой поверхности и описывается векторной функцией

$$\mathbf{r}(u, v) = \mathbf{b}(u, v) + h(u, v) \mathbf{m}(u, v), \quad (3.16.1)$$

где $\mathbf{b}(u, v)$ – базовая поверхность; $h(u, v)$ – функция деформации поверхности; $\mathbf{m}(u, v)$ – нормаль базовой поверхности. Область изменения параметров деформируемой поверхности совпадает с областью изменения параметров базовой поверхности. Нормаль $\mathbf{m}(u, v)$ имеет единичную длину и направление вектора $\mathbf{b}_u \times \mathbf{b}_v$. Функция $h(u, v)$ задает деформирование поверхности вдоль ее нормали. В тех точках, где $h(u, v)$ равна нулю, деформируемая поверхность совпадает с базовой поверхностью. Если $h(u, v) = \text{const}$, то деформируемая поверхность совпадает с эквидистантной поверхностью.

Функция деформации поверхности $h(u, v)$ определяется формулой

$$h(u, v) = \sum_{i=0}^n \sum_{j=0}^m N_i^i(v) N_j^k(u) h_{ij}, \quad (3.16.2)$$

где $N_i^i(v)$ и $N_j^k(u)$ – B -сплайны (2.8.1); h_{ij} , $i = 0, 1, \dots, n$, $j = 0, 1, \dots, m$ – совокупность отклонений поверхности в точках, условно расположенных в узлах прямоугольной сетки в виде $n + 1$ рядов по $m + 1$ точек в каждом ряду. Функция деформации $h(u, v)$ аналогична векторной функции (3.8.1), описывающей B -поверхность. Отличие функции деформации $h(u, v)$ от функции (3.8.1) заключается только в том, что функция (3.8.1) векторная и определяется контрольными точками \mathbf{p}_{ij} , а функция деформации скалярная и определяется отклонениями h_{ij} поверхности вдоль нормали в определенных точках.

B -сплайны $N_j^k(u)$ построены на неубывающей последовательности из $k + 2$ узлов $u_j, u_{j+1}, \dots, u_{j+k+1}$. B -сплайны $N_i^i(v)$ построены на неубывающей последовательности из $l + 2$ узлов $v_i, v_{i+1}, \dots, v_{i+l+1}$. Для незамкнутой по

первому (второму) параметру базовой поверхности первые и последние $k + 1$ ($l + 1$) узлов последовательности должны совпадать. Для циклически замкнутой базовой поверхности по первому (второму) параметру соответствующая последовательность узлов должна отражать замкнутость: первые $2k$ ($2l$) узлов должны идти через интервалы, повторяющие интервалы, через которые идут последние $2k$ ($2l$) узлов. Значения узлов должны быть согласованы с минимальными и максимальными параметрами базовой поверхности. Пусть нумерация узлов начинается с нуля, тогда $u_k = u_{\min}$, $u_{m+k+1} = u_{\max}$, $v_l = v_{\min}$, $v_{n+l+1} = v_{\max}$, где u_{\min} и u_{\max} – минимальное и максимальное значения первого параметра базовой поверхности, v_{\min} и v_{\max} – минимальное и максимальное значения второго параметра базовой поверхности.



Рис. 3.16.1



Рис. 3.16.2

На рис. 3.16.1 приведена цилиндрическая поверхность, а на рис. 3.16.2 приведена деформируемая поверхность, построенная на базе цилиндрической поверхности.

В общем случае деформируемая поверхность описывается векторной функцией

$$\mathbf{r}(u, v) = \mathbf{b}(u, v) + \sum_{i=0}^n \sum_{j=0}^m N_i^j(v) N_i^i(u) \mathbf{s}_{ij},$$

где \mathbf{s}_{ij} , $i = 0, 1, \dots, n$, $j = 0, 1, \dots, m$ – совокупность векторов отклонения поверхности в точках, условно расположенных в узлах прямоугольной сетки в виде $n + 1$ рядов по $m + 1$ точек в каждом ряду. Векторы отклонения \mathbf{s}_{ij} заданы в локальных системах координат в узлах прямоугольной сетки

с базисом $\mathbf{b}_1, \mathbf{b}_2, \mathbf{m}$, где $\mathbf{b}_1 = \frac{\partial \mathbf{b}(u, v)}{\partial u}$ и $\mathbf{b}_2 = \frac{\partial \mathbf{b}(u, v)}{\partial v}$ – частные производные поверхности в соответствующих узлах сетки, $\mathbf{m} = \mathbf{m}(u, v)$ – нормаль поверхности в соответствующих узлах сетки.

3.17. Поверхность на базе триангуляции

Предположим, что мы имеем множество контрольных точек в пространстве, по которым можно построить триангуляцию. Триангуляцией будем называть связную многогранную поверхность, образованную набором стыкующихся друг с другом по общим сторонам треугольных пластин, вершинами которых являются контрольные точки множества. Треугольные пластины не должны пересекать друг друга. К каждой стороне каждой треугольной пластины может стыковаться не более одной другой треугольной пластины. Стороны треугольных пластин, не имеющих соседей, будем называть краевыми.

После построения триангуляции по заданным контрольным точкам построим двумерную триангуляцию в некотором пространстве параметров создаваемой нами поверхности. Двумерная триангуляция должна быть синхронна с пространственной триангуляцией. Каждой контрольной точке должна соответствовать двумерная точка, каждой плоской пластине в пространстве должен соответствовать двумерный треугольник в пространстве параметров. Двумерные треугольники также должны стыковаться по общим сторонам и не пересекать друг друга. Поверхность на базе триангуляции будет представлять собой отображение в трехмерное пространство двумерной триангуляции в области параметров этой поверхности.

Далее заменим плоские треугольные пластины криволинейными треугольными поверхностями, стыкующимися друг с другом по общим кривым. Для этого в каждой контрольной точке определим общую для всех стыкующихся в ней треугольных пластин нормаль. За нормаль в контрольной точке можно принять усреднение нормалей стыкующихся в точке пластин. Стороны треугольных пластин искривим так, чтобы они подходили к своим контрольным точкам перпендикулярно нормальям этих точек.

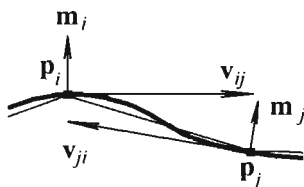


Рис. 3.17.1

На рис. 3.17.1 приведена искривленная сторона треугольника, соединяющая контрольные точки p_i и p_j . Искривленную сторону треугольника опишем кубическим сплайном Эрмита (2.3.10), который в рассматриваемом случае примет вид

$$\mathbf{r}_{ij}(t) = \mathbf{p}_i(1 - 3t^2 + 2t^3) + \mathbf{p}_j(3t^2 - 2t^3) + \mathbf{v}_{ij}(t - 2t^2 + t^3) - \mathbf{v}_j(-t^2 + t^3),$$

$$0 \leq t \leq 1. \quad (3.17.1)$$

В (3.17.1) векторы \mathbf{v}_{ij} и \mathbf{v}_j равны векторам $\mathbf{p}_j - \mathbf{p}_i$ и $\mathbf{p}_i - \mathbf{p}_j$, из которых вычтены их составляющие вдоль нормалей \mathbf{m}_i и \mathbf{m}_j соответственно:

$$\mathbf{v}_{ij} = (\mathbf{p}_j - \mathbf{p}_i) - \mathbf{m}_i (\mathbf{m}_i^\top (\mathbf{p}_j - \mathbf{p}_i)),$$

$$\mathbf{v}_j = (\mathbf{p}_i - \mathbf{p}_j) - \mathbf{m}_j (\mathbf{m}_j^\top (\mathbf{p}_i - \mathbf{p}_j)).$$

Плоские треугольные пластины заменим криволинейными треугольными поверхностями (3.10.11), каждая из которых построена на трех кривых (3.17.1). На рис. 3.17.2 приведена треугольная поверхность

$$\mathbf{r}_{jk}(a, b, c) = a(\mathbf{r}_{ki}(1-b) + \mathbf{r}_{ij}(c) - \mathbf{p}_i) + \quad (3.17.2)$$

$$+ b(\mathbf{r}_{jk}(1-c) + \mathbf{r}_{ki}(a) - \mathbf{p}_k) +$$

$$+ c(\mathbf{r}_{ij}(1-a) + \mathbf{r}_{jk}(b) - \mathbf{p}_j),$$

$$a \in [0, 1], b \in [0, 1], c \in [0, 1], a + b + c = 1,$$

построенная на трех кривых $\mathbf{r}_{ij}(t)$, $\mathbf{r}_{jk}(t)$ и $\mathbf{r}_{ki}(t)$, описывающих стороны криволинейного треугольника $\mathbf{p}_i, \mathbf{p}_k, \mathbf{p}_j$.

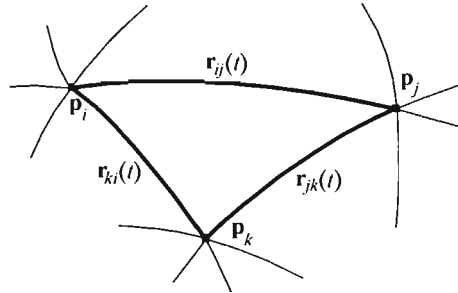


Рис. 3.17.2

Аргументы a, b, c треугольной поверхности (3.17.2) являются барицентрическими координатами (3.10.4) двумерной точки параметрической области поверхности. Барицентрические координаты a, b, c двумерной точки \mathbf{p} вычислены по двумерным точкам $\mathbf{p}_i = [u_i \ v_i]^\top$, $\mathbf{p}_k = [u_k \ v_k]^\top$ и $\mathbf{p}_j = [u_j \ v_j]^\top$, соответствующим контрольным точкам $\mathbf{p}_i, \mathbf{p}_k$ и \mathbf{p}_j , так что

$$\mathbf{p} = a \mathbf{p}_i + b \mathbf{p}_k + c \mathbf{p}_j.$$

Поверхность на базе триангуляции состоит из криволинейных треугольников (3.17.2), стыкующихся друг с другом по общим кривым. Внутренними данными поверхности являются множество контрольных точек в пространстве, соответствующее ему множество нормалей поверхности в этих контрольных точках, соответствующее контрольным

точкам множество двумерных точек в параметрической области поверхности и множество треугольников. Каждый треугольник описан тремя номерами вершин из множества контрольных точек.

Для любой поверхности с прямоугольной областью определения параметров можно построить поверхность на базе триангуляции. Для этого прямоугольную область определения параметров заполним двумерными треугольниками, стыкующимися по общим ребрам и вершинам и не пересекающимися друг друга. Для каждой вершины двумерной триангуляции вычислим контрольную точку и нормаль для нее. По перечисленным данным построим поверхность на базе триангуляции, аппроксимирующую исходную поверхность.

На рис. 3.17.3 приведена поверхность, построенная на базе триангуляции тора, а на рис. 3.17.4 приведена эта же поверхность после перемещения некоторых ее контрольных точек.



Рис. 3.17.3



Рис. 3.17.4

Рис. 3.17.4 демонстрирует возможности данной поверхности изменять форму. Для построения рассматриваемой поверхности по облаку точек, полученному замером реального объекта, необходимо построить трехмерную триангуляцию и соответствующую ей двумерную триангуляцию параметров поверхности.

3.18. Поверхность с произвольной границей

В общем случае область определения параметров поверхности опишем двумерными замкнутыми составными кривыми – двумерными *контурами*. Сегментами контура могут быть любые двумерные кривые за исключением составных кривых. Начало каждого последующего сегмента контура совпадает с концом предыдущего сегмента. В общем случае в местах стыковки сегментов производные контура терпят разрыв по длине и направлению.

Пусть двумерный контур $c(t)$ содержит n сегментов $r_i(w_i)$, $w_{i\min} \leq w_i \leq w_{i\max}$, $i = 1, 2, \dots, n$. Начальное значение параметра контура положим равным нулю. Параметрическую длину контура положим равной сумме параметрических длин составляющих его сегментов:

$$t_{\min} = 0, \quad t_{\max} = \sum_{i=1}^n (w_{i\max} - w_{i\min}).$$

При вычислении радиуса-вектора контура сначала необходимо определить сегмент, которому соответствует значение параметра контура. Номер сегмента k , которому соответствует заданный параметр контура t , найдем из условия

$$\sum_{i=1}^{k-1} (w_{i\max} - w_{i\min}) \leq t < \sum_{i=1}^k (w_{i\max} - w_{i\min}).$$

Соответствующее значение собственного параметра w_k найденного сегмента определим по формуле

$$w_k = w_{k\min} + t - \sum_{i=1}^{k-1} (w_{i\max} - w_{i\min}).$$

Радиус-вектор двумерного контура для заданного параметра t равен радиусу-вектору найденного сегмента:

$$r(t) = r_k(w_k).$$

Контуром можно описать границу двумерной области любой требуемой формы и размера.

Поверхность с границей, заданной двумерными контурами, будем называть **ограниченной контурами поверхностью**. Ограниченная контурами поверхность представляет собой наиболее общий вид поверхности и состоит из некоторой поверхности и совокупности двумерных контуров. Ограниченную контурами поверхность можно построить на основе любой другой поверхности. Исходную поверхность будем называть *базовой поверхностью*. Пусть $\mathbf{b}(u, v)$ есть радиус-вектор базовой поверхности. Радиус-вектор ограниченной контурами поверхности описывается тем же законом, что и базовая поверхность, но имеет другую область определения параметров

$$\mathbf{r}(u, v) = \mathbf{b}(u, v), \quad u, v \in \Omega, \quad (3.18.1)$$

где Ω — область определения параметров, которая представляет собой двумерную связную область. В общем случае область определения параметров Ω ограниченной контурами поверхности $\mathbf{r}(u, v)$ может выходить за область определения параметров базовой поверхности $\mathbf{b}(u, v)$. Вне области определения параметров базовой поверхности радиус-вектор $\mathbf{r}(u, v)$

вычисляется в соответствии с правилами построения продолженной поверхности.

Пусть границы области определения поверхности Ω описываются с помощью m двумерных контуров $c_j(t_j) = [u_j(t_j) \ v_j(t_j)]^T$, $t_{j,\min} \leq t_j \leq t_{j,\max}$, $j=1, 2, \dots, m$. Потребуем, чтобы граничные контуры удовлетворяли следующим условиям. Контур не должен пересекать сами себя и друг друга. Один из контуров должен содержать внутри охватываемой им области все остальные контуры. Охватывающий контур будем называть *внешним*, а лежащие внутри него контуры будем называть *внутренними*. Внутренние контуры не могут быть вложены друг в друга. Для быстрого определения положения некоторой двумерной точки относительно области определения параметров поверхности контуры ориентируют так, чтобы при движении вдоль границы поверхность всегда находилась бы с одной и той же стороны. Пусть внешний контур ориентирован так, что обход по нему осуществляется против движения часовой стрелки, если смотреть на поверхность навстречу ее нормали, тогда внутренние контуры должны быть ориентированы в противоположном направлении.

Простым примером ограниченной контурами поверхности может служить плоскость, ограниченная одним контуром $c(t)$. Плоскость, проходящую через точку \mathbf{p} и параллельную ортонормированным векторам \mathbf{i}_x и \mathbf{i}_y , опишем векторной функцией

$$\mathbf{r}(u, v) = \mathbf{p} + u \mathbf{i}_x + v \mathbf{i}_y, \\ u, v \in \Omega.$$

Границу области определения Ω опишем одним двумерным контуром $c(t) = [u(t) \ v(t)]^T$, $t_{\min} \leq t \leq t_{\max}$. Границу рассмотренной плоскости будет описывать трехмерная кривая

$$\mathbf{r}(t) = \mathbf{p} + u(t) \mathbf{i}_x + v(t) \mathbf{i}_y, \quad t_{\min} \leq t \leq t_{\max}.$$

На рис. 3.18.1 приведена сплайновая поверхность, имеющая прямоугольную область определения параметров, и замкнутые составные кривые на ней. На рис. 3.18.2 приведена ограниченная контурами поверхность, построенная на базе сплайновой поверхности.

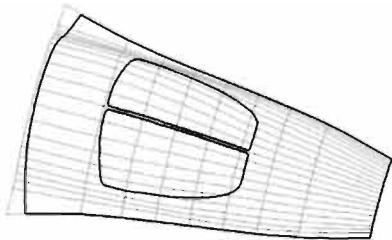


Рис. 3.18.1

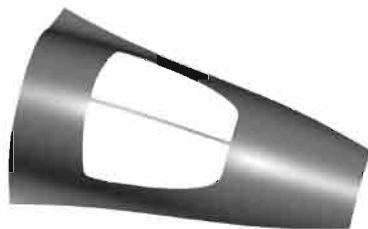


Рис. 3.18.2

В качестве базовой поверхности для ограниченной контурами поверхности не должна использоваться другая ограниченная контурами поверхность, а должна использоваться базовая поверхность последней. Ограниченные контурами поверхности должны создаваться в последнюю очередь. Например, если требуется построить эквидистантную поверхность для ограниченной контурами поверхности, то сначала должна быть построена эквидистантная поверхность на основе базовой поверхности $\mathbf{b}(u, v)$, а потом на базе эквидистантной поверхности должна быть создана ограниченная контурами поверхность.

Упражнения

1. Вычислите длину одного витка цилиндрической спирали.
2. Вычислите с помощью теоремы Менье кривизну окружностей широты на глобусе.
3. Опишите векторной функцией поверхность тора с эллиптическим сечением.
4. Выразите поверхность тора с эллиптическим сечением в виде B -поверхности.
5. Опишите векторной функцией часть плоскости, ограниченную эллипсом.

Глава 4. Проекции и пересечения

Рассмотренные кривые и поверхности выдают геометрическую информацию, если мы обратимся к ним с некоторыми параметрами. При построении геометрической модели необходимо решать обратные задачи — получать параметры кривой или поверхности по сформулированным геометрическим условиям. Решение таких задач требуется для вычисления проекций, для определения точек пересечения, для построения на основе кривых и поверхностей новых кривых и поверхностей, необходимых в геометрическом моделировании.

4.1. Проекция точки на кривую

Пусть требуется найти проекцию точки \mathbf{p} на прямую

$$\mathbf{c}(t) = \mathbf{c}_0 + t \mathbf{c}', \quad (4.1.1)$$

где \mathbf{c}_0 — точка прямой; \mathbf{c}' — вектор направления ненулевой длины. На рис. 4.1.1 приведены вектор \mathbf{c}' , точка \mathbf{c}_0 и проекция точки \mathbf{p} . Построим вектор из точки прямой \mathbf{c}_0 в точку \mathbf{p} и вычислим скалярное произведение этого вектора и вектора \mathbf{c}' . Если разделим это скалярное произведение на длину вектора \mathbf{c}' , то получим длину проекции вектора $\mathbf{p} - \mathbf{c}_0$ на прямую линию. Если же разделим это скалярное произведение на квадрат длины вектора \mathbf{c}' , то получим длину проекции вектора $\mathbf{p} - \mathbf{c}_0$ на прямую в единицах длины вектора \mathbf{c}' , то есть получим параметр t для проекции точки \mathbf{p} на прямую линию.

Напомним, что за каждой буквой латинского алфавита, выделенной полужирным шрифтом, стоят три координаты. Когда мы складываем, вычитаем или умножаем радиусы-векторы, то для каждой координаты выполняются соответствующие действия. Эти действия однотипные, что дает нам возможность записывать действия в векторном виде. На рисунках мы не показываем глобальную систему координат и радиусы-векторы точек, а показываем только относительное положение рассматриваемых объектов. Повторим подробнее шаги вычисления проекции точки на прямую.

Построим вектор $\mathbf{p} - \mathbf{c}_0$ из начальной точки \mathbf{c}_0 прямой в проецируемую точку \mathbf{p} . Вычислим скалярное произведение $(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'$, которое равно длине вектора производной \mathbf{c}' прямой, умноженной на длину вектора $\mathbf{p} - \mathbf{c}_0$. Разделив скалярное произведение на длину вектора \mathbf{c}' , получим число $\frac{(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'}{\sqrt{\mathbf{c}' \cdot \mathbf{c}'}}$, равное расстоянию (со знаком) от начальной точки

прямой \mathbf{c}_0 до точки \mathbf{p}_p , являющейся проекцией на прямую точки \mathbf{p} . Разделив полученное число еще раз на длину вектора \mathbf{c}' , получим число $\frac{(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'}{\mathbf{c}' \cdot \mathbf{c}'}$, равное длине (со знаком) вектора $\mathbf{p}_p - \mathbf{c}_0$, выраженной в единицах длины производной прямой (4.1.1). Отрицательный знак $\frac{(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'}{\mathbf{c}' \cdot \mathbf{c}'}$ будет говорить о том, что точка \mathbf{p}_p лежит на прямой со стороны отрицательных значений параметра t . Для начальной точки \mathbf{c}_0 параметр t прямой равен нулю. Для точки \mathbf{p}_p параметр t_p прямой будет равен $\frac{(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'}{\mathbf{c}' \cdot \mathbf{c}'}$. По выражению радиуса-вектора прямой (4.1.1) и параметру t_p проекции точки \mathbf{p} найдем радиус-вектор проекции \mathbf{p}_p . Таким образом, параметр t_p проекции точки \mathbf{p} на прямую линию и радиус-вектор проекции \mathbf{p}_p вычисляются по формулам

$$t_p = \frac{(\mathbf{p} - \mathbf{c}_0) \cdot \mathbf{c}'}{|\mathbf{c}'|^2}, \quad \mathbf{p}_p = \mathbf{c}(t_p) = \mathbf{c}_0 + t_p \mathbf{c}'. \quad (4.1.2)$$

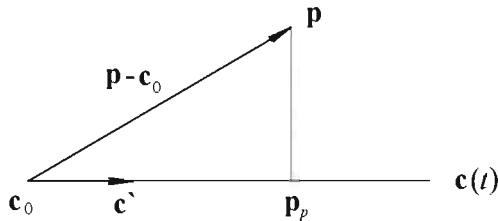


Рис. 4.1.1

Расстояние от точки до ее проекции на кривую равно длине вектора $\mathbf{p}_p - \mathbf{p}$. Расстояние от точки до ее проекции на прямую линию можно определить, не вычисляя проекцию точки, а воспользовавшись формулой

$$l_p = \frac{|(\mathbf{p} - \mathbf{c}_0) \times \mathbf{c}'|}{|\mathbf{c}'|}.$$

Аналогично найдем проекцию точки \mathbf{p} на отрезок прямой:

$$\mathbf{c}(t) = (1 - t) \mathbf{p}_0 + t \mathbf{p}_1.$$

Параметр t_p проекции точки \mathbf{p} на отрезок вычислим по формуле (4.1.2), в которой точку \mathbf{c}_0 заменим точкой \mathbf{p}_0 , а вектор \mathbf{c}' заменим вектором $\mathbf{p}_1 - \mathbf{p}_0$:

$$t_p = \frac{(\mathbf{p} - \mathbf{p}_0) \cdot (\mathbf{p}_1 - \mathbf{p}_0)}{|\mathbf{p}_1 - \mathbf{p}_0|^2}.$$

Пусть требуется найти все проекции точки \mathbf{p} на кривую $\mathbf{c}(t)$. Каждая искомая точка кривой $\mathbf{c}(t)$ удовлетворяет уравнению

$$(\mathbf{p} - \mathbf{c}(t)) \cdot \frac{d\mathbf{c}}{dt} = 0. \quad (4.1.3)$$

Это уравнение содержит одну неизвестную величину — параметр t . Задачу определения проекций точки на кривую решим за два шага. На первом шаге определим начальные приближения параметров проекций точки на кривую, а на втором шаге найдем с заданной точностью значения параметров кривой, определяющие проекции заданной точки \mathbf{p} на кривую $\mathbf{c}(t)$.

Для поиска начальных приближений в некоторых точках кривой с параметрами t_i вычислим левую часть равенства (4.1.3). Алгоритмы выбора параметров пробных точек рассмотрены в конце главы. Если в двух соседних точках кривой с параметрами t_i и t_{i+1} указанное скалярное произведение имеет разные знаки, то значение $0,5(t_i + t_{i+1})$ может служить начальным приближением для поиска корня уравнения (4.1.3). Корни уравнения (4.1.3) можно найти одним из методов решения нелинейных алгебраических уравнений. Например, с помощью метода Ньютона $(r + 1)$ -е приближение параметра проекции точки на кривую вычисляется по формуле

$$t^{(r+1)} = t^{(r)} - \frac{(\mathbf{p} - \mathbf{c}) \cdot \frac{d\mathbf{c}}{dt}}{(\mathbf{p} - \mathbf{c}) \cdot \frac{d^2\mathbf{c}}{dt^2} - \frac{d\mathbf{c}}{dt} \cdot \frac{d\mathbf{c}}{dt}} \Bigg|_{t^{(r)}}.$$

Процесс уточнения параметра закончим, когда на очередной итерации $|t^{(r+1)} - t^{(r)}|$ станет меньше заданной величины.

Проекция точки на многие аналитические кривые может быть найдена без привлечения численных методов. Например, чтобы найти проекцию точки \mathbf{p} на окружность

$$\mathbf{r}(t) = \mathbf{p}_0 + r \cos t \mathbf{i}_x + r \sin t \mathbf{i}_y, \quad 0 \leq t \leq 2\pi,$$

где \mathbf{i}_x и \mathbf{i}_y — два взаимно ортогональных вектора единичной длины, определяющих положение плоскости окружности, нужно перевести проектируемую точку в местную систему координат с началом в точке \mathbf{p}_0 и базисными векторами $\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z = \mathbf{i}_x \times \mathbf{i}_y$, спроецировать эту точку на плоскость окружности и найти параметр двумерной проекции заданной точки. Эти действия описываются формулой

$$t_p = \arctg \left(\frac{\mathbf{w} \cdot \mathbf{i}_v}{\mathbf{w} \cdot \mathbf{i}_r} \right), \quad (4.1.4)$$

где $\mathbf{w} = (\mathbf{p} - \mathbf{p}_0) - \mathbf{i}_z((\mathbf{p} - \mathbf{p}_0) \cdot \mathbf{i}_z)$ – проекция вектора $\mathbf{p} - \mathbf{p}_0$ на плоскость, ортогональную вектору \mathbf{i}_z .

4.2. Проекция точки на поверхность

Пусть требуется найти проекцию точки \mathbf{p} на плоскость:

$$\mathbf{s}(u, v) = \mathbf{s}_0 + u \mathbf{s}_1 + v \mathbf{s}_2, \quad (4.2.1)$$

где \mathbf{s}_0 – точка плоскости; \mathbf{s}_1 и \mathbf{s}_2 – непараллельные векторы ненулевой длины. Точка \mathbf{p} имеет единственную проекцию на плоскость $\mathbf{s}(u, v)$. Построим единичную нормаль к плоскости:

$$\mathbf{m} = \frac{\mathbf{s}_1 \times \mathbf{s}_2}{|\mathbf{s}_1| |\mathbf{s}_2|}. \quad (4.2.2)$$

Далее построим вектор $\mathbf{p} - \mathbf{s}_0$ из начальной точки \mathbf{s}_0 плоскости в проецируемую точку \mathbf{p} . Вычислим скалярное произведение $(\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m}$, которое равно длине проекции вектора $\mathbf{p} - \mathbf{s}_0$ на нормаль \mathbf{m} . Умножив вектор \mathbf{m} на $(\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m}$, получим вектор $\mathbf{m}((\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m})$, параллельный вектору нормали, но длиной, равной расстоянию от проецируемой точки до плоскости (4.2.1), рис. 4.2.1.

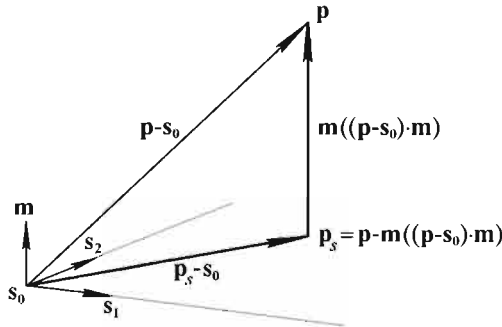


Рис. 4.2.1

Радиус-вектор проекции \mathbf{p}_s точки \mathbf{p} на плоскость равен разности радиуса-вектора проецируемой точки и вектора $\mathbf{m}((\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m})$:

$$\mathbf{p}_s = \mathbf{p} - \mathbf{m}((\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m}).$$

Зная радиус-вектор \mathbf{p}_s , проекции точки \mathbf{p} на плоскость (4.2.1), найдем параметры u_p и v_p плоскости для точки \mathbf{p}_s . Параметры u_p и v_p являются координатами точки \mathbf{p}_s в локальной системе координат плоскости с началом в точке \mathbf{s}_0 и базисными векторами $\mathbf{s}_1, \mathbf{s}_2, \mathbf{m}$. Для общности рассмотрим случай аффинной системы координат. Пусть векторы \mathbf{s}_1 и \mathbf{s}_2 не ортогональны и имеют произвольную ненулевую длину. Для вычисления параметров u_p и v_p точки \mathbf{p}_s найдем длины проекций вектора $\mathbf{p}_s - \mathbf{s}_0$ на векторы \mathbf{s}_1 и \mathbf{s}_2 , которые обозначим через p_1 и p_2

$$p_1 = \frac{(\mathbf{p}_s - \mathbf{s}_0) \cdot \mathbf{s}_1}{|\mathbf{s}_1|}, \quad p_2 = \frac{(\mathbf{p}_s - \mathbf{s}_0) \cdot \mathbf{s}_2}{|\mathbf{s}_2|}.$$

На рис. 4.2.2 приведены векторы \mathbf{s}_1 и \mathbf{s}_2 , точка \mathbf{s}_0 и проекция \mathbf{p}_s заданной точки на плоскость.

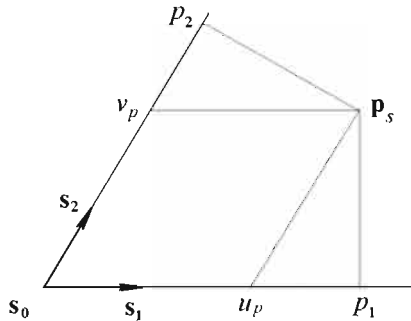


Рис. 4.2.2

Параметры u_p и v_p и длины проекций p_1 и p_2 связаны уравнениями

$$\begin{aligned} p_1 &= u_p |\mathbf{s}_1| + v_p |\mathbf{s}_2| \cos \phi, \\ p_2 &= v_p |\mathbf{s}_2| + u_p |\mathbf{s}_1| \cos \phi, \end{aligned}$$

где $\cos \phi = \frac{\mathbf{s}_1 \cdot \mathbf{s}_2}{|\mathbf{s}_1| |\mathbf{s}_2|}$ – косинус угла между векторами \mathbf{s}_1 и \mathbf{s}_2 . Из системы уравнений найдем параметры проекции точки на плоскость:

$$u_p = \frac{p_1 - p_2 \cos \phi}{(1 - \cos^2 \phi) |\mathbf{s}_1|} = \frac{g_{22} \mathbf{s}_1 \cdot (\mathbf{p}_s - \mathbf{s}_0) - g_{12} \mathbf{s}_2 \cdot (\mathbf{p}_s - \mathbf{s}_0)}{g_{11} g_{22} - g_{12}^2}, \quad (4.2.3)$$

$$v_p = \frac{p_2 - p_1 \cos \phi}{(1 - \cos^2 \phi) |\mathbf{s}_2|} = \frac{g_{11} \mathbf{s}_2 \cdot (\mathbf{p}_s - \mathbf{s}_0) - g_{12} \mathbf{s}_1 \cdot (\mathbf{p}_s - \mathbf{s}_0)}{g_{11} g_{22} - g_{12}^2}, \quad (4.2.4)$$

где $g_{11} = \mathbf{s}_1 \cdot \mathbf{s}_1$, $g_{22} = \mathbf{s}_2 \cdot \mathbf{s}_2$, $g_{12} = \mathbf{s}_1 \cdot \mathbf{s}_2$ — коэффициенты первой основной квадратичной формы плоскости. Если векторы \mathbf{s}_1 и \mathbf{s}_2 ортогональные, то $g_{12} = 0$ и формулы (4.2.3) и (4.2.4) примут вид

$$u_p = \frac{(\mathbf{p}_s - \mathbf{s}_0) \cdot \mathbf{s}_1}{\mathbf{s}_1 \cdot \mathbf{s}_1}, \quad v_p = \frac{(\mathbf{p}_s - \mathbf{s}_0) \cdot \mathbf{s}_2}{\mathbf{s}_2 \cdot \mathbf{s}_2}.$$

Расстояние от точки до ее проекции в общем случае вычисляется как длина вектора $\mathbf{p}_s - \mathbf{p}$. Расстояние от точки до ее проекции на плоскость можно определить, не вычисляя проекцию точки, а вычислив проекцию вектора $\mathbf{p} - \mathbf{s}_0$ на нормаль к плоскости:

$$l_s = (\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m}.$$

Переход от пространственных координат точки к параметрам поверхности и обратно является довольно частой операцией. Найдем матрицу преобразования из локальной системы координат плоскости в глобальную систему координат. Итак, мы получили, что точка \mathbf{p} в локальной системе координат плоскости с началом в точке \mathbf{s}_0 и базисными векторами \mathbf{s}_1 , \mathbf{s}_2 , \mathbf{m} имеет координаты u_p , v_p , l_s , следовательно, справедливо равенство:

$$\mathbf{p} = \mathbf{s}_0 + u_p \mathbf{s}_1 + v_p \mathbf{s}_2 + l_s \mathbf{m}.$$

Запишем это равенство в матричном виде:

$$\mathbf{p} = \mathbf{s}_0 + [\mathbf{s}_1 \ \mathbf{s}_2 \ \mathbf{m}] \cdot \begin{bmatrix} u_p \\ v_p \\ l_s \end{bmatrix} = \mathbf{s}_0 + \mathbf{A} \cdot \begin{bmatrix} u_p \\ v_p \\ l_s \end{bmatrix},$$

где столбцы матрицы \mathbf{A} составлены из компонент базисных векторов \mathbf{s}_1 , \mathbf{s}_2 , \mathbf{m} локальной системы координат. Пусть вектор \mathbf{s}_1 описывается компонентами s_{11} , s_{12} , s_{13} , вектор \mathbf{s}_2 описывается компонентами s_{21} , s_{22} , s_{23} , вектор \mathbf{m} описывается компонентами m_1 , m_2 , m_3 , тогда

$$\mathbf{A} = \begin{bmatrix} s_{11} & s_{21} & m_1 \\ s_{12} & s_{22} & m_2 \\ s_{13} & s_{23} & m_3 \end{bmatrix}.$$

Таким образом, переход от пространственных координат точки к параметрам плоскости можно выполнить с помощью равенства

$$\begin{bmatrix} u_p \\ v_p \\ l_s \end{bmatrix} = \mathbf{A}^{-1} \cdot (\mathbf{p} - \mathbf{s}_0),$$

где \mathbf{A}^{-1} — обратная матрица для матрицы \mathbf{A} .

Определение параметров проекции точки на плоскость используется при поиске начальных приближений для **проекции точки на поверхность**. В итерационных методах решения систем уравнений поиск начальных приближений играет важную роль. Начальное приближение должно находиться в области схождения решения к искомой точке. Неправильное задание начального приближения может привести к тому, что решение системы уравнений будет уходить от искомой точки, то есть будет расходиться. Кроме того, поиск начальных приближений требует затрат времени на порядок больше, чем итерационный процесс решения системы уравнений при известном начальном приближении.

Пусть требуется найти все проекции точки \mathbf{p} на поверхность $\mathbf{s}(u, v)$. Каждая искомая точка поверхности $\mathbf{s}(u, v)$ удовлетворяет системе двух уравнений:

$$\begin{aligned}(\mathbf{p} - \mathbf{s}(u, v)) \cdot \frac{\partial \mathbf{s}}{\partial u} &= 0, \\(\mathbf{p} - \mathbf{s}(u, v)) \cdot \frac{\partial \mathbf{s}}{\partial v} &= 0.\end{aligned}\tag{4.2.5}$$

Система уравнений (4.2.5) содержит две неизвестные величины – параметры u и v . Задачу определения проекций точки на поверхность решим за два шага. На первом шаге определим начальные приближения параметров проекций точки на поверхность, а на втором шаге найдем с заданной точностью значения параметров поверхности, определяющие проекции заданной точки \mathbf{p} на поверхность $\mathbf{s}(u, v)$.

Для поиска начальных приближений в окрестности некоторых точек поверхности вычислим левые части равенств (4.2.5). Пусть пробные точки поверхности имеют параметры u_i и v_j , а их окрестности определяются приращениями параметров Δu и Δv . Вычислим величины:

$$\begin{aligned}a_j &= (\mathbf{p} - \mathbf{s}(u_i - \Delta u, v_j)) \cdot \frac{\partial \mathbf{s}}{\partial u}, & b_j &= (\mathbf{p} - \mathbf{s}(u_i + \Delta u, v_j)) \cdot \frac{\partial \mathbf{s}}{\partial u}, \\c_j &= (\mathbf{p} - \mathbf{s}(u_i, v_j - \Delta v)) \cdot \frac{\partial \mathbf{s}}{\partial v}, & d_j &= (\mathbf{p} - \mathbf{s}(u_i, v_j + \Delta v)) \cdot \frac{\partial \mathbf{s}}{\partial v}.\end{aligned}$$

Если $a_j b_j \leq 0$ и $c_j d_j \leq 0$, то в окрестности пробной точки может находиться решение системы (4.2.5), а параметры u_i , v_j могут быть приняты за начальное приближение. Начиная с начального приближения параметров $u^{(0)} = u_i$, $v^{(0)} = v_j$, одним из методов решения нелинейных алгебраических уравнений найдем решение системы уравнений с заданной точностью. Например, в методе Ньютона на r -й итерации приращения параметров $\Delta u^{(r)}$ и $\Delta v^{(r)}$ проекции найдем из системы линейных уравнений:

$$((\mathbf{p}-\mathbf{s}) \cdot \mathbf{s}_{11} - \mathbf{s}_1 \cdot \mathbf{s}_1) \Delta u^{(r)} + ((\mathbf{p}-\mathbf{s}) \cdot \mathbf{s}_{12} - \mathbf{s}_2 \cdot \mathbf{s}_1) \Delta v^{(r)} = (\mathbf{s}-\mathbf{p}) \cdot \mathbf{s}_1,$$

$$((\mathbf{p}-\mathbf{s}) \cdot \mathbf{s}_{21} - \mathbf{s}_1 \cdot \mathbf{s}_2) \Delta u^{(r)} + ((\mathbf{p}-\mathbf{s}) \cdot \mathbf{s}_{22} - \mathbf{s}_2 \cdot \mathbf{s}_2) \Delta v^{(r)} = (\mathbf{s}-\mathbf{p}) \cdot \mathbf{s}_2,$$

где $\mathbf{s} = \mathbf{s}(u^{(r)}, v^{(r)})$, $\mathbf{s}_1 = \frac{\partial \mathbf{s}}{\partial u}$, $\mathbf{s}_2 = \frac{\partial \mathbf{s}}{\partial v}$, $\mathbf{s}_{11} = \frac{\partial^2 \mathbf{s}}{\partial u^2}$, $\mathbf{s}_{12} = \mathbf{s}_{21} = \frac{\partial^2 \mathbf{s}}{\partial u \partial v}$, $\mathbf{s}_{22} = \frac{\partial^2 \mathbf{s}}{\partial v^2}$ — частные производные радиуса-вектора поверхности при $u = u^{(r)}$ и $v = v^{(r)}$. Следующее приближение параметров проекции точки равно $u^{(r+1)} = u^{(r)} + \Delta u^{(r)}$, $v^{(r+1)} = v^{(r)} + \Delta v^{(r)}$. Процесс уточнения параметров закончим, когда на очередной итерации значения $|\Delta u^{(r)}|$ и $|\Delta v^{(r)}|$ станут меньше заданной величины.

Проекции точки на некоторые аналитические поверхности могут быть найдены без привлечения численных методов. Например, чтобы найти проекции точки на поверхность кругового цилиндра, конуса, сферы или тора, нужно перевести проецируемую точку в местную систему координат поверхности, где легко найти параметры проекций.

4.3. Точки пересечения кривых

Найдем точку пересечения двумерных прямых:

$$l_1(w_1) = \mathbf{p}_1 + w_1 \mathbf{c}_1, \quad (4.3.1)$$

$$l_2(w_2) = \mathbf{p}_2 + w_2 \mathbf{c}_2. \quad (4.3.2)$$

Прямые имеют единственную точку пересечения, если они не параллельны. Через \mathbf{n}_1 и \mathbf{n}_2 обозначим единичные нормали к прямым $l_1(w_1)$ и $l_2(w_2)$. Напомним, что за каждой буквой латинского алфавита, выделенной полужирным наклонным шрифтом, стоят две координаты двумерного пространства. Если касательный вектор $\mathbf{c}' = [x \ y]^T$ кривой имеет координаты x и y , то ортогональный ему вектор \mathbf{n} имеет координаты $-ky$ и kx , где $k = \frac{1}{\sqrt{x^2 + y^2}}$. На рис. 4.3.1 показаны прямые $l_1(w_1)$ и $l_2(w_2)$, норма-

ли к ним и точка пересечения \mathbf{p} . Из рисунка видно, что параметр второй прямой для точки пересечения может быть найден как отношение проекции вектора $\mathbf{p}_1 - \mathbf{p}_2$ на нормаль первой прямой к проекции вектора \mathbf{c}_2' на ту же нормаль.

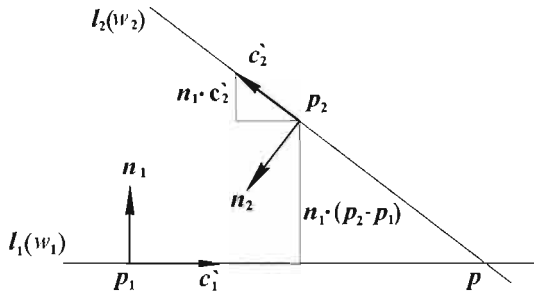


Рис. 4.3.1

Таким образом, параметры точки пересечения для прямых равны

$$w_1 = \frac{n_2 \cdot (p_2 - p_1)}{n_2 \cdot c_1^*}, \quad (4.3.3)$$

$$w_2 = \frac{n_1 \cdot (p_1 - p_2)}{n_1 \cdot c_2^*}. \quad (4.3.4)$$

Радиус-вектор точки пересечения получим, подставив найденный параметр (4.3.3) или (4.3.4) в формулу для соответствующей прямой:

$$p = p_1 + \frac{n_2 \cdot (p_2 - p_1)}{n_2 \cdot c_1^*} c_1^* = p_2 + \frac{n_1 \cdot (p_1 - p_2)}{n_1 \cdot c_2^*} c_2^*.$$

Пусть даны **двумерные отрезки**

$$l_1(t_1) = (1 - t_1) p_1 + t_1 q_1, \quad l_2(t_2) = (1 - t_2) p_2 + t_2 q_2,$$

построенные по точкам $p_1 = [x_1 \ y_1]^T$, $p_2 = [x_2 \ y_2]^T$, $q_1 = [a_1 \ b_1]^T$, $q_2 = [a_2 \ b_2]^T$. Параметры точки пересечения найдем из векторного уравнения

$$(1 - t_1) p_1 + t_1 q_1 = (1 - t_2) p_2 + t_2 q_2,$$

которое представляет собой систему двух скалярных уравнений:

$$\begin{aligned} (x_1 - a_1) t_1 - (x_2 - a_2) t_2 &= x_1 - x_2, \\ (y_1 - b_1) t_1 - (y_2 - b_2) t_2 &= y_1 - y_2. \end{aligned}$$

Если $(x_2 - a_2)(y_1 - b_1) - (x_1 - a_1)(y_2 - b_2) \neq 0$, то по формуле Крамера параметры точки пересечения равны

$$t_1 = \frac{(x_2 - a_2)(y_1 - y_2) - (x_1 - x_2)(y_2 - b_2)}{(x_2 - a_2)(y_1 - b_1) - (x_1 - a_1)(y_2 - b_2)},$$

$$t_2 = \frac{(x_1 - a_1)(y_2 - y_1) - (x_2 - x_1)(y_1 - b_1)}{(x_1 - a_1)(y_2 - b_2) - (x_2 - a_2)(y_1 - b_1)}.$$

Отрезки пересекаются, если $0 \leq t_1 \leq 1$ и $0 \leq t_2 \leq 1$, в противном случае пересекаются продолжения отрезков.

Пусть требуется найти все **точки пересечения двумерных кривых**, радиусы-векторы которых описываются функциями $c_1(t_1)$ и $c_2(t_2)$. В каждой точке пересечения удовлетворяется векторное уравнение

$$c_1(t_1) - c_2(t_2) = 0. \quad (4.3.5)$$

Неизвестными являются параметры кривых t_1 и t_2 . Система уравнений в общем случае является нелинейной и решается итерационными методами.

Для нахождения начальных приближений выберем набор пробных точек для каждой кривой. В каждом сочетании пробных точек построим касательные прямые (4.3.1) и (4.3.2), где $p_1 = c_1(t_1)$, $p_2 = c_2(t_2)$ – пробные точки кривых, $c_1' = dc_1/dt_1$, $c_2' = dc_2/dt_2$ – производные кривых в пробных точках. Далее по формулам (4.3.3) и (4.3.4) определим параметры w_1 и w_2 касательных прямых в точке их пересечения. Вблизи точки пересечения параметры w_1 и w_2 приблизительно соответствуют приращениям параметров кривых. Приращения w_1 и w_2 будем считать малыми, если в окрестности пробной точки кривой кривую можно приближенно аппроксимировать касательной прямой. Если приращения малы, то величины $t_1^{(0)} = t_1 + w_1$ и $t_2^{(0)} = t_2 + w_2$ могут быть приняты в качестве начального приближения при решении системы уравнений (4.3.5).

Точки пересечения некоторых аналитических кривых могут быть найдены без привлечения численных методов. Например, путем решения квадратного уравнения могут быть найдены точки пересечения прямой с кривыми второго порядка.

Точки касания кривых $c_1(t_1)$ и $c_2(t_2)$ могут быть найдены путем решения системы уравнений

$$\begin{aligned} (c_1(t_1) - c_2(t_2)) \cdot \frac{dc_1}{dt_1} &= 0, \\ (c_2(t_2) - c_1(t_1)) \cdot \frac{dc_2}{dt_2} &= 0. \end{aligned} \quad (4.3.6)$$

Кроме точек пересечения и касания этой системе удовлетворяют также точки, в которых касательные прямые параллельны друг другу и отрезок, соединяющий точки касания, ортогонален касательным прямым.

Найдем ближайшие точки трехмерных прямых:

$$\begin{aligned} c_1(t_1) &= p_1 + t_1 d_1, \\ c_2(t_2) &= p_2 + t_2 d_2. \end{aligned}$$

Каждая прямая определена точкой \mathbf{p}_i и вектором \mathbf{d}_i , $i = 1, 2$. Пусть ближайшим точкам соответствуют параметры прямых w_1 и w_2 . Отрезок, соединяющий ближайшие точки, будет ортогонален обоим прямым. То есть для искомым точек должны выполняться равенства:

$$\begin{aligned}(\mathbf{p}_1 + w_1 \mathbf{d}_1 - \mathbf{p}_2 - w_2 \mathbf{d}_2) \cdot \mathbf{d}_1 &= 0, \\(\mathbf{p}_1 + w_1 \mathbf{d}_1 - \mathbf{p}_2 - w_2 \mathbf{d}_2) \cdot \mathbf{d}_2 &= 0.\end{aligned}$$

Если прямые не параллельны, то система приведенных уравнений имеет единственное решение:

$$\begin{aligned}w_1 &= \frac{(\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{d}_1 a_{22} - (\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{d}_2 a_{12}}{a_{11} a_{22} - a_{12} a_{21}}, \\w_2 &= \frac{(\mathbf{p}_1 - \mathbf{p}_2) \cdot \mathbf{d}_2 a_{11} - (\mathbf{p}_1 - \mathbf{p}_2) \cdot \mathbf{d}_1 a_{12}}{a_{11} a_{22} - a_{12} a_{21}},\end{aligned}$$

где $a_{11} = \mathbf{d}_1 \cdot \mathbf{d}_1$; $a_{22} = \mathbf{d}_2 \cdot \mathbf{d}_2$; $a_{12} = \mathbf{d}_1 \cdot \mathbf{d}_2$. Если прямые пересекаются, то w_1 и w_2 соответствуют точке их пересечения. Если прямые не пересекаются, то w_1 и w_2 соответствуют ближайшим точкам прямых.

Для определения параметров t_1 и t_2 **точек пересечения трехмерных кривых** $\mathbf{c}_1(t_1)$ и $\mathbf{c}_2(t_2)$ будем использовать проекцию векторного уравнения

$$\mathbf{c}_1(t_1) - \mathbf{c}_2(t_2) = 0$$

на плоскость, определяемую касательными к кривым векторами $\mathbf{c}_1' = d\mathbf{c}_1/dt_1$ и $\mathbf{c}_2' = d\mathbf{c}_2/dt_2$. Для этого построим локальную систему координат с началом в точке $0,5(\mathbf{c}_1(t_1) + \mathbf{c}_2(t_2))$ и базисными векторами $\mathbf{i}_1 = \mathbf{c}_1'$, $\mathbf{i}_2 = \mathbf{c}_2'$, $\mathbf{i}_3 = \mathbf{i}_1 \times \mathbf{i}_2$ и переведем в эту систему векторное уравнение. Решение будем искать для системы уравнений, составленной из равенства первых двух локальных координат точек кривых. При решении системы уравнений итерационными методами на каждой итерации строить локальную систему координат лучше заново. Если касательные к кривым векторы \mathbf{c}_1' и \mathbf{c}_2' параллельны, то следует перейти к системе уравнений (4.3.6), в которой двумерные кривые нужно заменить трехмерными кривыми.

При определении параметров пересечения пространственных, кривых кроме точек пересечения и касания, мы получим также точки, в которых касательные прямые ортогональны отрезку, соединяющему найденные точки. После нахождения решения следует проверить найденные точки кривых на равенство всех трех координат.

4.4. Точки пересечения поверхности и кривой

Пусть требуется найти пересечение прямой (4.1.1) и плоскости (4.2.1). Если $\mathbf{c}'\mathbf{m} \neq 0$, где \mathbf{m} – нормаль (4.2.2) плоскости, то прямая не

параллельна плоскости, точка пересечения существует и является единственной. Параметр прямой, соответствующий точке ее пересечения с плоскостью, вычислим по формуле

$$t_p = \frac{\mathbf{m} \cdot (\mathbf{s}_0 - \mathbf{c}_0)}{\mathbf{m} \cdot \mathbf{c}'}. \quad (4.4.1)$$

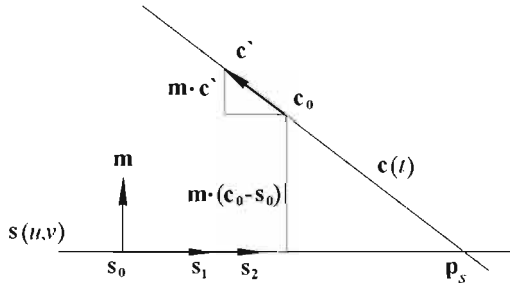


Рис. 4.4.1

На рис. 4.4.1 приведены прямая и плоскость при взгляде параллельно плоскости. В числителе правой части (4.4.1) стоит расстояние от начальной точки прямой \mathbf{c}_0 до плоскости, а в знаменателе стоит нормальная к плоскости составляющая производной прямой. Отношение этих величин даст параметр пересечения прямой с плоскостью. Подставим (4.4.1) в (4.1.1) и получим радиус-вектор \mathbf{p}_s точки пересечения плоскости и прямой:

$$\mathbf{p}_s = \mathbf{c}_0 + \mathbf{c}' \frac{\mathbf{m} \cdot (\mathbf{s}_0 - \mathbf{c}_0)}{\mathbf{m} \cdot \mathbf{c}'}. \quad (4.4.2)$$

Далее по формулам (4.2.3) и (4.2.4) вычислим параметры плоскости, соответствующие точке пересечения.

Определение параметров пересечения прямой и плоскости используется при поиске начальных приближений для точек **пересечения поверхности и кривой**. Пусть требуется найти все точки пересечения поверхности, радиус-вектор которой описывается векторной функцией $\mathbf{s}(u, v)$, и кривой, радиус-вектор которой описывается векторной функцией $\mathbf{c}(t)$. В каждой точке пересечения должно удовлетворяться векторное уравнение

$$\mathbf{s}(u, v) - \mathbf{c}(t) = 0, \quad (4.4.2)$$

представляющее собой систему трех скалярных уравнений. В качестве неизвестных выступают три параметра — u , v и t . Решив эту систему уравнений, получим параметры u , v и t , по которым при необходимости можно вычислить точки на поверхности и на кривой. Система уравне-

ний в общем случае является нелинейной и решается итерационными методами. Точек пересечения может быть несколько.

Для нахождения начальных приближений решения выберем набор пробных точек поверхности и кривой. Для каждого сочетания пробных точек выполним следующую проверку. Пусть в пробной точке поверхности значения параметров равны u_i и v_i , а в пробной точке кривой значение параметра равно t_j . Построим касательную плоскость (4.2.1) в пробной точке поверхности и касательную прямую (4.1.1) в пробной точке кривой, где $s_0 = s(u_i, v_i)$, $c_0 = c(t_j)$, $s_1 = ds / du$ и $s_2 = ds / dv$ в пробной точке поверхности, $c' = dc / dt$ в пробной точке кривой. Если пробные точки поверхности и кривой являются обычными и если касательная прямая не параллельна касательной плоскости, то по формулам (4.4.1), (4.2.3) и (4.2.4) найдем параметр t_p касательной прямой и параметры u_p и v_p касательной плоскости, определяющие точку их пересечения. Если пробные точки находятся вблизи точки пересечения кривой и поверхности, то значения u_p и v_p приблизительно соответствуют приращениям параметров поверхности, а значение t_p приблизительно соответствует приращению параметра кривой. Приращения u_p и v_p будем считать малыми, если в окрестности точки поверхности с параметрами u_i, v_i поверхность можно приближенно аппроксимировать касательной плоскостью. Приращение t_p будем считать малым, если в окрестности точки кривой с параметром t_j , кривую можно приближенно аппроксимировать касательной прямой. Если приращения малы, то параметры $u^{(0)} = u_i + u_p, v^{(0)} = v_i + v_p, t^{(0)} = t_j + t_p$ могут быть приняты в качестве начального приближения для решения системы уравнений (4.4.2). На рис. 4.4.2 приведены поверхность, касательная к ней плоскость, кривая и касательная к ней прямая.

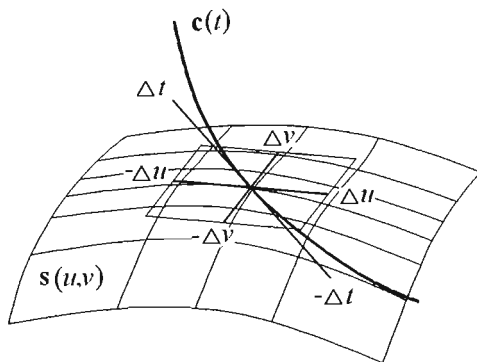


Рис. 4.4.2

Точки касания кривой и поверхности могут быть найдены путем решения системы трех уравнений:

$$(\mathbf{s}(u, v) - \mathbf{c}(t)) \cdot \frac{\partial \mathbf{s}}{\partial u} = 0,$$

$$(\mathbf{s}(u, v) - \mathbf{c}(t)) \cdot \frac{\partial \mathbf{s}}{\partial v} = 0,$$

$$(\mathbf{s}(u, v) - \mathbf{c}(t)) \cdot \frac{d\mathbf{c}}{dt} = 0.$$

Кроме точек пересечения и касания, этой системе удовлетворяют также точки, в которых касательная прямая параллельна касательной плоскости и отрезок, соединяющий точки касания, ортогонален касательной прямой и касательной плоскости.

4.5. Точки пересечения трех поверхностей

Пусть даны три плоскости

$$\begin{aligned} \mathbf{r}(u, v) &= \mathbf{r}_0 + u\mathbf{r}_1 + v\mathbf{r}_2, \\ \mathbf{s}(a, b) &= \mathbf{s}_0 + a\mathbf{s}_1 + b\mathbf{s}_2, \\ \mathbf{q}(x, y) &= \mathbf{q}_0 + x\mathbf{q}_1 + y\mathbf{q}_2 \end{aligned} \quad (4.5.1)$$

с единичными нормальными векторами \mathbf{m}_r , \mathbf{m}_s , \mathbf{m}_q соответственно. Плоскости имеют единственную точку пересечения, если смешанное произведение нормалей отлично от нуля: $(\mathbf{m}_r \times \mathbf{m}_s) \cdot \mathbf{m}_q \neq 0$. Пусть искомая точка пересечения имеет радиус-вектор \mathbf{p} . Тогда вектор $\mathbf{p} - \mathbf{r}_0$ параллелен плоскости $\mathbf{r}(u, v)$, вектор $\mathbf{p} - \mathbf{s}_0$ параллелен плоскости $\mathbf{s}(a, b)$, вектор $\mathbf{p} - \mathbf{q}_0$ параллелен плоскости $\mathbf{q}(x, y)$. Следовательно, должны выполняться равенства

$$(\mathbf{p} - \mathbf{r}_0) \cdot \mathbf{m}_r = 0, \quad (\mathbf{p} - \mathbf{s}_0) \cdot \mathbf{m}_s = 0, \quad (\mathbf{p} - \mathbf{q}_0) \cdot \mathbf{m}_q = 0. \quad (4.5.2)$$

Обозначим через p_1, p_2, p_3 координаты точки \mathbf{p} , через m_1, m_2, m_3 — компоненты вектора нормали \mathbf{m}_r , через n_1, n_2, n_3 — компоненты вектора нормали \mathbf{m}_s , через l_1, l_2, l_3 — компоненты вектора нормали \mathbf{m}_q . Тогда система уравнений (4.5.1) примет вид

$$\begin{aligned} m_1 p_1 + m_2 p_2 + m_3 p_3 &= \mathbf{r}_0 \cdot \mathbf{m}_r, \\ n_1 p_1 + n_2 p_2 + n_3 p_3 &= \mathbf{s}_0 \cdot \mathbf{m}_s, \\ l_1 p_1 + l_2 p_2 + l_3 p_3 &= \mathbf{q}_0 \cdot \mathbf{m}_q. \end{aligned}$$

Решив эту систему линейных алгебраических уравнений относительно p_1, p_2, p_3 , найдем общую точку трех плоскостей. Параметры пересечения плоскостей найдем путем проецирования точки \mathbf{p} на эти плоскости описанным выше способом.

В общем случае **пересечения трех поверхностей** $\mathbf{r}(u,v)$, $\mathbf{s}(a,b)$ и $\mathbf{q}(x,y)$ каждая точка пересечения должна удовлетворять системе векторных уравнений

$$\begin{aligned}\mathbf{r}(u,v) - \mathbf{q}(x,y) &= 0, \\ \mathbf{s}(a,b) - \mathbf{q}(x,y) &= 0.\end{aligned}$$

Эта система векторных уравнений содержит шесть скалярных алгебраических уравнений относительно шести параметров u , v , a , b , x , y поверхностей.

Для нахождения начальных приближений решения выберем набор пробных точек поверхностей. Для каждого сочетания трех пробных точек построим три касательные плоскости (4.5.1), где $\mathbf{r}_0 = \mathbf{r}(u_i, v_i)$, $\mathbf{r}_1 = \partial \mathbf{r} / \partial u$, $\mathbf{r}_2 = \partial \mathbf{r} / \partial v$ вычислены в пробной точке $[u_i, v_i]^T$ поверхности $\mathbf{r}(u,v)$; $\mathbf{s}_0 = \mathbf{s}(a_j, b_j)$, $\mathbf{s}_1 = \partial \mathbf{s} / \partial a$, $\mathbf{s}_2 = \partial \mathbf{s} / \partial b$ вычислены в пробной точке $[a_j, b_j]^T$ поверхности $\mathbf{s}(a,b)$; $\mathbf{q}_0 = \mathbf{q}(x_k, y_k)$, $\mathbf{q}_1 = \partial \mathbf{q} / \partial x$, $\mathbf{q}_2 = \partial \mathbf{q} / \partial y$ вычислены в пробной точке $[x_k, y_k]^T$ поверхности $\mathbf{q}(x,y)$. Найдем точку пересечения трех касательных плоскостей и оценим ее близость к пробным точкам и, соответственно, возможность использовать рассматриваемые пробные точки в качестве начального приближения.

4.6. Кривые на поверхностях

Плоская кривая может быть построена с помощью двумерной кривой

$$\mathbf{c}(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}, \quad t_{\min} \leq t \leq t_{\max}$$

и локальной системы координат в трехмерном пространстве. Пусть начало локальной декартовой системы координат расположено в точке \mathbf{p} , а векторы \mathbf{s}_1 , \mathbf{s}_2 , \mathbf{s}_3 являются ее базисными векторами. Тогда радиус-вектор кривой, построенной в плоскости первых двух базисных векторов, вычислим по формуле

$$\mathbf{r}(t) = \mathbf{p} + u(t) \mathbf{s}_1 + v(t) \mathbf{s}_2.$$

Плоская кривая имеет ту же область определения, что и двумерная кривая. Производные радиуса-вектора плоской кривой вычисляются как производные сложной функции:

$$\frac{d\mathbf{r}}{dt} = \mathbf{s}_1 \frac{du}{dt} + \mathbf{s}_2 \frac{dv}{dt}, \quad \frac{d^2\mathbf{r}}{dt^2} = \mathbf{s}_1 \frac{d^2u}{dt^2} + \mathbf{s}_2 \frac{d^2v}{dt^2}, \quad \frac{d^3\mathbf{r}}{dt^3} = \mathbf{s}_1 \frac{d^3u}{dt^3} + \mathbf{s}_2 \frac{d^3v}{dt^3}.$$

Двумерная кривая $\mathbf{c}(t) = [u(t) \ v(t)]^T$ в пространстве параметров плоскости (4.2.1) образует плоскую кривую

$$\mathbf{r}(t) = \mathbf{s}_0 + u(t) \mathbf{s}_1 + v(t) \mathbf{s}_2$$

в трехмерном пространстве.

Кривая на поверхности описывается векторной функцией (3.1.3) и представляет собой совокупность поверхности $\mathbf{s}(u,v)$ и двумерной кривой $\mathbf{c}(t) = [u(t) \ v(t)]^T$, $t_{\min} \leq t \leq t_{\max}$ в пространстве параметров этой поверхности. Двумерная кривая отображает некоторый участок числовой оси в пространство параметров поверхности, а поверхность отображает соответствующие точки в трехмерное пространство. Кривая на поверхности является трехмерной кривой, ее радиус-вектор вычисляется по формуле

$$\mathbf{r}(t) = \mathbf{s}(u(t), v(t)). \quad (4.6.1)$$

Кривая на поверхности имеет ту же область определения, что и двумерная кривая. Производные радиуса-вектора кривой на поверхности вычисляются как производные сложной функции:

$$\frac{d\mathbf{r}}{dt} = \mathbf{s}_1 \frac{du}{dt} + \mathbf{s}_2 \frac{dv}{dt}, \quad (4.6.2)$$

$$\frac{d^2\mathbf{r}}{dt^2} = \mathbf{s}_{11} \left(\frac{du}{dt}\right)^2 + 2\mathbf{s}_{12} \frac{du}{dt} \frac{dv}{dt} + \mathbf{s}_{22} \left(\frac{dv}{dt}\right)^2 + \mathbf{s}_1 \frac{d^2u}{dt^2} + \mathbf{s}_2 \frac{d^2v}{dt^2}, \quad (4.6.3)$$

$$\begin{aligned} \frac{d^3\mathbf{r}}{dt^3} = & \mathbf{s}_{111} \left(\frac{du}{dt}\right)^3 + 3\mathbf{s}_{112} \left(\frac{du}{dt}\right)^2 \frac{dv}{dt} + 3\mathbf{s}_{122} \frac{du}{dt} \left(\frac{dv}{dt}\right)^2 + \mathbf{s}_{222} \left(\frac{dv}{dt}\right)^3 + \\ & + 3\mathbf{s}_{11} \frac{d^2u}{dt^2} \frac{du}{dt} + 3\mathbf{s}_{12} \left(\frac{d^2u}{dt^2} \frac{dv}{dt} + \frac{d^2v}{dt^2} \frac{du}{dt}\right) + 3\mathbf{s}_{22} \frac{d^2v}{dt^2} \frac{dv}{dt} + \mathbf{s}_1 \frac{d^3u}{dt^3} + \mathbf{s}_2 \frac{d^3v}{dt^3}, \end{aligned} \quad (4.6.4)$$

где $\mathbf{s}_1 = \frac{\partial \mathbf{s}}{\partial u}$, $\mathbf{s}_2 = \frac{\partial \mathbf{s}}{\partial v}$, $\mathbf{s}_{11} = \frac{\partial^2 \mathbf{s}}{\partial u^2}$, $\mathbf{s}_{12} = \frac{\partial^2 \mathbf{s}}{\partial u \partial v}$, $\mathbf{s}_{22} = \frac{\partial^2 \mathbf{s}}{\partial v^2}$, $\mathbf{s}_{111} = \frac{\partial^3 \mathbf{s}}{\partial u^3}$, $\mathbf{s}_{112} = \frac{\partial^3 \mathbf{s}}{\partial u^2 \partial v}$, $\mathbf{s}_{122} = \frac{\partial^3 \mathbf{s}}{\partial u \partial v^2}$, $\mathbf{s}_{222} = \frac{\partial^3 \mathbf{s}}{\partial v^3}$ — частные производные радиуса-вектора поверхности в точке $[u(t) \ v(t)]^T$.

Кривая на поверхности будет периодической, если двумерная кривая периодическая или если поверхность периодическая по одному из параметров и двумерная кривая проходит через точки, отстоящие на период поверхности по периодическому параметру при неизменном другом параметре.

При выходе двумерной точки $[u \ v]^T$ за пределы области определения параметров поверхности радиус-вектор $\mathbf{r}(u,v)$ вычисляется в соответствии с правилами построения продолженной поверхности.

Пусть граница поверхности $s(u, v)$ описывается с помощью n двумерных контуров $c_i(t_i) = [u_i(t_i) \ v_i(t_i)]^T$, $t_{\min} \leq t_i \leq t_{\max}$, $i=1, 2, \dots, n$. В трехмерном пространстве границу этой поверхности можно описать с помощью совокупности кривых на поверхности, двумерными кривыми которых являются двумерные контуры, или с помощью трехмерных замкнутых составных кривых, сегментами которых служат кривые на поверхности $c_i(t_i) = s(u_i(t_i), v_i(t_i))$.

Проекционная кривая представляет собой проекцию трехмерной кривой на поверхность. Пусть дана трехмерная кривая $c(t)$, $t_{\min} \leq t \leq t_{\max}$ и поверхность $s(u, v)$. Двумерную кривую $p(t) = [u(t) \ v(t)]^T$, каждая точка $[u(w) \ v(w)]^T$ которой совпадает с проекцией точки $c(w)$ на поверхность $s(u, v)$, называют *двумерной проекционной кривой*, рис. 4.6.1. В совокупности с заданной поверхностью $s(u, v)$ двумерная проекционная кривая $p(t)$ образует кривую на поверхности (4.6.1) – *трехмерную проекционную кривую* $s(u(t), v(t))$.

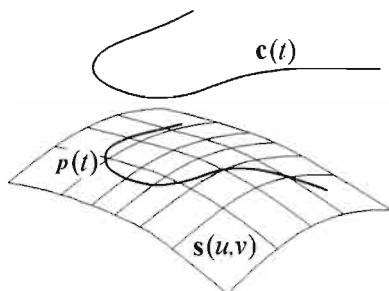


Рис. 4.6.1

В общем случае построить двумерную кривую $p(t)$ не всегда возможно, поэтому аппроксимируем ее некоторым сплайном. Для построения сплайна выберем некоторую неубывающую последовательность параметров t_i , $i = 0, 1, \dots, n$, где $t_i < t_{i+1}$, $t_0 = t_{\min}$, $t_n = t_{\max}$. Для каждой точки кривой $c(t_i)$, $i = 0, 1, \dots, n$ найдем параметры u_i и v_i проекции этой точки на поверхность $s(u, v)$ и построим некоторый двумерный сплайн $r(t)$, проходящий через точки $p = [u_i \ v_i]^T$ при значениях параметра $t = t_i$, $i = 0, 1, \dots, n$. Сплайн $r(t)$ имеет ту же параметрическую длину, что и кривая $c(t)$, и совпадает с проекционной кривой в отдельных точках. При наличии такого сплайна точку проекционной кривой для некоторого заданного параметра w найдем по следующему алгоритму: вычислим точку сплайна $r(w)$ для того же параметра и используем ее в качестве начального приближения в процессе поиска проекции точки $c(w)$ на поверхность $s(u, v)$.

Проекционную кривую построим из трех объектов: трехмерной кривой $c(t)$, поверхности $s(u, v)$ и двумерной кривой начальных приближе-

ний $r(t)$, $t_{\min} \leq t \leq t_{\max}$. Эти объекты вместе с алгоритмом проецирования точки на поверхность позволяют определять точки проекционной кривой с заданной точностью. Если $s(u, v)$ является плоскостью, то с заданной точностью можно определить производные радиуса-вектора проекционной кривой.

Линия очерка является разновидностью кривой на поверхности. Пусть дана поверхность $s(u, v)$ и некоторая точка наблюдения w . Линию, проходящую через точку наблюдения и точку поверхности, будем называть *линией визирования*. Точки, в которых нормаль поверхности ортогональна линии визирования, расположены на кривых, называемых *линиями очерка*. При переходе через линию очерка нормаль поверхности меняет направление по отношению к линии визирования. Линия очерка разделяет видимую из точки наблюдения часть поверхности и невидимую из точки наблюдения часть поверхности.

На рис. 4.6.2 приведены линии очерка поверхности вращения.

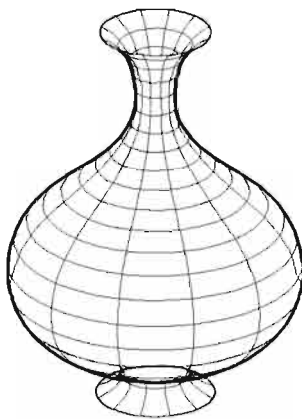


Рис. 4.6.2

Расположение линий очерка на поверхности зависит от взаимного положения поверхности и точки наблюдения. В общем случае у поверхности может быть несколько линий очерка. Каждая линия очерка или замкнута, или оканчивается на границе поверхности.

По определению точки линии очерка должны удовлетворять уравнению

$$\mathbf{i}(u, v) \cdot \mathbf{m}(u, v) = 0, \quad (4.6.5)$$

где $\mathbf{m}(u, v)$ – нормаль поверхности $s(u, v)$; $\mathbf{i}(u, v) = s(u, v) - w$ – вектор направления линии визирования. Точка наблюдения w и вектор $\mathbf{i}(u, v)$ оп-

ределяют линию визирования для рассматриваемой точки поверхности. Если точка наблюдения находится бесконечно далеко от поверхности, то вектор $\mathbf{i}(u, v)$ не будет зависеть от параметров поверхности и будет общим для всех точек поверхности.

Уравнение (4.6.5) содержит два искомых параметра u, v . Если задать один из параметров, то другой можно найти из уравнения (4.6.5), то есть один из параметров является функцией другого параметра. Для упрощения параметров представим их в виде функций $u = u(t)$ и $v = v(t)$ некоторого общего параметра t . Результатом решения уравнения (4.6.5) является кривая на поверхности $\mathbf{c}(t) = \mathbf{s}(u(t), v(t))$, которая является линией очерка. В общем случае найти функции $u(t)$ и $v(t)$ не всегда возможно, поэтому аппроксимируем их некоторым двумерным сплайном $\mathbf{r}(t)$, $t_{\min} \leq t \leq t_{\max}$. Точки сплайна $\mathbf{r}(t)$ служат начальными приближениями для итерационного решения уравнения (4.6.5).

Линию очерка построим из трех объектов: точки наблюдения \mathbf{w} или вектора направления линии визирования \mathbf{i} (при бесконечно далеком положении точки наблюдения), поверхности $\mathbf{s}(u, v)$ и двумерной кривой начальных приближений $\mathbf{r}(t)$. Эти объекты вместе с алгоритмом решения уравнения (4.6.5) позволяют определять точки линии очерка с заданной точностью.

Алгоритм построения линий очерка условно разобьем на два шага: на первом шаге найдем набор отдельных точек линии очерка, на втором шаге построим сами линии очерка. Для поиска отдельных точек линии очерка вычислим скалярное произведение \mathbf{im} в некоторых выбранных точках поверхности. Если для некоторой пары соседних точек \mathbf{im} меняет знак, то это свидетельствует о переходе через линию очерка. Взяв в качестве начального приближения средние значения параметров этих соседних точек, одним из численных методов найдем параметры u, v точки линии очерка. Таким способом получим набор отдельных точек линий очерка. Эти точки никак не связаны друг с другом и могут принадлежать различным линиям очерка. Важно только, чтобы от каждой линии очерка в наборе присутствовала хотя бы одна точка.

Далее возьмем любую точку из имеющегося набора и, двигаясь от нее с некоторым шагом сначала в одну сторону, потом в другую, точку за точкой найдем искомую совокупность точек линии очерка. Направление движения дает вектор

$$\mathbf{t} = \pm \frac{\mathbf{m}_1 \times \mathbf{m}}{|\mathbf{s}_1|} \pm \frac{\mathbf{m}_2 \times \mathbf{m}}{|\mathbf{s}_2|},$$

где \mathbf{m}_1 и \mathbf{m}_2 – частные производные нормали $\mathbf{m}(u, v)$; \mathbf{s}_1 и \mathbf{s}_2 – частные производные радиуса-вектора поверхности $\mathbf{s}(u, v)$ по параметрам u и v . Знак перед слагаемым $(\mathbf{m}_j \times \mathbf{m}) / |\mathbf{s}_j|$, $j = 1, 2$ совпадает со знаком скалярного

произведения $\mathbf{i} \mathbf{m}$, $j = 1, 2$. Шаг движения вычислим в соответствии с кривизной поверхностей в текущей точке. Если

$$\frac{|\mathbf{t} \cdot \mathbf{r}_1|}{|\mathbf{r}_1|} > \frac{|\mathbf{t} \cdot \mathbf{r}_2|}{|\mathbf{r}_2|},$$

то дадим приращение параметру u и найдем соответствующий ему параметр v поверхности. В противном случае дадим приращение параметру v и найдем соответствующий ему параметр u поверхности. Движение по линии очерка закончим, когда дойдем до края поверхности или когда линия замкнется.

В процессе построения точек линии очерка будем проверять, не лежат ли вблизи них точки из набора, полученного в начале. Если расстояние до какой-либо точки набора соизмеримо с текущим шагом движения, то эту точку удалим из набора как более ненужную. Так получим последовательность точек $\mathbf{p}_i = [u_i, v_i]^T$, $i = 0, 1, 2, \dots, n$ одной линии очерка. При этом в наборе точек, полученном в начале, не будет содержаться ни одной точки данной линии очерка. Если в наборе останутся еще точки, то данная поверхность имеет, по крайней мере, еще одну линию очерка. Совокупность ее точек найдем, взяв любую точку из набора и повторив описанные выше действия.

По последовательности точек построим сплайн $\mathbf{r}(t)$, проходящий через точки $\mathbf{p} = [u_i, v_i]^T$ при значениях параметра $t = t_i$, $i = 0, 1, \dots, n$. Имея информацию о поверхности $\mathbf{s}(u, v)$, точке наблюдения \mathbf{w} или векторе направления линии визирования \mathbf{i} , по отдельным точкам линии очерка, расположенным в порядке их следования, всегда можно найти любую другую точку линии. Например, для нахождения точки, лежащей между двумя заданными соседними точками линии очерка, проведем плоскость перпендикулярно соединяющему соседние точки отрезку и найдем общую точку для поверхности и плоскости, решив систему уравнений пересечения совместно с уравнением (4.6.5).

Линии очерка используются для построения проекций поверхности $\mathbf{s}(u, v)$ на плоскость. Проекция линии очерка может иметь изломы и точки возврата, но сама линия очерка не имеет особых точек, если они отсутствуют у поверхности в области линии очерка. Точки излома у проекции возникают там, где касательная линии очерка параллельна линии визирования.

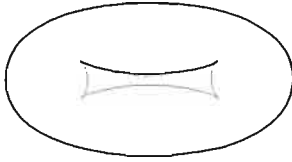


Рис. 4.6.3

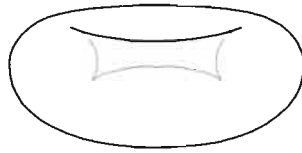


Рис. 4.6.4

На рис. 4.6.3 приведена проекция линий очерка тора, когда точка наблюдения находится бесконечно далеко от поверхности, а на рис. 4.6.4 для сравнения приведена проекция линий очерка тора, когда точка наблюдения находится на конечном расстоянии от поверхности тора.

4.7. Кривая пересечения поверхностей

В общем случае невозможно оперировать кривой пересечения отдельно от породивших ее поверхностей, поэтому кривую пересечения опишем двумя кривыми (4.6.1) на пересекаемых поверхностях. Кривую пересечения поверхностей $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$ построим из двух кривых $\mathbf{r}(u(t), v(t))$ и $\mathbf{s}(a(t), b(t))$ и будем записывать в виде

$$\mathbf{c}(t) = \begin{cases} \mathbf{r}(u(t), v(t)) \\ \mathbf{s}(a(t), b(t)) \end{cases}, \quad t_{\min} \leq t \leq t_{\max}. \quad (4.7.1)$$

Радиус-вектор $\mathbf{c}(t)$ кривой пересечения вычислим как среднее арифметическое радиусов-векторов точек кривых на пересекаемых поверхностях.

Для кривых $\mathbf{r}(u(t), v(t))$ и $\mathbf{s}(a(t), b(t))$ должны соблюдаться следующие два правила.

Правило 1. Кривые $\mathbf{r}(u(t), v(t))$ и $\mathbf{s}(a(t), b(t))$ имеют одинаковые области определения $t_{\min} \leq t \leq t_{\max}$.

Правило 2. Кривые $\mathbf{r}(u(t), v(t))$ и $\mathbf{s}(a(t), b(t))$ дают один и тот же радиус-вектор и его производные при одинаковых параметрах, то есть справедливо равенство $\mathbf{r}(u(t), v(t)) = \mathbf{s}(a(t), b(t))$. В определенных случаях допустимо приближенное выполнение этого равенства.

В частных случаях можно найти аналитические функции $u(t)$, $v(t)$, $a(t)$, $b(t)$, для которых выполняется равенство $\mathbf{r}(u(t), v(t)) = \mathbf{s}(a(t), b(t))$. Такие кривые пересечения будем называть *обыкновенными*. В общем случае функции $u(t)$, $v(t)$, $a(t)$, $b(t)$, аппроксимируем некоторыми сплайнами, совпадающими с этими функциями в отдельных точках. Такие кривые пересечения будем называть *особенными*. Второе правило у особенных кривых выполняется приближенно, тем не менее, мы всегда будем

вычислять точное значение радиуса-вектора кривой пересечения путем решения уравнения

$$\mathbf{r}(u, v) - \mathbf{s}(a, b) = 0. \quad (4.7.2)$$

У особенных кривых аппроксимации функций $u(t)$, $v(t)$, $a(t)$, $b(t)$ служат начальными приближениями для решения уравнения (4.7.2). В дальнейшем посредством $u(t)$, $v(t)$, $a(t)$, $b(t)$ для кривых пересечения мы будем обозначать функции, удовлетворяющие уравнению (4.7.2), или их аппроксимации.

Функции $u(t)$, $v(t)$ образуют двумерную кривую $\mathbf{c}_r(t) = [u(t) \ v(t)]^T$ в пространстве параметров поверхности $\mathbf{r}(u, v)$, а функции $a(t)$, $b(t)$ образуют двумерную кривую $\mathbf{c}_s(t) = [a(t) \ b(t)]^T$ в пространстве параметров поверхности $\mathbf{s}(a, b)$. Кривые $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$ никак не ограничены областью определения поверхностей и могут выходить за их пределы. Вне области определения параметров радиус-вектор поверхности вычисляется в соответствии с правилами построения продолженной поверхности (3.15.1)–(3.15.3).

Найдем линию пересечения плоскостей

$$\begin{aligned} \mathbf{r}(u, v) &= \mathbf{r}_0 + u \mathbf{r}_1 + v \mathbf{r}_2, \\ \mathbf{s}(a, b) &= \mathbf{s}_0 + a \mathbf{s}_1 + b \mathbf{s}_2. \end{aligned}$$

Пусть \mathbf{r}_3 и \mathbf{s}_3 – единичные нормали плоскостей. Направление линии пересечения плоскостей определяется вектором $\mathbf{z} = \mathbf{r}_3 \times \mathbf{s}_3$. Если длина вектора \mathbf{z} отлична от нуля, то плоскости пересекаются, в противном случае они параллельны или совпадают. Для пересекающихся плоскостей найдем точку на линии пересечения. Построим вектор $\mathbf{w} = \mathbf{r}_0 - \mathbf{s}_0$ и вычислим его составляющую, параллельную плоскости $\mathbf{s}(a, b)$: $\mathbf{w}' = \mathbf{w} - \mathbf{s}_3(\mathbf{w} \cdot \mathbf{s}_3)$. Далее определим точку пересечения линии, проходящей через точку \mathbf{s}_0 в направлении вектора \mathbf{w}' , с плоскостью $\mathbf{r}(u, v)$:

$$\mathbf{p} = \mathbf{s}_0 + \mathbf{w}' \frac{\mathbf{r}_3 \cdot (\mathbf{r}_0 - \mathbf{s}_0)}{\mathbf{r}_3 \cdot \mathbf{w}'}$$

Плоскости пересекаются по линии

$$\mathbf{l}(t) = \mathbf{p} + t \mathbf{z}.$$

Параметры плоскостей, соответствующие точке \mathbf{p} , определим по формулам

$$\begin{aligned} u_p &= \frac{\mathbf{r}_1 \cdot (\mathbf{p} - \mathbf{r}_0)}{\mathbf{r}_1 \cdot \mathbf{r}_1}, & v_p &= \frac{\mathbf{r}_2 \cdot (\mathbf{p} - \mathbf{r}_0)}{\mathbf{r}_2 \cdot \mathbf{r}_2}, \\ a_p &= \frac{\mathbf{s}_1 \cdot (\mathbf{p} - \mathbf{s}_0)}{\mathbf{s}_1 \cdot \mathbf{s}_1}, & b_p &= \frac{\mathbf{s}_2 \cdot (\mathbf{p} - \mathbf{s}_0)}{\mathbf{s}_2 \cdot \mathbf{s}_2}, \end{aligned}$$

при условии ортогональности пар векторов $\mathbf{r}_1, \mathbf{r}_2$ и $\mathbf{s}_1, \mathbf{s}_2$ или по формулам (4.2.3) и (4.2.4) в общем случае. Аналогичным образом найдем параметры для еще одной точки линии пересечения, например, точки $\mathbf{q} = \mathbf{p} + t_q \mathbf{z}$. Пусть параметры u_q и v_q соответствуют точке \mathbf{q} на плоскости $\mathbf{r}(u, v)$, а параметры a_q и b_q соответствуют этой же точке на плоскости $\mathbf{s}(a, b)$. Функции

$$\begin{aligned} u(t) &= u_p(1-t) + u_q t, \\ v(t) &= v_p(1-t) + v_q t, \\ a(t) &= a_p(1-t) + a_q t, \\ b(t) &= b_p(1-t) + b_q t \end{aligned}$$

являются решением уравнения (4.7.2) и позволяют построить обыкновенную кривую пересечения плоскостей (4.7.1).

Рассмотрим **частный случай** пересечения поверхности выдавливания и плоскости. Пусть поверхность $\mathbf{c}(t) = \mathbf{w}(x(t), y(t))$, полученная выдавливанием кривой на плоскости $\mathbf{c}_w(t) = [x(t) \ y(t)]^T$, $t_{\min} \leq t \leq t_{\max}$ и плоскостью $\mathbf{w}(x, y) = \mathbf{p}_0 + x\mathbf{i}_1 + y\mathbf{i}_2$. Поверхность выдавливания будет описываться радиусом-вектором

$$\begin{aligned} \mathbf{r}(u, v) &= \mathbf{p}_0 + x(u)\mathbf{i}_1 + y(u)\mathbf{i}_2 + v\mathbf{d}, \\ u &\in [t_{\min}, t_{\max}], \ v \in [v_{\min}, v_{\max}]. \end{aligned}$$

Кривую пересечения поверхности выдавливания и плоскости $\mathbf{s}(a, b) = \mathbf{q}_0 + a\mathbf{j}_1 + b\mathbf{j}_2$, параллельной плоскости $\mathbf{w}(x, y)$, построим из трансформированной копии кривой $\mathbf{c}_w(t)$ на плоскости $\mathbf{s}(a, b)$ и отрезка $\mathbf{c}_s(t) = [t \ v_s]^T$ на поверхности $\mathbf{r}(u, v)$. Значение второго параметра v_s поверхности $\mathbf{r}(u, v)$, соответствующее положению отрезка $\mathbf{c}_s(t)$, равно

$$v_s = \frac{(\mathbf{q}_0 - \mathbf{p}_0) \cdot \mathbf{d}}{|\mathbf{d}|^2}.$$

Трансформированную копию кривой $\mathbf{c}_w(t)$ на плоскости $\mathbf{s}(a, b)$ обозначим через $\mathbf{c}_s(t) = [a(t) \ b(t)]^T$. Копия кривой $\mathbf{c}_w(t)$ должна быть преобразована в соответствии с взаимным положением базисных векторов и начальных точек плоскостей $\mathbf{s}(a, b)$ и $\mathbf{w}(x, y)$. Это преобразование описывается равенством

$$\begin{bmatrix} a(t) \\ b(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix},$$

где $a_{11} = \frac{\mathbf{j}_1 \cdot \mathbf{i}_1}{\mathbf{j}_1 \cdot \mathbf{j}_1}$, $a_{12} = \frac{\mathbf{j}_1 \cdot \mathbf{i}_2}{\mathbf{j}_1 \cdot \mathbf{j}_1}$, $a_{21} = \frac{\mathbf{j}_2 \cdot \mathbf{i}_1}{\mathbf{j}_2 \cdot \mathbf{j}_2}$, $a_{22} = \frac{\mathbf{j}_2 \cdot \mathbf{i}_2}{\mathbf{j}_2 \cdot \mathbf{j}_2}$, $a_1 = \frac{\mathbf{j}_1 \cdot (\mathbf{p}_0 - \mathbf{q}_0)}{\mathbf{j}_1 \cdot \mathbf{j}_1}$, $a_2 = \frac{\mathbf{j}_2 \cdot (\mathbf{p}_0 - \mathbf{q}_0)}{\mathbf{j}_2 \cdot \mathbf{j}_2}$. Отрезок $\mathbf{c}_s(t)$ и кривая $\mathbf{c}_s(t)$, расположенные на разных

поверхностях, имеют одинаковую параметрическую длину $t_{\min} \leq t \leq t_{\max}$ и представляют обыкновенную кривую пересечения

$$c(t) = \begin{cases} \mathbf{r}(t, v) \\ \mathbf{s}(a(t), b(t)) \end{cases}$$

поверхности $\mathbf{r}(u, v)$ и плоскости $\mathbf{s}(a, b)$.

Аналогично можно построить кривые пересечения и некоторых других поверхностей в частном случае их взаимного положения, например, кривую пересечения поверхности, полученной вращением кривой на плоскости вокруг оси, принадлежащей этой плоскости, и плоскости, проходящей через ось вращения. Еще пример: кривую пересечения поверхности вращения и плоскости, ортогональной оси вращения, построим из окружности на плоскости и отрезка на поверхности.

В общем случае кривую пересечения поверхностей представим в виде последовательности отдельных точек кривой. Отдельными точками в данном случае будем называть пары параметров u_i, v_i и a_i, b_i , для которых выполняется равенство (4.7.2). Оформим последовательности отдельных точек в виде двух кривых в пространствах параметров поверхностей $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$, например, в виде двумерных сплайнов Эрмита (3.3.10), проходящих через точки $[u_i, v_i]^T$ и $[a_i, b_i]^T$ соответственно при значении параметра кривой пересечения $t_i, i = 0, 1, 2, \dots, n$. Значения параметров t_i , соответствующие точкам $[u_i, v_i]^T$ и $[a_i, b_i]^T$, не должны убывать. Изменение параметра при переходе от точки к следующей точке положим пропорциональным длине участка кривой пересечения между соседними точками. Такая параметризация кривой пересечения близка к натуральной параметризации, когда параметром является длина кривой.

Имея последовательность отдельных точек кривой пересечения, всегда можно найти любую другую точку кривой. Действительно, пусть параметры u_1, v_1, a_1, b_1 и параметры u_2, v_2, a_2, b_2 поверхностей соответствуют соседним точкам \mathbf{p}_1 и \mathbf{p}_2 последовательности точек кривой пересечения поверхностей $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$. Пусть первой точке соответствует параметр кривой пересечения t_1 , а второй точке соответствует параметр кривой пересечения t_2 . Требуется найти точку кривой пересечения с параметром t , где $t_1 < t < t_2$. Построим плоскость

$$\mathbf{q}(x, y) = \mathbf{p} + x \mathbf{q}_1 + y \mathbf{q}_2,$$

где $\mathbf{p} = \frac{t_2 - t}{t_2 - t_1} \mathbf{p}_1 + \frac{t - t_1}{t_2 - t_1} \mathbf{p}_2$, векторы \mathbf{q}_1 и \mathbf{q}_2 ортогональны отрезку, построенному из точки \mathbf{p}_1 в точку \mathbf{p}_2 , и не параллельны друг другу, рис. 4.7.1.

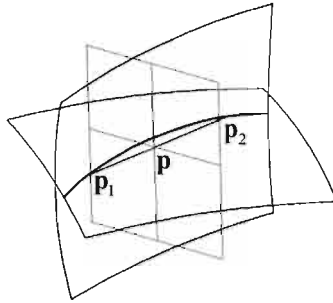


Рис. 4.7.1

Одним из итерационных методов найдем параметры точки пересечения поверхностей $\mathbf{r}(u,v)$, $\mathbf{s}(a,b)$ и плоскости $\mathbf{q}(x,y)$ из системы шести уравнений, которые в векторной записи имеют вид

$$\begin{aligned}\mathbf{r}(u,v) - \mathbf{q}(x,y) &= 0; \\ \mathbf{s}(a,b) - \mathbf{q}(x,y) &= 0.\end{aligned}$$

За начальные приближения примем параметры

$$\begin{aligned}u^{(0)} &= \frac{t_2 - t}{t_2 - t_1} u_1 + \frac{t - t_1}{t_2 - t_1} u_2, \quad v^{(0)} = \frac{t_2 - t}{t_2 - t_1} v_1 + \frac{t - t_1}{t_2 - t_1} v_2, \\ a^{(0)} &= \frac{t_2 - t}{t_2 - t_1} a_1 + \frac{t - t_1}{t_2 - t_1} a_2, \quad b^{(0)} = \frac{t_2 - t}{t_2 - t_1} b_1 + \frac{t - t_1}{t_2 - t_1} b_2, \\ x^{(0)} &= 0, \quad y^{(0)} = 0.\end{aligned}$$

Найденные параметры u , v , a , b определяют искомую точку кривой пересечения. Таким образом, последовательность точек $[u, v]^T$ и $[a, b]^T$ позволяет найти любую другую точку кривой пересечения поверхностей $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$.

4.8. Алгоритм построения кривых пересечения

Две поверхности могут пересекаться по нескольким кривым. Поставим своей целью построить все кривые пересечения, хотя бы частично принадлежащие области определения параметров обеих поверхностей. На рис. 4.8.1 приведен пример пересечения поверхностей по двум кривым.

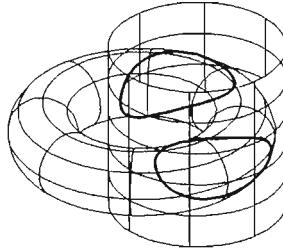


Рис. 4.8.1

Векторное уравнение (4.7.2) содержит три скалярных уравнения для координат радиусов-векторов поверхностей и четыре искомого параметра: u , v , a , b . Если определить один из четырех параметров, то остальные три можно найти из системы уравнений (4.7.2). На этом основан алгоритм построения кривой пересечения поверхностей. Алгоритм построения кривой пересечения поверхностей разобьем на два шага.

На первом шаге, двигаясь по поверхностям и исследуя их близость, найдем набор отдельных точек кривых пересечения, рис. 4.8.2. Точки из полученного на первом шаге набора никак не связаны друг с другом и могут принадлежать различным кривым пересечения. Важно только, чтобы от каждой кривой пересечения в наборе присутствовала хотя бы одна точка.

На втором шаге берем любую точку из имеющегося набора и, двигаясь от нее с некоторым шагом сначала в одну сторону, затем в другую, найдем точку за точкой искомую последовательность точек кривой пересечения. Направление движения дает векторное произведение нормалей к поверхностям. Шаг движения вычислим в соответствии с кривизной поверхностей в текущей точке. Движение по кривой пересечения закончим, когда дойдем до края одной из поверхностей или когда кривая замкнется, то есть новая точка окажется на расстоянии текущего шага от точки старта. В процессе движения будем проверять, не лежат ли вблизи пути следования точки из набора, полученного на первом шаге. Для этого по пути следования будем вычислять расстояние от текущей точки кривой пересечения до каждой точки из набора, полученного на первом шаге. Если вычисленное расстояние до какой-либо точки набора соизмеримо с текущим шагом движения, то эту точку удалим из набора как ненужную. Так получим последовательность отдельных точек $[u_i, v_i]^T$ и $[a, b]^T$, $i = 0, 1, 2, \dots, n$ одной кривой пересечения, рис. 4.8.3.

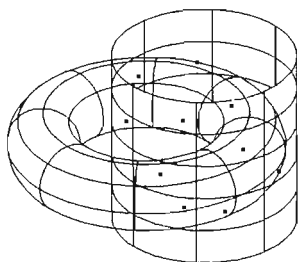


Рис. 4.8.2

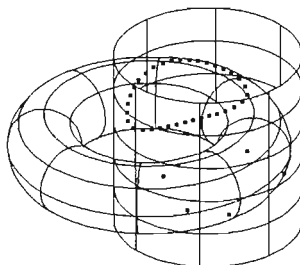


Рис. 4.8.3

В пространствах параметров поверхностей $r(u,v)$ и $s(a,b)$ построим два двумерных сплайна, проходящих через точки $[u, v]^T$ и $[a, b]^T$, рис. 4.8.4. После этого набор точек, полученный на первом шаге, не будет содержать ни одной точки данной кривой пересечения. Если в наборе останутся еще точки, то данные поверхности имеют, по крайней мере, еще одну кривую пересечения. Последовательность ее точек найдем, взяв любую точку из набора и повторив все сказанное о втором шаге построения, рис. 4.8.5. Построение кривых закончим, когда в наборе не останется ни одной точки.

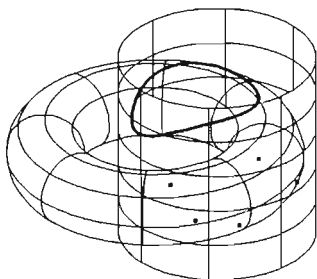


Рис. 4.8.4

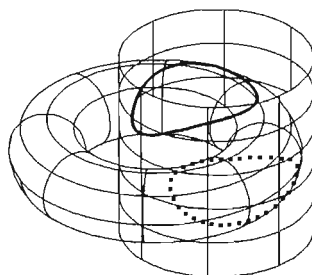


Рис. 4.8.5

Наиболее трудоемким является первый шаг. Для определения начальных точек пересечения могут использоваться различные стратегии. Их смысл сводится к тому, чтобы представить каждую поверхность набором некоторых простых объектов и с их помощью принять решение о запуске итерационного процесса решения системы уравнений (4.7.2) с одним фиксированным параметром. Простыми объектами могут быть участки плоскости, отрезки прямых, габаритные кубы частей поверхности или другие элементы. Для ускорения работы объекты могут быть отсортированы тем или иным способом. Таким образом на первом шаге получим на-

бор точек, принадлежащих в общем случае различным кривым пересечения заданных поверхностей. В этом наборе нам понадобится по одной точке для построения каждой кривой пересечения.

Рассмотрим пример использования простых объектов на первом шаге. Охватим каждую из двух пересекаемых поверхностей направленными прямоугольными параллелепипедами. Центр параллелепипеда расположим в точке поверхности с известными параметрами поверхности. Две стороны параллелепипеда направим по производным поверхности в центральной точке параллелепипеда, а их длины свяжем с приращениями параметров. Третью сторону параллелепипеда свяжем с прогибом поверхности в центральной точке. Таким образом, параллелепипед будет определен параметрами поверхности u_i, v_i и приращениями параметров du_i, dv_i , рис. 4.8.6.

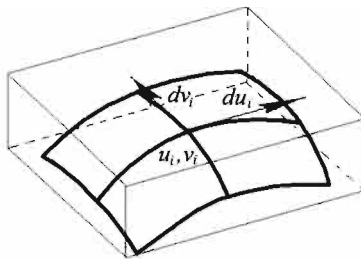


Рис. 4.8.6

Начальное количество параллелепипедов может быть небольшим. Важно, чтобы в начале работы алгоритма поверхность полностью попала во внутренний объем своего набора параллелепипедов. Далее проверим набор параллелепипедов одной поверхности на пересечение с набором параллелепипедов другой поверхности. В каждом наборе отметим только параллелепипеды, которые пересекаются с параллелепипедами другой поверхности. Далее каждый пересекающийся параллелепипед поверхности разобьем на четыре части путем смещения параметров центральной точки и делением пополам приращений параметров поверхности. Проверим на пересечение измельченный набор отмеченных параллелепипедов одной поверхности с измельченным набором отмеченных параллелепипедов другой поверхности и отметим только параллелепипеды, которые пересекаются с параллелепипедами другой поверхности. Продолжим процесс деления параллелепипедов, пока приращение параметров поверхности, определяющее размер стороны параллелепипеда, не достигнет заданной минимальной величины. Далее запустим итерационный процесс пересечения поверхностей для пересекающихся па-

раллелепипедов с начальными приближениями, определяющими центральные точки параллелепипедов.

Второй шаг при всей громоздкости описания более быстр по сравнению с первым шагом. Возьмем из полученного на первом шаге набора произвольную точку пересечения поверхностей. Пусть ее параметры на первой поверхности равны u_k, v_k , а на второй поверхности равны a_k, b_k . Построим нормали к поверхностям $\mathbf{m}_1(u_k, v_k)$ и $\mathbf{m}_2(a_k, b_k)$. Кривая пересечения в каждой точке ортогональна нормальям обеих поверхностей, поэтому направление кривой пересечения параллельно вектору $\mathbf{t} = \mathbf{m}_1(u_k, v_k) \times \mathbf{m}_2(a_k, b_k)$. Выберем в качестве направления движения к следующей точке кривой пересечения положительное направление вектора \mathbf{t} . Далее определим возможный шаг смещения по каждому из четырех параметров $\Delta u, \Delta v, \Delta a, \Delta b$ из условия, что нормали поверхностей отклонятся на угол, не превосходящий заданную величину $\Delta\alpha$. Возможные шаги смещения вычислим по формулам

$$\Delta u = \frac{|\Delta\alpha | \mathbf{r}_1 |}{\mathbf{m}_1 \cdot \mathbf{r}_1}, \quad \Delta v = \frac{|\Delta\alpha | \mathbf{r}_2 |}{\mathbf{m}_2 \cdot \mathbf{r}_2}, \quad \Delta a = \frac{|\Delta\alpha | \mathbf{s}_1 |}{\mathbf{m}_1 \cdot \mathbf{s}_1}, \quad \Delta b = \frac{|\Delta\alpha | \mathbf{s}_2 |}{\mathbf{m}_2 \cdot \mathbf{s}_2},$$

где $\mathbf{r}_1 = \frac{\partial \mathbf{r}}{\partial u}$, $\mathbf{r}_2 = \frac{\partial \mathbf{r}}{\partial v}$, $\mathbf{r}_{11} = \frac{\partial^2 \mathbf{r}}{\partial u^2}$, $\mathbf{r}_{22} = \frac{\partial^2 \mathbf{r}}{\partial v^2}$, $\mathbf{s}_1 = \frac{\partial \mathbf{s}}{\partial a}$, $\mathbf{s}_2 = \frac{\partial \mathbf{s}}{\partial b}$, $\mathbf{s}_{11} = \frac{\partial^2 \mathbf{s}}{\partial a^2}$,

$\mathbf{s}_{22} = \frac{\partial^2 \mathbf{s}}{\partial b^2}$ – производные в точке с параметрами u_k, v_k и a_k, b_k . Если одному

из параметров дать приращение и сделать его неподвижным, то из системы уравнений (4.7.2) можно найти остальные параметры. Важно только, чтобы изменение каждого параметра по модулю не превышало найденные возможные шаги. В качестве фиксированного параметра, которому будет дано приращение, выберем тот, направление движения вдоль которого ближе к направлению кривой пересечения, то есть составляет меньший угол с вектором \mathbf{t} .

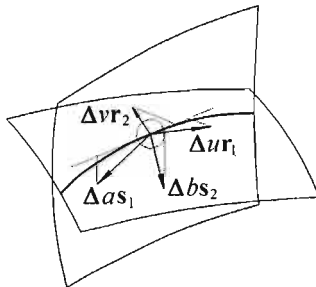


Рис. 4.8.7

На рис. 4.8.7 показаны отправная точка $\mathbf{p}_k = \mathbf{r}(u_k, v_k) = \mathbf{s}(a_k, b_k)$, векторы производных $\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2$ и сфера смещения, внутри которой приращения всех параметров не превышают, соответственно, $\Delta u, \Delta v, \Delta a, \Delta b$. Следующая точка кривой пересечения не должна выйти за сферу смещения. Радиус сферы смещения r_0 равен наименьшей по абсолютной величине проекции векторов $\Delta u \mathbf{r}_1, \Delta v \mathbf{r}_2, \Delta a \mathbf{s}_1, \Delta b \mathbf{s}_2$ на вектор \mathbf{t} :

$$r_0 = \min(|\Delta u \mathbf{r}_1 \cdot \mathbf{t}|, |\Delta v \mathbf{r}_2 \cdot \mathbf{t}|, |\Delta a \mathbf{s}_1 \cdot \mathbf{t}|, |\Delta b \mathbf{s}_2 \cdot \mathbf{t}|) \frac{1}{|\mathbf{t}|}.$$

Обращающиеся в ноль проекции векторов $\Delta u \mathbf{r}_1, \Delta v \mathbf{r}_2, \Delta a \mathbf{s}_1, \Delta b \mathbf{s}_2$ на вектор \mathbf{t} не учитываем. Для следующей точки кривой пересечения в качестве начальных приближений примем значения параметров

$$\begin{aligned} u_{k+1}^{(0)} &= u_k + r_0 \frac{\mathbf{t} \cdot \mathbf{r}_1}{|\mathbf{t}|(\mathbf{r}_1 \cdot \mathbf{r}_1)}, & v_{k+1}^{(0)} &= v_k + r_0 \frac{\mathbf{t} \cdot \mathbf{r}_2}{|\mathbf{t}|(\mathbf{r}_2 \cdot \mathbf{r}_2)}, \\ a_{k+1}^{(0)} &= a_k + r_0 \frac{\mathbf{t} \cdot \mathbf{s}_1}{|\mathbf{t}|(\mathbf{s}_1 \cdot \mathbf{s}_1)}, & b_{k+1}^{(0)} &= b_k + r_0 \frac{\mathbf{t} \cdot \mathbf{s}_2}{|\mathbf{t}|(\mathbf{s}_2 \cdot \mathbf{s}_2)}. \end{aligned}$$

Тот параметр, производная по которому составляет меньший по абсолютной величине угол с направлением кривой пересечения, сделаем неподвижным. Далее, решив систему уравнений (4.7.2), в которой неподвижный параметр будет считаться известным и неизменным, получим остальные три параметра. Таким образом, найдем параметры u_{k+1}, v_{k+1} первой поверхности и параметры a_{k+1}, b_{k+1} второй поверхности для следующей точки кривой пересечения. Продолжим этот процесс, пока не дойдем до края одной из поверхностей или не окажемся вблизи точки старта.

Если в процессе движения мы оказались вблизи точки старта, то кривая пересечения является циклически замкнутой. Близость точек определяется с помощью шагов $\Delta u, \Delta v, \Delta a, \Delta b$.

Если в процессе движения мы достигли края одной из поверхностей, то после определения крайней точки кривой пересечения поверхностей продолжим движение и поиск параметров кривой пересечения в другом направлении от точки старта, то есть в противоположном вектору \mathbf{t} направлении. Такая кривая не может быть циклически замкнутой, поэтому движение будем продолжать до тех пор, пока не будут найдены параметры другого края кривой пересечения поверхностей.

Если кривая пересечения проходит через точку касания поверхностей, то в этой точке поверхности имеют коллинеарные нормали. Матрица линеаризованной системы уравнений (4.7.2) в этой точке будет иметь нулевой определитель, и из этой системы уравнений нельзя найти

параметры точки касания поверхностей. Параметры точки касания поверхностей можно найти из решения системы уравнений:

$$(\mathbf{r}(u, v) - \mathbf{s}(a, b)) \cdot \frac{\partial \mathbf{r}}{\partial u} = 0,$$

$$(\mathbf{r}(u, v) - \mathbf{s}(a, b)) \cdot \frac{\partial \mathbf{r}}{\partial v} = 0,$$

$$(\mathbf{s}(a, b) - \mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{s}}{\partial a} = 0,$$

$$(\mathbf{s}(a, b) - \mathbf{r}(u, v)) \cdot \frac{\partial \mathbf{s}}{\partial b} = 0.$$

Этой системе удовлетворяют не только точки касания, а еще и другие точки, в которых касательные к поверхностям плоскости параллельны друг другу и ортогональны отрезку, соединяющему точки касания. Поэтому после решения системы уравнений следует проверить, что полученные точки являются точками касания.

4.9. Поверхности сопряжения

Поверхности сопряжения сглаживают переход от одной поверхности к другой в районе пересечения поверхностей. К поверхностям сопряжения относят поверхности скругления и поверхности фаски. Поверхности сопряжения представляют собой поверхности движения изменяющейся образующей кривой по двум направляющим кривым.

Пусть имеются две пересекающиеся (возможно на продолжениях) поверхности. Вблизи кривой пересечения пространство делится поверхностями на четыре сектора. Нормали поверхностей и перпендикуляры, восстановленные от поверхностей к точкам сектора, могут совпадать или иметь противоположные направления. Пусть в первом секторе перпендикуляры, восстановленные от поверхностей к точкам сектора, совпадают с нормалью обеих поверхностей. Во втором секторе перпендикуляр, восстановленный от первой поверхности к точкам сектора, совпадает с нормалью первой поверхности, а перпендикуляр, восстановленный от второй поверхности к точкам сектора, противоположен по направлению нормали второй поверхности. В третьем секторе перпендикуляры, восстановленные от поверхностей к точкам сектора, противоположны по направлению нормальям обеих поверхностей. В четвертом секторе перпендикуляр, восстановленный от первой поверхности к точкам сектора, противоположен по направлению нормали первой поверхности, а перпендикуляр, восстановленный от второй поверхности к точкам сектора, совпадает с нормалью второй поверхности.

Поверхность скругления может быть получена движением сферы, касающейся одновременно двух сопрягаемых поверхностей. Сфера движется вдоль кривой пересечения поверхностей в одном из четырех упомянутых секторов. Поверхность скругления располагается между кривыми касания сферы поверхностей.

Пусть требуется построить поверхность скругления между поверхностями $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$. Представим поверхность скругления в виде рациональной поверхности (3.4.5), построенной по трем кривым. Для этого построим кривые касания сферы с сопрягаемыми поверхностями $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$. Точки третьей кривой $\mathbf{c}(t)$ должны располагаться на пересечении трех плоскостей: двух касательных к сопрягаемым поверхностям в точках $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$ и плоскости, проходящей через точки касания и центр сферы. Кривые $\mathbf{c}_1(t)$, $\mathbf{c}_2(t)$ и $\mathbf{c}(t)$ должны иметь одинаковую параметрическую длину. Радиусы-векторы кривых $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$ должны определять соответствующие друг другу точки касания сферы.

Кривые $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$ будем называть *опорными кривыми*, кривую $\mathbf{c}(t)$ будем называть *средней кривой*. Поверхность скругления опишем функцией

$$\mathbf{q}(t, z) = \frac{(1-z)^2 \mathbf{c}_1(t) + 2(1-z)zw(t)\mathbf{c}(t) + z^2 \mathbf{c}_2(t)}{(1-z)^2 + 2(1-z)zw(t) + z^2}, \quad (4.9.1)$$

$$t \in [t_{\min}, t_{\max}], z \in [0, 1].$$

где $w(t) = \cos\left(\frac{\alpha}{2}\right)$ – вес точек средней кривой; α – угол между нормальными сопрягаемых поверхностей. В общем случае угол α является функцией первого параметра поверхности скругления. Первый параметр поверхности скругления (4.9.1) совпадает с параметром кривых $\mathbf{c}_1(t)$, $\mathbf{c}_2(t)$ и $\mathbf{c}(t)$. При изменении второго параметра при фиксированном первом параметре поверхности скругления движение будет выполняться по дуге окружности. Эта дуга окружности представлена в виде рациональной кривой (2.5.10).

Продолжение нормалей к поверхностям в точках касания пересекутся в центре движущейся сферы. Параметры точек касания сферы и, следовательно, кривые $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$ связаны уравнением

$$\mathbf{r}(u, v) + \rho_r \mathbf{m}_r(u, v) = \mathbf{s}(a, b) + \rho_s \mathbf{m}_s(a, b), \quad (4.9.2)$$

где $\mathbf{m}_r(u, v)$ и $\mathbf{m}_s(a, b)$ – нормали поверхностей $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$ в точках касания; ρ_r и ρ_s – соответственно, длины проекций на эти нормали векторов, построенных из точек касания в центр сферы. Величины ρ_r и ρ_s по модулю равны радиусу сферы ρ , но имеют знак, характеризующий один из упомянутых секторов. На рис. 4.9.1 приведено сечение поверхностей $\mathbf{r}(u, v)$, $\mathbf{s}(a, b)$ и сферы плоскостью, проходящей через точки касания и центр сферы.

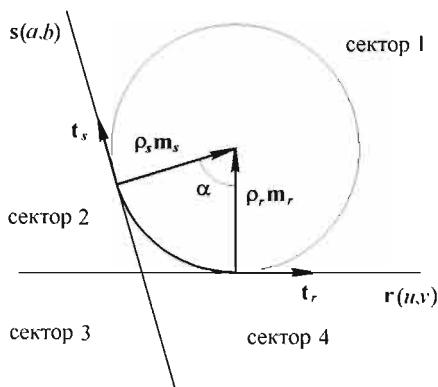


Рис. 4.9.1

Записанные в координатном виде три скалярных уравнения (4.9.2) содержат четыре искомого параметра u, v, a, b . Если определить один из четырех параметров, то остальные три можно найти из системы уравнений (4.9.2). Построение поверхности скругления по уравнению (4.9.2) сходно с задачей построения кривой пересечения поверхностей. В обоих случаях результатом решения являются функции $u(t), v(t), a(t), b(t)$, определяющие две кривые на поверхностях $c_r(t) = \mathbf{r}(u(t), v(t))$ и $c_s(t) = \mathbf{s}(a(t), b(t))$. Функции $u(t), v(t)$ образуют двумерную кривую $c_r(t) = [u(t) \ v(t)]^T$ в пространстве параметров поверхности $\mathbf{r}(u, v)$, а функции $a(t), b(t)$ образуют двумерную кривую $c_s(t) = [a(t) \ b(t)]^T$ в пространстве параметров поверхности $\mathbf{s}(a, b)$. Кривые $c_r(t)$ и $c_s(t)$ никак не ограничены областью определения поверхностей и могут выходить за их пределы. Вне области определения параметров радиус-вектор поверхности вычисляется в соответствии с правилами построения продолженной поверхности (3.15.1) – (3.15.3).

В частных случаях можно найти аналитические функции $u(t), v(t), a(t), b(t)$, для которых выполняется уравнение (4.9.2). Например, при сопряжении двух плоскостей угол α между нормальными поверхностями не меняется, функции $u(t), v(t), a(t), b(t)$ являются линейными, а средняя кривая $c(t)$ совпадает с линией пересечения сопрягаемых поверхностей. Аналогично можно найти аналитические функции $u(t), v(t), a(t), b(t)$ при сопряжении цилиндрической поверхности и плоскости, ортогональной или параллельной ее оси, при сопряжении цилиндрических, конических, поверхностей, поверхностей торов и поверхностей вращения с совпадающими осями. В частных случаях следует использовать аналитические поверхности. Так, при сопряжении двух плоскостей поверхность (4.9.1) следует заменить цилиндрической поверхностью, а при сопряже-

нии двух поверхностей вращения с совпадающими осями поверхность (4.9.1) следует заменить поверхностью тора.

В общем случае для сопряжения произвольных поверхностей двумерные кривые $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$ представим в виде сплайнов, проходящих через точки $[u_i, v_i]^T$ и $[a_i, b_i]^T$, при значении параметра t , $i = 0, 1, 2, \dots, n$. Параметры $u_i, v_i, a_i, b_i, i = 0, 1, 2, \dots, n$ аппроксимируют функции $u(t), v(t), a(t), b(t)$ и для них выполняется уравнение (4.9.2). Изменение параметра $t_{i+1} - t_i$ при переходе от точки к следующей точке положим пропорциональным средней длине участка трехмерных кривых $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$ между соседними точками.

Если при построении поверхности (4.9.1) имеется кривая пересечения поверхностей $\mathbf{c}_0(s)$, то уравнение (4.9.2) можно дополнить уравнением

$$\left(\mathbf{c}_0(s) - \frac{1}{2}(\mathbf{r}(u, v) + \rho_r \mathbf{m}_r(u, v) + \mathbf{s}(a, b) + \rho_s \mathbf{m}_s(a, b)) \right) \cdot \frac{d\mathbf{c}_0}{ds} = 0$$

и искать параметры u_i, v_i, a_i, b_i путем решения системы четырех уравнений относительно четырех неизвестных, двигаясь с некоторым шагом вдоль кривой пересечения.

Средняя кривая $\mathbf{c}(t)$ и функция весов $w(t)$ вычисляются по опорным кривым $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$. Текущее значение функции весов равно косинусу половины угла между векторами $\rho_r \mathbf{m}_r$ и $\rho_s \mathbf{m}_s$. Косинус угла между этими векторами равен их скалярному произведению, деленному на произведение длин векторов: $\cos \alpha = \frac{\rho_r \mathbf{m}_r \cdot \rho_s \mathbf{m}_s}{\rho^2}$. Используя формулу

$2\cos^2(\alpha/2) = 1 + \cos \alpha$, получим функцию весов для соответствующих точек средней кривой:

$$w(t) = \cos(\alpha/2) = \frac{1}{\sqrt{2}} \sqrt{1 + \frac{(\rho_r \mathbf{m}_r \cdot \rho_s \mathbf{m}_s)}{\rho^2}}. \quad (4.9.3)$$

При известных нормалях \mathbf{m}_r и \mathbf{m}_s в точках опорных кривых и функции весов радиус-вектор средней кривой вычислим по формуле

$$\begin{aligned} \mathbf{c}(t) &= \frac{1}{2}(\mathbf{c}_r(t) + \rho_r \mathbf{m}_r + \mathbf{c}_s(t) + \rho_s \mathbf{m}_s) - \frac{1}{2} \frac{\rho_r \mathbf{m}_r + \rho_s \mathbf{m}_s}{\cos^2(\alpha/2)} \\ &= \frac{1}{2} \left(\mathbf{c}_r(t) + \mathbf{c}_s(t) - (1 - w^2) \frac{\rho_r \mathbf{m}_r + \rho_s \mathbf{m}_s}{w^2} \right). \end{aligned} \quad (4.9.4)$$

Если опорные кривые $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$ циклически замкнуты, то поверхность скругления (4.9.1) также циклически замкнута по первому параметру. На рис. 4.9.2 приведена поверхность скругления в общем случае.

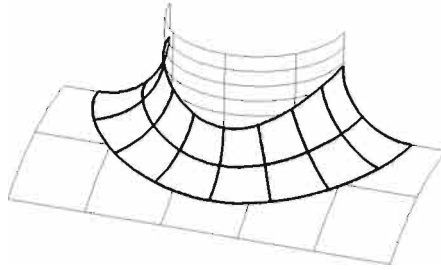


Рис. 4.9.2

Пусть требуется построить **поверхность скругления переменного радиуса**. Для этого потребуется кривая пересечения поверхностей. Величины радиуса скругления $\rho_r = \rho_r(s)$ и $\rho_s = \rho_s(s)$ будем считать функциями длины дуги s кривой пересечения $\mathbf{c}_0(s)$ сопрягаемых поверхностей. В данном случае движущаяся сфера будет иметь переменный радиус. Кроме того, положение центра сферы связано с точкой на кривой пересечения. Расположим центр движущейся сферы в нормальной плоскости кривой пересечения сопрягаемых поверхностей. Нормальная плоскость ортогональна касательному вектору кривой. Параметры точек касания сферы связаны системой уравнений

$$\mathbf{r}(u, v) + \rho_r(s) \mathbf{m}_r(u, v) = \mathbf{s}(a, b) + \rho_s(s) \mathbf{m}_s(a, b),$$

$$\left(\mathbf{c}_0'(s) - \frac{1}{2} (\mathbf{r}(u, v) + \rho_r(s) \mathbf{m}_r(u, v) + \mathbf{s}(a, b) + \rho_s(s) \mathbf{m}_s(a, b)) \right) \cdot \frac{d\mathbf{c}_0}{ds} = 0. \quad (4.9.5)$$

Эти уравнения содержат четыре скалярных уравнения относительно четырех искомых параметров u, v, a, b . Параметр s кривой пересечения является известной величиной. По текущему параметру s вычислим радиусы $\rho_r(s)$ и $\rho_s(s)$, точку $\mathbf{c}_0(s)$ и касательный вектор кривой в ней $\mathbf{c}_0'(s)$. Решив систему уравнений (4.9.5), получим параметры u, v, a, b касания движущейся сферы и поверхностей в текущем положении. По последовательности полученных точек построим опорные кривые, среднюю кривую, функцию весов соответствующих точек средней кривой и функцию радиусов, зависящую от первого параметра поверхности сопряжения. Поверхность (4.9.1), построенная по таким данным, будет иметь переменный радиус сопряжения.

На рис. 4.9.3 приведен пример поверхностей скругления переменного радиуса.

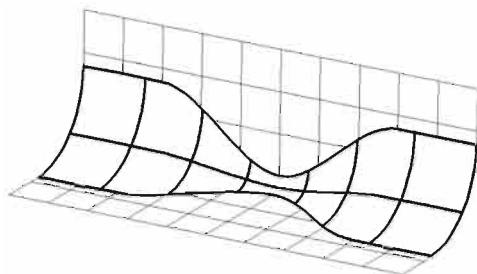


Рис. 4.9.3

Поверхность скрутения, сохраняющую кромку, строят в случаях, когда требуется, чтобы опорные кривые не выходили за границу сопрягаемых поверхностей. При достижении одной из опорных кривых границы сопрягаемой поверхности требуется, чтобы далее опорная кривая проходила по границе поверхности. В этих случаях можно построить поверхность, полученную движением сферы, касающейся одной поверхности и упирающейся в граничную кромку другой поверхности. Поверхность скрутения, сохраняющая кромку, приведена на рис 4.9.4.

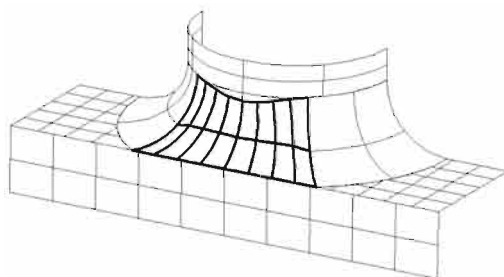


Рис. 4.9.4

Пусть для поверхности (4.9.1) граница поверхности $s(a,b)$ лежит ближе к кривой пересечения сопрягаемых поверхностей, чем опорная кривая $c_s(t)$. Пусть этот участок границы описывается кривой $h(s)$. Для построения опорной кривой $c_s(t) = r(u(t), v(t))$ вместо системы уравнений (4.9.2) используем систему уравнений

$$\begin{aligned}
 |r(u, v) + \rho \cdot m_r(u, v) - h(s)| &= |\rho_r|, \\
 (r(u, v) + \rho \cdot m_r(u, v) - h(s)) \cdot \frac{dh}{ds} &= 0.
 \end{aligned}
 \tag{4.9.6}$$

Искомыми являются параметры u, v . При движении по участку границы $\mathbf{h}(s)$ параметр s является известной величиной. Результатом решения системы уравнений (4.9.6) является кривая $[u(s) v(s)]^T$ на поверхности $\mathbf{r}(u, v)$. Она должна быть согласована с кривой $\mathbf{h}(s)$.

По двум опорным кривым $\mathbf{c}_r(s) = \mathbf{r}(u(s), v(s))$ и $\mathbf{h}(s)$ построим поверхность скругления

$$\mathbf{q}(s, z) = \frac{(1-z)^2 \mathbf{c}_r(s) + 2(1-z)zw(s)\mathbf{c}(s) + z^2 \mathbf{h}(s)}{(1-z)^2 + 2(1-z)zw(s) + z^2}, \quad (4.9.7)$$

$$s \in [s_{\min}, s_{\max}], z \in [0, 1],$$

проходящую через кромку $\mathbf{h}(s)$. Среднюю кривую $\mathbf{c}(s)$ и функцию веса $w(s)$ построим по формулам (4.9.4) и (4.9.3), заменив ρ, \mathbf{m}_s на вектор

$$\rho, \mathbf{m}_s = -\mathbf{h}(s) + \mathbf{r}(u, v) + \rho, \mathbf{m}_r(u, v).$$

Поверхность скругления с заданной хордой является разновидностью поверхности (4.9.1), в которой расстояние между опорными кривыми $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$ (хорда дуги вдоль параметра v) равно заданной величине h . Для построения опорных кривых вместо системы уравнений (4.9.5) используем систему уравнений

$$\mathbf{r}(u, v) + \rho, \mathbf{m}_r(u, v) = \mathbf{s}(a, b) + \rho, \mathbf{m}_s(a, b),$$

$$\left(\mathbf{c}_o(s) - \frac{1}{2}(\mathbf{r}(u, v) + \rho, \mathbf{m}_r(u, v) + \mathbf{s}(a, b) + \rho, \mathbf{m}_s(a, b)) \right) \cdot \frac{d\mathbf{c}_o}{ds} = 0,$$

$$|\mathbf{r}(u, v) - \mathbf{s}(a, b)| = h.$$

Эта система содержит пять скалярных уравнений относительно пяти искомых параметров u, v, a, b и $\rho = |\rho_r| = |\rho_s|$. На рис. 4.9.5 приведена поверхность скругления с заданной хордой, а на рис. 4.9.6 для сравнения приведена поверхность скругления с постоянным радиусом.

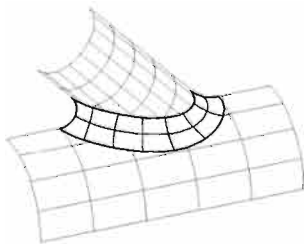


Рис. 4.9.5

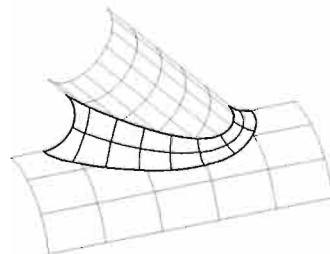


Рис. 4.9.6

Рассмотрим **сечение поверхности скругления** (4.9.1) при некотором фиксированном значении t первого параметра. Пусть $\mathbf{p}_0 = \mathbf{c}_r(t)$, $\mathbf{p}_1 = \mathbf{c}(t)$, $\mathbf{p}_2 = \mathbf{c}_s(t)$. Тогда сечение поверхности (4.9.1) опишется радиусом-вектором

$$\mathbf{q}(z) = \frac{(1-z)^2 \mathbf{p}_0 + 2(1-z)zw\mathbf{p}_1 + z^2 \mathbf{p}_2}{(1-z)^2 + 2(1-z)zw + z^2}, \quad (4.9.8)$$

который совпадает с (2.5.10). Сечение поверхности скругления при неподвижном первом параметре приведено на рис. 4.9.7.

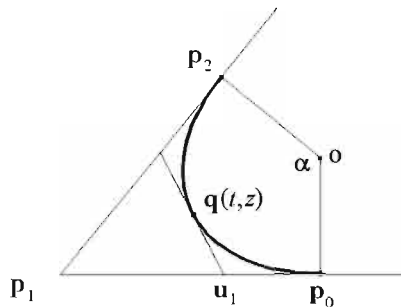


Рис. 4.9.7

Параметризация кривой (2.5.10) не является равномерной. Из-за этого отношение длины некоторой части дуги к параметрической длине этой же части дуги не является постоянным. Покажем, как получить равномерную параметризацию поверхности по второму параметру.

Приравняем правые части равенств (2.5.6) и (2.5.9) для z_1

$$z_1 = \frac{u_1}{1-u_1}, \quad z_1 = \frac{1}{w} \frac{1-v}{v}$$

и найдем из полученного равенства величину $1-u_1$

$$1-u_1 = \frac{vw}{1-v+vw}.$$

Длина отрезка $|\mathbf{p}_0\mathbf{p}_1|$ равна

$$l = \rho \cdot \operatorname{tg} \frac{\alpha}{2} = \rho \frac{\sqrt{1-w^2}}{w},$$

где ρ – радиус скругления; α – угол $\mathbf{p}_0\mathbf{O}\mathbf{P}_2$; w – косинус половины угла $\mathbf{p}_0\mathbf{O}\mathbf{P}_2$. Длина отрезка $|\mathbf{p}_0\mathbf{u}_1|$ пропорциональна величине $1-u_1$ и равна

$$l_1 = l(1-u) = \rho \frac{\sqrt{1-w^2}}{w} \frac{zw}{1-z+zw}.$$

Поделив l_1 на ρ , мы получим тангенс половины угла $\rho_0 \mathbf{oq}$:

$$\operatorname{tg} \frac{\beta}{2} = \frac{z\sqrt{1-w^2}}{1-z+zw}.$$

Введем новый параметр x , $0 \leq x \leq 1$, который равен отношению текущего угла β к полному углу дуги α . Тогда последнее равенство примет вид

$$\operatorname{tg} \left(\frac{\alpha}{2} x \right) = \frac{z\sqrt{1-w^2}}{1-z+zw}.$$

Решим последнее равенство относительно параметра поверхности z и получим функцию

$$z(x) = \frac{\operatorname{tg} \left(\frac{\alpha}{2} x \right)}{(1-w)\operatorname{tg} \left(\frac{\alpha}{2} x \right) + \sqrt{1-w^2}}. \quad (4.9.9)$$

Заменим в формуле (4.9.8) параметр z параметром x с помощью функции $z(x)$ (4.9.9) и получим формулу

$$\mathbf{q}(t, x) = \frac{(1-z(x))^2 \mathbf{c}_r(t) + 2(1-z(x))z(x)w(t)\mathbf{c}(t) + (z(x))^2 \mathbf{c}_s(t)}{(1-z(x))^2 + 2(1-z(x))z(x)w(t) + (z(x))^2}$$

$$t \in [t_{\min}, t_{\max}], x \in [0, 1]$$

для радиуса-вектора поверхности скругления с равномерной параметризацией по второму параметру.

Гладко сопрячь две поверхности можно множеством способов. Если в уравнениях (4.9.2) или (4.9.5) параметры радиусов ρ_r и ρ_s по модулю будут не равны, то получим поверхность, которую можно назвать **эллиптической поверхностью сопряжения**, рис. 4.9.8.

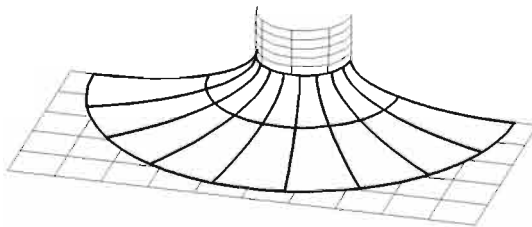


Рис. 4.9.8

поверхности Ω вычислим по скалярному произведению двумерных век-

Действительно, если сопрягаемые поверхности пересекаются под прямым углом, то линии $\mathbf{q}(\text{const}, v)$ сечения поверхности (4.9.1) будут являться дугами эллипсов. В общем случае всегда можно рассчитать параметры радиусов $\rho_r(t)$ и $\rho_s(t)$, чтобы поверхность сопряжения в сечении $\mathbf{q}(\text{const}, v)$ представляла собой дугу эллипса.

При подстановке постоянной функции веса $w(t) = 1$ в поверхность (4.9.1) получим **параболическую поверхность сопряжения**:

$$\mathbf{q}(t, z) = \frac{(1-z)^2 \mathbf{c}_r(t) + 2(1-z)z \mathbf{c}(t) + z^2 \mathbf{c}_s(t)}{(1-z)^2 + 2(1-z)z + z^2},$$

$$t \in [t_{\min}, t_{\max}], z \in [0, 1].$$

Используя для поверхности (4.9.1) функции веса, отличные от (4.9.3), также можно получить различные поверхности сопряжения. При подстановке постоянной функции веса $w(t) > 1$ получим **гиперболическую поверхность сопряжения**.

4.10. Поверхность фаски

Наряду с поверхностями гладкого сопряжения требуется выполнять построение поверхностей сопряжения в виде фасок. Поверхность фаски строится аналогично поверхности скругления.

Пусть имеются две пересекающиеся поверхности, описываемые радиусами-векторами $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$. В общем случае поверхность фаски представим в форме линейчатой поверхности

$$\mathbf{q}(t, z) = (1-z) \mathbf{c}_r(t) + z \mathbf{c}_s(t), \quad t \in [t_{\min}, t_{\max}], z \in [0, 1], \quad (4.10.1)$$

построенной по опорным кривым $\mathbf{c}_r(t) = \mathbf{r}(u(t), v(t))$ и $\mathbf{c}_s(t) = \mathbf{s}(a(t), b(t))$. Функции $u(t)$, $v(t)$, $a(t)$, $b(t)$, образующие двумерную кривую $\mathbf{c}_r(t) = [u(t) \ v(t)]^T$ в пространстве параметров поверхности $\mathbf{r}(u, v)$ и двумерную кривую $\mathbf{c}_s(t) = [a(t) \ b(t)]^T$ в пространстве параметров поверхности $\mathbf{s}(a, b)$, найдем путем решения векторного уравнения

$$\mathbf{r}(u, v) + d_r \mathbf{t}_r(u, v) = \mathbf{s}(a, b) + d_s \mathbf{t}_s(a, b), \quad (4.10.2)$$

где d_r и d_s – расстояния от точек опорных кривых до точки пересечения касательных плоскостей; $\mathbf{t}_r(u, v) = \pm \frac{\mathbf{m}_r \times (\mathbf{m}_r \times \mathbf{m}_s)}{|\mathbf{m}_r \times \mathbf{m}_s|}$, $\mathbf{t}_s(a, b) = \mp \frac{\mathbf{m}_s \times (\mathbf{m}_s \times \mathbf{m}_r)}{|\mathbf{m}_s \times \mathbf{m}_r|}$ –

векторы, лежащие в касательных к поверхностям $\mathbf{r}(u, v)$ и $\mathbf{s}(a, b)$ плоскостях и ортогональны опорным кривым $\mathbf{c}_r(t)$ и $\mathbf{c}_s(t)$. На рис. 4.10.1 приведено сечение поверхностей $\mathbf{r}(u, v)$, $\mathbf{s}(a, b)$ плоскостью, ортогональной ли-

нии пересечения касательных к поверхностям $\mathbf{r}(u,v)$ и $s(a,b)$ плоскостей в рассматриваемых точках.

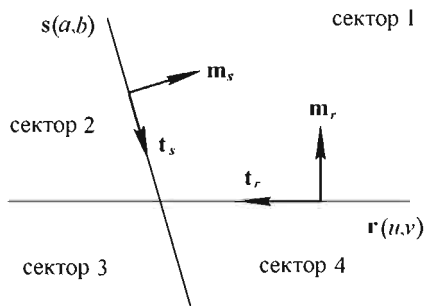


Рис. 4.10.1

Знаки величин d_r и d_s определяют сектор, в котором лежит поверхность фаски. Если в области пересечения поверхности плоские и ортогональны друг другу, то абсолютные величины d_r и d_s равны катетам фаски.

Записанные в координатном виде три скалярных уравнения (4.10.2) содержат четыре искомого параметра u , v , a , b . Если определить один из четырех параметров, то остальные три можно найти из системы уравнений (4.10.2). При наличии кривой пересечения $\mathbf{c}_0(s)$ поверхностей $\mathbf{r}(u,v)$ и $s(a,b)$ уравнение (4.10.2) может быть дополнено уравнением

$$\left(\mathbf{c}_0(s) - \frac{1}{2} (\mathbf{r}(u,v) + d_r \mathbf{t}_r(u,v) + \mathbf{s}(a,b) + d_s \mathbf{t}_s(a,b)) \right) \cdot \frac{d\mathbf{c}_0}{ds} = 0.$$

Полученная система содержит четыре уравнения относительно четырех параметров: u , v , a , b . В этой системе величины d_r и d_s можно сделать переменными, поставив в зависимость от длины дуги s кривой $\mathbf{c}_0(s)$: $d_r = d_r(s)$, $d_s = d_s(s)$, тогда (4.10.1) опишет поверхность фаски с катетами переменной величины.

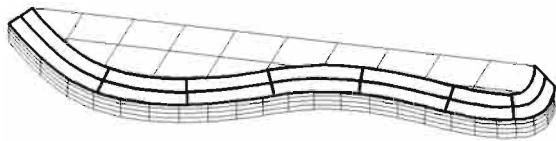


Рис. 4.10.2

На рис. 4.10.2 приведен пример поверхности фаски.

4.11. Положение точки относительно поверхности

Относительно заданной поверхности пространственная точка может располагаться над поверхностью, под поверхностью или на поверхности. Для определения положения точки \mathbf{p} относительно поверхности $\mathbf{s}(u, v)$ найдем проекцию $\mathbf{s}(u_p, v_p)$ точки \mathbf{p} на поверхность или ее продолжение. За пределами области определения параметров поверхность продолжим по закону ее построения или в соответствии с правилами построения продолженной поверхности (3.15.1) – (3.15.3) в случае отсутствия такого закона. Далее построим вектор $\mathbf{v} = \mathbf{p} - \mathbf{s}(u_p, v_p)$ из проекции в точку \mathbf{p} , вычислим нормаль $\mathbf{m}(u_p, v_p)$ поверхности в точке проекции. Положение точки относительно поверхности определим по скалярному произведению вектора \mathbf{v} и нормали \mathbf{m} . Если длина вектора \mathbf{v} равна нулю, то $\mathbf{v} \cdot \mathbf{m} = 0$ и точка \mathbf{p} располагается на поверхности, если $\mathbf{v} \cdot \mathbf{m} > 0$, то точка \mathbf{p} располагается над поверхностью, если $\mathbf{v} \cdot \mathbf{m} < 0$, то точка \mathbf{p} располагается под поверхностью.

В определенных случаях требуется ответить на вопрос: принадлежит ли двумерная точка $\mathbf{r} = [u_p \ v_p]^T$, описывающая проекцию точки \mathbf{p} , области определения параметров поверхности Ω или располагается вне области Ω . Пусть поверхность $\mathbf{s}(u, v)$ имеет сложную границу, которая описывается одним внешним двумерным контуром и несколькими внутренними двумерными контурами, целиком расположенными внутри внешнего контура. Граничные контуры поверхности ориентированы так, что при движении вдоль каждого контура поверхность всегда находится с одной и той же стороны. Пусть внешний контур ориентирован против движения часовой стрелки, если смотреть на поверхность навстречу ее нормали, а внутренние контуры ориентированы по движению часовой стрелки.

Для определения положения двумерной точки \mathbf{r} относительно области определения параметров поверхности Ω найдем ближайшую к точке \mathbf{r} точку двумерной границы области Ω . Пусть ближайшей двумерной точкой границы к точке \mathbf{r} оказалась точка $\mathbf{c}(t)$ граничного контура $\mathbf{c}(t) = [u(t) \ v(t)]^T$. Построим двумерный вектор $\mathbf{w} = \mathbf{r} - \mathbf{c}(t)$ из ближайшей граничной точки $\mathbf{c}(t)$ в точку \mathbf{r} и вычислим нормаль $\mathbf{n}(t)$ двумерного контура в ближайшей точке. Ближайшей точкой граничного контура может быть или проекция точки \mathbf{r} на один из сегментов контура, или точка стыковки двух сегментов контура. В точках стыковки двух сегментов контура за нормаль примем нормализованную сумму нормалей стыкующихся сегментов, то есть среднюю нормаль. Так как нормаль двумерной кривой направлена влево от первой производной этой кривой, то при принятой выше ориентации граничных контуров нормаль \mathbf{n} будет направлена от границы вглубь области определения параметров поверхности. Положение двумерной точки \mathbf{r} относительно области определения

поверхности Ω вычислим по скалярному произведению двумерных векторов w и n . Если длина вектора w равна нулю, то $w \cdot n = 0$ и точка r располагается на границе области Ω , если $w \cdot n > 0$, то точка r располагается внутри области Ω , если $w \cdot n < 0$, то точка r располагается вне области Ω .

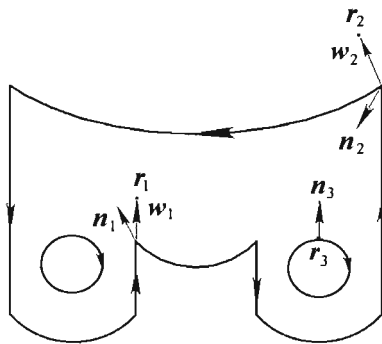


Рис. 4.11.1

На рис. 4.11.1 точка r_1 располагается внутри области определения поверхности ($w_1 \cdot n_1 > 0$), точка r_2 располагается снаружи области определения поверхности ($w_2 \cdot n_2 < 0$), точка r_3 располагается на границе области определения поверхности ($|w_3| = 0$).

4.12. Поиск начальных приближений

Для определения начальных приближений выберем в параметрической области кривой или поверхности пробные точки и проверим в пробных точках выполнение некоторого условия, характерного для данного построения. Пробные точки должны быть выбраны таким образом, чтобы было найдено начальное приближение для каждого решения. Характерным условием для некоторого построения, как правило, является смена знака невязки уравнений рассматриваемой системы или уменьшение невязки уравнений ниже известного уровня. Начальное приближение может быть найдено довольно грубо, важно только, чтобы оно лежало в области сходимости решения системы уравнений к одному из точных решений.

Пробные точки в параметрической области должны располагаться достаточно часто, чтобы не пропустить решение, и в то же время достаточно редко, чтобы потратить меньше времени. Можно избрать различные стратегии работы с кривыми и поверхностями при выборе пробных точек. Например, можно двигаться от минимального параметра к мак-

симальному параметру с некоторым вычисляемым шагом и проверять выполнение некоторого условия. Или можно провести предварительную разбивку кривой или поверхности на элементы и в дальнейшем использовать ее в работе. Стратегия работы с кривыми и поверхностями сводится к тому, чтобы представить каждую кривую и поверхность набором некоторых простых объектов и с их помощью принять решение о запуске итерационного процесса. Простыми объектами могут быть отрезки прямых, участки плоскости, габаритные кубы частей объектов или другие элементы. Для ускорения работы простые объекты могут быть отсортированы тем или иным способом. При разбивке кривых и поверхностей требуется определять участки в области определения параметров кривых и поверхностей, которые можно аппроксимировать простыми объектами.

В данной схеме проверок выполнения характерного условия остается неясным, какие точки можно считать близкими и какие точки следует выбрать для проверки следующими. Для кривых и поверхностей эти величины удобнее измерять в параметрических единицах. Расстояние близости и длина перемещения для следующей проверки в соседнюю точку являются связанными величинами. Действительно, если близким считать некоторое параметрическое расстояние Δ , то следующий поиск следует выполнять на параметрическом расстоянии Δ от данной точки. Длину параметрического шага будем связывать с искривленностью геометрического объекта. Приращения параметров Δ будем называть *параметрическими шагами*. Шаги перемещения по параметрической области кривой через кривизну связаны с допустимым углом отклонения касательной к ней. Шаги перемещения по параметрической области поверхности через соответствующую кривизну связаны с допустимым углом отклонения нормали поверхности. Максимально допустимую величину угла отклонения касательной или нормали обозначим $\Delta\alpha$.

Шаг Δt разбивки кривой $c(t)$ вычислим исходя из условия, что при переходе от точки к точке касательный вектор кривой отклонялся бы на угол, не превышающий $\Delta\alpha$. Приращение параметра или параметрический шаг Δt меняется от точки к точке и зависит от кривизны кривой. Пусть параметр t_0 описывает текущее положение точки в процессе движения по кривой. Следующее значение параметра $t_0 + \Delta t$ должно быть таким, чтобы касательный вектор в новой точке отклонился бы от касательного вектора кривой в предыдущей точке на угол, не превышающий $\Delta\alpha$. Для этого определим кривизну кривой в некоторой окрестности точки t_0 и в соответствии с найденной кривизной выберем следующую точку. Предположим, что кривизна кривой сохраняется постоянной в окрестности точки t_0 , тогда кривая отклонится на угол $\Delta\alpha$, если точка сместится по дуге на расстояние

$$\Delta s = \frac{\Delta \alpha}{k},$$

где k — кривизна кривой в данной точке. Предположим, что и длина производной кривой сохраняется постоянной в окрестности рассматриваемой точки, тогда смещению Δs по дуге будет соответствовать приращение параметра

$$\Delta t = \frac{\Delta s}{|\mathbf{c}'|} = \frac{\Delta \alpha}{k |\mathbf{c}'|}, \quad (4.12.1)$$

где \mathbf{c}' — производная кривой. Подставив в последнее равенство выражение для кривизны кривой, получим шаг по параметру

$$\Delta t = \frac{\Delta \alpha |\mathbf{c}'|^2}{|\mathbf{c}' \times \mathbf{c}''|}, \quad (4.12.2)$$

где \mathbf{c}'' — вторая производная кривой. Для некоторых кривых нет необходимости прибегать к вычислениям для определения шага Δt . Например, для окружности или ее дуги $\Delta t = \Delta \alpha$, если параметром является центральный угол, для отрезка прямой можно положить $\Delta t = t_{\max} - t_{\min}$.

Приведенные рассуждения базируются на том, что кривую в рассматриваемой окрестности можно аппроксимировать дугой окружности конечного радиуса. При стремлении кривизны в рассматриваемой точке к нулю длина параметрического шага стремится к бесконечности. Кривизна кривой в рассматриваемой точке может быть очень малой или вообще равной нулю, хотя в достаточно близкой к рассматриваемой точке области кривизна может иметь гораздо большее значение. Для учета подобных ситуаций можно ввести ограничение максимального параметрического шага или сделать небольшой шаг по параметру и повторить вычисления в новой точке. Если в рассматриваемой точке длина первой производной кривой стремится к нулю, то длина параметрического шага тоже стремится к нулю. Длина первой производной кривой в рассматриваемой точке может быть очень малой, хотя в достаточно близкой к рассматриваемой точке области длина производной может иметь гораздо большее значение. Для учета подобных ситуаций введем ограничение минимального параметрического шага. Для разных кривых это ограничение может быть различным.

Шаги Δu и Δv разбивки поверхности $s(u, v)$ вычислим исходя из условия, что при переходе от точки к точке нормаль поверхности отклонялась бы на угол, не превышающий некоторую величину $\Delta \alpha$. Шаги по параметрам поверхности будем вычислять как шаги для соответствующих координатных кривых поверхности. Положим один из параметров постоянным и будем двигаться вдоль другого параметра. Для вычисле-

ния шага по параметрам поверхности в правую часть (4.12.1) подставим кривизны вдоль координатных линий поверхности. При $v = \text{const}$ или $u = \text{const}$ кривизны линий соответственно равны $\mu_u = b_{11}/g_{11}$ или $\mu_v = b_{22}/g_{22}$, где g_{11} и g_{22} – коэффициенты первой квадратичной формы поверхности (3.1.4), b_{11} и b_{22} – коэффициенты второй квадратичной формы поверхности (3.1.5). В результате шаг по параметрам поверхности вычислим по формулам

$$\Delta u = \frac{\Delta\alpha\sqrt{g_{11}}}{b_{11}} = \frac{\Delta\alpha|s_1|}{|\mathbf{m} \cdot \mathbf{s}_1|}, \quad (4.12.3)$$

$$\Delta v = \frac{\Delta\alpha\sqrt{g_{22}}}{b_{22}} = \frac{\Delta\alpha|s_2|}{|\mathbf{m} \cdot \mathbf{s}_2|}. \quad (4.12.4)$$

Так же как для кривых, предполагаем, что рассматриваемая окрестность точки поверхности сохраняет постоянной нормальную кривизну вдоль параметрических направлений в пределах шага. В областях резкого изменения формы поверхности можно ввести ограничение максимального шага и минимального шага по соответствующему параметру или сделать небольшой шаг по соответствующему параметру и повторить вычисления в новой точке.

Параметрические шаги, вычисленные по формулам (4.12.2), (4.12.3), (4.12.4), не зависят от размеров кривых и поверхностей.

4.13. Точность геометрических построений

В прикладных задачах важно знать, с какой точностью были выполнены геометрические построения. Для описания кривых и поверхностей мы используем параметрическое представление, поэтому поиск некоторого решения заключается в поиске соответствующих ему параметров. Параметры решения не изменяются при одновременном перемещении, повороте или трансформировании взаимодействующих кривых и поверхностей. Во многих случаях параметры кривых и поверхностей нам известны с некоторой точностью. Будем различать *пространственную погрешность* и *параметрическую погрешность*. Для различных кривых и поверхностей одна и та же пространственная погрешность будет соответствовать различной параметрической погрешности. Поэтому для нормального взаимодействия всего многообразия кривых и поверхностей следует отталкиваться от общей пространственной погрешности.

Пусть предъявлено требование: выполнять геометрические построения с пространственной точностью ξ . Это означает, что две точки в пространстве считаются совпадающими, если расстояние между ними не

превосходит ξ (или, в более удобном виде, если координаты точек совпадают с точностью ξ). Пусть некоторой точке P_0 кривой $\mathbf{c}(t)$ соответствует параметр t_0 . Точка P_1 этой же кривой с параметром $t_1 = t_0 \pm \varepsilon$, где $\varepsilon > 0$ и ε во много раз меньше области изменения параметра кривой $t_{\max} - t_{\min}$, будет отстоять от точки P_0 на расстоянии

$$r = \varepsilon |\mathbf{c}'(t_0)|, \quad 0 < \varepsilon < \langle t_{\max} - t_{\min} \rangle,$$

где $\mathbf{c}'(t_0)$ – производная кривой в точке P_0 . Если будет выполняться неравенство

$$\varepsilon \leq \frac{\xi}{|\mathbf{c}'(t_0)|}, \quad (4.13.1)$$

то точки P_0 и P_1 кривой $\mathbf{c}(t)$ будут неразличимы для заданной пространственной точности ξ . В (4.13.1) ε является параметрической точностью для кривой $\mathbf{c}(t)$. Аналогично точка поверхности $\mathbf{r}(u, v)$ с параметрами u_0, v_0 будет считаться совпадающей с точками этой же поверхности, параметры которых $u_1 = u_0 \pm \varepsilon_1, v_1 = v_0 \pm \varepsilon_2$, если будут выполняться неравенства

$$\varepsilon_1 \leq \frac{\xi}{|\mathbf{r}_1(u_0, v_0)|} \quad \text{и} \quad \varepsilon_2 \leq \frac{\xi}{|\mathbf{r}_2(u_0, v_0)|}, \quad (4.13.2)$$

где $\mathbf{r}_1(u_0, v_0)$ и $\mathbf{r}_2(u_0, v_0)$ – частные производные поверхности по первому и второму параметру, соответственно, в рассматриваемой точке; ξ – пространственная точность; ε_1 и ε_2 – параметрические точности для первого и второго параметров поверхности $\mathbf{r}(u, v)$.

В итерационных методах геометрических построений параметры кривых и поверхностей мы находим с некоторой погрешностью. Процесс уточнения параметров следует продолжать до тех пор, пока на очередной итерации изменение каждого параметра по абсолютной величине не будет удовлетворять неравенствам (4.13.1), (4.13.2). Но, кроме выполнения неравенств (4.13.1), (4.13.2), необходимо добиться выполнения равенств заданной системы уравнений с точностью не ниже ξ . Но это еще не гарантирует, что нами будет достигнута требуемая точность. Вблизи точек касания решение сходится очень медленно. Поэтому в неравенства (4.13.1), (4.13.2) можно ввести коэффициент k , который может быть равен нескольким десяткам или даже сотням. Итерационные процессы геометрических построений следует продолжать, пока изменения параметров не станут меньше ε / k и пока уравнения не будут выполнены с точностью ξ .

В тех случаях, когда нам не известны параметры кривой или поверхности, но требуется указать для них погрешность, можно использовать среднюю погрешность. Для этого нам придется привлечь некоторый геометрический размер геометрического объекта. В качестве такого гео-

метрического размера можно взять длину кривой или диагональ габаритного куба поверхности. Предположим, что параметрическая длина рассматриваемой кривой или поверхности не меньше единицы и размеры соизмеримы с величиной d . Тогда вместо погрешностей (4.13.1), (4.13.2) можно использовать средние погрешности

$$\varepsilon = \frac{\xi}{d} |t_{\max} - t_{\min}|$$

для кривой $\mathbf{c}(t)$, $t_{\min} \leq t \leq t_{\max}$,

$$\varepsilon_1 = \frac{\xi}{d} |u_{\max} - u_{\min}| \text{ и } \varepsilon_2 = \frac{\xi}{d} |v_{\max} - v_{\min}|$$

для поверхности $\mathbf{r}(u, v)$, $u \in [u_{\min}, u_{\max}]$, $v \in [v_{\min}, v_{\max}]$.

Там, где это возможно геометрические построения следует выполнять без привлечения численных методов. В этих случаях точность геометрических построений будет зависеть от точности диагностики частных случаев. Например, при пересечении цилиндрической поверхности с ортогональной ее оси плоскостью точность построения линии пересечения будет определяться точностью, при которой вектор нормали плоскости и ось цилиндра считаются параллельными с учетом радиуса цилиндра.

В то же время величины ξ и ε не должны быть меньше 10^{-n} , где n — число значащих цифр представления координат и параметров в памяти компьютера. В противном случае разность чисел, отличающихся на относительную величину, меньшую 10^{-n} , будет неразличима и мы не сможем в принципе обеспечить точность ξ или ε .

Кроме величин ξ и ε , в геометрических построениях играют роль минимально возможный размер d_{\min} и максимально возможный размер d_{\max} модели. Это означает, что должно обеспечиваться пересечение, например, поверхности или кривой размера d_{\min} с поверхностью или кривой размера d_{\max} . Пусть область изменения параметров $|t_{\max} - t_{\min}|$, $|u_{\max} - u_{\min}|$, $|v_{\max} - v_{\min}|$ составляет несколько единиц. В таком случае для больших моделей вычисленные параметрические погрешности будут очень малы и могут оказаться меньше 10^{-n} , где n — число точных чисел представления параметров. В этом случае относительную погрешность вычисления параметров, равную ε , обеспечить будет невозможно. С другой стороны, если d_{\min} будет соизмерима с величиной ξ , то для маленьких моделей ε будет соизмерима со всей параметрической областью. Построения на моделях таких размеров окажутся невозможными.

Вышесказанное показывает, что общее количество значащих цифр представления действительных чисел в памяти компьютера является

отправной точкой для определения максимальной точности, которую можно достичь численными методами.

Упражнения

1. Опишите алгоритм вычисления проекции точки на окружность.
2. Опишите алгоритм вычисления ближайших точек прямой и цилиндрической поверхности.
3. Опишите алгоритм построения окружности заданного радиуса, касающейся двух прямых.
4. Опишите алгоритм построения кривой пересечения плоскости и цилиндрической поверхности.

Глава 5. Тела

Главной частью геометрической модели является описание формы моделируемого объекта. Одной поверхностью в общем случае невозможно описать форму объекта, но это можно сделать с помощью набора стыкующихся определенным образом друг с другом поверхностей. Граничное представление с помощью криволинейных граней, стыкующихся друг с другом по общим ребрам, описывает граничную поверхность, отделяющую внутренний объем моделируемого объекта от остальной части пространства. Описание граничной поверхности моделируемого объекта приводит к построению тел.

5.1. Оболочка

Геометрическую форму моделируемого объекта опишем набором стыкующихся друг с другом поверхностей. На рис. 5.1.1 слева приведено несколько плоских, цилиндрических поверхностей и несколько поверхностей тора, обрезанных определенным образом.

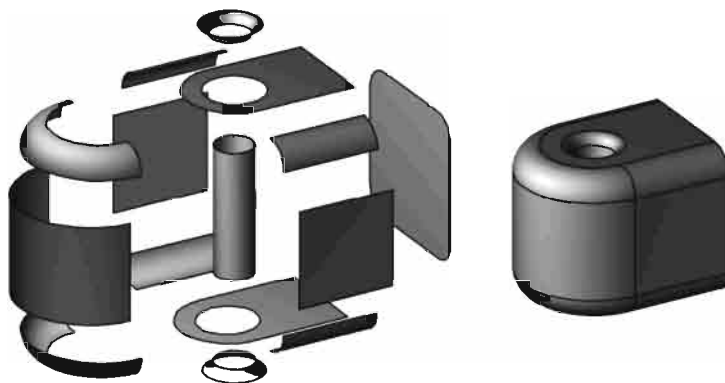


Рис. 5.1.1

По краям поверхностей построены кривые. На рис. 5.1.1 справа приведены те же поверхности, состыкованные друг с другом. Поверхности стыкуются друг с другом по кривым. С помощью приведенных на рис. 5.1.1 поверхностей можно описать поверхность некоторого моделируемого объекта. Это описание будет составлять главную часть геометрической модели.

Для описания геометрической формы используются оболочки, которые строятся из таких объектов, как грань, ребро, вершина и цикл.

Гранью будем называть конечную поверхность, которой приписано направление нормали. Направление нормали грани указывает специальный признак. Будем говорить, что грань базируется на поверхности и имеет признак направления нормали. Признак принимает положительное значение, если нормали грани и поверхности совпадают, в противном случае признак принимает отрицательное значение. Сторону грани, при взгляде на которую мы смотрим навстречу нормали грани, будем называть *внешней*, другую сторону грани будем называть *внутренней*.

Грань будем обозначать так же, как и поверхность. Стороны поверхности не обладают равноправием относительно ее нормали, так как у поверхности одна сторона всегда внешняя, а другая сторона всегда внутренняя. В отличие от поверхности грань имеет возможность связать направление нормали с любой из двух своих сторон.

Пусть параметры грани принадлежат двумерной связной области Ω . Границы грани представляют собой отображение в трехмерное пространство двумерной границы области Ω . Грань имеет одну внешнюю границу и несколько внутренних границ, если грань имеет внутренние вырезы. Каждая граница грани замкнута. Каждую границу разобьем на участки и выберем на них параметры, то есть участки границ представим кривыми, которые будем называть *граничными кривыми*.

Ребром будем называть кривую, которой приписано направление. Направление ребра указывает специальный признак. Будем говорить, что ребро базируется на кривой и имеет признак направления. Признак принимает положительное значение, если направления ребра и кривой совпадают, и принимает отрицательное значение, если ребро направлено против кривой. В отличие от кривой ребро можно направить требуемым образом.

Мы будем использовать ребра, построенные на граничных кривых граней. По граничной кривой грань может стыковаться с другой гранью. В таком случае ребро будем строить по кривой пересечения соседних граней. Грани, имеющие общее ребро, будем называть *смежными*. Ребро стыковки граней опишем кривой пересечения (4.4.1).

У циклически замкнутой по одному или обоим параметрам грани присутствуют участки границы, по которым грань стыкуется сама с собой. Участок границы и построенное на нем ребро, по которому грань стыкуется сама с собой, будем называть *швом*.

Ребро, построенное на свободном краю грани, опишем кривой на поверхности (4.3.1). Такое ребро будем называть *краевым ребром*.

Участок границы грани в особых точках может представлять собой точку, хотя в области параметров этому участку соответствует некоторая кривая. Построенное по такому участку границы ребро будем называть *полюсным ребром*, а соответствующую пространственную точку — *полюсом*.

Если ребро начинается и заканчивается в одной и той же точке, то такое ребро называется *замкнутым*.

Вершиной будем называть точку, в которой стыкуются ребра. Вершины могут лежать только на концах ребер. В каждой вершине может стыковаться любое конечное число ребер.

Циклом грани будем называть последовательность ребер, исчерпывающую некоторую границу этой грани. Цикл, так же как и граница грани, всегда замкнут. Цикл направим так, чтобы грань всегда для нас находилась бы слева, если мы движемся вдоль цикла с внешней стороны грани. Направление ребра цикла может совпадать с направлением цикла или не совпадать с ним. Признак, характеризующий совпадение направления ребра с направлением цикла будем называть *флагом*. Ребру, направление которого совпадает с направлением цикла, припишем положительный флаг, а ребру, направление которого не совпадает с направлением цикла, припишем отрицательный флаг. Таким образом, цикл состоит из списка ребер в порядке их следования и списка соответствующих им флагов.

Оболочкой будем называть связную поверхность, образованную конечной совокупностью граней, стыкующихся друг с другом по общим ребрам. Оболочками в геометрическом моделировании описывают граничную поверхность, отделяющую моделируемый объект от остального пространства. Оболочка может описывать всю граничную поверхность моделируемого объекта или только некоторую часть его граничной поверхности. Оболочка, описывающая только часть граничной поверхности моделируемого объекта, имеет края. Оболочку будем называть *замкнутой*, если она не имеет края. В противном случае оболочку будем называть *незамкнутой*.

5.2. Характеристики оболочки

Свойства оболочек, не зависящие от количественных характеристик (длин и углов), отражают непрерывную связь между точками оболочки и называются топологическими.

Эйлерова характеристика оболочки связывает количество ее граней, ребер, вершин и циклов. Пусть оболочка содержит F граней, E ребер, V вершин и L циклов. Число вершин, ребер, граней и циклов оболочки связаны между собой соотношением

$$F - E + V + (F - L) = H, \quad (5.2.1)$$

где величина H называется *эйлеровой характеристикой* оболочки. Формула (5.2.1) называется *формулой Эйлера*. Если каждая грань оболочки имеет один цикл, то $F - L = 0$ и слагаемое в скобках в левой части (5.2.1) может быть опущено.

Эйлерова характеристика не зависит от количественных характеристик (длин, площади, объема и углов) оболочки, а также от разбивки оболочки на грани, ребра и вершины. В этом можно убедиться с помощью рис. 5.2.1, на котором приведены различные способы разбивки участка оболочки на грани, ребра и вершины.

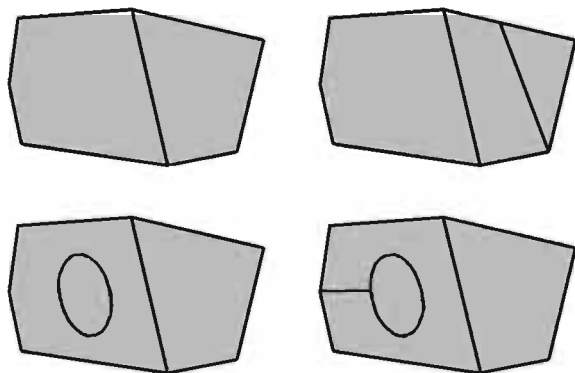


Рис. 5.2.1

Эйлерова характеристика зависит от свойств оболочки, которые изучает топология. С помощью Эйлеровой характеристики можно судить о возможности установить между оболочками взаимно однозначное соответствие. Рассмотрим две оболочки: оболочку в форме четырехугольной призмы и оболочку в форме четырехугольной призмы с четырехугольным отверстием в середине, рис. 5.2.2.

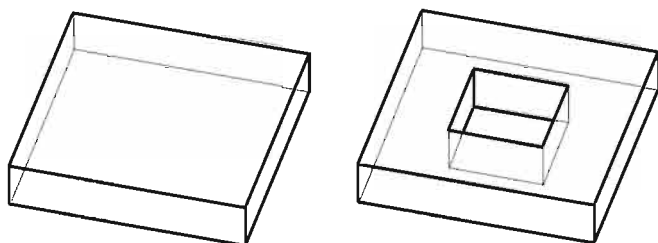


Рис. 5.2.2

Эйлерова характеристика первой оболочки равна двум. Эйлерова характеристика второй оболочки равна нулю. Вторая оболочка неэквивалентна первой оболочке, поэтому оболочки имеют разные эйлеровы характеристики.

Исследуем топологические свойства оболочки, которые сохраняются при всевозможных ее деформациях. Предположим, что оболочка выполнена из эластичного материала. *Деформацией* будем называть изменение формы оболочки путем растяжения, сжатия, сдвига или изгиба ее поверхности, не приводящее к разрывам и не требующее склеивания поверхности оболочки. Будем говорить, что две оболочки имеют одинаковую топологию, если с помощью деформации из одной оболочки можно получить другую. Из первой оболочки, показанной на рис. 5.2.2, путем деформаций можно получить сферу, из второй оболочки путем деформаций можно получить тор, рис. 5.2.3. Никакими деформациями невозможно из сферы получить тор, так как они имеют разную топологию.

Оболочку, имеющую топологию круга, будем называть *простейшей*. Простейшая оболочка незамкнута, имеет одну границу и может быть представлена одной гранью. Ответим на вопрос: какое минимальное число кривых можно провести на поверхности оболочки, чтобы по этим кривым ее можно было разрезать на две простейшие оболочки?



Рис. 5.2.3

Связность оболочки определим минимальным числом кривых, по которым ее можно разрезать на две простейшие оболочки. Кривые должны быть замкнутыми или начинаться и оканчиваться на границах оболочки. Связность обозначим через h . Если связность оболочки равна h , то $h - 1$ разрезов достаточно, чтобы ими раскроить оболочку, превратив ее в простейшую. Оболочку связности h с помощью h разрезов можно разделить на две отдельные части.

Простейшая оболочка делится на две отдельные части любой кривой, проведенной от одной точки границы до другой точки, не совпадающей с первой точкой. Простейшую оболочку называют *односвязной*.

На рис. 5.2.4 приведены незамкнутые оболочки, имеющие конечное число границ, которые путем деформирования можно сделать плоскими.

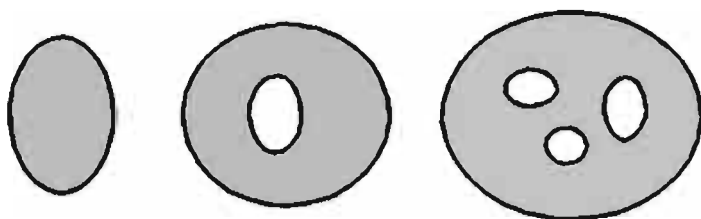


Рис. 5.2.4

Если плоская оболочка имеет L границ, то на ней можно провести $L - 1$ кривых, не разрезающих оболочку на две отдельные части, например, от внешней границы к каждой внутренней границе. Любая следующая кривая, начинающаяся и оканчивающаяся на границе оболочки, разрежет ее на две части. Плоская оболочка с L границами имеет связность, равную L .

Оболочку, имеющую топологию сферы, по любой замкнутой кривой на ее поверхности можно разрезать на две отдельные части, представляющие собой односвязные оболочки. Для оболочки, имеющей топологию тора, потребуется, как минимум, две замкнутые кривые на ней для раскроя ее в односвязную оболочку. Способ раскроя поверхности тора показан на рис. 5.2.5. Используя определение, получим, что связность сферы равна единице, а связность тора равна трем. Цилиндрическая оболочка конечной длины имеет связность, равную двум. Способ раскроя цилиндрической оболочки конечной длины в односвязную оболочку приведен на рис. 5.2.6.



Рис. 5.2.5



Рис. 5.2.6

На рис. 5.2.7 приведены незамкнутые оболочки, полученные из замкнутых оболочек.

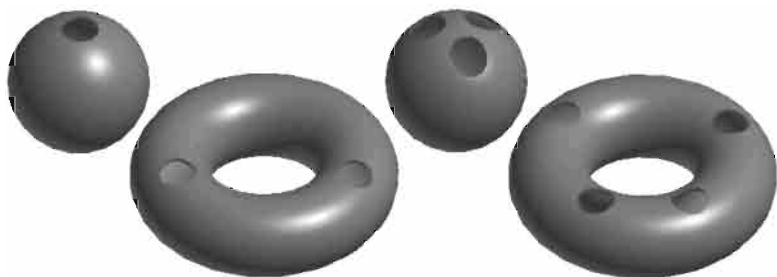


Рис. 5.2.7

Если в замкнутой оболочке связности h выполнить одно отверстие с одной замкнутой границей, то связность полученной незамкнутой оболочки будет равна связности исходной замкнутой оболочки. Каждое последующее отверстие с одной замкнутой границей будет увеличивать связность полученной оболочки на единицу. Связность приведенных на рис. 5.2.7 незамкнутых оболочек соответственно равна: один, четыре, три, шесть. Система разрезов на этих оболочках строится аналогично системе разрезов на соответствующей замкнутой оболочке, только часть кривых проводится от одной границы к другой границе.

Ориентируемость оболочки характеризует ее свойство иметь две стороны. Оболочки, имеющие две стороны, называют *двусторонними* или *ориентируемыми*. Оболочки, имеющие только одну сторону, называют *односторонними* или *неориентируемыми*.

Примером односторонней оболочки является лист Мебиуса, рис. 5.2.8. Лист Мебиуса можно получить из бумажной полосы, если ее дальние края повернуть друг относительно друга на 180° и склеить. До склеивания краев полосы ее стороны можно покрасить двумя разными цветами. Если покраску проводить после склеивания, то окажется, что мы окрасим одним цветом всю оболочку.

Примером замкнутой односторонней оболочки является бутылка Клейна. Бутылка Клейна имеет одну замкнутую кривую самопересечения. На рис. 5.2.9. приведена полупрозрачная бутылка Клейна.



Рис. 5.2.8

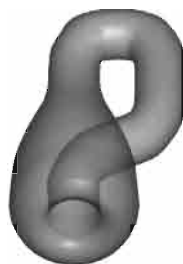


Рис. 5.2.9

Бутылка Клейна не может служить сосудом. Система линий, разрезающих бутылку Клейна на две односвязные части, аналогична системе линий тора. Связность бутылки Клейна равна трем. Если бутылку Клейна разрезать плоскостью ее симметрии, то получим две незамкнутые самопересекающиеся оболочки, из которых путем деформирования можно получить два листа Мебиуса. Еще одним примером односторонней оболочки может служить гептаэдр.

Односторонние оболочки называют неориентируемыми, так как у таких оболочек, задав для малой окрестности некоторой точки ориентацию обхода и двигаясь по оболочке, мы попадем в исходную точку, но с противоположным направлением обхода.

Топологический тип характеризует оболочку, так же как и связность. Замкнутые оболочки имеют нечетную связность. Для таких оболочек эйлерова характеристика H связана с ее связностью h соотношением

$$H = 3 - h.$$

Для замкнутых ориентируемых оболочек формулу Эйлера можно представить в виде

$$F - E + V + (F - L) = 3 - h.$$

Из тора путем его деформирования можно получить оболочку, по форме напоминающую гирию, которую будем называть сферой с ручкой. В общем случае любой замкнутой оболочке путем деформирования можно придать форму сферы с G ручками. Так, если взять толстую плиту, пробить в ней G отверстий и скруглить все ребра, то получим объект, оболочка которого имеет топологию сферы с G ручками. Величина G (genus) характеризует *топологический тип* оболочки. На рис. 5.2.10 приведена сфера с четырьмя ручками.

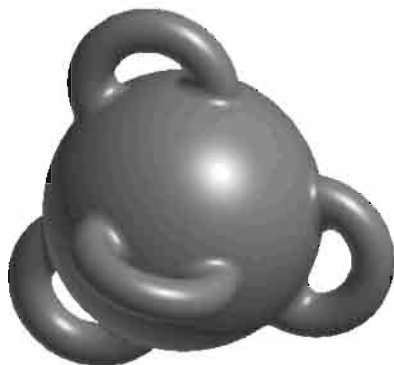


Рис. 5.2.10

Более наглядной, чем связность, характеристикой топологии оболочки может служить число ручек G сферы, к которой путем деформирования можно привести замкнутую оболочку. На оболочке с топологией сферы с G ручками можно провести $2G$ замкнутых кривых, по которым она раскраивается в простейшую оболочку. Любая следующая замкнутая кривая разрежет оболочку на две простейшие оболочки. Таким образом, сфера с G ручками имеет связность $h = 2G + 1$. Подставим значение связности, выраженное через величину G , в формулу (5.2.1) и получим формулу, связывающую число граней F , ребер E , вершин V и циклов L с характерной величиной G

$$F - E + V + (F - L) = 2(1 - G).$$

Данная формула называется *формулой Эйлера–Пуанкаре*. Формула Эйлера–Пуанкаре позволяет определить топологический тип оболочки, если известно число ее граней, ребер, вершин и циклов.

5.3. Однородная оболочка

В общем случае оболочка может пересекать сама себя, иметь или не иметь края, уходить в бесконечность. Среди множества различных оболочек выделим оболочки, отражающие геометрические свойства реальных объектов.

Однородной оболочкой (Manifold shell) называется оболочка, удовлетворяющая следующим пяти требованиям.

1. Оболочка должна быть ограниченной.
2. Оболочка не должна пересекать сама себя.
3. Оболочка должна быть двусторонней.

4. В каждом ребре оболочки должны стыковаться не более двух граней, причем две грани должны стыковаться так, чтобы внешняя сторона одной грани переходила бы во внешнюю сторону другой грани.
5. При обходе любой вершины оболочки по поверхности необходимо посетить все примыкающие к данной вершине грани и пересечь все выходящие из нее ребра.

Так как внешняя сторона любой грани переходит во внешнюю сторону соседней грани, то однородная оболочка также имеет внешнюю и внутреннюю стороны.

В однородной оболочке ребро, разделяющее две грани, входит в два цикла. В цикл одной из граней ребро входит с положительным флагом, в цикл другой грани — с отрицательным флагом. Если ребро проходит по краю незамкнутой оболочки, то ребро входит только в один цикл.

В геометрическом моделировании используют преимущественно однородные оболочки, так как ими можно просто и полно описать геометрическую форму моделируемого объекта.

Неоднородными оболочками являются уходящие в бесконечность оболочки, самопересекающиеся оболочки, односторонние оболочки.

Неоднородной является оболочка, приведенная на рис. 5.3.1, так как в ребре ED оболочки стыкуется более двух граней. Неоднородной является оболочка, приведенная на рис. 5.3.2, так как вершину V можно обойти вокруг, не посетив все примыкающие к вершине грани.

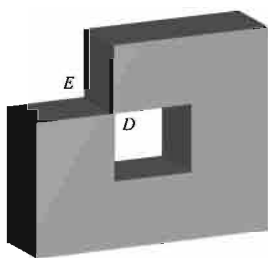


Рис. 5.3.1

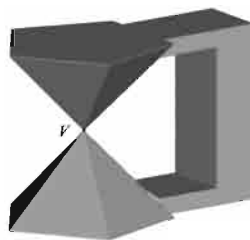


Рис. 5.3.2

Если оболочки, приведенные на рис. 5.3.1 и 5.3.2, разрезать по ребру ED и вершине V , то оболочки станут однородными.

5.4. Тело

Замкнутая однородная оболочка делит трехмерное пространство на две части, одна из которых конечная. Назовем замкнутую однородную оболочку *внешней*, если ее внешняя сторона направлена вне конечной

части пространства. Назовем замкнутую однородную оболочку *внутренней*, если ее внешняя сторона направлена внутрь конечной части пространства.

Телом в геометрическом моделировании называется связанное множество точек, располагающихся с внутренней стороны одной внешней и нескольких внутренних оболочек, расположенных внутри внешней оболочки, в совокупности с точками этих оболочек.

Тело описывается совокупностью своих оболочек. В большинстве случаев тело описывается одной внешней оболочкой. Внутренние оболочки тела представляют пустоты внутри него. Тела являются одними из основных элементов геометрической модели, но не единственными. Незамкнутые или не являющиеся однородными оболочки имеют вполне равные с телами права на отражение геометрических свойств моделируемых объектов.

С помощью тел мы можем описать граничную поверхность моделируемого объекта, то есть его геометрическую форму. Тела составляют основную часть геометрической модели. Кроме тел, геометрическая модель может содержать и другие компоненты. Другие компоненты модели предназначены для ее редактирования. Если модель редактировать не требуется, то она может состоять только из тел. Все остальные компоненты модели не обязательны.

5.5. Структуры данных

Геометрическая модель в граничном представлении в простейшем случае состоит из тел. Каждое тело состоит из набора оболочек. Каждая оболочка представляет собой набор граней, построенных на кривых и поверхностях, которые в свою очередь описываются точками, векторами, целыми числами и символьными выражениями. В конечном итоге геометрическая модель представляет собой последовательность чисел, организованных определенным образом в заранее определенные структуры данных.

Структуры данных отражают состав описываемых ими объектов. Приведем один из возможных вариантов структур данных граней, ребер, вершин, оболочек и тел. Данный вариант позволяет с любого ребра или с любой грани перейти на любой другой элемент оболочки и получить о нем всю информацию.

Структура данных вершины содержит точку.

Структура данных ребра содержит кривую и признак совпадения направления ребра с этой кривой. Кроме этого для удобства работы ребро содержит информацию о начальной и конечной вершинах и о том, какая грань находится слева, а какая – справа от ребра, если смотреть с внеш-

ней стороны граней вдоль направления ребра. На рис. 5.5.1 приведена структура данных ребра, начинающегося в вершине А, оканчивающегося в вершине В.

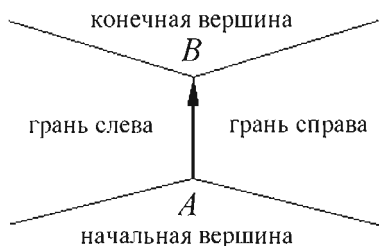


Рис. 5.5.1

Структура данных ориентированного ребра содержит ребро и флаг направления движения вдоль ребра. Ориентированные ребра вводятся для удобства описания циклов грани.

Структура данных цикла грани содержит ориентированные ребра в порядке их следования, то есть содержит ребра в порядке их следования и соответствующие им флаги совпадения направлений ребер с направлением цикла.

Структура данных грани содержит поверхность, признак направления нормали грани с направлением нормали поверхности и совокупность циклов грани. Циклы грани наследуют свойства границы области определения поверхности. Количество циклов грани равно количеству границ поверхности грани. Одна из границ грани является внешней и содержит внутри границы внутренних вырезов. Цикл грани, содержащий внутри себя остальные циклы, будем называть *внешним*, а остальные – *внутренними*.

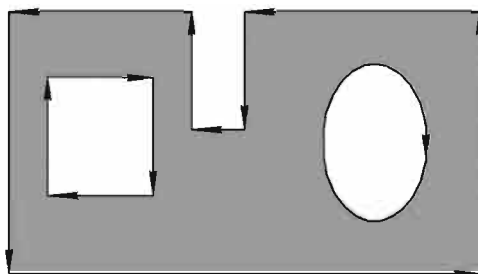


Рис. 5.5.2

Внешний цикл грани ориентирован против часовой стрелки, а внутренние циклы ориентированы по часовой стрелке, если смотреть на встречу нормали грани. Циклы одной грани не могут пересекать друг друга и сами себя. На рис. 5.5.2 стрелки указывают направления внешне-го и двух внутренних циклов грани.

Таким образом, грань содержит информацию обо всех вершинах и ребрах, соединяющих ее с другими гранями. От ребер можно получить информацию о смежных гранях или их отсутствии. Такая структура данных ребра и грани устанавливает двунаправленную связь между ними. Оказавшись на одном из объектов, можно последовательно пройти всю оболочку и получить информацию обо всех объектах оболочки.

Структура данных оболочки содержит составляющие ее грани и признак замкнутости оболочки.

Структура данных тела содержит составляющие его оболочки, а также может содержать последовательность и способы построения этих оболочек.

Проиллюстрируем общие принципы описания геометрической формы моделируемых объектов на примере элементарных тел.

5.6. Элементарные тела

К элементарным телам относят призму, цилиндр, конус, шар и тор. Элементарные тела состоят из одной оболочки. Для каждого элементарного тела выберем локальную декартову прямоугольную систему координат, радиус-вектор начала которой обозначим через \mathbf{p} , а базисные векторы обозначим через $\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z$.

Прямоугольная призма состоит из шести граней. Начало местной декартовой системы координат поместим в одну из вершин призмы, а ее базисные векторы направим по ребрам, стыкующимся в этой вершине, рис. 5.6.1. Пусть в направлении базисного вектора \mathbf{i}_x тело имеет длину, равную x , в направлении базисного вектора \mathbf{i}_y — y , в направлении базисного вектора \mathbf{i}_z — z . Каждая грань призмы представляет собой часть плоскости с признаком ориентации нормали плоскости наружу тела и одним циклом. Грани призмы опишем поверхностями

$$\begin{aligned}
 \mathbf{r}_1(u_1, v_1) &= \mathbf{p} + u_1 \mathbf{i}_x + v_1 \mathbf{i}_y, & u_1 \in [0, x], v_1 \in [0, y], \\
 \mathbf{r}_2(u_2, v_2) &= \mathbf{p} + u_2 \mathbf{i}_y + v_2 \mathbf{i}_z, & u_2 \in [0, y], v_2 \in [0, z], \\
 \mathbf{r}_3(u_3, v_3) &= \mathbf{p} + u_3 \mathbf{i}_x + v_3 \mathbf{i}_z, & u_3 \in [0, x], v_3 \in [0, z], \\
 \mathbf{r}_4(u_4, v_4) &= \mathbf{p} + u_4 \mathbf{i}_x + v_4 \mathbf{i}_y + z \mathbf{i}_z, & u_4 \in [0, x], v_4 \in [0, y], \\
 \mathbf{r}_5(u_5, v_5) &= \mathbf{p} + u_5 \mathbf{i}_y + v_5 \mathbf{i}_z + x \mathbf{i}_x, & u_5 \in [0, y], v_5 \in [0, z], \\
 \mathbf{r}_6(u_6, v_6) &= \mathbf{p} + u_6 \mathbf{i}_x + v_6 \mathbf{i}_z + y \mathbf{i}_y, & u_6 \in [0, x], v_6 \in [0, z].
 \end{aligned}$$

Нормали поверхностей r_3 , r_4 и r_5 совпадают с нормальными соответствующих граней, а нормали поверхностей r_1 , r_2 и r_6 направлены против нормалей соответствующих граней. Эта информация содержится в грани в виде признака совпадения нормалей.

Прямоугольная призма имеет двенадцать ребер. Каждое ребро состоит из кривых пересечения поверхностей соседних граней и признака совпадения направления ребра с направлением кривой пересечения. Каждая кривая пересечения состоит из двух кривых на поверхности: одна на поверхности одной грани, другая на поверхности второй грани. Обе кривые на поверхности имеют одинаковую геометрическую и параметрическую длину и полностью совпадают в пространстве. Каждая кривая на поверхности представляет собой совокупность поверхности и двумерной кривой в пространстве параметров.

Приведем описание цикла первой грани, базирующейся на поверхности r_1 . Ребра для цикла построим на кривых пересечения поверхности r_1 с поверхностями r_3 , r_5 , r_6 , r_2 :

$$e_{13}(t) = \begin{Bmatrix} r_1(xt, 0) \\ r_3(xt, 0) \end{Bmatrix}, \quad e_{15}(t) = \begin{Bmatrix} r_1(x, yt) \\ r_5(yt, 0) \end{Bmatrix}, \quad e_{16}(t) = \begin{Bmatrix} r_1(xt, y) \\ r_6(xt, 0) \end{Bmatrix}, \quad e_{12}(t) = \begin{Bmatrix} r_1(0, yt) \\ r_2(yt, 0) \end{Bmatrix},$$

$$T \in [0, 1].$$

Пусть направление каждого ребра совпадает с направлением соответствующей кривой, что зафиксируем в признаках совпадения направлений ребер и кривых, на которых они базируются. Цикл первой грани опишем списком кривых, на которых базируются ребра, со знаками плюс или минус, соответствующими флагам ребер: $+e_{12}$, $+e_{16}$, $-e_{15}$, $-e_{13}$. Если ребро входит в цикл с положительным флагом, то при движении по циклу ребро проходится от начала к концу. Если ребро входит в цикл с отрицательным флагом, то при движении по циклу ребро проходится от конца к началу. Для цикла не имеет значения, с какого ребра начинается список. Список ребер циклически замкнут. Имеют значение последовательность ребер и совпадение направлений ребер с направлением цикла. Напомним, что при движении по циклу грань всегда лежит слева, если двигаться по внешней стороне грани.

Остальные грани и ребра призмы имеют аналогичное устройство. В углах призмы располагаются восемь вершин. Ориентация циклов граней прямоугольной призмы приведена на рис. 5.6.2. Линиями со стрелками показано движение вблизи циклов трех граней призмы.

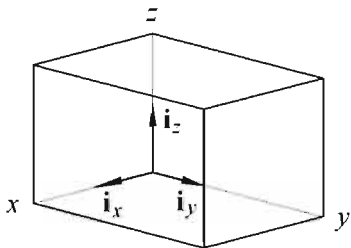


Рис. 5.6.1

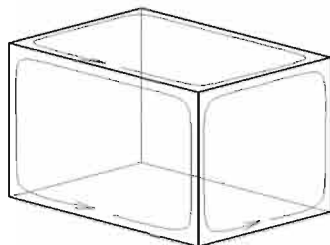


Рис. 5.6.2

Цилиндр имеет три грани, три ребра и две вершины. Начало местной системы координат \mathbf{p} поместим в центр одного из торцов цилиндра, а базисный вектор \mathbf{i}_z направим вдоль его оси. Пусть цилиндр имеет радиус r и длину h .

Пусть боковая грань тела базируется на цилиндрической поверхности:

$$\mathbf{r}_1(u_1, v_1) = \mathbf{p} + r \cos u_1 \mathbf{i}_x + r \sin u_1 \mathbf{i}_y + h v_1 \mathbf{i}_z, \\ u_1 \in [0, 2\pi], v_1 \in [0, 1].$$

Нормали поверхности \mathbf{r}_1 и ее грани совпадают по направлению. Пусть грани основания базируются на плоскостях

$$\mathbf{r}_2(u_2, v_2) = \mathbf{p} + u_2 \mathbf{i}_x + v_2 \mathbf{i}_y, \\ \mathbf{r}_3(u_3, v_3) = \mathbf{p} + u_3 \mathbf{i}_x + v_3 \mathbf{i}_y + h \mathbf{i}_z,$$

где параметры u_1, v_1 лежат внутри окружности $u_1(t) = r \cos t, v_1(t) = r \sin t, 0 \leq t \leq 2\pi$, а параметры u_2, v_2 лежат внутри окружности $u_2(t) = r \cos t, v_2(t) = r \sin t, 0 \leq t \leq 2\pi$. Нормали поверхности \mathbf{r}_2 и ее грани противоположны по направлению, а нормали поверхности \mathbf{r}_3 и ее грани совпадают. Поверхность \mathbf{r}_1 пересекается с поверхностями \mathbf{r}_2 и \mathbf{r}_3 по кривым

$$\mathbf{e}_{12}(t) = \begin{cases} \mathbf{r}_1(t, 0) \\ \mathbf{r}_2(r \cos t, r \sin t) \end{cases}, \quad \mathbf{e}_{13}(t) = \begin{cases} \mathbf{r}_1(t, h) \\ \mathbf{r}_3(r \cos t, r \sin t) \end{cases}, \quad t \in [0, 2\pi].$$

Цилиндрическая поверхность \mathbf{r}_1 боковой грани является циклически замкнутой по первому параметру. Ребро, построенное на линии замыкания оболочки, является швом. Шов, так же как и любое другое ребро, базируется на кривой пересечения, которая в данном случае имеет вид

$$\mathbf{e}_{11}(t) = \begin{cases} \mathbf{r}_1(0, ht) \\ \mathbf{r}_1(2\pi, ht) \end{cases}, \quad t \in [0, 1].$$

Пусть направление каждого ребра цилиндра совпадает с направлением кривой пересечения, на которой оно базируется. Цикл цилиндрической грани опишем списком кривых, на которых базируются ребра, со знаками соответствующих этим ребрам флагов: $+e_{12}$, $+e_{11}$, $-e_{13}$, $-e_{11}$. Цикл второй грани состоит из одного ребра на базе кривой e_{12} с отрицательным флагом. Цикл третьей грани состоит из одного ребра на базе кривой e_{13} с положительным флагом. Вершины располагаются в точках стыковки кривых e_{12} и e_{13} с кривой e_{11} .

Цилиндр приведен на рис. 5.6.3. Тонкой линией показан шов цилиндрической грани, а линиями со стрелками показано движение вблизи циклов двух граней тела.

Конус строится аналогично цилиндрическому телу. Приведем отличия конического тела от цилиндрического тела. Боковая грань базируется на конической поверхности:

$$\mathbf{r}_1(u_1, v_1) = \mathbf{p} + (r + hv_1 \operatorname{tg} \gamma) (\cos u_1 \mathbf{i}_x + \sin u_1 \mathbf{i}_y) + hv_1 \mathbf{i}_z, \\ u_1 \in [0, 2\pi], v_1 \in [0, 1].$$

Одна из граней основания базируется на поверхности

$$\mathbf{r}_3(u_3, v_3) = \mathbf{p} + u_3 \mathbf{i}_x + v_3 \mathbf{i}_y + h \mathbf{i}_z,$$

где параметры u_3 , v_3 лежат внутри окружности $u_3(t) = (r + ht \operatorname{tg} \gamma) \cos t$, $v_3(t) = (r + ht \operatorname{tg} \gamma) \sin t$, $0 \leq t \leq 2\pi$. Поверхность \mathbf{r}_1 пересекается с поверхностью \mathbf{r}_3 по кривой

$$\mathbf{e}_{13}(t) = \begin{cases} \mathbf{r}_1(t, h) \\ \mathbf{r}_3((r + ht \operatorname{tg} \gamma) \cos t, (r + ht \operatorname{tg} \gamma) \sin t) \end{cases}, \quad t \in [0, 2\pi].$$

В общем случае конус имеет три грани. Если $r > 0$ и $r + ht \operatorname{tg} \gamma = 0$ или $r = 0$ и $\gamma > 0$, то одна из торцевых граней вырождена в точку. Вырожденную в точку грань можно исключить из модели тела, тогда оболочка тела будет иметь полюс. Конус с полюсом в верхней вершине приведен на рис. 5.6.4. Тонкой линией показан шов конической грани, а линией со стрелкой показано движение вблизи цикла этой грани.

Шар имеет одну грань, три ребра и две вершины. Начало местной системы координат \mathbf{p} поместим в центр шара и опишем тело поверхностью

$$\mathbf{r}(u, v) = \mathbf{p} + r \cos v (\cos u \mathbf{i}_x + \sin u \mathbf{i}_y) + r \sin v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [-\pi/2, \pi/2].$$

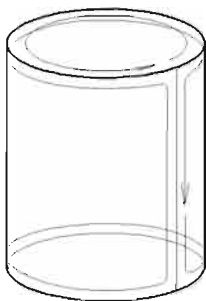


Рис. 5.6.3

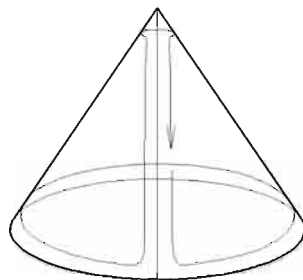


Рис. 5.6.4

Одно из ребер является швом и базируется на кривой

$$\mathbf{e}(t) = \begin{cases} \mathbf{r}(0, t) \\ \mathbf{r}(2\pi, t) \end{cases}, t \in [-\pi/2, \pi/2].$$

Два других ребра представляют полюса. Полюсные ребра базируются на кривых

$$\mathbf{e}_1(t) = \mathbf{r}\left(-\frac{\pi}{2}, t\right), \quad \mathbf{e}_2(t) = \mathbf{r}\left(\frac{\pi}{2}, t\right), t \in [0, 2\pi].$$

Полюсные ребра представляют собой точки, но двумерные кривые в пространстве параметров сферической поверхности в полюсах имеют ненулевую длину и представляют собой отрезки прямых.

Цикл грани шара опишем списком кривых, на которых базируются ребра, со знаками соответствующих этим ребрам флагов: $+e_1$, $+e$, $-e_2$, $-e$. Шов \mathbf{e} входит в список ребер цикла дважды, один раз с положительным флагом, второй — с отрицательным флагом.

Шар приведен на рис. 5.6.5. Тонкой линией показан шов сферической грани, а линией со стрелкой показано движение вблизи цикла этой грани.

Тор имеет одну грань, два ребра и одну вершину. Начало местной системы координат \mathbf{p} для тела в форме тора поместим в его центр, а базисный вектор \mathbf{i}_z направим по оси симметрии тела. Пусть больший радиус тора равен R , малый радиус тора равен r и $R > r$. Грань тора базируется на поверхности

$$\mathbf{r}(u, v) = \mathbf{p} + (R + r \cos v) (\cos u \mathbf{i}_x + \sin u \mathbf{i}_y) + r \sin v \mathbf{i}_z, \\ u \in [0, 2\pi], v \in [-\pi, \pi],$$

два ребра базируются на кривых

$$\mathbf{e}_1(t) = \begin{cases} \mathbf{r}(t, -\pi) \\ \mathbf{r}(t, \pi) \end{cases}, \quad t \in [0, 2\pi],$$

$$\mathbf{e}_2(t) = \begin{cases} \mathbf{r}(0, t) \\ \mathbf{r}(2\pi, t) \end{cases}, \quad t \in [-\pi, \pi].$$

Вершина располагается в точке пересечения ребер $\mathbf{p} + (R - r)\mathbf{i}_x$. Поверхность тора является циклически замкнутой по обоим параметрам, поэтому оба ребра оболочки тела замкнуты и являются швами. Нормаль поверхности и грани тора совпадают по направлению. Грань имеет один цикл. Цикл грани опишем списком кривых, на которых базируются ребра, со знаками соответствующих этим ребрам флагов: $+\mathbf{e}_1$, $+\mathbf{e}_2$, $-\mathbf{e}_1$, $-\mathbf{e}_2$. Тор, его швы и движение вблизи цикла грани показаны на рис. 5.6.6.

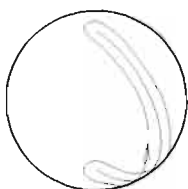


Рис. 5.6.5

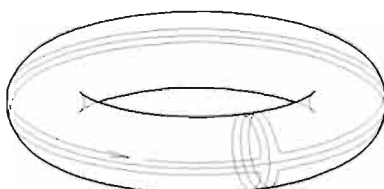


Рис. 5.6.6

Оболочки всех других тел построены по общим правилам и имеют такое же устройство, как и оболочки элементарных тел. Методы построения сложных оболочек тел рассмотрены в следующей главе.

5.7. Тела, полученные движением кривой

Построим тело путем движения составной кривой по заданной траектории. Пусть траектория движения описывается кривой $\mathbf{g}(v)$, которую будем называть *направляющей*. Движущуюся по траектории кривую будем называть *образующей* кривой. Направляющая кривая и образующая кривая не должны иметь точек самопересечения. Если образующая кривая не замкнута, то на ее основе в общем случае нельзя построить тело. Обычно из незамкнутой кривой создается замкнутая составная кривая путем «придания ей толщины» с помощью эквидистантных кривых. В общем случае образующая представляет собой замкнутую составную кривую. Если образующая является плоской кривой, то можно построить тело с плоскими торцами.

Пусть образующая кривая $\mathbf{c}(u)$ лежит на плоскости

$$\mathbf{r}(x, y) = \mathbf{p} + x \mathbf{i}_x + y \mathbf{i}_y \quad (5.7.1)$$

и представляет собой замкнутую составную кривую, состоящую из стыкующихся друг с другом k сегментов:

$$\mathbf{c}_i(u) = \mathbf{p} + x_i(u) \mathbf{i}_x + y_i(u) \mathbf{i}_y, \quad (5.7.2)$$

$$u_{\min} \leq u \leq u_{\max}, \quad i = 1, 2, \dots, k.$$

Пусть обход замкнутой образующей кривой происходит против часовой стрелки, если смотреть навстречу вектору производной направляющей кривой в ее начале. Если направляющая кривая $\mathbf{g}(v)$ не замкнута, то оболочка тела имеет торцевые грани. Торцевые грани тел движения представляют собой плоскости, область определения Ω которых описывается кривыми $x = x_i(u)$, $y = y_i(u)$. Если направляющая кривая замкнута, то оболочка тела не имеет торцевых граней. Боковые грани оболочки тела базируются на поверхностях движения сегментов образующей $\mathbf{c}_i(u)$ по направляющей $\mathbf{g}(v)$. Число боковых граней равно числу сегментов образующей кривой. В зависимости от типа траектории движения $\mathbf{g}(v)$ можно построить несколько типов тел.

Тело сдвига получим при плоскопараллельном движении образующей кривой вдоль монотонной направляющей кривой. Боковые грани тела сдвига описываются поверхностями (3.3.1), где образующими являются кривые (5.7.2). Первая торцевая грань тела будет описываться плоскостью (5.7.1) с областью определения Ω . Вторая торцевая грань будет базироваться на копии плоскости первой торцевой грани, перемещенной на вектор $\mathbf{d} = \mathbf{g}(v_{\max}) - \mathbf{g}(v_{\min})$. Чтобы тело сдвига не пересекало само себя, скалярное произведение нормали плоскости (5.7.1) и производной направляющей кривой не должно менять знак. Тело сдвига приведено на рис. 5.7.1.

Тело выдавливания мы получим, если направляющей кривой служит отрезок прямой $\mathbf{g}(v) = (1 - v)\mathbf{p}_1 + v\mathbf{p}_2$, $0 \leq v \leq 1$. Первая торцевая грань тела выдавливания будет описываться плоскостью (5.7.1) с областью определения Ω . Вторая торцевая грань будет базироваться на копии плоскости первой торцевой грани, перемещенной на вектор $\mathbf{d} = \mathbf{p}_2 - \mathbf{p}_1$. Пусть обход замкнутой образующей происходит против часовой стрелки, если смотреть навстречу вектору \mathbf{d} . Боковые грани тела выдавливания описываются поверхностями (3.3.2), где образующими являются кривые (5.7.2). Тело выдавливания, построенное по замкнутой составной кривой, приведено на рис. 5.7.2.



Рис. 5.7.1



Рис. 5.7.2

В некоторых случаях требуется построить тело выдавливания с уклоном боковых граней. В этом случае для плоской образующей кривой построим эквидистантную кривую и переместим ее на вектор \mathbf{d} . Если образующая кривая является составной, то эквидистантная кривая также должна быть составной и должна содержать такое же количество сегментов. В общем случае боковые грани тела выдавливания с уклоном являются линейчатыми поверхностями, каждая из которых построена на сегменте (5.7.2) и соответствующем сегменте эквидистантной кривой, перемещенном на вектор \mathbf{d} . Тангенс угла уклона боковых граней равен отношению эквидистанты к длине выдавливания.

Тело вращения мы получим, если направляющей кривой движения служит окружность или ее дуга. Пусть ось вращения проходит через точку \mathbf{p} , а ее направление характеризуется единичным вектором $\mathbf{i} = (i_1, i_2, i_3)$. Боковые грани тела вращения описываются поверхностями (3.3.3), где образующими являются кривые (5.7.2).

Ось вращения не должна пересекать боковые грани. Чтобы тело не получилось вывернутым наизнанку, направления оси вращения и обхода образующей кривой должны быть согласованы. Определим положительное направление оси вращения для замкнутой образующей кривой. Пусть мы находимся с той стороны плоскости (5.7.1), откуда движение вдоль образующей выглядит против часовой стрелки. За положительное направление оси вращения примем направление вектора \mathbf{i} , при взгляде вдоль которого образующая кривая находится слева. Если угол вращения равен 2π , то оболочка тела имеет топологию тора, в противном случае – топологию призмы. В первом случае оболочка не имеет торцевых граней. Тело вращения с топологией призмы приведено на рис. 5.7.3.

По незамкнутой составной плоской кривой можно построить тело вращения, приведенное на рис. 5.7.4. Из концов образующей на ось вращения опущены перпендикуляры, по которым построены соответствующие грани. В отличие от предыдущего примера тело, приведенное

на рис. 5.7.4, касается оси вращения. Такие тела будем называть *сфероидами*.



Рис. 5.7.3



Рис. 5.7.4

Кинематическое тело представляет собой обобщение тел выдавливания и вращения для произвольной направляющей. Кинематическое тело получим путем движения плоской замкнутой образующей кривой вдоль произвольной направляющей кривой при сохранении угла между плоскостью образующей и касательным вектором $g'(v)$ направляющей в текущей точке. Боковые грани кинематического тела описываются поверхностями (3.3.7), где образующими являются кривые (5.7.2). На рис. 5.7.5 приведено кинематическое тело.



Рис. 5.7.5

Если направляющая кривая замкнута, то получим кинематическое тело с топологией тора. Если направляющая кривая не замкнута, то получим кинематическое тело с топологией призмы. Боковые грани кинематического тела не должны пересекать сами себя. На рис. 5.7.6 приведено кинематическое тело, построенное по циклически замкнутой направляющей кривой.



Рис. 5.7.6

Опишем **цикл боковой грани** тел движения, базирующейся на поверхности $r_i(u, v)$, $u_{\min} \leq u \leq u_{\max}$, $v_{\min} \leq v \leq v_{\max}$. Для этого построим кривые пересечения поверхности $r_i(u, v)$ с соседними боковыми поверхностями

$$e_{i-1}(t) = \begin{cases} r_{i-1}(u_{i-1, \max}, t) \\ r_i(u_{\min}, t) \end{cases}, \quad e_i(t) = \begin{cases} r_i(u_{\max}, t) \\ r_{i+1}(u_{i+1, \min}, t) \end{cases}, \quad t \in [v_{\min}, v_{\max}].$$

При незамкнутой направляющей построим кривую пересечения поверхности $r_i(u, v)$ с плоскостью первого торца

$$e_{1i}(u) = \begin{cases} c_i(u) \\ r_i(u, v_{\min}) \end{cases}, \quad u \in [u_{\min}, u_{\max}]$$

и кривую пересечения поверхности $r_i(u, v)$ с плоскостью второго торца

$$e_{2i}(u) = \begin{cases} a_i(u) \\ r_i(u, v_{\max}) \end{cases}, \quad u \in [u_{\min}, u_{\max}],$$

где $a_i(u)$ граничный сегмент второго торца, соответствующий сегменту $c_i(u)$. Пусть направление каждого ребра совпадает с направлением кривой, на которой оно базируется. Ребра цикла i -й грани тела движения приведены на рис. 5.7.7.

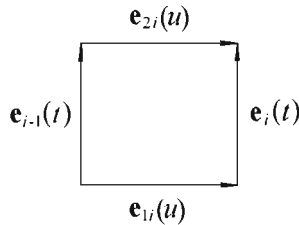


Рис. 5.7.7

Боковая грань тела движения имеет один цикл, который опишем списком ребер со знаками соответствующих этим ребрам флагов: $+e_1, +e_2, -e_3, -e_{k-1}$. При замкнутой направляющей вместо кривых e_1 и e_2 построим кривую шва:

$$e_{0_i}(u) = \begin{cases} r_i(u, v_{\min}) \\ r_i(u, v_{\max}) \end{cases}, \quad u \in [u_{\min}, u_{\max}].$$

Нормали боковой грани и поверхности совпадают.

Если направляющая кривая не замкнута, то тело имеет две торцевые грани. Цикл первой торцевой грани состоит из k ребер на базе кривых $e_i, i = k, k-1, \dots, 1$ с отрицательным флагом. Нормали первой торцевой грани и поверхности противоположно направлены. Цикл второй торцевой грани состоит из k ребер на базе кривых $e_i, i = 1, 2, \dots, k$ с положительным флагом. Нормали второй торцевой грани и поверхности совпадают.

5.8. Тела, построенные по сечениям

Построим тело по плоским сечениям, заданным семейством кривых $r_j(u), j = 0, 1, \dots, n$. Если кривые $r_j(u)$ являются составными, то они должны состоять из одинакового числа сегментов, тогда в каждой боковой грани оболочки будет использоваться по одному сегменту (5.7.2) от каждой кривой. Кривые, по которым строятся боковые грани, должны иметь одинаковую параметрическую длину. Если все кривые семейства являются замкнутыми, то по ним можно построить тело. Тело может иметь две торцевые грани, построенные на крайних кривых. Торцевые грани тела будут представлять собой плоскости, ограниченные этими кривыми. Боковыми гранями оболочки могут служить поверхности Лагранжа (3.4.2), поверхности Эрмита (3.4.3).

Сориентируем кривые семейства так, чтобы их обход происходил против часовой стрелки, если смотреть на кривую $e_j(u)$ из центра тяжести кривой $r_{j+1}(u)$. Каждой кривой $r_j(u)$ поставим в соответствие значение второго параметра $v = v_j, j = 0, 1, \dots, n$ поверхности боковой грани. Для поверхностей, проходящих через кривые семейства, при параметре v_i поверхность должна совпадать с кривой $r_j(u)$. Второй параметр поверхности v между кривыми $r_j(u)$ и $r_m(u)$ изменяется от v_j до v_m . Заметим, что форма поверхности зависит от значений параметров $v_j, j = 0, 1, \dots, n$. Корректная поверхность получается, когда разности значений параметров $v_j - v_{j-1}$ пропорциональны среднему расстоянию $|r_j(u) - r_{j-1}(u)|$ между кривыми.

На рис. 5.8.1 приведена поверхность Эрмита, построенная по кривым $\mathbf{r}_j(u)$, $j = 0, 1, \dots, n$. Тело, построенное по плоским сечениям, приведено на рис. 5.8.2.

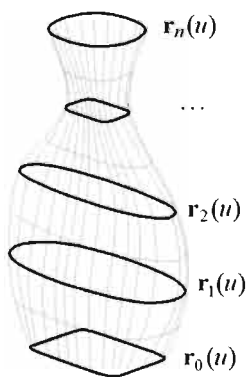


Рис. 5.8.1



Рис. 5.8.2

В зависимости от расположения плоских сечений тело может иметь топологию призмы или топологию тора. У тела с топологией призмы есть две торцевые грани. Циклы граней тела, построенного по сечениям, строятся аналогично циклам граней тел движения.

Если кроме плоских сечений нам дана направляющая кривая, то можно **построить тело по сечениям с направляющей**. Боковые грани такого тела базируются на поверхностях (3.4.9). На рис. 5.8.3 приведена поверхность (3.4.9), построенная по кривым $\mathbf{r}_0(u)$, $\mathbf{r}_1(u)$ и направляющей кривой $\mathbf{g}(v)$. Тело, построенное по плоским сечениям с направляющей кривой, приведено на рис. 5.8.4.

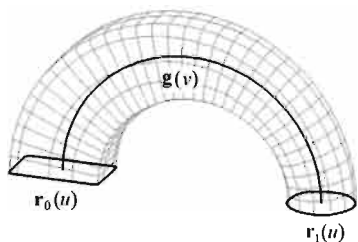


Рис. 5.8.3



Рис. 5.8.4

Все остальные параметры тела, построенного по плоским сечениям с направляющей, аналогичны параметрам тела, построенного по плоским сечениям без направляющей.

5.9. Тело, построенное по поверхности

На основе произвольной поверхности можно построить тело в форме листа конечной толщины. Пусть дана поверхность $\mathbf{b}(u, v)$, $u, v \in \Omega$. Выбрав толщину листа h , построим эквидистантную к ней поверхность:

$$\mathbf{r}(u, v) = \mathbf{b}(u, v) + h \mathbf{m}(u, v), \quad u, v \in \Omega,$$

где $\mathbf{m}(u, v)$ — нормаль поверхности $\mathbf{b}(u, v)$. На поверхностях $\mathbf{b}(u, v)$ и $\mathbf{r}(u, v)$ построим две основные грани листового тела. Грани, соединяющие эти основные грани, будем называть боковыми. Если область Ω определения параметров поверхности $\mathbf{b}(u, v)$ имеет прямоугольную форму, то листовое тело будет иметь четыре боковые грани. В общем случае листовое тело будет иметь столько боковых граней, сколько кривых содержит составная кривая, описывающая область определения параметров поверхности $\mathbf{b}(u, v)$. Пусть область Ω ограничена составной кривой, содержащей k сегментов:

$$\mathbf{c}_i(t) = [u_i(t) \ v_i(t)]^T, \quad t_{\min} \leq t \leq t_{\max}, \quad i = 1, 2, \dots, k.$$

Каждому сегменту соответствуют две пространственные кривые: $\mathbf{b}(u_i(t), v_i(t))$ и $\mathbf{r}(u_i(t), v_i(t))$, на которых можно построить боковую поверхность. В общем случае i -я боковая грань будет базироваться на линейчатой поверхности:

$$\mathbf{s}_i(t, w) = (1 - w) \mathbf{b}(u_i(t), v_i(t)) + w \mathbf{r}(u_i(t), v_i(t)), \\ t \in [t_{\min}, t_{\max}], \quad w \in [0, 1].$$

При $h > 0$ направление нормалей боковых граней совпадает с направлением нормалей их поверхностей. Направление нормали грани на основе эквидистантной поверхности при $h > 0$ совпадает с направлением нормали поверхности. Направление нормали грани на основе поверхности $\mathbf{b}(u, v)$ при $h > 0$ противоположно направлению нормали поверхности. При $h < 0$ направление нормалей граней изменится на противоположное направление.

Пусть $h > 0$. Тогда цикл первой основной грани состоит из k ребер на базе кривых

$$\mathbf{e}_{i,}(t) = \begin{cases} \mathbf{b}(u_i(t), v_i(t)) \\ \mathbf{s}_i(t, 0) \end{cases}, \quad t \in [t_{\min}, t_{\max}], \quad i = k, k - 1, \dots, 1$$

с отрицательным флагом. Нормали первой основной грани и поверхности $\mathbf{b}(u, v)$ противоположно направлены. Цикл второй основной грани состоит из k ребер на базе кривых

$$\mathbf{e}_{2i}(t) = \begin{cases} \mathbf{r}(u_i(t), v_i(t)) \\ \mathbf{s}_i(t, 1) \end{cases}, t \in [t_{\min}, t_{\max}], i = 1, 2, \dots, k$$

с положительным флагом. Нормали второй основной грани и поверхности $\mathbf{r}(u, v)$ совпадают. Для описания цикла боковой грани тела, базирующейся на поверхности $\mathbf{s}_i(t, w)$, построим кривые пересечения этой поверхности с соседними боковыми поверхностями:

$$\mathbf{e}_{i-1}(w) = \begin{cases} \mathbf{s}_{i-1}(t_{i-1, \max}, w) \\ \mathbf{s}_i(t_{i, \min}, w) \end{cases}, \quad \mathbf{e}_i(w) = \begin{cases} \mathbf{s}_i(t_{i, \max}, w) \\ \mathbf{s}_{i+1}(t_{i+1, \min}, w) \end{cases}, w \in [0, 1].$$

Пусть направление каждого ребра совпадает с направлением кривой, на которой оно базируется. Каждая боковая грань тела имеет один цикл, который опишем списком ребер со знаками соответствующих этим ребрам флагов: $+\mathbf{e}_{1i}$, $+\mathbf{e}_i$, $-\mathbf{e}_{2i}$, $-\mathbf{e}_{i-1}$. Нормали боковой грани и ее поверхности совпадают.

Если поверхность $\mathbf{b}(u, v)$ является замкнутой по одному из параметров, то листовое тело будет иметь форму трубы и топологию тора. Если поверхность замкнута по обоим параметрическим направлениям, то получим тело с пустотой внутри. Такое тело при $h > 0$ будет иметь внешнюю оболочку на основе поверхности $\mathbf{r}(u, v)$ и внутреннюю оболочку на основе поверхности $\mathbf{b}(u, v)$; при $h < 0$ внешняя оболочка будет базироваться на поверхности $\mathbf{b}(u, v)$, а внутренняя оболочка будет базироваться на поверхности $\mathbf{r}(u, v)$. На рис. 5.9.1 приведена сплайновая поверхность с границей в виде замкнутой составной кривой. Тело в форме листа, построенное по упомянутой поверхности, приведено на рис. 5.9.2.

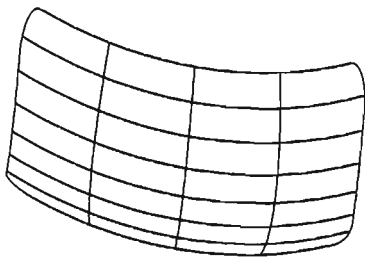


Рис. 5.9.1



Рис. 5.9.2

С помощью тел в форме листа можно моделировать детали кузова автомобиля и планера самолета. Для построения тела достаточно знать базовую поверхность $\mathbf{b}(u, v)$ и толщину тела h .

Упражнения

1. Какие представления могут использоваться для геометрической модели?
2. Постройте тело в форме тетраэдра.
3. Постройте тело в форме разрезанного вдоль оси цилиндра.
4. Постройте тело в форме разрезанного вдоль оси тора.

Глава 6. Методы построения тел

Процесс построения однородных оболочек в геометрическом моделировании похож на процесс изготовления моделируемого объекта. Сначала создаются оболочки простой формы, а далее выполняется набор действий, позволяющих из оболочек простой формы получить более сложные оболочки. При необходимости создаются вспомогательные объекты. Редактировать модель и создавать подобные модели можно путем изменения параметров модели с последующим повторением процесса построения оболочек или путем непосредственной модификации элементов уже построенных оболочек.

6.1. Методы геометрического моделирования

Телу сопоставлено множество расположенных с внутренней стороны оболочек точек и точек самих оболочек, над которыми можно выполнять различные действия, в том числе теоретико-множественные операции. Незамкнутой оболочке сопоставлено только множество точек самой оболочки, действия над которыми носят иной характер. Незамкнутые оболочки используются там, где достаточно воспроизвести лишь часть моделируемого объекта.

Совокупность действий над телами или оболочками, которая приводит к образованию тела или оболочки иной формы, будем называть *операцией*.

Особенности последовательности и методов построения модели в геометрическом моделировании привели к появлению терминов «поверхностное моделирование» (моделирование с помощью поверхностей) и «твердотельное моделирование» (моделирование с помощью твердых тел). В обоих случаях результатом моделирования является некоторая совокупность оболочек, описывающая поверхность моделируемого объекта, но процесс и методы моделирования в первом случае отличаются от процесса и методов моделирования во втором случае.

В поверхностном моделировании создают и модифицируют требуемым образом поверхности, описывающие отдельные элементы моделируемого объекта. Эти поверхности обрезают, деформируют, выглаживают, сопрягают друг с другом. Затем из полученных поверхностей может быть собрана оболочка путем образования из поверхностей граней и их сшивки. В поверхностном моделировании результирующая оболочка не обязательно должна быть замкнутой. Она может отражать лишь часть

моделируемого объекта. Поверхностное моделирование позволяет сосредоточить усилия на сложных формах моделируемого объекта.

В твердотельном моделировании с самого начала работа идет с оболочками тел, а не с отдельными поверхностями. Поверхности в оболочке представлены в виде граней, сохраняющих в процессе моделирования исходную вполне определенную связь друг с другом. Оболочки полностью описывают форму моделируемых объектов. Процесс построения тела начинается с создания тела простой формы. Далее тело изменяется необходимым образом. С помощью одних действий к телу можно добавить или удалить некоторый объем. С помощью других действий можно деформировать тело.

В геометрическом моделировании также применяется процесс построения оболочки, похожий на лепку. Для этого может использоваться модель любой стадии готовности: заготовка простой формы или почти готовая модель. К модели применяются методы прямого редактирования, позволяющие деформировать оболочку требуемым образом, перемещать отдельные элементы оболочки друг относительно друга, удалять или копировать отдельные элементы оболочки. Прямое редактирование модели называют также синхронным моделированием.

Построение оболочки или тела начинается с построения вспомогательных элементов, таких как местные системы координат, контрольные точки, отрезки прямых и прочие кривые. Далее с помощью кривых и других вспомогательных элементов строятся тела. Рассматриваемые ниже методы построения тел выполняют одновременно формирование границ граней и их стыковку друг с другом.

6.2. Булевы операции над телами

Булевыми операциями в геометрическом моделировании называются операции объединения, пересечения и вычитания множеств точек двух тел.

В порядке следования тел-операндов будем называть их первым телом и вторым телом. Результатом операции объединения двух тел является тело, которое содержит точки, принадлежащие или первому, или второму телу. Результатом операции пересечения двух тел является тело, которое содержит точки, принадлежащие и первому, и второму телу. Результатом операции вычитания двух тел является тело, которое содержит точки, принадлежащие первому телу и не располагающиеся внутри второго тела.

На рис. 6.2.1 приведены исходные тела до булевой операции. На рис. 6.2.2 приведен результат операции объединения тел, на рис. 6.2.3

приведен результат операции пересечения, на рис. 6.2.4 и рис. 6.2.5 приведены результаты операции вычитания тел.

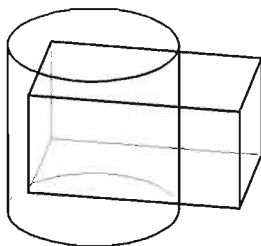


Рис. 6.2.1

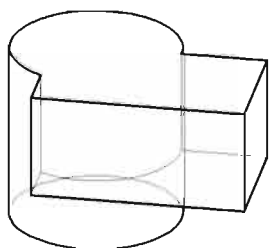


Рис. 6.2.2

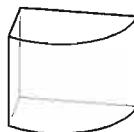


Рис. 6.2.3

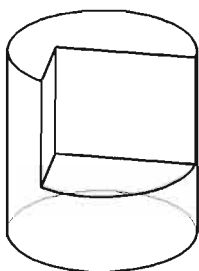


Рис. 6.2.4

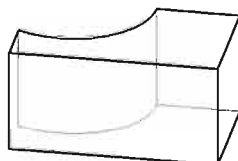


Рис. 6.2.5

Операцию вычитания тел можно свести к операции пересечения тел. Для этого нужно вывернуть второе тело наизнанку и найти точки его объема, одновременно принадлежащие и объему первого тела. При выворачивании тела наизнанку изменяются на противоположные признаки совпадения ориентации нормалей граней и их поверхностей, изме-

няются на противоположные направления циклов, в результате чего точками тела становятся те точки пространства, которые до этого находились снаружи тела. Все булевы операции содержат много общего и выполняются по единому алгоритму.

Операция объединения тел кратко может быть охарактеризована следующим образом: нужно найти ребра пересечения граней тел, разрезать ими пересекающиеся грани, взять части граней первого тела, которые не попали внутрь второго тела, и части второго тела, которые не попали внутрь первого тела, и построить из этих частей новое тело.

Перед началом операции найдем пары *подобных граней* в объединяемых телах. Каждая пара таких граней расположена на одинаковых или подобных поверхностях, но может иметь различные области определения и ориентации нормалей. Среди подобных граней выделим пары граней, имеющие одинаковые по направлению нормали в совпадающих точках. Такие грани будем называть *перекрывающимися гранями*. Примером перекрывающихся граней служат верхние грани тел, приведенных на рис. 6.2.1. Подобные грани не будем пересекать. Перекрывающиеся грани будем объединять.

Операцию начнем с построения кривых пересечения граней первого тела с гранями второго тела. На базе кривых пересечения граней первого и второго тел построим новые ребра. Новым ребрам дадим направление векторного произведения нормали грани первого тела с нормалью грани второго тела: $\mathbf{t}_e = \mathbf{m}_1 \times \mathbf{m}_2$. Вектор нормали грани первого тела \mathbf{m}_1 и вектор нормали грани второго тела \mathbf{m}_2 вычислены в точке пересечения граней. Направление ребра определяется знаком совпадения с направлением кривой пересечения. На рис. 6.2.6 приведены новое ребро пересечения грани первого тела с гранью второго тела.

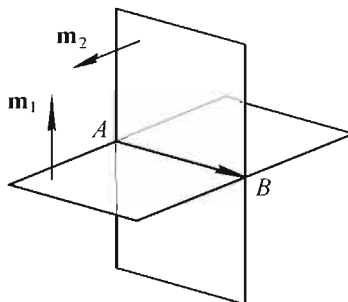


Рис. 6.2.6

Новые ребра построим так, чтобы они полностью лежали внутри циклов граней исходных тел. Новые ребра могут подходить к границам

грани только своими вершинами. Далее разрежем исходные ребра тела, к которым подходят новые ребра, так как в результате булевой операции войдет только часть исходной грани. В точках A и B , рис. 6.2.6, исходные ребра граней должны быть разрезаны на два ребра. Резка исходного ребра осуществляется путем рассечения кривой, на которой базируется ребро. Так как каждое ребро исходных тел входит в циклы двух смежных граней, то после резки исходных ребер тел необходимо произвести корректировку циклов с учетом разрезанных ребер.

Для того чтобы каждую из показанных на рис. 6.2.6 граней разрезать на части, нужно перестроить ее циклы и в соответствии с циклами изменить границу, описывающую область определения параметров поверхности грани. На рис. 6.2.7 показаны две пересекшиеся грани (тонкими линиями со стрелками показано направление циклов граней исходных тел) и ребро пересечения. На рис. 6.2.8 показаны те части граней, которые войдут в объединение тел (тонкими линиями со стрелками показано направление циклов граней результата булевой операции).

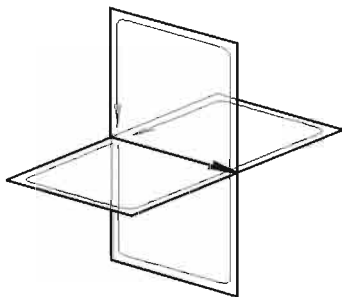


Рис. 6.2.7

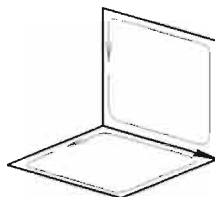


Рис. 6.2.8

На рис. 6.2.8 видно, с какими флагами должны войти ребра пересечения в циклы граней первого и второго тел.

Лемма 1. *В булевой операции объединения двух тел новые ребра, направленные вдоль $m_1 \times m_2$, войдут в циклы граней первого тела с отрицательным флагом, а в циклы граней второго тела — с положительным флагом.*

В самом деле, результат операции должен содержать часть грани первого тела, лежащую вне второго тела, и часть грани второго тела, лежащую вне первого тела. Пусть мы движемся по некоторому новому ребру, находясь с внешней стороны граней тел. Часть грани первого тела, лежащая вне второго тела, будет располагаться для нас справа от нового ребра, следовательно, новое ребро войдет в цикл грани первого тела с отрицательным флагом. Часть грани второго тела, лежащая вне первого

тела, будет располагаться для нас слева от нового ребра, следовательно, новое ребро войдет в цикл грани второго тела с положительным флагом.

Исходные ребра тел, которые сохраняются в результате булевой операции, войдут в перестроенные циклы, сохранив свои флаги.

После пересечения граней тел получим совокупность новых ребер, ориентированных определенным образом, стыкующихся друг с другом и с ребрами исходных тел в вершинах. Пример построения новых ребер приведен на рис. 6.2.9. Не все новые ребра пересекшихся граней войдут в результат операции, некоторые новые ребра должны быть отброшены (или не должны строиться).

В операции объединения тел потребуются только те новые ребра, смежные грани которых имеют продолжение вне одного из тел.

Пример определения нужных новых ребер приведен на рис 6.2.10.

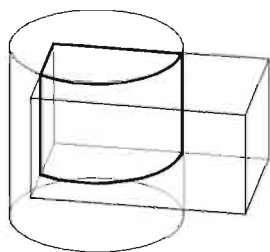


Рис. 6.2.9

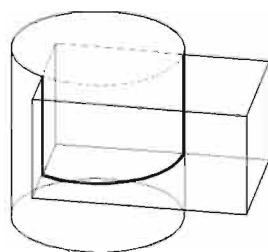


Рис. 6.2.10

Далее перестроим циклы перекрывающихся граней. Грани могут перекрываться полностью или частично. Нормали \mathbf{m}_1 и \mathbf{m}_2 перекрывающихся граней одинаково направлены в общих точках. Перекрывающиеся являются грани, которые можно перенести на одну общую поверхность. Таких граней может быть больше двух. Перед сборкой циклов перекрывающиеся грани и их ребра должны быть приведены к одному носителю — к поверхности одной из перекрывающихся граней. В операции объединения тел исходные ребра перекрывающихся граней первого тела, лежащие внутри второго тела, и исходные ребра перекрывающихся граней второго тела, лежащие внутри первого тела, не войдут в оболочку, полученную в результате булевой операции.

Построение некоторого цикла перекрывающейся грани начнем с ребра, для которого известно, что оно должно войти в результат операции. Для продолжения цикла среди новых ребер и среди исходных ребер грани найдем все ребра, стыкующиеся с данным ребром в соответствующей вершине. Среди найденных ребер выберем то ребро, которое образует наибольший угол по сравнению с углами других стыкующихся ребер, если смотреть с внешней стороны грани и отсчитывать угол по

часовой стрелке. Выбранное ребро поставим в список цикла. Цикл будет построен, когда очередное ребро состыкуется с начальным ребром. Вновь построенные циклы перекрывающихся граней рассортируем по группам, каждая из которых состоит из внешнего цикла и входящих в него внутренних циклов. Пример перестроения перекрывающихся граней приведен на рис. 6.2.11.

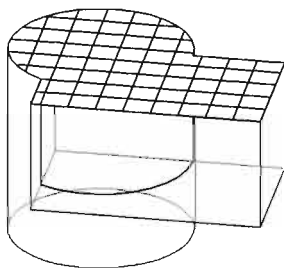


Рис. 6.2.11

Далее перестроим циклы пересекшихся граней. Построение некоторого цикла пересекшейся грани начнем с нового ребра. Для продолжения цикла среди новых ребер и среди исходных ребер грани найдем все ребра, стыкующиеся с данным ребром в соответствующей вершине. Среди найденных ребер выберем то ребро, которое образует наибольший угол по сравнению с углами других стыкующихся ребер, если смотреть с внешней стороны грани и отсчитывать угол *против часовой стрелки*. Например, с новым ребром BA в точке A стыкуются несколько ребер, рис. 6.2.12. Для продолжения цикла, начатого с ребра BA , следует выбрать ребро AC .

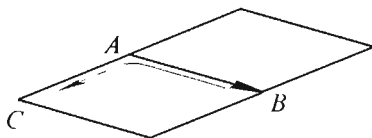


Рис. 6.2.12

Выбранное ребро поставим в список цикла. Цикл будет построен, когда очередное ребро состыкуется с начальным ребром. Циклы грани перестраиваем до тех пор, пока не используем все ее новые ребра. Таким способом построим в общем случае несколько новых циклов рассматриваемой грани первого тела.

Вновь построенные циклы перестраиваемой грани рассортируем по группам, каждая из которых состоит из внешнего цикла и входящих в него внутренних циклов. У рассматриваемой грани могут остаться нетронутыми один или несколько исходных циклов. Нетронутыми будем называть исходные циклы грани, ни одно ребро которых не вошло в перестроенные циклы. Среди исходных нетронутых циклов отберем те, которые необходимо включить в состав описания перестроенной грани. Ими являются исходные внутренние циклы грани, лежащие внутри новых внешних циклов. Если для некоторых новых внутренних циклов не найден новый внешний цикл и они лежат внутри исходного внешнего цикла, то следует включить в состав перестраиваемой грани исходный внешний нетронутый цикл грани.

Если в результате сортировки внешних циклов получилось больше одного, то это означает, что из исходной грани в результате операции образовалось несколько граней. Результат перестроения пересекшихся граней приведен на рис. 6.2.13.

Описанное перестроение циклов выполняется для каждой пересекшейся грани первого и второго тел. Новые ребра должны войти в перестроенные циклы граней первого тела с отрицательным флагом, в циклы граней второго тела новые ребра должны войти с положительным флагом.

Мы перестроили перекрывающиеся и пересекшиеся грани объединяемых тел. Все их включим в оболочку результата булевой операции. Остается к этим граням в оболочку тела добавить не пересекшиеся в операции грани, которые связаны с пересекшимися гранями. Для этого просмотрим ребра, входящие в оболочку результата операции, и включим в оболочку смежные грани ребер, если они там отсутствуют. Продолжив эти действия для ребер всех добавленных граней, получим оболочку результата булевой операции, рис. 6.2.14.

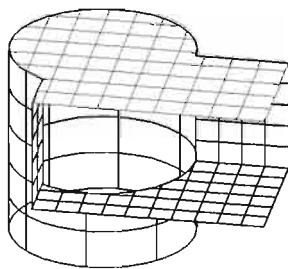


Рис. 6.2.13

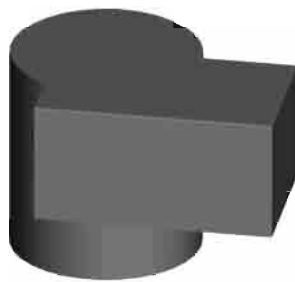


Рис. 6.2.14

Операция пересечения тел кратко может быть охарактеризована следующим образом: нужно найти ребра пересечения граней тел, разрезать ими пересекающиеся грани, взять части граней первого тела, которые попали внутрь второго тела, и части граней второго тела, которые попали внутрь первого тела, и построить из этих частей новое тело.

Операция пересечения тел имеет много общего с операцией объединения тел. Перед началом операции найдем пары подобных граней и среди них найдем пары перекрывающихся граней в пересекаемых телах. Так же как и в операции объединения тел, в операции пересечения тел выполнима построение новых ребер и резка ими исходных ребер тел. Новым ребрам так же дадим направление векторного произведения нормали грани первого тела с нормалью грани второго тела: $\mathbf{t}_c = \mathbf{m}_1 \times \mathbf{m}_2$.

На рис. 6.2.7 приведены пересекающиеся грани тел, которые режут друг друга. В пересечение тел войдет часть грани первого тела, лежащая внутри второго тела, и часть грани второго тела, лежащая внутри первого тела, рис. 6.2.15.



Рис. 6.2.15

На рис. 6.2.15 видно, с какими флагами должны войти ребра пересечения в циклы граней первого и второго тел.

Лемма 2. *В булевой операции пересечения двух тел новые ребра, направленные вдоль $\mathbf{m}_1 \times \mathbf{m}_2$, войдут в циклы граней первого тела с положительным флагом, а в циклы граней второго тела — с отрицательным флагом.*

В самом деле, результат операции должен содержать часть грани первого тела, лежащую внутри второго тела, и часть грани второго тела, лежащую внутри первого тела. Пусть мы движемся по некоторому новому ребру, находясь с внешней стороны граней тел. Часть грани первого тела, лежащая внутри второго тела, будет располагаться для нас слева от нового ребра, следовательно, новое ребро войдет в цикл грани первого тела с положительным флагом. Часть грани второго тела, лежащая внутри первого тела, будет располагаться для нас справа от нового ребра, следовательно, новое ребро войдет в цикл грани второго тела с отрицательным флагом.

Исходные ребра тел, которые сохраняются в результате булевой операции, войдут в перестроенные циклы, сохранив свои флаги.

Как и в операции объединения тел, в операции пересечения тел не все новые ребра пересекшихся граней войдут в результат операции, некоторые новые ребра должны быть отброшены (или не должны строиться).

В операции пересечения тел потребуются только те новые ребра, смежные грани которых имеют продолжение внутрь одного из тел.

Перед сборкой циклов перекрывающиеся грани и их ребра приведем к одному общему носителю — к поверхности одной из перекрывающихся граней. В операции пересечения тел исходные ребра перекрывающихся граней первого тела, лежащие вне второго тела, и исходные ребра перекрывающихся граней второго тела, лежащие вне первого тела, не войдут в оболочку, полученную в результате булевой операции.

Перестроение циклов граней в операции пересечения тел выполним аналогично перестроению циклов в операции объединения тел.

Построение некоторого цикла перекрывающейся грани начнем с ребра, для которого известно, что оно должно войти в результат операции. Для продолжения цикла среди новых ребер и среди исходных ребер грани найдем все ребра, стыкующиеся с данным ребром в соответствующей вершине. Среди найденных ребер выберем то ребро, которое образует наибольший угол по сравнению с углами других стыкующихся ребер, если смотреть с внешней стороны грани и отсчитывать угол *по часовой стрелке*. Выбранное ребро поставим в список цикла. Цикл будет построен, когда очередное ребро состыкуется с начальным ребром.

Построение некоторого цикла пересекшейся грани начнем с нового ребра. Для продолжения цикла среди новых ребер и среди исходных ребер грани найдем все ребра, стыкующиеся с данным ребром в соответствующей вершине. Среди найденных ребер выберем то ребро, которое образует наибольший угол по сравнению с углами других стыкующихся ребер, если смотреть с внешней стороны грани и отсчитывать угол *против часовой стрелки*. Выбранное ребро поставим в список цикла. Построение цикла закончим, когда очередное ребро состыкуется с начальным ребром.

Вновь построенные и нетронутые исходные циклы каждой перестраиваемой грани рассортируем по группам, каждая из которых состоит из внешнего цикла и входящих в него внутренних циклов, и с помощью этих групп перестроим грани. Если в результате сортировки внешних циклов получилось больше одного, то это означает, что из исходной грани в результате операции образовалось несколько граней.

Описанное перестроение циклов выполняется для каждой пересекшейся грани первого и второго тел. В операции пересечения тел новые ребра должны войти в перестроенные циклы граней первого тела с положительным флагом, в циклы граней второго тела новые ребра должны войти с отрицательным флагом.

Перестроенные перекрывающиеся и пересекающиеся грани пересекаемых тел включим в оболочку результата булевой операции. К этим граням в оболочку нового тела добавим связанные с ними грани.

Операция вычитания тел сводится к булевой операции пересечения первого тела и вывернутого наизнанку второго тела. Вывернутое наизнанку тело получим из исходного тела путем переориентации направлений нормалей граней и направлений циклов граней. Переориентация направления нормали грани производится изменением признака совпадения нормали поверхности и нормали ее грани. Переориентация направления цикла грани производится перестроением списка ребер (изменением на обратный порядок следования ребер в списке) и изменением значений флагов ребер в списках.

6.3. Алгоритм булевых операций

Все булевы операции подчинены общему алгоритму. Алгоритм разобьем на пять шагов и рассмотрим эти шаги на примере операции вычитания тел, приведенных на рис. 6.3.1. Перед операцией вычитания вывернем наизнанку вычитаемое тело.

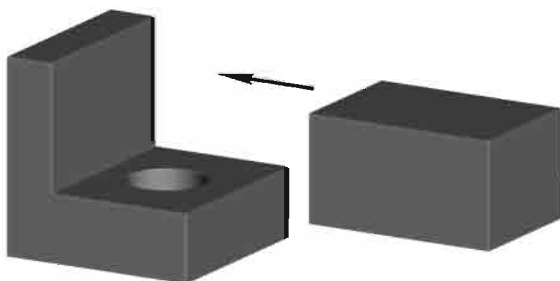


Рис. 6.3.1

Шаг первый – построение новых ребер, по которым пересекаются грани тел-операндов, рис. 6.3.2. Шаг второй – отбор среди новых ребер действительно необходимых для выполняемой операции, рис. 6.3.3. В операции объединения этих же тел набор действительно необходимых новых ребер будет другой.

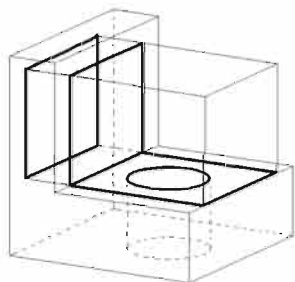


Рис. 6.3.2

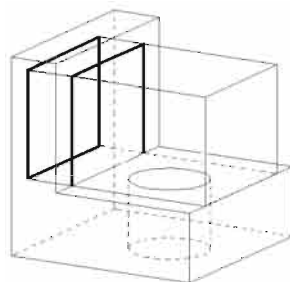


Рис. 6.3.3

Шаг третий – перестроение перекрывающихся граней, рис. 6.3.4.
Шаг четвертый – перестроение пересекающихся граней, рис. 6.3.5.

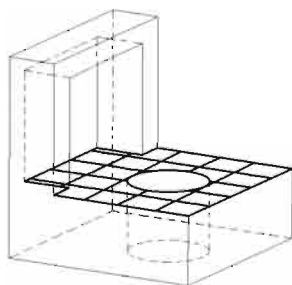


Рис. 6.3.4

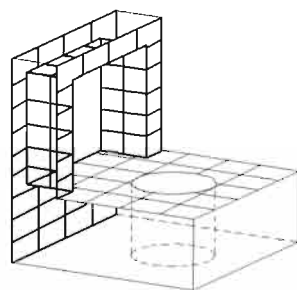


Рис. 6.3.5

Шаг пятый – формирование оболочки нового тела из перестроенных граней и связанных с ними исходных граней тел-операндов, рис. 6.3.6.



Рис. 6.3.6

Рассмотрим некоторые особенности перечисленных действий.

Наиболее трудоемким и требующим определенной точности действием в процессе выполнения булевых операций является построение новых ребер пересекающихся граней. Новые ребра должны стыковаться и друг с другом, и с исходными ребрами граней в вершинах. Новые ребра не должны пересекать друг друга вне крайних точек. На рис. 6.3.7 показаны два цилиндрических тела одинакового диаметра, оси которых пересекаются. В булевых операциях при пересечении цилиндрических граней должно быть построено не менее четырех новых ребер. В точках A и B (в точках касания цилиндров) должны располагаться вершины.

В булевых операциях возможен случай, когда какая-либо грань одного тела пересекает другое тело по его ребру. Возможен также случай, когда при построении новых ребер пересечения граней одного тела с гранями другого тела получим два новых ребра, совпадающих в пространстве друг с другом и с исходным ребром одного из тел. Все ребра являются разными, так как в них стыкуются разные грани. На рис. 6.3.8. приведены два тела, при выполнении булевой операции над которыми новые ребра совпадут с исходными ребрами одного из тел.

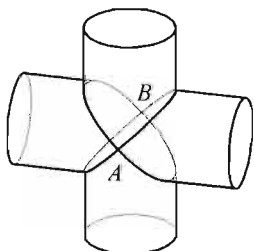


Рис. 6.3.7

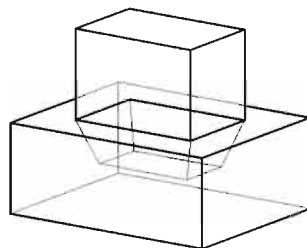


Рис. 6.3.8

Следуя общему алгоритму, в данном случае получим восемь новых ребер, половина из которых должна быть отброшена (или не должна быть построена).

При наличии новых ребер, совпадающих с исходными ребрами тел, будем выполнять следующую проверку. Построим в плоскости каждой из двух пересекаемых граней по два вектора, ортогональных новому ребру. На рис. 6.3.9 это векторы \mathbf{a}_1 и \mathbf{b}_1 для грани с нормалью \mathbf{m}_1 и векторы \mathbf{a}_2 и \mathbf{b}_2 для грани с нормалью \mathbf{m}_2 . Векторы \mathbf{a}_1 и \mathbf{a}_2 поворачивают влево от нового ребра, если двигаться вдоль нового ребра с внешней стороны соответствующей грани. Векторы \mathbf{b}_1 и \mathbf{b}_2 поворачивают вправо от нового ребра в соответствующих гранях, если двигаться вдоль нового ребра с внешней стороны соответствующей грани.

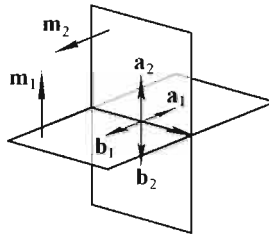


Рис. 6.3.9

Если грань не имеет продолжения от нового ребра в сторону какого-либо вектора a_1, a_2, b_1, b_2 (новое ребро частично или полностью совпадает с исходным ребром тела), то соответствующий вектор положим равным нулю, рис. 6.3.10, 6.3.11.

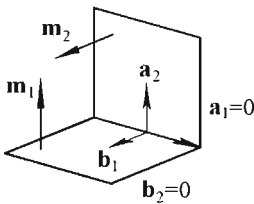


Рис. 6.3.10

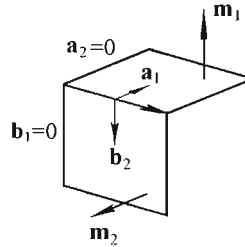


Рис. 6.3.11

Используя векторы a_1, a_2, b_1, b_2 и нормали m_1, m_2 , можно определить, будет ли данное новое ребро использовано в операции или оно должно быть отброшено. Для этого нужно выполнить следующие правила.

Правило 1. Пусть $r_1(u, v)$ и $r_2(a, b)$ – пересекающиеся грани первого и второго тел соответственно, для которых можно построить новое ребро. Если $r_1(u, v)$ имеет продолжение справа от нового ребра вне второго тела, то в операции объединения тел циклы грани $r_1(u, v)$ необходимо перестроить. Если $r_2(a, b)$ имеет продолжение слева от нового ребра вне первого тела, то в операции объединения тел циклы грани $r_2(a, b)$ необходимо перестроить.

Таким образом, в булевой операции объединения тел для нового ребра должны быть выполнены условия

$$m_1 \cdot a_2 > 0 \quad \text{и} \quad m_2 \cdot b_1 > 0. \quad (6.3.1)$$

В противном случае рассматриваемое новое ребро в булевой операции объединения тел строить не следует (если оно построено, то должно быть отброшено).

Правило 2. Пусть $r_1(u, v)$ и $r_2(a, b)$ – пересекающиеся грани первого и второго тел соответственно, для которых можно построить новое ребро. Если $r_1(u, v)$ имеет продолжение слева от нового ребра внутрь второго тела, то в булевой операции пересечения тел циклы грани $r_1(u, v)$ необходимо перестроить. Если $r_2(a, b)$ имеет продолжение справа от нового ребра внутрь первого тела, то в булевой операции пересечения тел циклы грани $r_2(a, b)$ необходимо перестроить.

Таким образом, в булевой операции пересечения тел для нового ребра должны быть выполнены условия

$$m_1 \cdot b_2 < 0 \quad \text{и} \quad m_2 \cdot a_1 < 0. \quad (6.3.2)$$

В противном случае рассматриваемое новое ребро в булевой операции пересечения тел строить не следует (если оно построено, то должно быть отброшено).

Для ответа на вопрос внутрь или наружу тела продолжается от нового ребра грань другого тела, нужно уметь определять положение точки \mathbf{p} относительно тела: принадлежит ли точка \mathbf{p} внутреннему пространству тела, лежит ли точка \mathbf{p} на поверхности тела или же точка \mathbf{p} находится вне тела. Для этого найдем ближайшую к точке \mathbf{p} точку \mathbf{p}_0 на грани одной из оболочек тела. Далее построим вектор $\mathbf{v} = \mathbf{p} - \mathbf{p}_0$ из ближайшей точки \mathbf{p}_0 в точку \mathbf{p} и вычислим нормаль оболочки \mathbf{m} в точке \mathbf{p}_0 . Способ вычисления нормали оболочки зависит от того, какая точка оболочки окажется ближайшей. Ближайшей точкой оболочки может быть или проекция точки \mathbf{p} на одну из граней, или проекция точки \mathbf{p} на одно из ребер, или одна из вершин оболочки. Если ближайшей точкой оболочки является проекция точки \mathbf{p} на одну из граней, то в качестве нормали оболочки \mathbf{m} примем нормаль этой грани в точке проекции. Если ближайшей точкой оболочки является проекция точки \mathbf{p} на одно из ребер, то в качестве нормали оболочки \mathbf{m} в ребре примем нормализованную сумму нормалей, стыкующихся в данном ребре граней в точке проекции. Если ближайшей точкой оболочки является вершина оболочки, то в качестве нормали оболочки \mathbf{m} в вершине примем нормализованную сумму нормалей, стыкующихся в данной вершине ребер в точке проекции. Для определения положения точки относительно тела используем принцип построения тел, в соответствии с которым нормаль поверхности тела направлена вне тела. Положение точки относительно оболочки тела определим по скалярному произведению вектора \mathbf{v} и нормали \mathbf{m} в ближайшей точке оболочек тела. Если длина вектора \mathbf{v} равна нулю, то $\mathbf{v} \cdot \mathbf{m} = 0$ и точка \mathbf{p} располагается на поверхности тела, если $\mathbf{v} \cdot \mathbf{m} > 0$, то точка \mathbf{p} располагается вне

тела, если $v \cdot m < 0$, то точка p принадлежит внутреннему пространству тела.

Если какое-либо из тел-операндов имеет пустоты и, соответственно, описывается несколькими оболочками, то нетронутые операцией оболочки и оболочки результата булевой операции должны быть проверены на вхождение друг в друга. В общем случае результат булевой операции также может иметь несколько оболочек.

6.4. Разрезанное тело

Алгоритм булевых операций позволяет резать тело незамкнутой оболочкой. Режущая оболочка должна полностью пересекать тело. Результатом такой операции является новое тело, представляющее собой часть исходного тела, лежащую с внешней или с внутренней стороны режущей оболочки. Незамкнутую режущую оболочку будем называть незаконченным телом.

При резке тела оболочкой мы выполняем операцию пересечения тела с частью пространства, лежащей по одну из сторон режущей оболочки. Если требуется построить часть тела, которая лежит с внешней стороны режущей оболочки, то выполним булеву операцию вычитания из заданного тела незаконченного тела. Если требуется построить ту часть тела, которая лежит с внутренней стороны режущей оболочки, то выполним булеву операцию пересечения заданного тела с незаконченным телом. На рис. 6.4.1 показаны тело в форме прямоугольного параллелепипеда, режущая оболочка, ее нормаль и ребра пересечения граней тела с оболочкой. На рис. 6.4.2 приведено разрезанное тело.

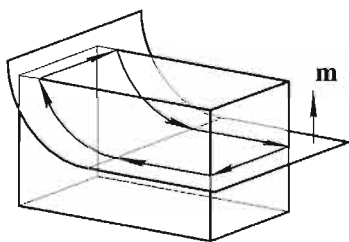


Рис. 6.4.1

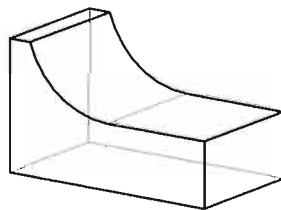


Рис. 6.4.2

В частном случае тело можно разрезать плоскостью. С помощью рассматриваемой операции можно получить сложный разрез тела. Для этого оболочка незаконченного тела должна состоять из нескольких стыкующихся между собой граней. Для получения сложного разреза возьмем

составную кривую на плоскости и построим по ней незаконченное тело выдавливания так, чтобы оно пересекало заданное тело требуемым образом. Далее выполним булеву операцию над заданным телом и незаконченным телом.



Рис. 6.4.3



Рис. 6.4.4

На рис. 6.4.3 приведено исходное тело, а на рис. 6.4.4 приведен результат сложного разреза тела.

6.5. Симметричное тело

Пусть имеется тело и плоскость. Построим тело, симметричное данному телу относительно данной плоскости. Симметричное тело будет зеркальной копией данного тела. Геометрия тела описывается точками, кривыми и поверхностями, которые построены из векторов и скаляров. Поэтому построение симметричного тела в конечном итоге сводится к преобразованию симметрии радиусов-векторов точек, свободных векторов и скаляров. Скалярные величины при преобразовании симметрии не изменяются, свободные векторы меняют свое направление, а точки — свое положение.

Пусть плоскость симметрии проходит через точку \mathbf{p}_0 , а нормаль к плоскости описывается вектором единичной длины $\mathbf{i} = [i_1 \ i_2 \ i_3]^T$, рис 6.5.1. Определим координаты точки \mathbf{r} , симметричной точки \mathbf{r}_0 относительно плоскости.

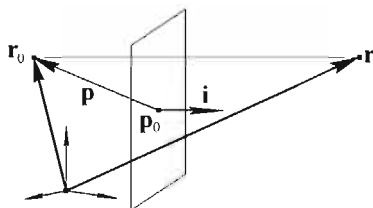


Рис. 6.5.1

Построим вектор $\mathbf{p} = \mathbf{r}_0 - \mathbf{p}_0$ и представим его в виде суммы двух векторов — проекции на плоскость симметрии и перпендикулярной плоскости составляющей \mathbf{n}

$$\mathbf{p} = (\mathbf{p} - \mathbf{n}) + \mathbf{n},$$

где $\mathbf{n} = (\mathbf{p} \cdot \mathbf{i})\mathbf{i}$. После зеркального отражения вектора \mathbf{p} его нормальная к плоскости составляющая \mathbf{n} изменится и будет равна $-\mathbf{n}$. Проекция $(\mathbf{p} - (\mathbf{p} \cdot \mathbf{i})\mathbf{i})$ вектора \mathbf{p} на плоскость симметрии не изменится. Положение точки, симметричной точке \mathbf{r}_0 , будет описываться радиусом-вектором

$$\mathbf{r} = \mathbf{p}_0 + (\mathbf{p} - (\mathbf{p} \cdot \mathbf{i})\mathbf{i}) - (\mathbf{p} \cdot \mathbf{i})\mathbf{i} = \mathbf{p}_0 + (\mathbf{E} - 2\mathbf{i}\mathbf{i}^T) \cdot \mathbf{p} = \mathbf{p}_0 + \mathbf{A} \cdot (\mathbf{r}_0 - \mathbf{p}_0),$$

где матрица преобразования симметрии \mathbf{A} определяется равенством

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} i_1 i_1 & i_1 i_2 & i_1 i_3 \\ i_2 i_1 & i_2 i_2 & i_2 i_3 \\ i_3 i_1 & i_3 i_2 & i_3 i_3 \end{bmatrix}. \quad (6.5.1)$$

Свободный вектор \mathbf{v}_0 после преобразования симметрии относительно этой плоскости будет описываться радиусом-вектором

$$\mathbf{v} = \mathbf{A} \mathbf{v}_0.$$

После преобразования симметрии все точки и кривые на поверхностях останутся неизменными. Нормали поверхностей после преобразования симметрии изменят свое направление на противоположное. Если выполнить только преобразование геометрических данных тела по матрице (6.5.1), то получим зеркальное отражение тела, вывернутое наизнанку. Таким образом, построение симметричного тела сводится к преобразованию его копии по матрице (6.5.1) и переориентации направлений нормалей граней и направлений циклов граней. Переориентация направления нормали грани производится изменением признака совпадения нормали поверхности и нормали ее грани. Переориентация направления цикла грани производится перестроением списка ребер (изменением на обратный порядок следования ребер в списке) и изменением значений флагов ребер в списках.

Пусть плоскость симметрии пересекает тело. Представим, что оболочка тела разрезана плоскостью, одна из отрезанных частей удалена, по другой части выполнена зеркальная копия и склеена с ней. Мы получим тело, симметричное относительно плоскости, состоящее из двух половинок, одна из которых совпадает с частью исходного тела. Плоскостью симметрии может служить плоская грань исходного тела. Построение такого симметричного тела имеет много общего с операцией объединения тел.

Не теряя общности, будем строить симметричное тело по его части, находящейся за плоскостью симметрии. Сначала построим кривые пересечения граней тела с плоскостью. Если грань описывается радиусом-вектором $\mathbf{r}(u, v)$, то кривую пересечения этой грани с плоскостью симметрии $\mathbf{p}(x, y)$ представим в виде

$$\mathbf{e}_p(t) = \begin{cases} \mathbf{r}(u(t), v(t)) \\ \mathbf{p}(x(t), y(t)) \end{cases}, \quad t \in [t_{\min}, t_{\max}].$$

Эта кривая содержит две двумерные кривые, из которых точка с координатами $u(t)$, $v(t)$ лежит на поверхности $\mathbf{r}(u, v)$, а точка с координатами $x(t)$, $y(t)$ лежит на плоскости симметрии. На базе кривых пересечения построим новые ребра. С помощью признака совпадения направления ребра и его кривой пересечения новым ребрам дадим направление векторного произведения нормали грани с нормалью плоскости: $\mathbf{t}_e = \mathbf{m}_r \times \mathbf{m}_p$. После построения новых ребер определим точки их пересечения с исходными ребрами тела, в этих точках построим вершины и этими вершинами разрежем ребра на части.

Далее построим симметричную относительно плоскости копию части оболочки тела, лежащей над плоскостью.

Остается сшить симметричные половинки тела по ребрам пересечения тела с плоскостью симметрии. Но прежде выполним замену кривых на плоскости, входящих в кривые пересечения новых ребер, на кривые на симметричной оболочке. Для этого вместо кривых \mathbf{e}_p в ребра положим кривые

$$\mathbf{e}(t) = \begin{cases} \mathbf{r}(u(t), v(t)) \\ \mathbf{r}'_p(u(t), v(t)) \end{cases}, \quad t \in [t_{\min}, t_{\max}].$$

Кривая $\mathbf{e}(t)$ содержит две двумерные кривые, из которых точка с координатами $u(t)$, $v(t)$ лежит на поверхности $\mathbf{r}(u, v)$, а ее копия лежит на симметричной копии $\mathbf{r}'_p(u, v)$ поверхности $\mathbf{r}(u, v)$. Теперь каждое ребро пересечения будет базироваться на двух симметричных половинках искомого тела.

Наконец перестроим циклы пересеченных плоскостью граней тела. Перестроение циклов грани подробно описано в булевой операции объединения тел. Каждое ребро пересечения должно войти в цикл исходной грани с отрицательным флагом, а в цикл ее симметричной части — с положительным флагом.

К пересеченным граням тела добавим непересеченные грани, которые связаны с первыми, и соответствующие симметричные копии. Так получим оболочку симметричного тела.

При построении симметричной оболочки следует учитывать все случаи совпадения новых ребер и исходных ребер тела и частичное перекрытие граней, описанные в булевых операциях над телами. Кроме того, следует учесть наличие нескольких оболочек у исходного тела.



Рис. 6.5.2



Рис. 6.5.3

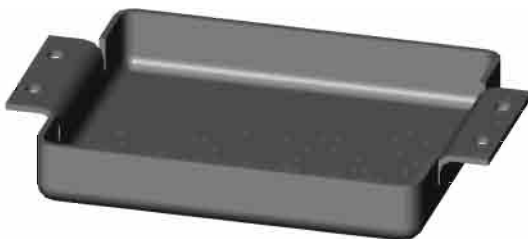


Рис. 6.5.4

На рис. 6.5.2 приведено исходное тело и плоскость симметрии. На рис. 6.5.3 приведен результат построения симметричного тела. На рис. 6.5.4 приведен результат последовательного выполнения двух операций построения симметричного тела относительно взаимно ортогональных плоскостей симметрии.

6.6. Тело с достраиваемыми элементами

На практике часто требуется выполнить булеву операцию, когда оболочка одного из операндов не полностью определена и должна быть достроена в процессе операции. Например, к заданному телу нужно добавить часть тела, полученного выдавливанием заданной замкнутой кривой в заданном направлении, лежащую между телом и кривой, рис. 6.6.1. В данном случае требуется «выдавить» заданную кривую до ближайших граней заданного тела. В разных местах заданной кривой ближайшими могут оказаться разные грани, поэтому точки кривой должны быть «выдавлены» на различные расстояния.

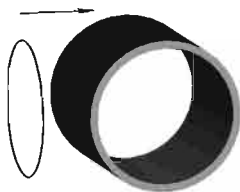


Рис. 6.6.1

Выполнить данную операцию можно следующим образом. По заданной кривой построим тело выдавливания достаточной глубины для пересечения с заданным телом. Далее из тела выдавливания вычтем заданное тело. В результате операции в общем случае получим несколько оболочек, рис. 6.6.2. Выберем из них ближайшую к заданной кривой, рис. 6.6.3. Тело с выбранной оболочкой объединим с заданным телом. Описанная операция содержит две булевых операции. При выполнении данных действий достаточно строить ребра пересечения только один раз. Результат такой операции приведен на рис. 6.6.4.



Рис. 6.6.2



Рис. 6.6.3



Рис. 6.6.4

Другим примером может служить операция вычитания из заданного тела части тела, полученного выдавливанием заданной замкнутой кривой в заданном направлении, лежащей между телом и кривой, рис. 6.6.5. В данном примере изменяются ближайшие к кривой грани тела и не трогаются грани, лежащие за ними. В разных местах точки кривой должны быть «выдавлены» на различные расстояния.

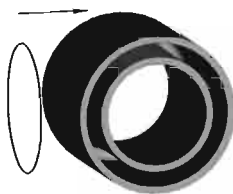


Рис. 6.6.5

Вырезать ближайшие грани заданного тела заданной кривой можно следующим образом. По заданной кривой построим тело выдавливания достаточной глубины для пересечения с заданным телом. Далее найдем пересечение тела выдавливания с заданным телом. В результате этих действий в общем случае получим несколько оболочек, рис. 6.6.6. Выберем из них ближайшую к заданной кривой оболочку, рис. 6.6.7. Тело с выбранной оболочкой вычтем из заданного тела. Описанная операция содержит две булевых операции. Результат такой операции приведен на рис. 6.6.8.



Рис. 6.6.6



Рис. 6.6.7



Рис. 6.6.8

Во многих конструкциях используются ребра жесткости. Операция построения ребер жесткости выполняется по той же схеме, что и операция «выдавливания» кривой до ближайших к ней граней тела. Тело с ребрами жесткости приведено на рис. 6.6.9.



Рис. 6.6.9

Когда оболочка приклеиваемого или вырезаемого тела должна быть построена в процессе выполнения операции, после выполнения операций следует проверить оболочку построенного тела на замкнутость.

6.7. Эквидистантное тело

По телу можно построить другое тело, все грани которого расположены на заданном расстоянии по нормали от поверхности исходного тела. Это расстояние будем называть *параметром эквидистанты* и обо-

значим через h . Тело, по которому строится новое тело, будем называть *базовым телом*, а новое тело будем называть *эквидистантным телом*. Параметр эквидистанты этой операции может принимать как положительные, так и отрицательные значения. Если $h > 0$, то базовое тело располагается внутри эквидистантного тела. Если $h < 0$, то эквидистантное тело располагается внутри базового тела. Недопустимыми значениями параметра являются такие значения, при которых оболочка нового тела получается вырожденной.

Процесс создания нового тела начнем с построения эквидистантных граней для каждой грани базового тела. Эквидистантная грань базируется на поверхности, эквидистантной к соответствующей поверхности базового тела. Радиус-вектор эквидистантной поверхности определяется формулой (3.15.4). На рис. 6.7.1 приведены три грани базового тела, имеющие общую вершину, и эквидистантные им поверхности. Как правило, эквидистантные поверхности необходимо усекать или продлевать до пересечения со смежными эквидистантными поверхностями для стыковки с ними. Продолжение эквидистантной поверхности выполним с помощью продолжения базовой поверхности в соответствии с правилами построения продолженной поверхности (3.15.1) – (3.15.3).

Для построения эквидистантного тела необходимо построить его вершины и ребра. Ребра будут базироваться на кривых пересечения эквидистантных поверхностей. Каждая кривая пересечения должна начинаться и оканчиваться в вершине тела. Поэтому далее построим вершины эквидистантной оболочки. Рассмотрим последовательно вершины базового тела. Вершине базового тела будут соответствовать одна или несколько вершин эквидистантного тела. Вершина эквидистантного тела базируется на точке пересечения эквидистантных поверхностей. Вычислив эту точку, найдем параметры эквидистантных поверхностей, которые будут служить в качестве начальных и конечных точек кривых пересечения поверхностей. На рис. 6.7.2 приведены точки пересечения продолженных эквидистантных поверхностей, на которых будут базироваться вершины.

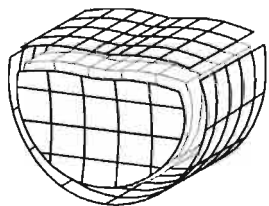


Рис. 6.7.1

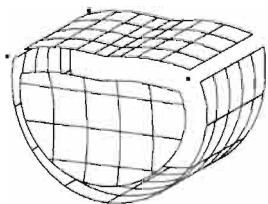


Рис. 6.7.2

Будем называть ребро *гладким*, если нормали соединяемых им граней совпадают на протяжении всего ребра. Гладкому ребру базового тела будет соответствовать гладкое ребро в эквидистантном теле. Более того, кривые пересечения таких ребер проходят по одним и тем же точкам $[u \ v]^T$ соответствующих поверхностей базового и эквидистантного тел.

Далее построим ребра эквидистантной оболочки. Каждое ребро будет базироваться на кривой пересечения эквидистантных поверхностей. Начальные и конечные точки ребер дадут вершины. Соответствующее ребро базового тела может служить направляющей кривой для построения ребра эквидистантной оболочки. На рис. 6.7.3 приведены кривые пересечения продолженных эквидистантных поверхностей, на которых базируются ребра.

Из ребер построим циклы эквидистантных граней, рис. 6.7.4. Этот процесс аналогичен процессу перестроения циклов граней в булевых операциях. После перечисленных построений состав эквидистантного тела совпадает с составом базового тела.

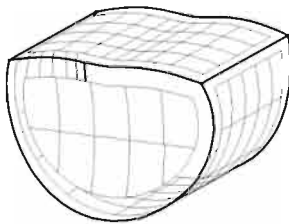


Рис. 6.7.3

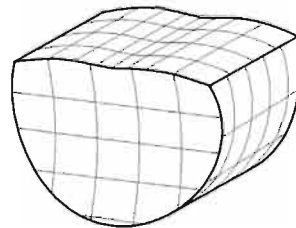


Рис. 6.7.4

Если в вершине базового тела стыкуется более трех негладких ребер, то в эквидистантном теле ей будут соответствовать несколько вершин и новых ребер. Все они могут быть найдены как точки пересечения троек поверхностей или как точки пересечения гладких ребер и поверхностей эквидистантного тела. На рис. 6.7.5. приведена пирамида и эквидистантное пирамиде тело с отрицательным параметром эквидистанты, а на рис. 6.7.6. приведена аналогичная пирамида и эквидистантное пирамиде тело с положительным параметром эквидистанты. В вершине A пирамиды стыкуются четыре ребра. В эквидистантном теле данной вершине соответствуют две вершины B и C и одно дополнительное ребро BC .

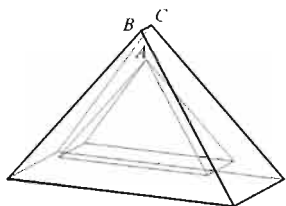


Рис. 6.7.5

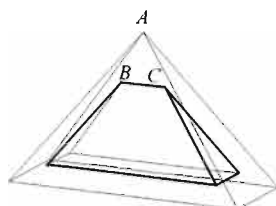


Рис. 6.7.6

Возможна ситуация, когда некоторой вершине, некоторому ребру или некоторой грани базового тела в эквидистантной оболочке не будет аналога. В таких телах после описанных выше действий возникнут самопересечения оболочки и самопересечения циклов некоторых граней. То есть в эквидистантной оболочке пересекутся грани, которые в базовой оболочке не пересекаются. Для обработки подобных ситуаций следует проверить на пересечение эквидистантные грани, которые не являются смежными. Если таковые имеются, то следует построить ребра пересечения и выполнить перестроение их циклов, как в булевой операции объединения тел.

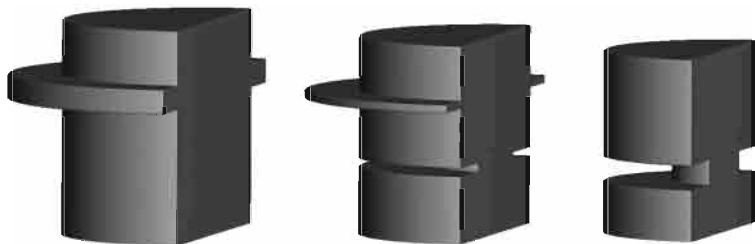


Рис. 6.7.7

На рис. 6.7.7 справа и слева показаны эквидистантные тела, построенные на основе базового тела, показанного в центре. Для тела, показанного на рис. 6.7.7 слева, $h > 0$. Для тела, показанного на рис. 6.7.7 справа, $h < 0$. Количество граней, ребер и вершин эквидистантных тел может отличаться от количества граней ребер и вершин базового тела.

6.8. Тонкостенное тело

Упрощенно построение тонкостенного тела можно описать следующим образом: у заданного тела удалим указанные грани, а оставшимся граням «придадим конечную толщину». «Придание конечной толщины» граням достигнем путем построения к оставшейся после удаления указанных граней незамкнутой оболочки эквидистантной оболочки и соединения этих незамкнутых оболочек частями удаленных граней. «Придаваемую оставшимся граням толщину» обозначим через h . Тонкостенное тело может быть построено как наружу ($h > 0$), так и внутрь ($h < 0$) от исходного тела.

На рис. 6.8.2 приведен пример тонкостенного тела ($h > 0$), построенного путем вскрытия одной грани тела, показанного на рис. 6.8.1.



Рис. 6.8.1



Рис. 6.8.2

Процесс создания тонкостенного тела начнем с сортировки граней базового тела на две группы: к первой группе отнесем вскрываемые грани базового тела, а ко второй группе отнесем остальные грани базового тела, которые будем называть сохраняемыми. Сгруппируем ребра базового тела: к первой группе отнесем ребра, по которым пересекаются между собой вскрываемые грани базового тела, а ко второй группе отнесем ребра сохраняемых граней базового тела. Процесс построения оболочки тонкостенного тела имеет много общего с процессом построения эквидистантного тела.

Для каждой сохраняемой грани базового тела построим эквидистантную грань. Поверхность каждой эквидистантной грани и каждой вскрываемой грани должна иметь возможность быть продолженной до пересечения с поверхностями соседних граней. Радиус-вектор эквидистантной поверхности определяется формулой (3.15.4).

Далее рассмотрим вершины сохраняемых граней базового тела. Каждой рассматриваемой вершине будут соответствовать одна или несколько вершин тонкостенного тела. В каждой вершине тонкостенного тела будут стыковаться несколько ребер. Вершина тонкостенного тела бази-

руется на точке пересечения эквидистантных поверхностей или на точке пересечения эквидистантных поверхностей с поверхностями вскрываемых граней. Вычислив эти точки, найдем параметры пересекающихся поверхностей, которые будут служить в качестве начальных и конечных точек кривых пересечения поверхностей.

Рассмотрим последовательно ребра сохраняемых граней базового тела (ребра второй группы). Для каждого ребра построим соответствующее ребро тонкостенного тела. Для этого найдем кривые пересечения эквидистантных поверхностей между собой и с поверхностями вскрываемых граней с использованием уже известных начальных и конечных точек.

По построенным вершинам перестроим ребра пересечения вскрываемых граней (ребра первой группы).

Далее вывернем наизнанку часть граней. Если $h > 0$, то изменим на противоположные направления циклов и нормалей сохраняемых граней базового тела и направления циклов вскрываемых граней базового тела. Если $h < 0$, то изменим на противоположные направления нормалей эквидистантных граней.

Наконец построим циклы эквидистантных граней и перестроим циклы вскрываемых граней. Таким образом, оболочка тонкостенного тела будет составлена из сохраняемых граней, эквидистантных к ним граней и частей вскрываемых граней.

При построении тонкостенного тела могут возникнуть самопересечения оболочки и самопересечения циклов эквидистантных граней. Для исправления таких ситуаций следует проверить на пересечение эквидистантные грани, которые не являются смежными, и если они пересекутся, то следует построить ребра пересечения и выполнить перестроение их циклов как в булевой операции объединения тел. Описанные действия должны быть выполнены при построении тонкостенного тела на основе тела, приведенного на рис. 6.8.3. Разрез этого тела показан на рис. 6.8.4.



Рис. 6.8.3



Рис. 6.8.4

На рис. 6.8.5 приведено тонкостенное тело, построенное путем удаления грани базового тела, показанного на рис. 6.8.3. На рис. 6.8.6 приведен разрез тонкостенного тела.



Рис. 6.8.5



Рис. 6.8.6

Если построить тонкостенное тело без удаления граней, то получится тело с двумя оболочками. Одна из них является внешней, а другая – внутренней. Если $h > 0$, то внешней оболочкой является эквидистантная оболочка, а оболочка базового тела должна быть вывернута наизнанку. Если $h < 0$, то внешней оболочкой является оболочка базового тела, а эквидистантную оболочку после построения следует вывернуть наизнанку. Оболочки отстоят на расстоянии $|h|$ друг от друга. Такое тело имеет внутри пустоту, отделенную от внешней части пространства.

6.9. Скругление ребер тела

Плавный переход от одной грани тела к другой называют сопряжением граней. Одним из способов сопряжения граней является скругление ребер тела. Скругление ребер обладает множеством разновидностей. В простейшем случае скругление ребра заключается в его замене на грань, полученную движением сферы заданного радиуса, касающейся одновременно двух смежных с ребром граней. Результат скругления двух выпуклых ребер и одного вогнутого ребра тела, показанного на рис. 6.9.1, приведен на рис. 6.9.2.

В частных случаях грань скругления может базироваться на цилиндрической поверхности или поверхности тора. Рассмотрим общий случай скругления произвольного криволинейного ребра. Пусть грань, располагающаяся справа от скругляемого ребра, при взгляде снаружи тела вдоль ребра базируется на поверхности $r(u, v)$, а грань, располагающаяся слева от скругляемого ребра, базируется на поверхности $s(a, b)$. Эти грани и их поверхности будем называть сопрягаемыми. Пусть скругляемое ребро базируется на кривой пересечения $c(s)$ поверхностей $r(u, v)$ и $s(a, b)$. По

кривой $\mathbf{c}(s)$ построим поверхность скругления (4.9.1), которая касается сопрягаемых поверхностей по кривым

$$\mathbf{c}_1(t) = \mathbf{r}(u(t), v(t)) \text{ и } \mathbf{c}_2(t) = \mathbf{s}(a(t), b(t)).$$

Параметр t является общим для кривых $\mathbf{c}_1(t)$, $\mathbf{c}_2(t)$ и поверхности скругления $\mathbf{q}(t, z)$.

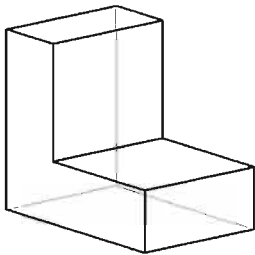


Рис. 6.9.1

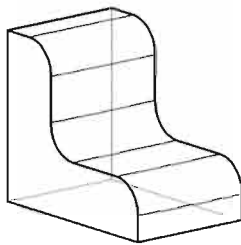


Рис. 6.9.2

Найдем пересечение опорных кривых $\mathbf{c}_1(t)$ и $\mathbf{c}_2(t)$ с ребрами сопрягаемых граней и изменим эти ребра. В одних случаях исходные ребра нужно обрезать, в других случаях их нужно продлить. Если скругляемое ребро циклически не замкнуто, то найдем также пересечение поверхности скругления с ребрами тела, стыкующимися со скругляемым ребром и не принадлежащими сопрягаемым граням. Эти ребра также следует обрезать поверхностью скругления или продлить до пересечения с поверхностью скругления.

Далее построим ребра, по которым поверхность скругления стыкуется с опорными гранями. Пусть кривая $\mathbf{c}_1(t)$ совпадает с $\mathbf{q}(t, z_{\min})$, а кривая $\mathbf{c}_2(t)$ совпадает с $\mathbf{q}(t, z_{\max})$. Построим два ребра вдоль краев поверхности скругления на базе кривых пересечения:

$$\mathbf{c}_1(t) = \begin{cases} \mathbf{r}(u(t), v(t)) \\ \mathbf{q}(t, z_{\min}) \end{cases}, \quad \mathbf{c}_2(t) = \begin{cases} \mathbf{s}(a(t), b(t)) \\ \mathbf{q}(t, z_{\max}) \end{cases}.$$

Эти ребра будем называть *продольными*, так как они направлены вдоль скругляемого ребра и имеют ту же ориентацию.

Если скругляемое ребро циклически замкнуто, то поверхность скругления также будет циклически замкнутой по параметру t . Тогда на базе кривой пересечения

$$\mathbf{c}_0(t) = \begin{cases} \mathbf{q}(t_{\min}, w) \\ \mathbf{q}(t_{\max}, w) \end{cases}, \quad w \in [z_{\min}, z_{\max}]$$

построим ребро, которое будет являться швом грани скругления.

Если скругляемое ребро циклически не замкнуто, то построим кривые, по которым поверхность скругления пересекается с другими гранями тела. На базе этих кривых пересечения создадим ребра, которые будем называть *поперечными*.

На базе поверхности скругления построим грань. Цикл этой грани будет состоять из построенных продольных и поперечных ребер или шва. Для грани скругления определим признак совпадения ее нормали с направлением нормали поверхности. Нормаль грани скругления должна быть направлена наружу тела, нормаль поверхности скругления может совпадать с ней или быть ей противоположной.

После этого перестроим циклы граней на сопрягаемых поверхностях $r(u, v)$ и $s(a, b)$. Вместо скругляемого ребра в цикл грани $r(u, v)$ войдет ребро на базе кривой пересечения c_1 , а в цикл грани $s(a, b)$ войдет ребро на базе кривой пересечения c_2 , рис. 6.9.3. Далее перестроим циклы остальных граней, пересекающихся с поверхностью скругления $q(t, z)$.

Скругление ребер требуется выполнять для различных ситуаций по-разному. На рис. 6.9.4 приведено тело со скруглением сопряженных ребер, у которых на концах продольные ребра грани скругления сходятся в одну точку. В данном примере между сопряженными гранями поперечные ребра стянуты в точку, а поверхности скругления в них вырождаются.



Рис. 6.9.3

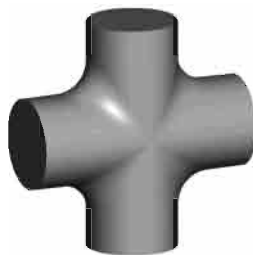


Рис. 6.9.4

Если скруглить три ребра, стыкующиеся в одной вершине, то в вершине получим картину, приведенную на рис. 6.9.5. Как правило, такую вершину скругляют. Скруглить вершину можно описанным выше способом: ребро пересечения двух граней скругления сопряжено с третьим ребром и может быть скруглено вместе с ним. На рис. 6.9.6 приведен пример скругления ребер и вершины призмы описанным способом.



Рис. 6.9.5



Рис. 6.9.6

Во многих случаях требуется скруглить сразу несколько стыкующихся друг с другом ребер. Будем называть стыкующиеся ребра *сопряженными*, если в точках стыковки они имеют общую касательную и одинаковые углы между нормальными соответствующих смежных граней. Скругление сопряженных ребер следует выполнять одновременно для всех сопряженных ребер.

В некоторых случаях требуется скруглить несколько ребер, стыкующихся в общей вершине. Если в вершине стыкутся более трех ребер, то такую комбинацию ребер и вершины будем называть *звездой*. Если все ребра звезды подлежат скруглению, то так же, как и для группы сопряженных ребер, различные стадии операции скругления следует выполнять одновременно для всех ребер звезды и сопряженных с ними ребер. При этом следует учесть, что звезда через сопряженные ребра может стыковаться с другими звездами. Поэтому перед началом операции следует найти группы гладко стыкующихся ребер, найти звезды для них, определить группы связанных звезд и далее описанным выше образом работать с этими группами звезд как с единым объектом. На рис. 6.9.7 приведен пример скругления звезды, содержащей 6 ребер.



Рис. 6.9.7

На одной из сопрягаемых граней может не хватить места для продольного ребра. В таком случае можно выполнить скругление с сохране-

нием кромки или с адаптацией грани скругления. Рассмотрим различные способы на примере скругления ребра у основания цилиндрической грани тела, приведенного на рис. 6.9.8.

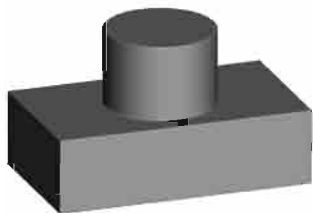


Рис. 6.9.8

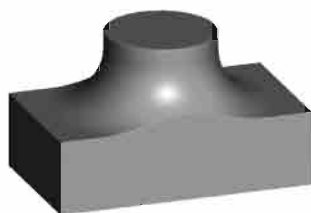


Рис. 6.9.9

У тела, приведенного на рис. 6.9.9, одна из опорных кривых поверхности скругления на определенном участке выходит за пределы сопрягаемой грани. Этот участок заменен кривой пересечения грани скругления и соответствующей грани тела. У тела, приведенного на рис. 6.9.10, в подобной ситуации участок грани скругления изменен так, чтобы сохранить исходную форму ребра и грани, в районе которых опорная кривая поверхности скругления выходит за пределы сопрягаемой грани. Поверхность скругления, сохраняющая кромку, описывается формулой (4.9.7). У тела, приведенного на рис. 6.9.11, опорная кривая поверхности скругления на определенном участке также выходит за пределы сопрягаемой грани. Сопрягаемая грань на этом участке плавно стыкуется с соответствующей соседней гранью, поэтому без изменения поверхности скругления построение невозможно. На упомянутом участке при построении поверхности скругления исходная сопрягаемая грань заменена соответствующей соседней гранью.

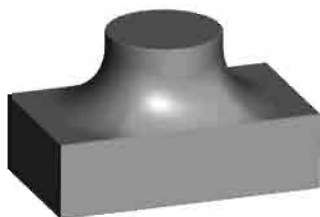


Рис. 6.9.10

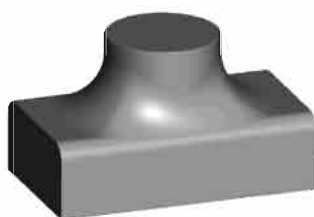


Рис. 6.9.11

Можно построить параболическое или эллиптическое сопряжение граней, при котором радиус кривизны грани скругления плавно изменя-

ется от одной сопрягаемой грани ко второй сопрягаемой грани, рис. 6.9.12.

Если радиус скругления будет меняться по длине скругляемого ребра, то получим грань скругления с переменным радиусом, рис. 6.9.13. При построении скруглений с переменным радиусом следует тем или иным способом задать функцию радиуса по длине ребра.

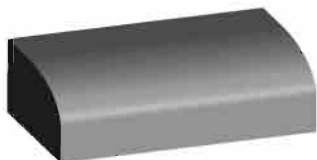


Рис. 6.9.12



Рис. 6.9.13

В верхней части рис. 6.9.14 приведена грань скругления с переменным радиусом. Радиус скругления изменяется так, чтобы хорда грани скругления имела постоянное значение. В нижней части рис. 6.9.14 для сравнения приведена грань скругления с постоянным радиусом.

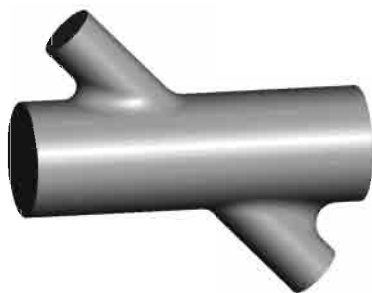


Рис. 6.9.14

Существует большое количество способов сопряжения граней тел. На перечисленных примерах мы проиллюстрировали основные идеи сопряжения граней тел путем обработки ребер. Способы построения опорных кривых для поверхностей скругления были рассмотрены выше.

6.10. Алгоритм скругления ребер тела

Проиллюстрируем алгоритм скругления ребер на модели, приведенной на рис. 6.10.1. В данном случае подлежащие скруглению ребра модели являются сопряженными. Перед началом операции скругления со-

ставим группы гладко стыкующихся ребер и далее будем работать с этими группами как с единым объектом описанным выше образом. На рис. 6.10.2 приведена группа из четырех сопряженных ребер.

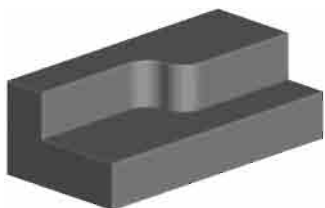


Рис. 6.10.1

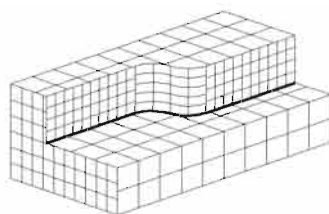


Рис. 6.10.2

Сначала для группы сопряженных ребер построим все поверхности скругления, рис. 6.10.3. Далее усечем этими поверхностями ребра тела, построим все продольные ребра, ребра стыковки поверхностей скругления между собой и ребра пересечения поверхностей скругления и граней тела, рис. 6.10.4.

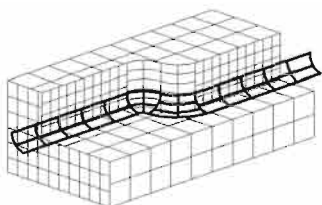


Рис. 6.10.3

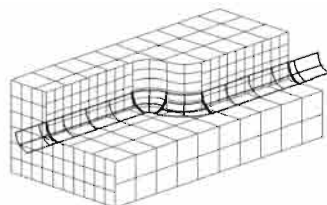


Рис. 6.10.4

После построения всех новых ребер построим грани скругления, рис. 6.10.5. В заключении операции перестроим циклы сопрягаемых граней тела и других граней, которые были затронуты операцией, рис. 6.10.6.

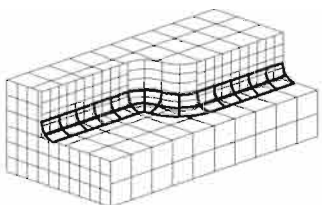


Рис. 6.10.5

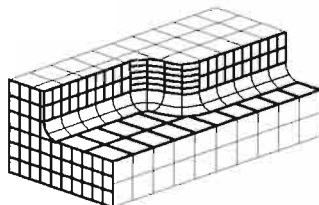


Рис. 6.10.6

После перестроения затронутых операцией граней добавим к ним другие связанные с ними грани и грани внутренних оболочек тела.

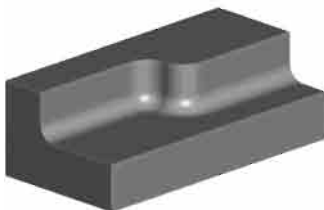


Рис. 6.10.7

Результат скругления четырех сопряженных ребер тела приведен на рис. 6.10.7.

6.11. Построение фаски ребер тела

Фаски ребер тела строятся аналогично скруглению ребер с той лишь разницей, что поверхности скругления заменяются поверхностями фасок. Фаску в общем случае опишем линейчатой поверхностью (4.9.10), построенной по двум опорным кривым $c_x(t)$, $c_y(t)$ на пересекающихся по ребру поверхностях. Результат построения фасок ребер тела, показанного на рис. 6.11.1, приведен на рис. 6.11.2.



Рис. 6.11.1



Рис. 6.11.2

Фаски могут быть построены по двум линейным размерам или по одному линейному размеру и одному угловому размеру. Линейные и угловые размеры фаски соответствуют параметрам треугольника, представляющего собой поперечное обрабатываемому ребру сечение срезаемой части тела. Способы построения опорных кривых для поверхностей фасок рассмотрены в предыдущей главе.

6.12. Прямое моделирование

Прямым моделированием называют модификацию геометрической модели путем изменения элементов, описывающих геометрическую форму модели, без полного перестроения модели. Методы прямого моделирования модифицируют оболочку без дополнительных построений непосредственным изменением граней оболочки. Непосредственное изменение элементов модели называют также *прямым редактированием* и *синхронным моделированием*. Прямое моделирование применяется в ситуациях, когда в результате переноса геометрической модели из одной системы в другую происходит потеря информации о последовательности и способах построения модели. Прямое моделирование позволяет редактировать модель без использования дополнительной информации.

Рассмотрим методы прямого моделирования на примере модификации тела, приведенного на рис. 6.12.1. Выберем группу граней тела и переместим эту группу граней относительно остальных граней. Результат операции, в которой мы переместили вверх ближайшую к нам группу граней, приведен на рис. 6.12.2.



Рис. 6.12.1



Рис. 6.12.2

Алгоритм данной операции прямого редактирования состоит из следующих действий. На первом шаге определим грани, стыкующиеся с перемещаемыми гранями. Если перемещаемые грани граничат с гранями скругления, то удалим скругления, запомнив сопрягаемые ими грани. В рассматриваемом нами примере следует перейти к телу, приведенному на рис. 6.12.3, и для него определить грани, стыкующиеся с перемещаемыми гранями. На втором шаге переместим выбранную группу граней. На третьем шаге построим новые вершины путем определения точек пересечения модифицированных граней и стыкующихся с ними граней. На четвертом шаге построим новые ребра, соединяющие новые вершины, путем построения кривых пересечения перечисленных граней. На пятом шаге построим циклы перемещенных

граней и перестроим циклы стыкующихся с ними граней. В результате получим модифицированное тело, но без скруглений, рис. 6.12.4.



Рис. 6.12.3



Рис. 6.12.4

На последнем шаге восстановим удаленные и запомненные сопряжения между гранями. Построение новых вершин и ребер, а также построение циклов затронутых операцией граней имеет много общего с алгоритмом построения эквидистантного тела.

Распространенный метод прямого моделирования изменяет выбранную группу граней требуемым образом и подстраивает под эти изменения другие грани. Выбранную группу граней можно не переместить, а заменить эквидистантными гранями. Для плоской грани эта модификация выглядит как перемещение в направлении ее нормали. Для цилиндрической, конической, сферической грани эта модификация выглядит как изменение радиуса соответствующей поверхности. У тела, приведенного на рис. 6.12.5, мы заменили эквидистантными гранями две горизонтальные грани. У тела, приведенного на рис. 6.12.6, мы заменили эквидистантными гранями грани двух отверстий.



Рис. 6.12.5



Рис. 6.12.6

Выбранную группу граней можно заменить уклоненными гранями. У тела, приведенного на рис. 6.12.7, мы уклонили коническую грань, в результате чего угол конусности изменил знак. Для изложенных методов требуется выполнить анализ оболочки, чтобы определить грани скругления. Перед модификацией скругления требуется удалить, а после модификации скругления нужно восстановить.

С помощью прямого моделирования можно удалить некоторую группу граней и построить тело из оставшихся граней оболочки. Результат операции, в которой мы удалили ближайшую к нам группу граней, приведен на рис. 6.12.8.



Рис. 6.12.7



Рис. 6.12.8



Рис. 6.12.9

Из удаляемой группы граней можно построить отдельное тело. Тело, приведенное на рис. 6.12.9, мы построили из группы удаленных в предыдущей операции граней.

Построенное из выбранной группы граней отдельное тело можно использовать для моделирования. На рис. 6.12.10 приведен результат булевой операции объединения тела, приведенного на рис. 6.12.7, и повернутого тела, приведенного на рис. 6.12.9.

При выборе для модификации грани скругления будем изменять радиус скругления. Для этого требуется собрать в одну группу стыкующиеся друг с другом грани скругления одного радиуса, запомнить сопрягаемые ими грани, удалить группу граней скругления, а затем построить сопряжение запомненных граней новым радиусом.



Рис. 6.12.10



Рис. 6.12.11

На рис. 6.12.11 приведен результат изменения радиусов скруглений тела, приведенного на рис. 6.12.10.

6.13. Деформирование оболочек

Рассмотрим два метода деформирования оболочек: масштабирование и локальное деформирование.

Масштабирование применимо ко всем без исключения однородным оболочкам. Масштабирование сводится к преобразованию по матрице данных, по которым построены вершины, ребра и грани оболочки. В качестве матрицы масштабирования используем расширенную матрицу (1.4.2), при работе с которой предполагается, что радиусы-векторы имеют вид (1.4.3), а векторы имеют вид (1.4.3). Матрица масштабирования описывается координатами p_1, p_2, p_3 неподвижной точки и масштабными коэффициентами по осям координат k_1, k_2, k_3 и имеет вид

$$\mathbf{M} = \begin{bmatrix} k_1 & 0 & 0 & t_1 \\ 0 & k_2 & 0 & t_2 \\ 0 & 0 & k_3 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

где $t_i = (1 - k_i)p_i, i = 1, 2, 3$. Если $k_1 = k_2 = k_3$, то масштабирование будет изотропным, в противном случае масштабирование будет анизотропным. Любое тело и любая оболочка могут быть отмасштабированы в соответствии с заданной матрицей. На рис. 6.13.1 приведено тело до масштабирования. На рис. 6.13.2 приведено тело после анизотропного масштабирования.



Рис. 6.13.1



Рис. 6.13.2

Локальное деформирование возможно только для некоторых поверхностей. Чтобы иметь возможность локально деформировать любое тело, заменим все или некоторые поверхности его граней деформируемыми поверхностями. В качестве деформируемой поверхности может использоваться поверхность (3.16.1), поверхность на базе триангуляции (3.17.2), а также NURBS-поверхность (3.8.1).

Деформируемая поверхность (3.16.1) может быть построена для любой поверхности. Для многих поверхностей могут быть построены NURBS-копии, геометрически совпадающие с копируемыми поверхностями. Для поверхностей, не имеющих NURBS-представления, можно построить аппроксимирующие NURBS-поверхности. Для любой поверхности может быть построена поверхность на базе триангуляции (3.17.2), которая с известной точностью аппроксимирует копируемую поверхность. После замены поверхностей кривые, на которых построены ребра тела, необходимо заменить кривыми пересечения соответствующих деформируемых поверхностей.

Форму тела, построенного на деформируемых поверхностях, можно локально модифицировать. На рис. 6.13.3 приведено тело до локального деформирования. На рис. 6.13.4 приведено тело после локального деформирования его граней.



Рис. 6.13.3



Рис. 6.13.4

Деформирование граней может быть значительным, поэтому особое внимание следует уделить управлению деформированием. Пример неуправляемого деформирования приведен на рис. 6.13.5.



Рис. 6.13.5

Алгоритм локального деформирования имеет много общего с алгоритмом построения эквидистантного тела и состоит из следующих действий. После деформирования поверхностей граней построим новые вершины путем определения точек пересечения граней и используя в

качестве начальных приближений информацию об исходных вершинах. Далее построим новые ребра, соединяющие новые вершины, взамен исходных путем построения кривых пересечения граней. Затем построим циклы граней. На последнем шаге перестроим грани по новым циклам и изменим области определения поверхностей граней.

Упражнения

1. Дайте определение булевых операций над телами.
2. Опишите алгоритм определения положения точки относительно тела.
3. Опишите алгоритм операции сверления отверстия в теле.
4. Опишите алгоритм деформирования плоской грани тела.

Глава 7. Геометрические ограничения

Сложная геометрическая модель может состоять из большого количества тел и других элементов, которые должны быть связаны между собой определенным образом. Геометрические ограничения моделируют реальные связи между элементами модели. Геометрические ограничения позволяют создавать подобные модели и управлять геометрической моделью. Геометрические ограничения являются частью геометрической модели.

7.1. Наложение геометрических ограничений

Каждый элемент геометрической модели имеет свою структуру данных. Скалярные величины, компоненты векторов, координаты точек, лежащие в структурах данных и подлежащие редактированию, могут служить параметрами модели. Параметрами модели могут служить специально созданные для этого переменные. Посредством параметров можно связать элементы модели друг с другом и управлять геометрической моделью. Например, детали в сборочной единице можно расположить на одной оси, в одной плоскости, можно соединить шарнирами или закрепить в определенных направлениях. Размеры детали можно связать с размерами других деталей или связать уравнениями с любыми другими параметрами.

Геометрическими ограничениями называют условия на элементы геометрической модели, которые сформулированы в виде уравнений и неравенств. Геометрические ограничения моделируют реальные зависимости элементов модели друг от друга. Геометрические ограничения называют также *вариационными связями*.

Геометрические ограничения необходимы для модели, приведенной на рис. 7.1.1, чтобы обеспечить движение элементов модели строго определенным образом.



Рис. 7.1.1

Геометрические ограничения позволяют упростить моделирование подобных деталей. Для этого достаточно создать зависимости размеров детали от одного или нескольких параметров. Задав набор значений управляющим параметрам, получим набор деталей определенного типа. На рис. 7.1.2 приведена модель, на размеры которой наложены зависимости от нескольких параметров. На рис. 7.1.3 приведена подобная модель с другим набором параметров.



Рис. 7.1.2



Рис. 7.1.3

Каждое геометрическое ограничение представим в виде объекта, описывающего одно или несколько алгебраических уравнений и содержащего перечень параметров, связанных каждым уравнением. Описание уравнений заключается в том, что при обращении к объекту можно получить информацию об уравнениях для текущих значений параметров, например, невязки уравнений и частные производные этих невязок по параметрам. Неизвестными уравнений геометрических ограничений служат изменения параметров модели. Одни и те же параметры могут участвовать в нескольких ограничениях. Геометрическое ограничение считается выполненным, если описываемые объектом алгебраические уравнения удовлетворяются и их невязки, соответственно, равны нулю.

Равновесным состоянием геометрической модели будем называть состояние, при котором все геометрические ограничения выполняются. В процессе моделирования приходится модифицировать элементы модели или добавлять новые элементы. Если в результате модификации модели часть геометрических ограничений перестает выполняться, то геометрические ограничения обеспечивают поиск новых значений параметров, при которых геометрические ограничения опять выполняются. Для этого геометрическими ограничениями формируется система уравнений, неизвестными которой являются изменения параметров модели, решается эта система уравнений и перестраивается модель в соответствии с новыми значениями параметров.

Имитировать движение модели, приведенной на рис. 7.1.1, можно повернув вал на небольшой угол и зафиксировав его. Геометрические ограничения перестанут выполняться. Решение системы уравнений геометрических ограничений найдет изменения параметров, при которых

геометрические ограничения опять будут выполняются. После изменения параметров элементы геометрической модели изменят свое положение относительно друг друга. Продолжая поворачивать вал, получим имитацию движения элементов модели.

Над геометрической моделью мы можем выполнять действия, связанные с редактированием модели, с созданием подобных элементов модели, с изменением положения элементов модели друг относительно друга и так далее. Если на параметры модели наложены геометрические ограничения, то этими действиями мы можем управлять.

Управление геометрической моделью заключается в модификации модели с помощью геометрических ограничений. Управление геометрической моделью происходит следующим образом. На некотором этапе геометрические ограничения выполняются. Далее одному или нескольким параметрам присваивается новое значение, и эти параметры фиксируются. Например, мы переместили некоторый элемент геометрической модели и хотим узнать, как это повлияет на другие связанные с ним элементы модели. Управление геометрической моделью может происходить путем изменения констант уравнений, что также выведет систему геометрических ограничений из состояния равновесия. Например, мы изменили значение некоторого размера при создании подобной модели. В результате для исходных значений параметров некоторые уравнения перестанут выполняться. Система геометрических ограничений выходит из равновесия также при добавлении к ней новых ограничений. Например, мы смоделировали две детали и при создании из них сборочной единицы хотим, чтобы некоторые две плоскости деталей были параллельны друг другу. При этом рождается новое геометрическое ограничение, которое добавляет в систему уравнений свои уравнения и свои неизвестные.

Пусть задан набор геометрических ограничений, приводящий к системе m уравнений

$$\begin{aligned} f_1(p_1, p_2, \dots, p_n) &= 0, \\ f_2(p_1, p_2, \dots, p_n) &= 0, \\ &\dots \dots \dots \\ f_m(p_1, p_2, \dots, p_n) &= 0 \end{aligned}$$

для n параметров p_1, p_2, \dots, p_n геометрической модели. Число уравнений, как правило, не равно числу связанных параметров. Выведенную из равновесия систему уравнений геометрических ограничений запишем в виде

$$\begin{aligned} f_1(p_1 + \Delta p_1, p_2 + \Delta p_2, \dots, p_n + \Delta p_n) &= 0, \\ f_2(p_1 + \Delta p_1, p_2 + \Delta p_2, \dots, p_n + \Delta p_n) &= 0, \\ &\dots \dots \dots \\ f_m(p_1 + \Delta p_1, p_2 + \Delta p_2, \dots, p_n + \Delta p_n) &= 0. \end{aligned} \tag{7.1.1}$$

Неизвестными этой системы являются изменения параметров модели $\Delta p_1, \Delta p_2, \dots, \Delta p_n$. После решения системы уравнений (7.1.1) следует перестроить модель в соответствии с новыми значениями параметров $p_i + \Delta p_i$, $i = 1, 2, \dots, n$. Если выведенную из равновесия систему уравнений геометрических ограничений решить не удастся, то следует вернуться к предыдущему равновесному состоянию геометрической модели и констант уравнений, при которых уравнения геометрических ограничений удовлетворяются.

Как правило, геометрические ограничения делятся на трехмерные и двумерные. Трехмерные геометрические ограничения связывают элементы отдельных сборочных единиц. Двумерные геометрические ограничения связывают элементы отдельных эскизов. Системы уравнений сборочных единиц и системы уравнений эскизов решаются независимо. Все системы могут быть связаны глобальными переменными модели, которые к моменту формирования отдельных систем уравнений должны быть известны.

7.2. Объекты геометрических ограничений

Геометрическими ограничениями можно связать любые параметры модели. Для этого система уравнений должна иметь возможность изменять эти параметры в процессе решения и получать отклик на эти изменения. В процессе решения системы уравнений, как правило, геометрические ограничения изменяют не непосредственно параметры модели, а параметры объектов, которые мы будем называть специальными. **Специальные объекты** заменяют элементы модели в геометрических ограничениях. Специальные объекты являются упрощенными копиями элементов модели. Количество параметров, которые предоставляет специальный объект системе уравнений для изменения, должно быть равно количеству степеней свободы объекта. После того как уравнения геометрических ограничений будут удовлетворены, специальные объекты передают информацию геометрической модели.

В качестве специального объекта может использоваться точка. Точка имеет три степени свободы и описывается тремя числами. Каждая специальная точка может управлять положением нескольких точек модели. Например, если некоторая специальная точка передает свои координаты центру некоторой сферы модели и контрольной точке некоторой сплайновой поверхности модели, то центр сферы и контрольная точка поверхности будут совпадать при любых модификациях модели. Каждая специальная точка после достижения системой уравнений равновесного состояния приравнивает координаты связанных с ней точек модели своим координатам.

В качестве специального объекта может использоваться плоскость. Плоскость имеет три степени свободы, и ее параметрами служат три числа. В качестве специального объекта может использоваться сфера. Сфера имеет четыре степени свободы, включая изменение радиуса, и описывается четырьмя числами. Прямую линию можно получить с помощью двух пересекающихся плоскостей. С помощью сферы и пересекающей ее плоскости можно получить окружность. В качестве специальных объектов могут использоваться и другие геометрические объекты.

Рассмотрим некоторые геометрические ограничения, которые можно создать с использованием точек в качестве специальных объектов. Пусть имеются точки \mathbf{p}_1 и \mathbf{p}_2 , глобальные координаты которых соответственно равны x_1, y_1, z_1 и x_2, y_2, z_2 .

Геометрическое ограничение, наложенное на точки \mathbf{p}_1 и \mathbf{p}_2 и представляющее собой линейный размер, описывается уравнением

$$\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2} - d = 0, \quad (7.2.1)$$

где d – требуемый размер; $\Delta x_1, \Delta y_1, \Delta z_1$ и $\Delta x_2, \Delta y_2, \Delta z_2$ – подлежащее определению изменение координат связанных точек. Линейный размер приведен на рис. 7.2.1.

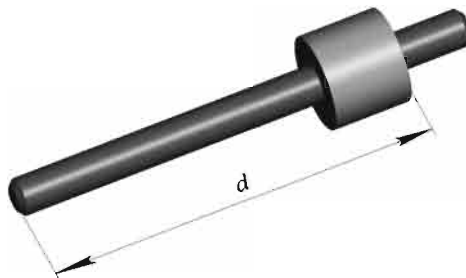


Рис. 7.2.1

Геометрическое ограничение, наложенное на точки \mathbf{p}_1 и \mathbf{p}_2 и представляющее собой размер вдоль одной координаты, описывается одним из уравнений

$$\begin{aligned} |x_2 + \Delta x_2 - x_1 - \Delta x_1| &= d_x, \\ |y_2 + \Delta y_2 - y_1 - \Delta y_1| &= d_y, \\ |z_2 + \Delta z_2 - z_1 - \Delta z_1| &= d_z, \end{aligned}$$

где d_x, d_y, d_z – требуемые размеры. Одно из указанных уравнений, примененное к отрезку $\mathbf{r}(t) = (1 - t)\mathbf{p}_1 + t\mathbf{p}_2$, может сделать отрезок параллельным соответствующей координатной плоскости. Два из указанных урав-

нений могут сделать отрезок параллельным одной из координатных осей.

Геометрическое ограничение, обеспечивающее совпадение точек \mathbf{p}_1 и \mathbf{p}_2 , описывается уравнениями

$$\begin{aligned}x_2 + \Delta x_2 - x_1 - \Delta x_1 &= 0, \\y_2 + \Delta y_2 - y_1 - \Delta y_1 &= 0, \\z_2 + \Delta z_2 - z_1 - \Delta z_1 &= 0.\end{aligned}$$

Для фиксации некоторого параметра геометрической модели построим уравнение

$$p + \Delta p - p_0 = 0,$$

где p – текущее значение параметра; Δp – подлежащее определению изменение параметра p_0 – заданное значение параметра, которое должно быть сохранено в дальнейшем.

С помощью фиксирующих уравнений можно закрепить тело или некоторый размер. Уравнения для фиксированных параметров могут быть выделены из общей системы в самостоятельную группу. Система уравнений для фиксированных параметров и система уравнений для остальных параметров не зависят друг от друга и могут решаться отдельно.

Геометрические ограничения можно устанавливать между любыми элементами модели, и эти ограничения также могут быть любыми, лишь бы их можно было представить в виде некоторых уравнений для специальных объектов. Набор специальных объектов расширяют так, чтобы можно было наложить все необходимые геометрические ограничения на элементы модели. Например, мы хотим установить линейный размер между определенной точкой кривой $\mathbf{r}_1(t_1)$ и определенной точкой кривой $\mathbf{r}_2(t_2)$. Пусть для простоты кривыми являются отрезки

$$\mathbf{r}_1(t_1) = (1 - t_1) \mathbf{p}_1 + t_1 \mathbf{p}_3, \quad \mathbf{r}_2(t_2) = (1 - t_2) \mathbf{p}_2 + t_2 \mathbf{p}_4,$$

опорные точки которых связаны со специальными точками, а в качестве связываемых точек мы рассмотрим середины отрезков. Середины отрезков определяются параметрами $t_1 = 1/2$, $t_2 = 1/2$. Линейный размер (7.2.1) для середины отрезков свяжет двенадцать параметров, которыми служат изменения координат специальных точек, управляющих опорными точками отрезков.

Параметры геометрической модели могут быть связаны алгебраическими уравнениями, например,

$$f(a + \Delta a, b + \Delta b, \dots, c + \Delta c) = 0,$$

где a, b, c – параметры модели; $\Delta a, \Delta b, \dots, \Delta c$ – искомые изменения параметров. Эти ограничения могут связывать не только описанные параметры модели, но и специальные переменные. **Специальные переменные**

создаются для управления геометрической моделью. Например, для того чтобы несколько параметров различных элементов геометрической модели сделать зависимыми от некоторого общего параметра, вводятся специальная переменная и совокупность алгебраических ограничений, описывающих требуемые зависимости параметров модели от введенной специальной переменной.

Параметры геометрической модели могут быть связаны алгебраическими неравенствами. В процессе решения системы уравнений геометрических ограничений неравенство заменяется соответствующим уравнением. Пусть, например, требуется связать параметры модели a и b неравенством $a > b$. Неравенство $a > b$ в процессе решения должно быть заменено равенством

$$a + \Delta a - b - \Delta b = e,$$

где $e = a + \Delta a - b - \Delta b$, если $a + \Delta a > b + \Delta b$ и $e = 0$, если $a + \Delta a \leq b + \Delta b$. Если на очередной итерации неравенство выполняется, то оно заменяется тождеством, если неравенство не выполняется, то оно заменяется строгим равенством. Геометрические ограничения, построенные на неравенствах, называют **нестрогими геометрическими ограничениями**.

7.3. Позиционирование набора тел

Для установки геометрических зависимостей между элементами модели некоторой сборочной единицы, детали которой представлены набором твердых тел, **в качестве специальных объектов геометрических ограничений удобно использовать локальные системы координат**. С каждым телом свяжем локальную систему координат, положение которой опишем в глобальной декартовой системе координат. Таким образом, позиционирование тел друг относительно друга сведем к изменению положения связанных с телами систем координат.

Каждое тело в трехмерном пространстве обладает шестью степенями свободы: тремя поступательными и тремя вращательными. Для каждого тела введем шесть параметров: три угла поворота и три перемещения. Через Δx_j , Δy_j , Δz_j обозначим перемещения j -го тела вдоль первой, второй и третьей глобальных осей соответственно, а через $\Delta \alpha_j$, $\Delta \beta_j$, $\Delta \gamma_j$ обозначим углы поворота j -го тела вокруг осей, параллельных первой, второй и третьей глобальным осям соответственно, и проходящих через начальную точку локальной системы координат. Назовем эти оси вращения первой, второй и третьей, соответственно. Если положение j -го тела не требует изменений, то $\Delta \alpha_j$, $\Delta \beta_j$, $\Delta \gamma_j$, Δx_j , Δy_j , Δz_j равны нулю. Если требуется изменить положение j -го тела, то это изменение выполним за четыре шага: повернем тело вокруг первой оси на угол $\Delta \alpha_j$, повернем тело вокруг

второй оси на угол $\Delta\beta$, повернем тело вокруг третьей оси на угол $\Delta\gamma$ и переместим тело на вектор $[\Delta x, \Delta y, \Delta z]^T$. Пример локальной системы координат и параметров тела приведен на рис. 7.3.1.

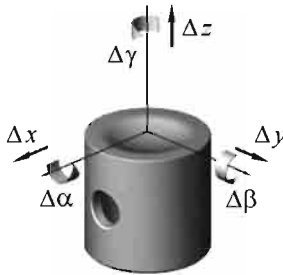


Рис. 7.3.1

После изменения положения j -го тела точка \mathbf{a} тела будет описываться радиусом-вектором $\mathbf{o}_j + \mathbf{M}_j(\mathbf{a} - \mathbf{o}_j)$, где \mathbf{o}_j – начальная точка j -ой локальной системы координат, а вектор \mathbf{n} j -го тела примет значение $\mathbf{M}_j\mathbf{n}$. Матрицу \mathbf{M}_j представим в расширенном виде (1.4.2), а радиусы-векторы точек и векторы – в виде (1.4.3) и (1.4.4) соответственно. Матрица \mathbf{M}_j описывает изменение положения j -го тела и выражается через величины $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$, Δx_j , Δy_j , Δz_j . Матрицу \mathbf{M}_j получим путем произведения трех матриц поворота и матрицы перемещения: $\mathbf{M}_j = \mathbf{M}_\Delta \mathbf{M}_\gamma \mathbf{M}_\beta \mathbf{M}_\alpha$, где

$$\mathbf{M}_\Delta = \begin{bmatrix} 1 & 0 & 0 & \Delta x_j \\ 0 & 1 & 0 & \Delta y_j \\ 0 & 0 & 1 & \Delta z_j \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ – матрица перемещения на вектор } [\Delta x_j, \Delta y_j, \Delta z_j]^T;$$

$$\mathbf{M}_\gamma = \begin{bmatrix} \cos \Delta\gamma_j & -\sin \Delta\gamma_j & 0 & 0 \\ \sin \Delta\gamma_j & \cos \Delta\gamma_j & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ – матрица поворота на угол } \Delta\gamma_j;$$

$$\mathbf{M}_\beta = \begin{bmatrix} \cos \Delta\beta_j & 0 & \sin \Delta\beta_j & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \Delta\beta_j & 0 & \cos \Delta\beta_j & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ – матрица поворота на угол } \Delta\beta_j;$$

$$\mathbf{M}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\Delta\alpha & -\sin\Delta\alpha & 0 \\ 0 & \sin\Delta\alpha & \cos\Delta\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \text{матрица поворота на угол } \Delta\alpha,$$

При наложении геометрических ограничений на тела модели будем использовать **маркеры**. Маркер помогает получить положения точек и направления векторов в глобальной системе координат. Маркер опишем начальной точкой, вектором нормали и вектором направления, ортогональным вектору нормали. Положение маркера в локальной системе координат, связанной с телом, должно быть известным и неизменным. Зная положение маркера в локальной системе координат, мы сможем получить координаты его начальной точки и компоненты его векторов в глобальной системе координат, в которой формируются уравнения геометрических ограничений. Для каждого геометрического ограничения, связывающего два тела, будем строить два маркера. Для каждого тела может быть построено столько маркеров, сколько геометрических ограничений связывают это тело с другими телами. С помощью маркеров будем составлять уравнения, комбинируя которые можно построить геометрические ограничения.

Для совмещения заданной точки \mathbf{a}_j j -го тела с заданной точкой \mathbf{a}_k k -го тела построим два маркера с началами в совмещаемых точках и потребуем выполнение векторного равенства

$$\mathbf{o}_j + \mathbf{M}_j(\mathbf{a}_j - \mathbf{o}_j) = \mathbf{o}_k + \mathbf{M}_k(\mathbf{a}_k - \mathbf{o}_k).$$

Данное уравнение снимает три степени свободы у одного из тел.

Для установки расстояния d между заданной точкой \mathbf{a}_j j -го тела и заданной точкой \mathbf{a}_k k -го тела построим два маркера с началами в совмещаемых точках и потребуем выполнение уравнения

$$|\mathbf{o}_j + \mathbf{M}_j(\mathbf{a}_j - \mathbf{o}_j) - \mathbf{o}_k - \mathbf{M}_k(\mathbf{a}_k - \mathbf{o}_k)| = d.$$

Для расположения заданной точки \mathbf{b}_j j -го тела на расстоянии d от заданной плоскости k -го тела построим маркер с началом в точке \mathbf{b}_j и маркер с началом и нормалью, совпадающими с началом и нормалью заданной плоскости k -го тела, и потребуем выполнение равенства

$$(\mathbf{o}_j + \mathbf{M}_j(\mathbf{b}_j - \mathbf{o}_j) - \mathbf{o}_k) \cdot (\mathbf{M}_k \mathbf{n}_k) = d,$$

где \mathbf{n}_k — нормаль заданной плоскости k -го тела. Данное уравнение снимает одну степень свободы у одного из тел. В частном случае при $d = 0$ приведенное уравнение описывает совпадение точки и плоскости.

Для совмещения заданной точки \mathbf{c}_j j -го тела с заданной осью k -го тела построим маркер с началом в точке \mathbf{c}_j и маркер с началом и нормалью,

совпадающими с началом и направлением заданной оси k -го тела, и потребуем выполнение равенства

$$(\mathbf{o}_j + \mathbf{M}_j(\mathbf{c}_j - \mathbf{o}_j) - \mathbf{o}_k) \cdot (\mathbf{M}_k \mathbf{d}_k) = 0,$$

где \mathbf{d}_k — вектор направления маркера k -го тела, ортогональный заданной оси. Данное уравнение снимает две степени свободы у одного из тел.

Для параллельности заданной оси j -го тела и заданной плоскости k -го тела построим два маркера. Начало и нормаль первого маркера совместим с началом и направлением заданной оси j -го тела, а начало и нормаль второго маркера совместим с началом и нормалью заданной плоскости k -го тела и потребуем выполнение равенства

$$(\mathbf{M}_j \mathbf{n}_j) \cdot (\mathbf{M}_k \mathbf{n}_k) = 0,$$

где \mathbf{n}_j — направление заданной оси j -го тела; \mathbf{n}_k — нормаль заданной плоскости k -го тела. Данное уравнение снимает одну степень свободы у одного из тел.

Для установки заданного угла θ , $0 < \theta < \pi$ между заданной плоскостью j -го тела и заданной плоскостью k -го тела построим два маркера с началами и нормальями, совпадающими с началами и нормальями заданных плоскостей, и потребуем выполнение равенства

$$(\mathbf{M}_j \mathbf{n}_j) \cdot (\mathbf{M}_k \mathbf{n}_k) = \cos \theta,$$

где \mathbf{n}_j — нормаль заданной плоскости j -го тела; \mathbf{n}_k — нормаль заданной плоскости k -го тела. Данное уравнение снимает одну степень свободы у одного из тел. В частном случае при $\cos \theta = 0$ приведенное уравнение описывает ортогональность плоскостей двух тел.

Для параллельности заданной плоскости j -го тела и заданной плоскости k -го тела построим два маркера с началами и нормальями, совпадающими с началами и нормальями заданных плоскостей, и потребуем выполнение равенств

$$(\mathbf{M}_j \mathbf{n}_j) \cdot (\mathbf{M}_k \mathbf{d}_k) = 0, \quad (\mathbf{M}_j \mathbf{n}_j) \cdot (\mathbf{M}_k (\mathbf{n}_k \times \mathbf{d}_k)) = 0,$$

где \mathbf{n}_j — нормаль заданной плоскости j -го тела; \mathbf{n}_k — нормаль заданной плоскости k -го тела; \mathbf{d}_k — вектор направления маркера k -го тела, ортогональный нормали \mathbf{n}_k . Данные уравнения снимают две степени свободы у одного из тел.

С помощью приведенных уравнений построим геометрические ограничения, устанавливающие зависимости между элементами модели. Решив систему уравнений геометрических ограничений, найдем изменения параметров модели. Скорректируем модель в соответствии с найденным решением. Например, если $\Delta \alpha_j = \Delta \beta_j = \Delta \gamma_j = \Delta x_j = \Delta y_j = \Delta z_j = 0$, то мат-

рица M_j равна единичной матрице и положение j -го тела не требует изменений. В противном случае повернем и переместим j -е тело в соответствии с матрицей M_j . В результате уравнения, наложенные геометрическими ограничениями, будут выполнены.

7.4. Решение уравнений геометрических ограничений

Особенностью системы уравнений геометрических ограничений является неравенство числа уравнений числу параметров. Число параметров может достигать нескольких тысяч. При этом геометрическая модель должна быстро реагировать на выход геометрических ограничений из равновесного состояния.

Решение системы уравнений геометрических ограничений осложняется неравенством числа уравнений и числа неизвестных. Если число уравнений меньше числа неизвестных, то система геометрических ограничений имеет множество решений, из которых следует выбрать наиболее подходящее. Если число уравнений превосходит число неизвестных, то следует найти геометрические ограничения, которые могут быть выполнены, а остальные ограничения следует отбросить. Система уравнений геометрических ограничений в общем случае решается численными методами. В определенных случаях система уравнений геометрических ограничений может быть решена без привлечения численных методов.

При использовании численных методов решение системы уравнений геометрических ограничений зависит от начального приближения, с которого начинается итерационный процесс. Чтобы начальное приближение находилось в области сходимости к искомому решению, следует не допускать сильных отклонений геометрических ограничений от их равновесных состояний. Например, если геометрические ограничения выведены из состояния равновесия путем перемещения некоторой точки геометрического объекта, то решение следует искать не для окончательного положения точки, а для нескольких промежуточных ее положений, постепенно сдвигая точку от исходного положения к конечному. Аналогично при значительном изменении значения некоторого размера следует последовательно искать решения для нескольких промежуточных значений этого размера, постепенно переходя от исходного значения к новому значению. При численном решении желательно, чтобы ненулевые приращения параметров были одного порядка. Для достижения этого используется нормирование параметров.

Для поиска реакции геометрических ограничений при их выходе из равновесного состояния используются методы, которые могут накладывать дополнительные геометрические ограничения, подчинять поведение параметров модели некоторым правилам, а также разбивать задачу

лей Лагранжа можно положить равными нулю. В процессе решения их также можно считать близкими к нулю, и это упростит решение системы уравнений.

Применим консервативный метод для определения изменения положений точки $\mathbf{p}_1 = [x_1, y_1, z_1]^T$ первого тела и точки $\mathbf{p}_2 = [x_2, y_2, z_2]^T$ второго тела при наложении геометрического ограничения (7.2.1), устанавливающего расстояние d между точками \mathbf{p}_1 и \mathbf{p}_2 . Сумма квадратов изменений перемещения связанных точек описывается функцией

$$\Psi = \frac{1}{2} (\Delta x_1^2 + \Delta y_1^2 + \Delta z_1^2 + \Delta x_2^2 + \Delta y_2^2 + \Delta z_2^2).$$

Используем метод неопределенных множителей Лагранжа для отыскания минимума функции ψ при условии (7.2.1). Необходимым условием минимума функции ψ при условии (7.2.1) является равенство нулю частных производных по параметрам $\Delta x_1, \Delta y_1, \Delta z_1, \Delta x_2, \Delta y_2, \Delta z_2$ функции

$$F = \frac{1}{2} (\Delta x_1^2 + \Delta y_1^2 + \Delta z_1^2 + \Delta x_2^2 + \Delta y_2^2 + \Delta z_2^2) + \lambda \left(\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2} - d \right),$$

где λ – подлежащий определению множитель. Искомые изменения координат точек и множитель λ найдем из системы уравнений

$$\begin{aligned} \sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2} - d &= 0, \\ \Delta x_1 + \lambda \frac{x_1 + \Delta x_1 - x_2 - \Delta x_2}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0, \\ \Delta y_1 + \lambda \frac{y_1 + \Delta y_1 - y_2 - \Delta y_2}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0, \\ \Delta z_1 + \lambda \frac{z_1 + \Delta z_1 - z_2 - \Delta z_2}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0, \\ \Delta x_2 + \lambda \frac{x_2 + \Delta x_2 - x_1 - \Delta x_1}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0, \\ \Delta y_2 + \lambda \frac{y_2 + \Delta y_2 - y_1 - \Delta y_1}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0, \\ \Delta z_2 + \lambda \frac{z_2 + \Delta z_2 - z_1 - \Delta z_1}{\sqrt{(x_2 + \Delta x_2 - x_1 - \Delta x_1)^2 + (y_2 + \Delta y_2 - y_1 - \Delta y_1)^2 + (z_2 + \Delta z_2 - z_1 - \Delta z_1)^2}} &= 0. \end{aligned}$$

Данная система уравнений имеет решение

$$\begin{aligned}\Delta x_1 &= \frac{d-g}{2g}(x_1 - x_2), \quad \Delta y_1 = \frac{d-g}{2g}(y_1 - y_2), \quad \Delta z_1 = \frac{d-g}{2g}(z_1 - z_2), \\ \Delta x_2 &= \frac{d-g}{2g}(x_2 - x_1), \quad \Delta y_2 = \frac{d-g}{2g}(y_2 - y_1), \quad \Delta z_2 = \frac{d-g}{2g}(z_2 - z_1), \\ \lambda &= \frac{g-d}{2},\end{aligned}$$

где $g = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. Можно доказать, что найденное решение является точкой минимума функции F . Точки переместятся вдоль прямой, проходящей через точки \mathbf{p}_1 и \mathbf{p}_2 , на одинаковое расстояние. В данной постановке задачи точки являются равноправными, а уравнения и решение системы уравнений являются симметричными.

Если в рассматриваемом примере закрепить точку \mathbf{p}_1 , то функция F примет вид

$$F = \frac{1}{2}(\Delta x_2^2 + \Delta y_2^2 + \Delta z_2^2) + \lambda \left(\sqrt{(x_2 + \Delta x_2 - x_1)^2 + (y_2 + \Delta y_2 - y_1)^2 + (z_2 + \Delta z_2 - z_1)^2} - d \right).$$

Эта функция достигает минимума при изменениях координат точки \mathbf{p}_2 и множителе λ :

$$\Delta x_2 = \frac{d-g}{g}(x_2 - x_1), \quad \Delta y_2 = \frac{d-g}{g}(y_2 - y_1), \quad \Delta z_2 = \frac{d-g}{g}(z_2 - z_1), \quad \lambda = g - d,$$

где $g = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$. Точка \mathbf{p}_2 переместится вдоль прямой, проходящей через точки \mathbf{p}_1 и \mathbf{p}_2 .

7.6. Метод декомпозиции

При использовании метода декомпозиции для решения системы уравнений геометрических ограничений исходная задача разбивается на ряд подзадач, последовательное решение которых позволяет решить исходную задачу. Это разбиение называют *декомпозицией*. Декомпозиция выполняется с помощью анализа графа задачи. В зависимости от цели декомпозиция может осуществляться по-разному. В общем случае результатом декомпозиции является разбиение исходной системы уравнений на подсистемы. Декомпозиция осуществляется с помощью анализа степеней свободы графа, построенного для элементов модели и геометрических ограничений. В одном из представлений геометрической модели в виде графа вершинами графа служат элементы модели и геомет-

рические ограничения. Ребро обозначает связь двух элементов модели посредством геометрических ограничений, наложенных на эти элементы. Каждому элементу геометрической модели сопоставляют число добавляемых степеней свободы g_i , а каждому геометрическому ограничению сопоставляют число снимаемых степеней свободы c_i . Для графа можно определить формальное число степеней свободы, которое равно величине $\Sigma g_i - \Sigma c_i$.

Пусть элементами модели являются твердые тела. Изменение положения тел выполняется путем поворотов и перемещений локальных систем координат, жестко связанных с телами. До закрепления каждое тело имеет шесть степеней свободы: три поступательных и три вращательных. Каждой степени свободы соответствует параметр, описывающий поворот или перемещение локальной системы координат тела. В системе уравнений геометрических ограничений каждому телу соответствуют шесть неизвестных. Модель, состоящая из n незакрепленных и несвязанных друг с другом тел, имеет $6n$ степеней свободы. Каждое геометрическое ограничение снимает некоторое число степеней свободы путем ограничения относительного движения тел модели. Геометрическое ограничение накладывается между некоторой парой тел. Если одно из тел рассматриваемой пары закрепить, то после наложения геометрических ограничений у другого тела снимется некоторое число степеней свободы. Если геометрические ограничения, наложенные между двумя телами, снимают все шесть степеней свободы, то пара тел будет жестко связана.

Минимальное число элементов модели, образующих нечто жесткое целое, равно двум. Максимальное число элементов модели, образующих нечто жесткое целое, может быть достаточно большим. Если определено, что некоторые элементы модели образуют нечто жесткое целое, то граф модели следует переписать, поместив в одну вершину жестко связанные элементы.

Пусть k_i — число геометрических ограничений, снимающих i степеней свободы модели, состоящей из n тел. Тогда формальное число степеней свободы модели, на которую наложены не противоречащие друг другу и не дублирующие друг друга геометрические ограничения, будет равно

$$m_n = 6n - k_1 - 2k_2 - 3k_3 - 4k_4 - 5k_5 - 6k_6. \quad (7.6.1)$$

Формальное число степеней свободы характеризует подвижность модели. Граф, построенный для модели, может содержать циклы. Если граф, построенный для некоторой совокупности элементов модели, представляет собой цикл и формальное число степеней свободы этих элементов меньше или равно шести, то данная совокупность элементов модели представляет нечто жесткое целое.

Минимальное число элементов цикла, не содержащего полностью закрепленных элементов и образующего нечто жесткое целое, равно трем. Если цикл не содержит полностью закрепленных элементов, то максимальное число элементов цикла, образующего нечто жесткое целое, равно шести. Действительно, в цикле число вершин равно числу ребер, поэтому $k_1 + k_2 + k_3 + k_4 + k_5 + k_6 = n$. Пусть цикл не содержит полностью закрепленных элементов, то есть $k_6 = 0$. Левая часть (7.6.1) принимает минимальное значение при $k_5 = n$ и $k_1 = k_2 = k_3 = k_4 = 0$. Таким образом, из равенства $b = 6n - 5n$ следует, что максимальное число элементов модели, образующих нечто жесткое целое, равно шести. Анализируя граф, можно найти группы элементов модели, представляющие собой жесткое целое, и упростить решение системы уравнений геометрических ограничений, выделив уравнения для найденных групп в отдельные подсистемы. Процесс поиска групп элементов модели, представляющих собой жесткое целое, входит в декомпозицию.

Подграф группы элементов модели и геометрических ограничений, образующих нечто жесткое целое, будем называть *кластером*. Результатом декомпозиции является выделение из системы уравнений кластеров.

С помощью формального анализа числа степеней свободы можно выделить подграфы, соответствующая подзадача которых может быть решена независимо. Для этого необходимо найти блоки и точки сочленения подграфов. Точкой сочленения может служить твердое тело в данном пространстве, через которое и только через которое одна группа объектов связана с другой группой объектов. Если для двух подграфов найдена точка сочленения, то задачу можно разбить на две независимые подзадачи и решить их по очереди.

В качестве примера построим граф для двумерного эскиза, приведенного на рис. 7.6.1. Точка p_1 эскиза закреплена. Точка p_2 связана с точкой p_1 горизонтальным размером a и горизонтальностью отрезка p_1p_2 . Точка p_3 связана с точкой p_2 вертикальным размером b и фиксированным углом $p_1p_2p_3$. Точка p_4 связана с точкой p_1 вертикальностью отрезка p_1p_4 и горизонтальностью отрезка p_3p_4 . Граф элементов эскиза и геометрических ограничений эскиза приведен на рис. 7.6.2.

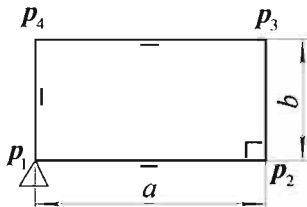


Рис. 7.6.1

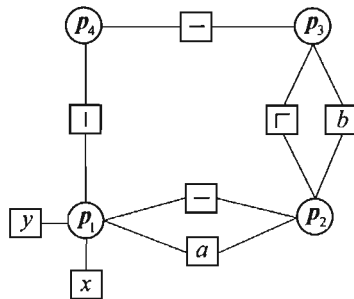


Рис. 7.6.2

Эскиз имеет восемь степеней свободы. Геометрические ограничения снимают все восемь степеней свободы. Формальное число степеней свободы эскиза равно нулю. Кластером в данном примере является подграф, состоящий из точки p_1 , точки p_2 , точки p_3 , горизонтального размера a , вертикального размера b и угла p, p, p_3 . Если убрать закрепление точки p_1 , то кластер получит две степени свободы и может быть перемещен в двумерном пространстве. Если убрать горизонтальность отрезка p, p_2 , то кластер получит еще одну степень свободы и может быть повернут. Мы видим, что указанный кластер ведет себя, как твердое тело двумерного пространства. В данном примере кластер зафиксирован. С помощью декомпозиции полная система уравнений может быть разбита на две подсистемы: система уравнений для кластера и система уравнений для остальных объектов. Сначала должна быть решена система уравнений для кластера, а затем должна быть решена система уравнений для точки p_4 .

Процедура декомпозиции применяется к системе уравнений итерационно, пока после очередного разбиения обнаруживается хотя бы один кластер. На каждой итерации происходит переписывание графа системы уравнений. Процедура прекращается, когда не удается найти больше ни одного кластера. Таким образом, декомпозиция может быть многоуровневой. На каждом уровне задача будет иметь размерность, меньшую размерности исходной задачи. Анализ графа является мощным средством декомпозиции, но он не использует геометрическую информацию и опирается только на дискретные данные, поэтому корректность результатов декомпозиции зависит от того, насколько правильно построен граф, отображающий геометрические ограничения и степени свободы элементов модели.

В результате декомпозиции исходная задача сводится к задаче меньшей размерности для кластеров и объектов, не вошедших в кластеры. Сначала определяется удовлетворяющее геометрическим ограничениям положение элементов модели для каждого кластера в отдельности, а затем решается соответствующая система уравнений для кластеров и элементов модели, не вошедших в кластеры. Размещение элементов модели внутри одного кластера осуществляется путем решения соответствующей системы уравнений. Для каждого кластера используется свой наиболее эффективный метод.

После решения системы уравнений геометрических ограничений производится перестроение модели в соответствии с найденным решением.

Упражнения

1. Как с помощью геометрических ограничений связать элементы модели?
2. Как с помощью геометрических ограничений построить подобную модель?
3. Опишите алгоритм решения системы уравнений геометрических ограничений.
4. Сформулируйте свой критерий поведения параметров геометрической модели.

Глава 8. Геометрическая модель

Геометрические модели нужны для проведения исследований и производства моделируемых объектов. С помощью геометрической модели можно увидеть, как будет выглядеть моделируемый объект, проверить его работоспособность и исследовать поведение моделируемого объекта в реальных условиях эксплуатации путем расчетов и постановки численных экспериментов. По геометрической модели можно подготовить производство и создать моделируемый объект. Кроме того, геометрическую модель можно многократно редактировать и строить на ее основе подобные модели с минимальными затратами времени и других ресурсов.

8.1. Состав геометрической модели

Геометрическая модель содержит **описание формы** моделируемого объекта. Для описания формы моделируемого объекта мы использовали граничное представление. Граничное представление точно описывает поверхность моделируемого объекта. Если геометрическая модель состоит из множества элементов, которые должны согласованно двигаться или редактироваться, то в геометрическую модель включают **геометрические ограничения**. Если требуется многократно повторять построение геометрической модели с целью ее модификации или создания подобных моделей, то в геометрическую модель включают описание процесса построения модели. Перечень действий, отражающий алгоритмическую последовательность построения модели, называется **деревом построения**. Дерево содержит все необходимые для построения модели данные. Если моделирование объекта выполняется для нескольких целей, то математическая модель описывает не только геометрические свойства объекта, но и другие его свойства. Геометрическая модель может содержать в виде **атрибутов** необходимые физические свойства своих элементов, например, плотность, модуль упругости, температуру, прозрачность.

Таким образом, геометрическая модель в общем случае включает в себя описание формы моделируемого объекта, геометрические ограничения, дерево построения, атрибуты, рис. 8.1.1.



Рис. 8.1.1

В определенных случаях геометрическая модель может содержать только описание формы моделируемого объекта. Описания формы достаточно, если модель не требуется редактировать или строить с ее помощью подобные модели, а также если не требуется использовать модель в качестве механизма. В таких случаях геометрические ограничения, дерево построения модели и атрибуты могут отсутствовать. Для описания геометрической формы существует несколько форматов данных, которые поддерживаются системами геометрического моделирования. При передаче данных из одной системы геометрического моделирования в другую систему в том или ином виде передается описание формы, а геометрические ограничения, дерево построения и атрибуты часто опускаются.

8.2. Применение геометрической модели

Геометрическая модель позволяет определить площадь поверхности, объем, инерционные характеристики проектируемого объекта, выполнить измерения длин и углов его элементов. Она дает возможность произвести расчет размерных цепей и определить собираемость проектируемого объекта. Если объект представляет собой механизм, то на модели можно проверить его работоспособность и выполнить расчет кинематических характеристик.

При моделировании процессов сплошной среды используют модели, состоящие из конечных элементов. Для этого внутренний объем модели разбивают на конечные элементы — пирамиды и призмы, стыкующиеся между собой по общим граням и вершинам. Конечные элементы наделяют необходимыми физическими и другими свойствами. Получить модель, состоящую из конечных элементов, можно на основе геометрической модели в виде оболочки. Таким образом, используя

геометрическую модель, можно поставить численный эксперимент по определению напряженно-деформированного состояния, частот и форм собственных колебаний, устойчивости элементов конструкции, тепловых, оптических и других свойств моделируемого объекта.

По геометрической модели можно вычислить траекторию режущего инструмента для механической обработки объекта. При выбранной технологии изготовления объекта геометрическая модель позволяет спроектировать оснастку и выполнить подготовку производства, а также проверить саму возможность изготовления объекта данным способом и качество этого изготовления. Кроме того, возможна графическая имитация процесса изготовления. Для того чтобы изготовить объект, кроме геометрической информации нужна информация о технологическом процессе, производственном оборудовании и многом другом, связанном с производством.

Многие из перечисленных проблем образуют самостоятельные разделы прикладной науки и по своей сложности не уступают проблеме создания геометрической модели. Геометрическая модель является отправной точкой для дальнейших действий. При построении геометрической модели мы не использовали физические законы, радиус-вектор каждой точки границы раздела внешней и внутренней частей моделируемого объекта является известным, поэтому при построении геометрической модели нам приходится составлять и решать алгебраические уравнения. Задачи, в которых используются физические законы, приводят к дифференциальным и интегральным уравнениям.

Геометрические модели используются в компьютерной графике. Увидеть модель можно, построив ее векторное или точечное изображение. Для построения изображения геометрической модели, как правило, используются упрощенные копии модели. Для построения векторного изображения на основе геометрической модели строится каркас, который проецируется на некоторую поверхность изображения. Для построения точечного изображения используется аппроксимация поверхности геометрической модели треугольными пластинами. Кроме того, можно смоделировать поток падающих и возвращающихся от поверхности модели лучей света. При этом элементам модели можно придать необходимый цвет, прозрачность, фактуру, отражающие и рассеивающие характеристики, а также другие физические свойства. Модель можно осветить с разных сторон светом различного цвета и интенсивности.

Мы рассмотрим применение геометрической модели для построения изображений моделируемого объекта и вычисления его инерционных характеристик. Мы увидим, что геометрическая модель может содержать несколько различных описаний формы моделируемого объекта, используемых для различных целей.

8.3. Построение векторного изображения

Для построения векторного изображения моделируемого объекта кроме геометрической модели необходимо знать взаимное расположение модели, точки наблюдения и плоскости изображения. При построении векторного изображения моделируемого объекта используют каркасную модель.

Каркасную модель можно получить из геометрической модели в виде набора тел и оболочек, если в последней удалить грани, сохранив ребра и вершины. Векторные изображения будут более информативны, если они дополнены видимыми линиями очерка модели и в них не показаны невидимые линии каркаса. Линии очерка проходят внутри граней и делят их на видимые из точки наблюдения части и невидимые из точки наблюдения части. В общем случае для одной поверхности может существовать несколько линий очерка. Каждая линия очерка или замкнута, или оканчивается на краях поверхности. Для разных направлений взгляда существует своя совокупность линий очерка, поэтому при повороте модели линии очерка необходимо строить заново.

Задача построения векторных изображений сводится к задаче определения видимых участков кривых каркасной модели. На рис. 8.3.1 приведено изображение модели с удалением невидимых линий. Линии плавного перехода между гранями имеют меньшую толщину, чем линии ребер и линии очерка.

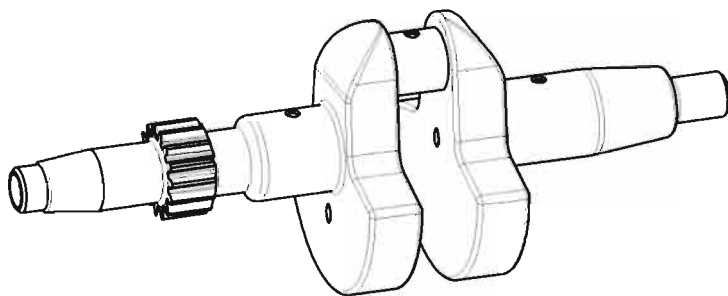


Рис. 8.3.1

С помощью линий очерка можно раскрыть поверхность отображаемого объекта на невидимые и условно видимые области. Пусть вектор \mathbf{i} направлен от точки наблюдения к рассматриваемой точке объекта. Тогда невидимыми будут те области поверхности, скалярное произведение нормали поверхности и вектора \mathbf{i} которых положительно. Области поверхности, скалярное произведение нормали поверхности и

вектора \mathbf{i} которых отрицательно, будут видимы при условии, что они не закрыты другими видимыми областями поверхности, более близкими к точке наблюдения. Аналогичным образом можно судить о видимости элементов модели. Если нормали стыкующихся в ребре граней на некотором участке направлены от точки наблюдения, то данный участок ребра является невидимым. В противном случае участок ребра может быть видим из точки наблюдения или может быть закрыт некоторой гранью. Таким образом, для некоторых элементов модели можно сразу сделать вывод о невидимости их из точки наблюдения, а для определения видимости других элементов модели нужно выполнить дополнительные действия.

Рассмотрим алгоритм определения видимости элементов геометрической модели. Построим каркасную модель объекта, состоящую из линий очерка и линий, на которых базируются ребра граней. Пусть каркасная модель состоит из пространственных кривых $\mathbf{c}_i(t)$, $i = 1, 2, \dots, n$. Для каждой пространственной кривой можно построить соответствующую ей параллельную или центральную проекцию на заданную плоскость. Далее определим точки кривых, в которых может меняться их видимость. Для этого найдем все точки, в которых пересекаются проекции пространственных кривых $\mathbf{c}_i(t)$, $i = 1, 2, \dots, n$. Этими точками разобьем кривые на участки. Теперь каждый участок кривой между точками разбивки будет или полностью видим или полностью невидим. Остается найти видимые участки кривых, построить для них проекции и удалить дублирующие проекции.

Для определения видимости участков кривых $\mathbf{c}_i(t)$, $i = 1, 2, \dots, n$ возьмем произвольную точку \mathbf{g} внутри каждого участка. Построим прямую $\mathbf{l}(t) = \mathbf{g} + t\mathbf{i}$, проходящую через точку \mathbf{g} и точку наблюдения и направленную от точки наблюдения. Для параллельной проекции вектор \mathbf{i} параллелен нормали плоскости проекции, для центральной проекции вектор \mathbf{i} пропорционален вектору $\mathbf{g} - \mathbf{w}$, где \mathbf{w} — точка наблюдения. Найдем точки пересечения этой прямой с поверхностями геометрической модели. Если значения параметра t прямой для всех точек пересечения неотрицательные, то рассматриваемый участок кривой является видимым. Если среди значений параметра t для точек пересечения найдется хотя бы одно отрицательное, то рассматриваемый участок кривой является невидимым, так как его закрывает одна из поверхностей модели, находящаяся ближе к наблюдателю, чем точка \mathbf{g} . Проведем эту процедуру с участками всех кривых, для которых сразу нельзя сделать заключение об их видимости из точки наблюдения.

При необходимости построим для видимых участков кривых каркасной модели плоские проекции.

Для построения векторного изображения кривой используют ломаные линии. Если для кривой построить набор точек, принадлежащих

кривой и следующих в порядке возрастания параметра, и соединить эти точки отрезками, то получим ломаную линию, аппроксимирующую заданную кривую. Чтобы ломаная линия выглядела плавной, ее участки должны быть достаточно малы.

Последовательность точек, по которым проходит аппроксимирующая кривую ломаная линия, называется **полигоном кривой**.

Проекция на плоскость пространственной ломаной линии также является ломаной линией. Поэтому кривые и поверхности, отображаемые компьютером, обычно сопровождаются полигонами, которые и используются в компьютерной графике.

Расстояния между точками полигона должны быть такими, чтобы полигон казался плавным. В качестве управляющего параметра примем максимальное расстояние между кривой и аппроксимирующей ее ломаной. Пусть задана кривая $c(t)$ и пусть мы находимся в некоторой ее точке t_0 , являющейся точкой полигона. Требуется найти параметр t_1 следующей точки полигона, так чтобы отклонение кривой от ее полигона не превышало заданную величину δ . Предположим, что в ближайшей окрестности кривая близка к своей соприкасающейся в этой точке окружности, рис. 8.3.2.

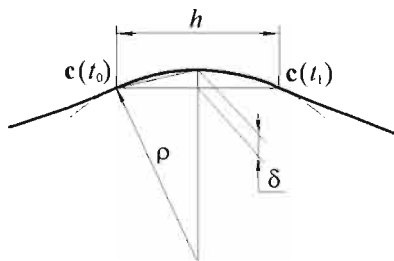


Рис. 8.3.2

Радиус ρ соприкасающейся окружности равен радиусу кривизны кривой и определяется формулой

$$\rho = \frac{|c'|^3}{|c \times c''|}.$$

По теореме Пифагора, квадрат половины длины хорды равен $\rho^2 - (\rho - \delta)^2 = \delta(2\rho - \delta)$. Длина всей хорды будет равна

$$h = 2\sqrt{\delta(2\rho - \delta)}.$$

Если кривая близка к дуге соприкасающейся окружности и угол дуги достаточно мал, то можно считать, что длина хорды примерно равна

длине дуги кривой, и приращение параметра Δl на длине дуги кривой можно принять равным

$$\Delta l \approx \frac{h}{|\mathbf{c}'|} = 2 \frac{\sqrt{\delta(2\rho - \delta)}}{|\mathbf{c}'|} = 2 \sqrt{\delta \left(\frac{2|\mathbf{c}'|}{|\mathbf{c}' \times \mathbf{c}''|} - \frac{\delta}{\mathbf{c}' \cdot \mathbf{c}''} \right)}. \quad (8.3.1)$$

Изменение параметра Δl будем называть параметрическим шагом. Таким образом, параметр следующей точки полигона может быть принят равным $t_1 = t_0 + \Delta l$. При резком изменении направления или длины второй производной в точке t_1 по сравнению с предыдущей точкой формула (8.3.1) дает ошибку, то есть прогиб может оказаться значительно больше δ . В этом случае шаг должен быть уточнен путем определения среднего радиуса кривизны кривой на рассматриваемом участке.

Формула вычисления параметрического шага между точками полигона (8.3.1) позволяет построить полигон кривой, отстоящий от самой кривой на величину, не превышающую δ . Проекция этого полигона будет отстоять от проекции кривой также на величину, не превышающую δ . Если требуется построить изображение в m раз больше натурального, величина δ должна быть уменьшена в m раз.

Вычислим шаг параметра кривой на поверхности для построения ее полигона. Пусть точка некоторого полигона поверхности $\mathbf{s}(u_0, v_0)$ определяется параметрами u_0 и v_0 . Найдем параметры u_1 и v_1 следующей точки поверхности так, чтобы пространственный отрезок из точки $\mathbf{s}(u_0, v_0)$ в точку $\mathbf{s}(u_1, v_1)$ отстоял от поверхности на расстоянии, не превышающем заданную величину δ . Для этого вычислим радиус кривизны линии на поверхности, имеющей направление вектора $[du \ dv]^T$. Радиус кривизны нормального сечения вдоль кривой на поверхности в направлении вектора $[du \ dv]^T$ определяется равенством

$$\rho = \frac{g_{11} du^2 + 2g_{12} dudv + g_{22} dv^2}{b_{11} du^2 + 2b_{12} dudv + b_{22} dv^2}, \quad (8.3.2)$$

где g_{11} , g_{22} и g_{12} — коэффициенты первой квадратичной формы поверхности; b_{11} , b_{22} и b_{12} — коэффициенты второй квадратичной формы поверхности. Подставим (8.3.2) в (8.3.1) и получим приращения для параметров соседней точки полигона:

$$\Delta u \approx 2 \frac{\sqrt{\delta(2\rho - \delta)}}{\left| \mathbf{s}_1 + \mathbf{s}_2 \frac{dv}{du} \right|}, \quad \Delta v \approx 2 \frac{\sqrt{\delta(2\rho - \delta)}}{\left| \mathbf{s}_1 \frac{du}{dv} + \mathbf{s}_2 \right|},$$

где $\mathbf{s}_1 = \partial \mathbf{s} / \partial u$, $\mathbf{s}_2 = \partial \mathbf{s} / \partial v$ — производные радиуса-вектора поверхности в рассматриваемой точке.

Радиус кривизны (8.3.2) нормального сечения поверхности в направлении изменения первого параметра равен $\rho_u = g_{11} / b_{11}$, а радиус кривизны (8.3.2) нормального сечения поверхности в направлении изменения второго параметра равен $\rho_v = g_{22} / b_{22}$. В соответствии с кривизной нормальных сечений поверхности приращение первого параметра поверхности для соседней точки полигона кривой на поверхности примем равным

$$\Delta u \approx 2 \frac{\sqrt{\delta \left(2 \frac{g_{11}}{b_{11}} - \delta \right)}}{\sqrt{g_{11}}}, \quad (8.3.3)$$

а приращение второго параметра поверхности для соседней точки полигона кривой на поверхности примем равным

$$\Delta v \approx 2 \frac{\sqrt{\delta \left(2 \frac{g_{22}}{b_{22}} - \delta \right)}}{\sqrt{g_{22}}}. \quad (8.3.4)$$

Изменения параметров Δu и Δv будем называть параметрическими шагами. При резком изменении направления или длины вторых производных радиуса-вектора поверхности формулы (8.3.3) и (8.3.4) дают ошибку, и шаги должны быть уточнены путем определения средней кривизны поверхности на рассматриваемом участке.

Для удобства полигоны, принадлежащие одному объекту, объединяют в сетки полигонов. Совокупность полигонов геометрической модели называют **сеткой полигонов**. Сетка полигонов кривой состоит из одного полигона. Сетка полигонов поверхности состоит из полигонов ее границ и, возможно, нескольких полигонов кривых на ней. Сетка полигонов тела состоит из полигонов поверхностей его граней и полигонов его ребер.

8.4. Моделирование оптических свойств

Для построения реалистичного изображения моделируемого объекта кроме геометрической модели необходимо знать ее излучающую способность, прозрачность, отражающие и рассеивающие характеристики для различных участков светового спектра. Кроме этого необходима информация о взаимном расположении модели, точки наблюдения, плоскости изображения, источников света, их спектре и интенсивности. Световой поток мы будем характеризовать интенсивностью — плотностью потока энергии световой волны.

Пришедший в точку наблюдения свет от моделируемого объекта можно условно разделить на *излученный* объектом свет, *отраженный* и *пропущенный* им свет от других источников. Реальные объекты некоторую часть падающего на них света поглощают, некоторую часть света пропускают и некоторую часть падающего света отражают. Распределение падающего на объект света на поглощенную часть, пропущенную часть и отраженную часть зависит от длины волны и от физических свойств объекта.

Большинство объектов, в пришедшем от которых свете преобладает излученный свет, излучают его равномерно во все стороны. Если излучающий объект является протяженным, то видимость его отдельных частей зависит от потока энергии, излучаемой соответствующей частью. Особенности формы таких объектов мы наблюдаем в виде контура, ограничивающего светящееся пятно.

Восприятие объектов, в пришедшем от которых свете преобладает отраженный и пропущенный свет, зависит от положения точки наблюдения, положения источников света, потока энергии и от оптических свойств поверхности объекта. При моделировании отраженного света его условно делят на две части: *диффузно отраженный* и *зеркально отраженный*. Это деление связано с законами, по которым описывается направление лучей отраженного света. Если большую часть отраженного поверхностью света можно описать законом диффузного отражения, то поверхность называют *матовой*, если же большую часть отраженного поверхностью света можно описать законом зеркального отражения, то ее называют *зеркальной*. Объекты, поверхности которых способны пропускать большую часть падающего на них света вглубь, называют *прозрачными*. Пропущенный через поверхность свет также условно делят на две части: *диффузно пропущенный* и *направленно пропущенный*. Направленно пропущенный свет претерпевает преломление. Пропущенный свет может быть отражен от других поверхностей.

Будем считать, что доля излученного и поглощенного моделируемым объектом света незначительна по сравнению с долей отраженного и пропущенного света. В этом случае приходящий в точку наблюдения свет разделим на четыре составляющие части: диффузно отраженный свет, зеркально отраженный свет, рассеянный свет, пропущенный свет.

Для диффузного отражения света используют закон косинусов Ламберта, устанавливающим соответствие между количеством отраженного света и косинусом угла θ между направлением на точечный источник света интенсивности I_p и нормалью поверхности. Интенсивность диффузно отраженного света определяется формулой

$$I_d = k_d I_p \cos \theta,$$

где k_d – коэффициент диффузного отражения, зависящий от материала поверхности. Количество диффузно отраженного света не зависит от положения наблюдателя. Матовые поверхности отражают одинаковое количество световой энергии во всех направлениях, но это количество пропорционально $\cos\theta$.

От зеркальной поверхности большая часть падающего света отражается в направлении, лежащем в одной плоскости с направлением падающего света и нормалью \mathbf{m} в точке падения и составляющем с нормалью \mathbf{m} угол θ , равный углу падения света, рис. 8.4.1. Это направление называется направлением отраженного света. При отклонении на некоторый угол α от этого направления интенсивность отраженного света резко уменьшается. Изменение интенсивности зеркально отраженного света около направления отраженного света Фонг предложил описывать функцией $\cos^n \alpha$, где n лежит в диапазоне от единицы до двух сотен.

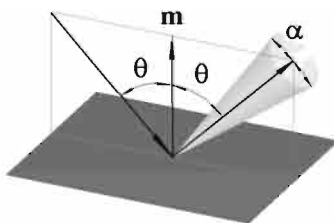


Рис. 8.4.1

Количество отраженного света зависит также и от угла падения θ , но эта зависимость является слабой. В рассматриваемой модели зависимость отраженного света от угла падения заменяют константой $f(\theta) = k_s$, которую выбирают опытным путем так, чтобы результаты были приемлемы с эстетической точки зрения. Если интенсивность источника света равна I_p , то интенсивность зеркально отраженного света определим формулой Фонга

$$I_s = k_s I_p \cos^n \alpha, \quad (8.4.1)$$

где k_s будем называть коэффициентом зеркального отражения, а n будем называть коэффициентом блеска. Чем больше n , тем четче и уже световой блик.

Интенсивность рассеянного света одинакова во всех направлениях, поэтому в точку наблюдения от каждой точки поверхности приходит одинаковое количество рассеянного света. При наличии только рассеянного света интенсивности I_a в точку наблюдения от каждой точки объ-

екта, независимо от ориентации в ней, поверхности придет свет интенсивности

$$I = k_a I_0,$$

где k_a – коэффициент, определяющий и диффузно, и зеркально отраженную долю рассеянного света.

Для простоты будем считать, что доля диффузно пропущенного света мала по сравнению с долей направленно пропущенного света, и опустим из рассмотрения диффузно пропущенный свет. Направленно пропущенный свет претерпевает преломление на поверхности раздела двух сред. Преломленный луч лежит в одной плоскости с падающим лучом и нормалью \mathbf{m} в точке падения, рис. 8.4.2.

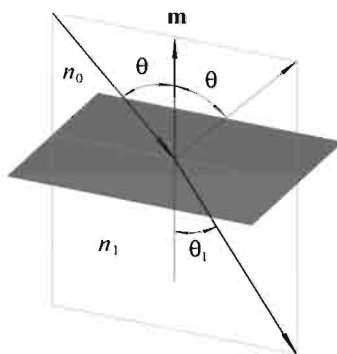


Рис. 8.4.2

Его направление составляет угол θ_1 с нормалью и описывается законом преломления

$$n_0 \sin \theta = n_1 \sin \theta_1,$$

где n_0 – показатель преломления среды со стороны падающего света; n_1 – показатель преломления среды с другой стороны поверхности раздела. Для интенсивности направленно пропущенного света примем закон изменения от направления, аналогичный закону (8.4.1) для направленно отраженного света

$$I_u = k_r I_i \cos^2 \beta,$$

где β – угол отклонения направления пропущенного света от направления в точку наблюдения.

От точки поверхности прозрачного объекта в точку наблюдения придет не только отраженный свет, но и свет, пропущенный поверхностью из глубины объекта. Пропущенный поверхностью свет может прийти в

результате отражения от другой поверхности или от источника света через прозрачный объект. В обоих случаях свет пройдет через вещество и будет им частично поглощен. Интенсивность света при прохождении через вещество убывает по экспоненциальному закону и описывается законом Бугера

$$I = I_0 e^{-\mu l},$$

где I_0 – интенсивность света на входе в поглощающее вещество; l – пройденное расстояние; μ – коэффициент поглощения света данным веществом.

Интенсивность света, пришедшего в точку наблюдения от неизлучающего и непрозрачного объекта, в рассматриваемой модели света определится суммой трех составляющих:

$$I = k_a I_a + I_p k_d \cos \theta + I_s k_s \cos^n \alpha. \quad (8.4.2)$$

Результаты закраски, полученные с помощью формулы (8.4.2), обладают большой реалистичностью, но они имеют существенный недостаток. Если две плоские поверхности одного цвета параллельны друг другу и их проекции перекрываются, то поверхности будут закрашены одинаково и их изображения сольются, независимо от их расстояния до точки наблюдения. Для устранения этого эффекта интенсивность пришедшего от поверхности света будем считать зависимой от расстояния и запишем формулу вычисления интенсивности света в виде

$$I = \frac{k_a I_a}{1 + k_r (r - r_0)} + I_p (k_d \cos \theta + k_s \cos^n \alpha), \quad (8.4.3)$$

где r – расстояние от точки наблюдения до точки поверхности; r_0 – расстояние от точки наблюдения до ближайшей точки объекта; k_r – коэффициент влияния удаленности точки объекта на интенсивность пришедшего от нее света. При $k_r = 0$ мы получим формулу (8.4.2). Интенсивность света от ближайшей точки также будет совпадать с (8.4.2). Расстояние r_0 введено из-за того, что для параллельных проекций точка наблюдения находится бесконечно далеко. Для центральных проекций можно положить $r_0 = 0$, тогда при удалении всего объекта от точки наблюдения будет наблюдаться уменьшение его яркости.

Если моделируемый объект является прозрачным, то в точку наблюдения кроме отраженного света придет пропущенный свет и формула (8.4.3) примет вид

$$I = \frac{k_a I_a}{1 + k_r (r - r_0)} + I_p (k_d \cos \theta + k_s \cos^n \alpha) + I_t k_t \cos^n \beta, \quad (8.4.4)$$

где I_t – интенсивность света, проникающего с другой стороны поверхности.

При наличии нескольких источников света каждый из них даст свой вклад в интенсивность пришедшего в точку наблюдения света. Для нескольких источников света, с интенсивностью каждого I_{ρ_i} , формула (8.4.3) примет вид

$$I = \frac{k_a I_a}{1 + k_r(r - r_0)} + \sum_i I_{\rho_i} (k_s \cos \theta_i + k_x \cos \alpha_i) .$$

Аналогичным образом изменится формула (8.4.4) вычисления интенсивности света, пришедшего от прозрачной поверхности. Она примет вид

$$I = \frac{k_a I_a}{1 + k_r(r - r_0)} + \sum_i I_{\rho_i} (k_s \cos \theta_i + k_x \cos \alpha_i) + \sum_j I_{\rho_j} k_t \cos \beta_j ,$$

где I_{ρ_j} – интенсивность j -го пропущенного луча света (проникающего с другой стороны поверхности).

Свет и его восприятие человеком представляют собой сложные явления, которые изучены не полностью. Из экспериментальных исследований известно, что свет с длиной волны из определенной части спектра воспринимается глазом в виде света, определенного цвета. Воспринимаемый цвет зависит от источника света, от свойств поверхности объекта и от зрительной системы человека. Зависимость воспринимаемой глазом интенсивности света от зрительной системы человека показана на рис. 8.4.3.

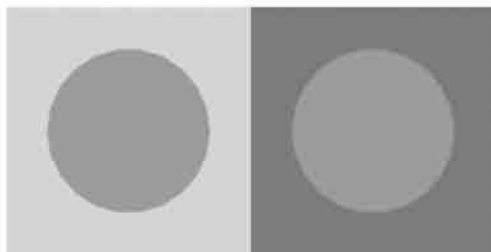


Рис. 8.4.3

Яркость круглых областей внутри двух квадратов одинакова, но круг на светлом фоне кажется более темным, чем такой же круг на темном фоне. Это явление называется *одновременным контрастом*. Подобное явление существует и при восприятии цветных изображений. В зависимости от цвета и яркости окружения цветные области могут иметь разный оттенок (казаться меняющимися цвет). Кроме того, границы областей

постоянной интенсивности кажутся более яркими, чем центральная часть областей. Это явление называется *эффектом полос Маха*.

Для получения цветных изображений используется представление света в виде трех компонент. Чувствительность глаза не одинакова к разным цветам: наибольшая она для зеленого цвета, наименьшая — для синего цвета. Кроме того, глаз воспринимает не абсолютное значение интенсивности того или иного цвета, а ее относительное значение. Следовательно, линейное распределение уровней интенсивности приходящего света в зрительной системе человека становится логарифмическим.

Интенсивность падающего света будем представлять в виде красной, зеленой и синей компонент. Положим, что в белом цвете все три компоненты имеют равные интенсивности. Такая цветовая модель является аддитивной системой смешивания цветов и называется *RGB* моделью (*Red, Green, Blue*). Перечисленные цвета являются основными цветами данной модели. Соответствующими дополнительными цветами являются бирюзовый (*Cyan*), розовый (*Magenta*) и желтый (*Yellow*). Если основной цвет сложить в равной пропорции с его дополнительным цветом, то получим белый цвет. Таким образом, бирюзовый цвет можно представить в виде суммы зеленого и синего цветов, розовый цвет можно представить в виде суммы синего и красного цветов, желтый цвет можно представить в виде суммы красного и зеленого цветов. Остальные оттенки получаются в виде смешения основных цветов в различных пропорциях. Эта модель не единственная, но она наиболее распространена, так как представление света в виде взвешенной суммы красного, зеленого и синего цветов согласуется с экспериментальными данными и легко реализуема.

В соответствии с *RGB* моделью интенсивность падающего и отраженного объектом света должна быть представлена в виде суммы интенсивностей красного, зеленого и синего цветов:

$$I = I_R + I_G + I_B.$$

Интенсивность каждого цвета определяется по одной из приведенных формул, в которой каждая цветовая компонента описывается своими коэффициентами, например:

$$\begin{aligned} I_R &= \frac{k_{aR} I_a}{1 + k_r (r - r_0)} + I_p (k_{sR} \cos \theta + k_{sR} \cos^2 \alpha) + I_l k_{sR} \cos^2 \beta, \\ I_G &= \frac{k_{aG} I_a}{1 + k_r (r - r_0)} + I_p (k_{sG} \cos \theta + k_{sG} \cos^2 \alpha) + I_l k_{sG} \cos^2 \beta, \\ I_B &= \frac{k_{aB} I_a}{1 + k_r (r - r_0)} + I_p (k_{sB} \cos \theta + k_{sB} \cos^2 \alpha) + I_l k_{sB} \cos^2 \beta. \end{aligned} \quad (8.4.5)$$

Значения интенсивностей должны быть нормированы при $r = r_0$. Варьируя направлением падающего света и значениями коэффициентов диффузного отражения, зеркального отражения и пропускания для разных компонент света, можно подобрать желаемое изображение объекта. Если источников света несколько, то каждый из них должен быть учтен в формулах (8.4.5) соответствующими слагаемыми.

8.5. Построение точечного изображения

При построении точечных изображений требуется не только определять видимую часть геометрической модели, но еще цвет и яркость каждой точки изображения. Если моделируемый объект не прозрачный, то цвет точки зависит от цвета ближайшей к наблюдателю грани в этой точке, а яркость точки зависит от освещенности объекта, оптических свойств его поверхности и от ориентации нормали поверхности по отношению к лучам падающего света и к линии, соединяющей рассматриваемую точку и точку наблюдения. Если моделируемый объект прозрачный, то цвет точки изображения зависит от цвета нескольких граней, их оптических свойств и от ориентации их нормалей.

В общем случае для каждой точки изображения следует построить линию, проходящую через рассматриваемую точку изображения и точку наблюдения и найти точки пересечения этой линии с отображаемой моделью. В точках пересечения потребуется знать направление нормали поверхности модели и такие оптические свойства поверхности, как ее способность излучать, пропускать и отражать свет различного спектрального диапазона. По расположению источников света, оптическим свойствам точки модели и направлению ее нормали можно определить цвет и яркость точки изображения.



Рис. 8.5.1

Для плоских поверхностей определение точек пересечения с линиями наблюдения существенно проще, чем для криволинейных поверхностей, по-

этому для построения точечных изображений используют упрощенную модель, в которой поверхность модели аппроксимируют набором треугольных пластин. На рис. 8.5.1 приведено точечное изображение непрозрачного объекта, построенное по упрощенной модели.

Упрощенная модель состоит из треугольных пластин, в вершинах которых вычислены нормали поверхности модели. Треугольные пластины стыкуются по общим сторонам. Нормали в вершинах имеют истинное направление, а в пределах каждого треугольника принимается линейный закон их изменения от вершины к вершине. Пусть в вершинах $\mathbf{p}_a, \mathbf{p}_b, \mathbf{p}_c$ некоторого треугольника нормали описываются единичными векторами $\mathbf{m}_a, \mathbf{m}_b, \mathbf{m}_c$, тогда нормали на поверхности треугольника припишем направление вектора:

$$\mathbf{m}(a,b,c) = a\mathbf{m}_a + b\mathbf{m}_b + c\mathbf{m}_c, \quad (8.5.1)$$

где a, b, c – барицентрические координаты точки \mathbf{p} треугольной пластины, рис. 8.5.2.

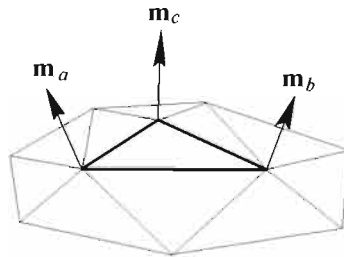


Рис. 8.5.2

Барицентрические координаты определим из системы уравнений

$$\begin{aligned} a x_a + b x_b + c x_c &= x, \\ a y_a + b y_b + c y_c &= y, \\ a + b + c &= 1, \end{aligned}$$

где x_a, y_a – координаты проекции точки \mathbf{p}_a ; x_b, y_b – координаты проекции точки \mathbf{p}_b ; x_c, y_c – координаты проекции точки \mathbf{p}_c ; x, y – координаты проекции точки \mathbf{p} на некоторую плоскость, например, плоскость изображения. Вектор $\mathbf{m}(a,b,c)$ в общем случае не имеет единичной длины и его следует нормировать.

При аппроксимации (8.5.1) нормали плавно меняют свое направление при переходе от одного треугольника к другому, поэтому на точечных изображениях поверхность выглядит гладкой. Данный подход используется в методе Фонга закраски точечных изображений. Если в пре-

делах каждой треугольной пластины использовать истинное направление нормали пластины, то изображение выглядит граненым вследствие полос Маха, рис. 8.5.3. Метод Фонга полностью устраняет полосы Маха.



Рис. 8.5.3

Рассмотрим построение точечного изображения модели на плоскости. Поиск точек пересечения линии и треугольных пластин упрощается, если его выполнять в местной системе координат, связанной с плоскостью изображения, поэтому упрощенную модель переведем в местную систему координат, связанную с плоскостью изображения. Пусть построение точечного изображения модели выполняется на плоскости

$$\mathbf{p}(x,y) = \mathbf{q} + x \mathbf{i}_x + y \mathbf{i}_y$$

с началом в точке $\mathbf{q} = [q_1 \ q_2 \ q_3]^T$ и ортогональными единичными осями координат $\mathbf{i}_x = [x_1 \ x_2 \ x_3]^T$, $\mathbf{i}_y = [y_1 \ y_2 \ y_3]^T$. С плоскостью $\mathbf{p}(x,y)$ свяжем местную систему координат $Qxyz$. Вектор взгляда направим перпендикулярно плоскости навстречу базисному вектору $\mathbf{i}_z = \mathbf{i}_x \times \mathbf{i}_y = [z_1 \ z_2 \ z_3]^T$. Глобальные координаты $r_1 \ r_2 \ r_3$ точки модели и ее местные координаты x, y, z связаны соотношением

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} \cdot \begin{bmatrix} r_1 - q_1 \\ r_2 - q_2 \\ r_3 - q_3 \end{bmatrix}. \quad (8.5.2)$$

Компоненты нормали m_1, m_2, m_3 в глобальной системе координат и компоненты нормали x_m, y_m, z_m в системе координат плоскости изображения связаны соотношением

$$\begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}. \quad (8.5.3)$$

С помощью формул (8.5.2) и (8.5.3) переведем в местную систему координат упрощенную модель и источники света.

При бесконечном удалении точки наблюдения от плоскости изображения линии, проходящие через точки изображения и точку наблюдения, будут параллельны друг другу и ортогональны плоскости изображения.

При построении перспективного изображения упрощенную модель подвергнем дополнительному преобразованию: местные координаты x и y у вершин треугольных пластин умножим на коэффициент

$$k_p = \frac{w}{w-z},$$

где w — расстояние от точки наблюдения до плоскости изображения, а нормали в вершинах треугольных пластин повернем на угол

$$\eta = \arccos\left(-\frac{(\mathbf{r}-\mathbf{w}) \cdot \mathbf{i}_z}{|\mathbf{r}-\mathbf{w}|}\right) = \arccos\left(-\frac{z-w}{\sqrt{xx+yy+(z-w)(z-w)}}\right)$$

вокруг вектора

$$\mathbf{v} = -\frac{(\mathbf{r}-\mathbf{w}) \times \mathbf{i}_z}{|(\mathbf{r}-\mathbf{w}) \times \mathbf{i}_z|} = \frac{1}{\sqrt{xx+yy}} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix},$$

где \mathbf{w} — радиус-вектор точки наблюдения; \mathbf{r} — радиус-вектор точки поверхности, в которой вычислена нормаль; \mathbf{i}_z — нормаль плоскости изображения. Поворот вектора \mathbf{m} описывается формулой

$$\mathbf{m}_p = (\mathbf{v}\mathbf{v}^T) \cdot \mathbf{m} + \cos \eta (\mathbf{E} - \mathbf{v}\mathbf{v}^T) \mathbf{m} + \sin \eta \mathbf{v} \times \mathbf{m}.$$

Таким образом, перспективное изображение будем строить так же, как параллельное, но для искаженной соответствующим образом упрощенной модели.

Для определения цвета и яркости точки изображения с координатами x и y найдем пересечение ортогональной к плоскости изображения прямой, проходящей через рассматриваемую точку изображения, с упрощенной моделью. В местной системе координат этот поиск сводится к определению треугольных пластин, в проекции которых лежит точка с координатами x и y . Последняя задача является двумерной. Точка изображения будет располагаться внутри проекции пластины с вершинами $\mathbf{p}_a, \mathbf{p}_b, \mathbf{p}_c$, если ее барицентрические координаты a, b, c , вычисленные относительно вершин пластины, принимают неотрицательные значения. Если хотя бы одна из барицентрических координат отрицательна, то это означает, что треугольная пластина не пересекается прямой, проходя-

шей через точку изображения и ортогональной плоскости изображения. Для пересекаемых треугольных пластин в точках пересечения вычислим направление нормали и оптические свойства. Если моделируемый объект не прозрачен, то нас будет интересовать только одна треугольная пластина, у которой точка пересечения имеет наибольшую координату z . Описанным способом найдем для рассматриваемой точки изображения точки поверхности модели, оптические свойства поверхности, направление нормали поверхности, направление на каждый источник света, интенсивность и цветовой диапазон каждого источника света. По перечисленным сведениям можно определить цвет и яркость точки изображения.

Для создания неровностей на поверхности модели могут использоваться случайные отклонения нормали. На рис. 8.5.4 приведена модель, нормальям поверхности которой сообщены случайные отклонения.



Рис. 8.5.4

На поверхность модели может быть нанесен рисунок, подобный рисунку на стеновых обоях. Для нанесения рисунка на поверхность нужно установить функциональную зависимость между плоскостью рисунка и плоскостью параметров поверхности.

В общем случае, при несовпадении точки наблюдения с источником света некоторая часть поверхности модели будет закрыта от источника, то есть будет находиться в тени. Алгоритм определения затенения точек аналогичен алгоритму определения видимости точек. Если точка поверхности видна из источника света, то она освещена. Точки поверхности модели, которые видны из точки наблюдения и из источника света, будут передавать отраженный свет источника на точечном изображении. Точки поверхности модели, которые видны из точки наблюдения, но не видны из источника света, находятся в тени и передают только рассеянный и излученный свет. На рис. 8.5.5 приведено изображение модели с учетом тени.



Рис. 8.5.5

Четкие тени дают точечные источники света. При освещении модели распределенными источниками света возникают области полутеней. Полутени возникают в изображении точек модели, из которых видна только часть распределенного источника света.

Для прозрачных моделей в точку наблюдения приходит свет, отраженный не только от ближайшей поверхности, но и от других поверхностей, которые видны сквозь нее. Если не учитывать преломление света, то в точку наблюдения придет свет от всех поверхностей, с которыми пересекается линия наблюдения. На рис. 8.5.6 приведено изображение прозрачного объекта.



Рис. 8.5.6

При моделировании преломления света линия наблюдения для за-слоненных поверхностей должна быть смещена. Интенсивность света, пришедшего от каждой поверхности, будет ослаблена из-за поглощения света веществом. Чтобы определить, какие из световых лучей попадают в точку наблюдения, используют метод трассировки лучей.

Трассирование начинается из точки наблюдения, и лучи света отслеживаются в обратном направлении к каждому источнику света. Луч света, падающий на поверхность, разделяется на четыре основные части: диффузно пропущенный свет, направленно пропущенный преломленный свет, диффузно отраженный свет и зеркально отраженный свет.

Аналогично луч света, исходящий от поверхности модели, является суммой двух отраженных составляющих света и двух составляющих света, пришедших из-за поверхности. Таким образом, каждый раз, когда трассируемый из точки наблюдения луч подходит к поверхности модели, возможно появление нескольких новых лучей, то есть трассируемый луч расщепляется на несколько составляющих. Диффузно отраженный свет и диффузно пропущенный свет могут прийти в общем случае из бесконечного числа направлений, поэтому обычно трассируют только лучи света, появляющиеся в результате зеркального отражения и преломления. В результате при подходе трассируемого луча к поверхности прозрачной модели он расщепляется на две части. Рассеянный свет и диффузно отраженный свет можно получить от ближайшей к точке наблюдения поверхности. Для каждого трассируемого луча нужно получить его пересечение со всеми поверхностями модели. В точках пересечения по свойствам поверхности определяется, нужно ли расщеплять световой луч.

8.6. Триангуляция

Для построения точечных изображений, вычисления инерционных характеристик, определения соударений элементов модели, генерации конечных элементов и для других целей используют геометрическую модель, в которой поверхность моделируемого объекта аппроксимируют набором треугольных пластин. В простейшем случае такая модель может быть представлена массивом точек в пространстве и массивом треугольников, вершины которого ссылаются на точки первого массива. Для получения изображения модели без полос Маха к массиву точек должен быть добавлен синхронный с ним массив нормалей в этих точках. Для быстрого доступа к элементам описанной модели точки и треугольники могут быть разбиты на группы и отсортированы требуемым образом. В более сложных случаях модель может быть дополнена ребрами, содержащими информацию о связываемых ими точках и треугольных пластинах.

Геометрическую модель, состоящую из набора треугольных пластин, можно получить из геометрической модели в виде оболочки. Для этого области определения поверхностей модели разбивают на треугольные и четырехугольные подобласти. Для каждой треугольной подобласти можно построить треугольную пластину. Для каждой четырехугольной подобласти можно построить четырехугольную пластину или две треугольные пластины, разбив подобласть диагональю на две треугольные подобласти. Совокупность этих пластин аппроксимирует

поверхность модели с определенной точностью, которой можно управлять.

Триангуляцией называют процесс и результат построения конечного числа не перекрывающихся друг друга треугольников, стыкующихся между собой по общим сторонам.

С помощью триангуляции мы можем аппроксимировать поверхность модели треугольными пластинами. Эта задача решается в плоскости параметров поверхности и является двумерной.

В простейшем случае задача триангуляции формулируется следующим образом: дана совокупность точек на плоскости, по которой требуется построить максимально возможное число не перекрывающихся друг друга треугольников с вершинами в этих точках. Отрезки, соединяющие вершины треугольников и являющиеся сторонами треугольников, называют ребрами. В общем случае задача построения триангуляции заключается в соединении заданных вершин непересекающимися ребрами так, чтобы в полученной триангуляции нельзя было построить новое ребро без пересечения с уже имеющимися ребрами. Поставленная задача может иметь несколько решений, поэтому существуют различные типы триангуляции, обладающие характерными свойствами.

Триангуляция называется **оптимальной**, если сумма длин всех ребер минимальна среди всех возможных триангуляций, построенных при тех же исходных данных.

Оптимальная триангуляция используется довольно редко ввиду большой трудоемкости ее построения. На практике чаще применяются другие триангуляции, не являющиеся оптимальными. К ним относятся жадная триангуляция и триангуляция Делоне.

Триангуляция называется **жадной**, если она построена с помощью жадного алгоритма. Жадными называют алгоритмы, которые никогда не отменяют то, что было сделано ранее.

Жадный алгоритм триангуляции выполняет следующие действия: вычисляет расстояния между всеми парами заданных вершин будущей триангуляции, сортирует их по возрастанию и, начиная с минимального расстояния, по сортированным парам вершин строит ребра треугольников, проверяя их на пересечение с уже имеющимися ребрами. Новое ребро строится при условии, что оно не пересекает ни одно уже построенное ребро. Если все расстояния между вершинами различные, то жадная триангуляция однозначна, в противном случае она зависит от последовательности вставки ребер одинаковой длины.

Триангуляция называется **триангуляцией Делоне**, если внутри описанной вокруг любого треугольника окружности отсутствуют вершины других треугольников.

Триангуляция Делоне не допускает построение неоправданно вытянутых треугольников. Триангуляция Делоне будет уникальной, если никакие четыре вершины не лежат на одной окружности. Триангуляция Делоне обладает полезными свойствами и невысокой трудоемкостью, поэтому наиболее часто применяется на практике.

Рассмотрим построение триангуляции Делоне для заданного набора двумерных точек. Для произвольно выбранной точки найдем ближайшую точку. Построим между точками ребро. Попытаемся построить на данном ребре треугольник, третья вершина которого лежит слева или справа от ребра. Пусть требуется построить треугольник, стороной которого является ребро AB , рис. 8.6.1.

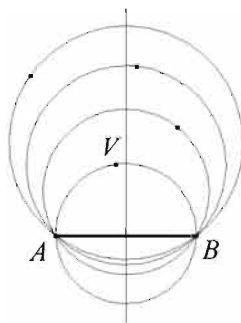


Рис. 8.6.1

Через вершины A , B и любую, не лежащую с ними на одной прямой вершину V можно провести окружность. Центр окружности, проходящей через точки A , B и V , лежит на пересечении перпендикуляров к серединам отрезков AB , VB и AV . Пусть вершины A , B и V описываются двумерными радиусами-векторами $\mathbf{a} = [x_a \ y_a]^T$, $\mathbf{b} = [x_b \ y_b]^T$, $\mathbf{v} = [x_v \ y_v]^T$ соответственно. Построим прямую

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_a + x_b \\ y_a + y_b \end{bmatrix} + t \begin{bmatrix} y_a - y_b \\ x_b - x_a \end{bmatrix}, \quad (8.6.1)$$

перпендикулярную отрезку AB , проходящую через его середину и прямую

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_v + x_b \\ y_v + y_b \end{bmatrix} + w \begin{bmatrix} y_v - y_b \\ x_b - x_v \end{bmatrix}, \quad (8.6.2)$$

перпендикулярную отрезку VB , проходящую через середину отрезка VB . В точке пересечения прямых правые и левые части выражений (8.3.1) и

(8.3.2) равны друг другу. Пересечение прямых (8.6.1) и (8.6.2) найдем, решив систему уравнений

$$\begin{aligned} \frac{1}{2}(x_v + x_b) + t(y_v - y_b) &= \frac{1}{2}(x_v + x_a) + w(y_v - y_a), \\ \frac{1}{2}(y_v + y_b) + t(x_b - x_a) &= \frac{1}{2}(y_v + y_b) + w(x_b - x_a) \end{aligned}$$

относительно параметров прямых t и w , определяющих положение точки на прямых. Нас будет интересовать параметр t , по которому из (8.6.1) можно получить положение центра окружности, проходящей через вершины A , B и V , лежащего на перпендикуляре к середине отрезка AB . Из решения системы уравнений этот параметр равен:

$$t = \frac{1}{2} \frac{(x_v - x_b)(x_v - x_a) + (y_v - y_b)(y_v - y_a)}{(y_v - y_b)(x_b - x_a) - (x_v - x_b)(y_b - y_a)}. \quad (8.6.3)$$

Для центра окружности, лежащего на середине отрезка AB , значение параметра t равно нулю. Для центров окружностей, лежащих слева от отрезка AB , параметр $t > 0$. Для центров окружностей, лежащих справа от отрезка AB , параметр $t < 0$. В качестве третьей вершины треугольника Делоне выберем вершину V , соответствующая которой окружность не содержит других вершин с той же стороны относительно отрезка AB , с которой лежит вершина V . Будем называть такую вершину ближайшей. Для ближайшей слева к отрезку AB вершины параметр t имеет минимальное значение, а для ближайшей справа к отрезку AB вершины параметр t имеет максимальное значение.

В общем случае для ребра можно найти две ближайшие вершины: одну слева от ребра, другую справа от ребра, и таким образом построить два смежных треугольника Делоне. Для ребра, относительно которого все заданные точки лежат только с одной стороны, можно найти только одну ближайшую вершину. Ребро последнего типа будем называть *граничным*. Алгоритм построения триангуляции Делоне работает до тех пор, пока для каждого уже имеющегося внутреннего ребра не будут построены два смежных треугольника, а для каждого граничного ребра не будет построен один треугольник. В результате получим триангуляцию внутренней области выпуклого многоугольника, содержащего все заданные вершины, рис. 8.6.2.

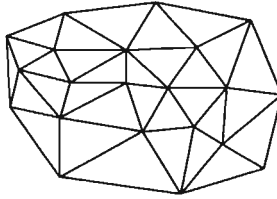


Рис. 8.6.2

Триангуляцию Делоне можно получить из любой другой триангуляции, последовательно перестраивая смежные пары треугольников ABC и ADB , не удовлетворяющие условию Делоне, в пары треугольников ADC и BCD , удовлетворяющие условию Делоне, рис. 8.6.3.

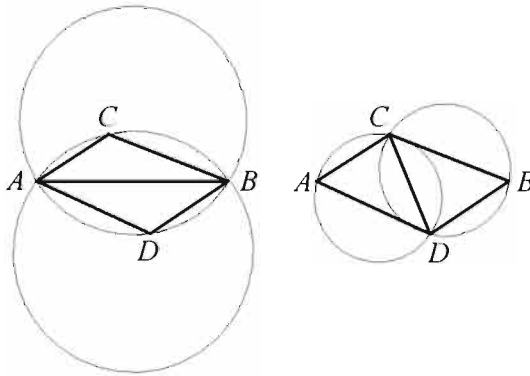


Рис. 8.6.3

Задачу триангуляции можно сформулировать следующим образом: имеется частично построенная триангуляция, к которой добавляется новая вершина; требуется достроить триангуляцию.

Построение триангуляции путем последовательного добавления точек в частично построенную триангуляцию называется **итеративной триангуляцией**.

В данном подходе сначала определяется положение новой вершины относительно имеющейся триангуляции, затем в зависимости от результата выполняются те или иные действия. Если она совпадает с имеющейся вершиной, ничего не выполняется. Если новая вершина попала на некоторое ребро, то оно разбивается на два ребра и смежные с ребром треугольники также делятся на два меньших треугольника. Если новая вершина попала внутрь какого-либо треугольника, то он удаляется и внутри образовавшегося многоугольника выполняется построение но-

вых треугольников с участием новой вершины. На рис. 8.6.4 приведен процесс вставки новой вершины V , попадающей внутрь триангуляции.

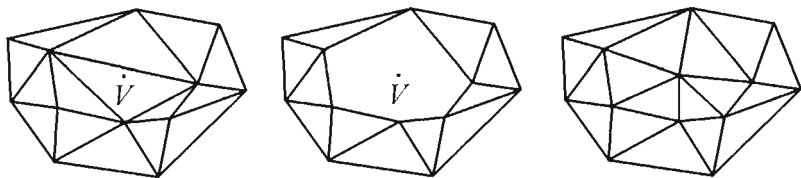


Рис. 8.6.4

Если новая вершина не попадает внутрь триангуляции, то находятся граничные ребра, с которыми она может образовать новые треугольники. Во всех случаях, когда появляются новые треугольники, производится проверка на выполнение условия Делоне для новых треугольников и соседних с ними. При нарушении условия Делоне производится локальное перестроение триангуляции, рис. 8.6.3.

В практических задачах триангуляции кроме заданных вершин присутствуют дополнительные условия, которые, как правило, осложняют задачу. Чаще всего дополнительными условиями являются некоторые кривые, которые должны тем или иным образом соотноситься с триангуляцией. В одних случаях кривые должны быть вписаны в триангуляцию, в других случаях триангуляция должна вписаться в заданные кривые.

Пусть требуется выполнить триангуляцию двумерной области параметров поверхности, ограниченной замкнутой ломаной линией, содержащей набор точек внутри этой области. Для решения этой задачи ориентируем ограничивающую область ломаную линию так, чтобы область лежала слева от нее. На каждом участке ломаной линии построим ребро. Каждую заданную точку и вершину ломаной будем использовать в качестве вершины триангуляции. Ребра и вершины, участвующие в алгоритме триангуляции, будем называть *активными*. Вначале активными являются все ребра и вершины. Построение треугольников начнем с произвольного граничного ребра. Для каждого ребра будем искать вершину, которая лежит слева от него и может быть использована для построения треугольника. Найдем для выбранного ребра ближайшую вершину, соответствующая окружность которой не содержит других вершин. Пусть для граничного ребра AB найдена ближайшая вершина V . Для этой вершины правая часть (8.6.3) принимает наименьшее значение. Проверим, не пересекают ли отрезки AV и BV активные ребра, то есть можно ли построить треугольник ABV .

Если построение треугольника невозможно, то перейдем к следующему ребру, в противном случае построим треугольник ABV и переведем ребро AB в разряд неактивных. *Неактивными* будем называть ребра и вершины, которые не участвуют в алгоритме триангуляции. Если среди граничных ребер отсутствует ребро BV , то на отрезке VB построим новое граничное ребро. Если же среди граничных ребер есть ребро BV , то переведем его и вершину B в разряд неактивных. Если среди граничных ребер отсутствует ребро VA , то на отрезке AV построим новое граничное ребро. Если же среди граничных ребер есть ребро VA , то переведем его и вершину A в разряд неактивных. Процесс триангуляции приведен на рис. 8.6.5, где активные ребра показаны толстыми линиями, а неактивные ребра показаны тонкими линиями.

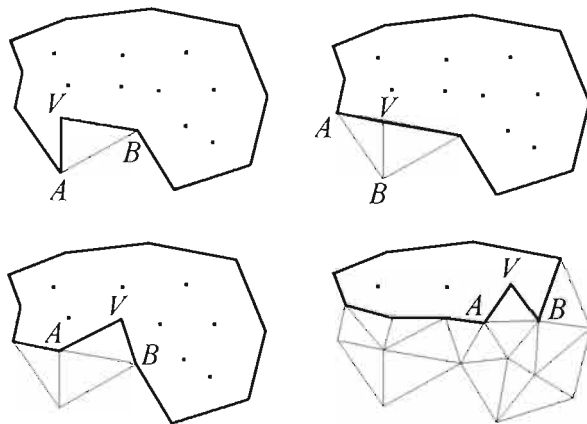


Рис. 8.6.5

В процессе триангуляции происходит изменение ломаной линии, образованной активными ребрами. При этом область, подлежащая триангуляции, уменьшается. Описанный алгоритм можно использовать и при отсутствии вершин внутри ограниченной области, и для областей, внутри которых есть не принадлежащие им вырезы, описываемые ломаными линиями. В этом случае по границе вырезов также построим ребра на участках ломаных линий, сориентировав их так, чтобы область лежала слева от ребер. Эти ребра участвуют в триангуляции наравне с ребрами внешней границы.

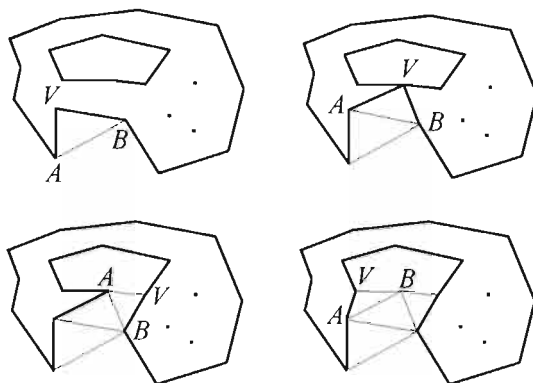


Рис. 8.6.6

Процесс триангуляции области с одним внутренним вырезом приведен на рис 8.6.6. В процессе триангуляции рано или поздно произойдет слияние внешней и внутренней граничной ломаной, образованной активными ребрами. Триангуляцию закончим, когда будет построен последний треугольник и все вершины и ребра станут неактивными.

8.7. Триангуляция поверхности

Для аппроксимации поверхности плоскими треугольными пластинами, стыкующимися между собой по общим сторонам, выполним триангуляцию области определения параметров поверхности и отобразим ее в трехмерное пространство, построив для каждой треугольной подобласти треугольную пластину.

В общем случае граница поверхности описывается двумерными замкнутыми составными кривыми в пространстве параметров поверхности. Аппроксимируем границы области определения параметров поверхности двумерными замкнутыми ломаными линиями. Сориентируем ломаные так, чтобы область определения параметров располагалась слева от сегментов ломаных. Построим набор двумерных точек внутри области. В качестве точек внутри области примем точки пересечения семейства линий $u_i = const$, $i = 1, 2, \dots, m$ и семейства линий $v_j = const$, $j = 1, 2, \dots, n$. Размерами сторон треугольников и расстоянием между точками набора внутри области будем управлять с помощью параметрических шагов граничных кривых и параметрических шагов поверхности. В

качестве примера рассмотрим триангуляцию поверхности, приведенной на рис. 8.7.1.

Построим ломаные линии, аппроксимирующие граничные кривые $c_j(t) = s(u_j(t), v_j(t))$ поверхности $s(u, v)$, рис 8.7.2. Если триангуляция поверхности выполняется для построения точечного изображения, то параметрические шаги вдоль границы вычислим по формуле (7.2.1), управляемой величиной прогиба δ . Если триангуляция поверхности выполняется для вычисления инерционных характеристик, то параметрические шаги вдоль границы вычислим по формуле (4.12.2), управляемой величиной углового отклонения $\Delta\alpha$.



Рис. 8.7.1

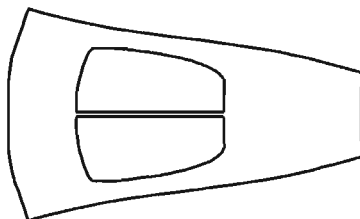


Рис. 8.7.2

Пусть граничные кривые области определения параметров поверхности $s(u, v)$ располагаются внутри прямоугольника, определяемого максимальными и минимальными значениями параметров $u_{\min}, v_{\min}, u_{\max}, v_{\max}$. Разобьем прямоугольную область $u_{\min} \leq u \leq u_{\max}, v_{\min} \leq v \leq v_{\max}$ поверхности на прямоугольные подобласти двумерными линиями $u_i = \text{const}_i$ и $v_j = \text{const}_j, i = 1, 2, \dots, m, j = 1, 2, \dots, n$. Если триангуляция поверхности выполняется для построения точечного изображения, то параметрические шаги $\Delta u_i = u_{i+1} - u_i$ между соседними линиями $u_i = \text{const}$ в соответствии с формулой (8.3.3) возьмем равными

$$\Delta u_i = \min \left(2 \sqrt{\delta \left(\frac{2}{b_{11}(u, v)} - \frac{\delta}{g_{11}(u, v)} \right)} \right) \text{ для всех } v_{\min} \leq v \leq v_{\max},$$

а параметрические шаги $\Delta v_j = v_{j+1} - v_j$ между соседними линиями $v_j = \text{const}$ в соответствии с формулой (8.3.4) возьмем равными

$$\Delta v_j = \min \left(2 \sqrt{\delta \left(\frac{2}{b_{22}(u, v_j)} - \frac{\delta}{g_{22}(u, v_j)} \right)} \right) \text{ для всех } u_{\min} \leq u \leq u_{\max}.$$

Если триангуляция поверхности выполняется для вычисления инерционных характеристик, то параметрические шаги $\Delta u_i = u_{i+1} - u_i$ между соседними линиями $u_i = \text{const}$ в соответствии с формулой (4.12.3) возьмем равными

$$\Delta u_i = \min \left(\frac{\Delta \alpha \sqrt{g_{11}(u, v)}}{b_{11}(u, v)} \right) \text{ для всех } v_{\min} \leq v \leq v_{\max},$$

а параметрические шаги $\Delta v_j = v_{j+1} - v_j$ между соседними линиями $v_j = \text{const}$ в соответствии с формулой (4.12.4) возьмем равными

$$\Delta v_j = \min \left(\frac{\Delta \alpha \sqrt{g_{22}(u, v_j)}}{b_{22}(u, v_j)} \right) \text{ для всех } u_{\min} \leq u \leq u_{\max}.$$

На рис. 8.7.3 приведено разбиение описанным способом области $u_{\min} \leq u \leq u_{\max}$, $v_{\min} \leq v \leq v_{\max}$ рассматриваемой поверхности. Определим прямоугольники, располагающиеся внутри области определения поверхности, а остальные прямоугольники отбросим, рис. 8.7.4.

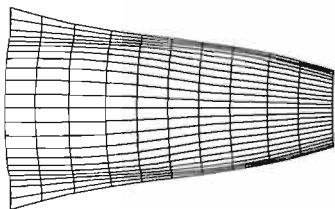


Рис. 8.7.3

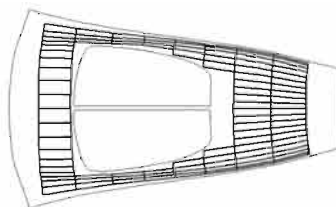


Рис. 8.7.4

Выполним триангуляцию части области определения поверхности, не вошедшую в прямоугольные подобласти, описанным выше алгоритмом Делоне для ограниченной области. На рис. 8.7.5. приведено конечное разбиение поверхности на четырехугольные и треугольные подобласти.

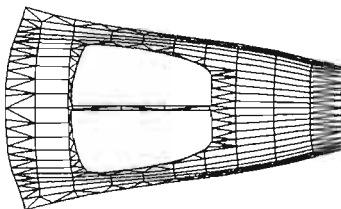


Рис. 8.7.5

Разбив четырехугольники диагональю на два треугольника, получим триангуляцию поверхности.

8.8. Триангуляция оболочки

На основе геометрической модели в виде оболочки можно получить геометрическую модель, состоящую из набора треугольных пластин. Для этого построим совокупность стыкующихся между собой по общим сторонам треугольников путем триангуляции граней оболочки.

Набор треугольных пластин с общими нормальными в общих точках должен быть построен для каждой грани. Наборы треугольных пластин, аппроксимирующих соседние грани оболочки, связаны между собой. Если выполнить триангуляцию отдельных граней независимо друг от друга, то на краях граней пластины триангуляции не будут стыковаться по общим сторонам, рис. 8.8.1. Триангуляция отдельных поверхностей отличается от триангуляции граней оболочки тем, что в последнем случае должны быть согласованы общие стороны и вершины крайних треугольников для смежных граней, рис. 8.8.2.

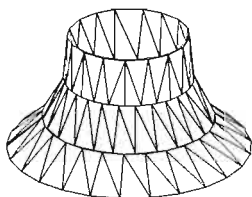


Рис. 8.8.1

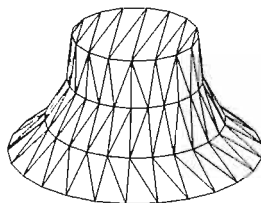


Рис. 8.8.2

Стыковка треугольников смежных граней должна осуществляться по совпадающим в пространстве сторонам. В противном случае триангуляция будет содержать щели. Если триангуляция выполняется для построения точечных изображений, то каждая вершина треугольника содержит нормаль к поверхности модели. В совпадающих вершинах соседних треугольников, принадлежащих двум разным граням, должны быть вычислены две разные нормали, одна — для одной грани, вторая — для другой смежной грани. Если при переходе от одного треугольника к другому треугольнику этой же грани нормаль изменяется плавно, то при переходе от треугольника одной грани к смежному треугольнику другой грани нормаль изменяется скачкообразно.

8.9. Инерционные характеристики

Для вычисления инерционных характеристик моделируемого объекта кроме геометрической модели необходимо знать плотность тел, моделирующих отдельные элементы объекта. Будем считать, что плотность каждого тела постоянна по всему его объему. Это позволит нам свести интегралы по объему тела к интегралам по поверхностям граней тела. Для вычисления объемных интегралов воспользуемся формулой Остроградского–Гаусса

$$\iint_S \mathbf{m} \cdot \mathbf{F}(\mathbf{r}) dS = \iiint_V \nabla \cdot \mathbf{F}(\mathbf{r}) dV, \quad (8.9.1)$$

где \mathbf{m} – нормаль поверхности тела; $\mathbf{F}(\mathbf{r})$ – векторная функция точек тела; \mathbf{r} – радиус-вектор точки тела; ∇ – набла-оператор или оператор Гамильтона. Напомним, что нормаль \mathbf{m} всегда направлена вне тела. Инерционные характеристики вычислим в декартовой прямоугольной системе, координаты которой обозначим через x, y, z .

Площадь поверхности тела равна сумме площадей граней и может быть вычислена по формуле

$$S = \sum_i \iint_{\Omega_i} \sqrt{g_{11}g_{22} - g_{12}^2} du_i dv_i,$$

где u_i, v_i – параметры поверхности i -й грани тела; $g_{11}(u_i, v_i), g_{12}(u_i, v_i), g_{21}(u_i, v_i), g_{22}(u_i, v_i)$ – коэффициенты первой основной квадратичной формы поверхности; Ω_i – область определения параметров поверхности i -й грани тела.

В дальнейших формулах инерционных характеристик тела опустим знак суммирования и индексы граней, считая, что суммирование, естественно, выполняется.

Положим в (8.9.1) $\mathbf{F}(\mathbf{r}) = \mathbf{r}$. Тогда из теоремы о дивергенции получим

$$\iint_S \mathbf{m} \cdot \mathbf{r} dS = \iiint_V \nabla \cdot (\mathbf{r}) dV = 3 \iiint_V dV = 3V,$$

$$\text{где } \nabla \cdot \mathbf{r} = \left[\frac{\partial}{\partial x} \quad \frac{\partial}{\partial y} \quad \frac{\partial}{\partial z} \right] \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{\partial}{\partial x} x + \frac{\partial}{\partial y} y + \frac{\partial}{\partial z} z = 1 + 1 + 1 = 3.$$

Объем тела вычислим по формуле

$$V = \frac{1}{3} \iint_S \mathbf{m} \cdot \mathbf{r} dS = \frac{1}{3} \iint_{\Omega} \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv.$$

Масу тела получим, умножив объем тела на его плотность ρ :

$$M = \rho V = \frac{\rho}{3} \iiint_{\Omega} \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv.$$

Координаты центра масс тела x_c, y_c, z_c определяются путем деления соответствующих статических моментов на массу тела:

$$x_c = \frac{1}{M} \iiint_V \rho x dV, \quad y_c = \frac{1}{M} \iiint_V \rho y dV, \quad z_c = \frac{1}{M} \iiint_V \rho z dV,$$

где ρ – плотность тела. Для вычисления статических моментов тела положим в (8.9.1) последовательно $\mathbf{F}(\mathbf{r}) = x\mathbf{r}$, $\mathbf{F}(\mathbf{r}) = y\mathbf{r}$, $\mathbf{F}(\mathbf{r}) = z\mathbf{r}$. Тогда из теоремы о дивергенции получим

$$\begin{aligned} \iint_S \mathbf{m} \cdot (x\mathbf{r}) dS &= \iiint_V \nabla \cdot (x\mathbf{r}) dV = 4 \iiint_V x dV, \\ \iint_S \mathbf{m} \cdot (y\mathbf{r}) dS &= \iiint_V \nabla \cdot (y\mathbf{r}) dV = 4 \iiint_V y dV, \\ \iint_S \mathbf{m} \cdot (z\mathbf{r}) dS &= \iiint_V \nabla \cdot (z\mathbf{r}) dV = 4 \iiint_V z dV, \end{aligned}$$

где

$$\begin{aligned} \nabla \cdot (x\mathbf{r}) &= \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} \cdot \begin{bmatrix} xx \\ xy \\ xz \end{bmatrix} = \frac{\partial}{\partial x} x^2 + \frac{\partial}{\partial y} xy + \frac{\partial}{\partial z} xz = 2x + x + x = 4x, \\ \nabla \cdot (y\mathbf{r}) &= \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} \cdot \begin{bmatrix} yx \\ yy \\ yz \end{bmatrix} = \frac{\partial}{\partial x} yx + \frac{\partial}{\partial y} y^2 + \frac{\partial}{\partial z} yz = y + 2y + y = 4y, \\ \nabla \cdot (z\mathbf{r}) &= \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} \cdot \begin{bmatrix} zx \\ zy \\ zz \end{bmatrix} = \frac{\partial}{\partial x} zx + \frac{\partial}{\partial y} zy + \frac{\partial}{\partial z} z^2 = z + z + 2z = 4z. \end{aligned}$$

Координаты центра масс тела вычислим по формулам

$$\begin{aligned} x_c &= \frac{1}{4V} \iint_S \mathbf{m} \cdot (x\mathbf{r}) dS = \frac{1}{4V} \iint_{\Omega} \mathbf{m} \cdot (x\mathbf{r}) \sqrt{g_{11}g_{22} - g_{12}^2} dudv, \\ y_c &= \frac{1}{4V} \iint_S \mathbf{m} \cdot (y\mathbf{r}) dS = \frac{1}{4V} \iint_{\Omega} \mathbf{m} \cdot (y\mathbf{r}) \sqrt{g_{11}g_{22} - g_{12}^2} dudv, \\ z_c &= \frac{1}{4V} \iint_S \mathbf{m} \cdot (z\mathbf{r}) dS = \frac{1}{4V} \iint_{\Omega} \mathbf{m} \cdot (z\mathbf{r}) \sqrt{g_{11}g_{22} - g_{12}^2} dudv. \end{aligned}$$

Осевые моменты инерции тела определяются формулами

$$J_{xx} = \iiint_V \rho(y^2 + z^2) dV, \quad J_{yy} = \iiint_V \rho(z^2 + x^2) dV, \quad J_{zz} = \iiint_V \rho(x^2 + y^2) dV.$$

Центробежные моменты инерции тела определяются формулами

$$J_{xy} = J_{yx} = \iiint_V \rho xy dV, \quad J_{yz} = J_{zy} = \iiint_V \rho yz dV, \quad J_{zx} = J_{xz} = \iiint_V \rho xz dV.$$

Моменты инерции вводятся при рассмотрении вращения твердого тела вокруг неподвижной точки. При совмещении начала координат с неподвижной точкой скорость некоторой точки тела определяется соотношением $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, где $\boldsymbol{\omega}$ – вектор мгновенной угловой скорости тела, \mathbf{r} – радиус-вектор точки тела. Вектор момента количества движения тела равен

$$\mathbf{L} = \iiint_V (\mathbf{r} \times \mathbf{v}) \rho dV = \iiint_V (\mathbf{r}^2 \boldsymbol{\omega} - (\mathbf{r} \cdot \boldsymbol{\omega}) \mathbf{r}) \rho dV = \mathbf{J} \cdot \boldsymbol{\omega},$$

где величина \mathbf{J} называется *тензором инерции* и определяется матрицей, элементами которой являются осевые моменты инерции и центробежные моменты инерции, взятые с противоположным знаком:

$$\mathbf{J} = \begin{bmatrix} J_{xx} & -J_{xy} & -J_{xz} \\ -J_{xy} & J_{yy} & -J_{yz} \\ -J_{xz} & -J_{yz} & J_{zz} \end{bmatrix}.$$

Для вычисления моментов инерции положим в (8.9.1) последовательно $\mathbf{F}(\mathbf{r}) = x^2 \mathbf{r}$, $\mathbf{F}(\mathbf{r}) = y^2 \mathbf{r}$, $\mathbf{F}(\mathbf{r}) = z^2 \mathbf{r}$, $\mathbf{F}(\mathbf{r}) = yz \mathbf{r}$, $\mathbf{F}(\mathbf{r}) = zx \mathbf{r}$, $\mathbf{F}(\mathbf{r}) = xy \mathbf{r}$. Тогда из теоремы о дивергенции получим

$$\begin{aligned} \iint_S \mathbf{m} \cdot (x^2 \mathbf{r}) dS &= 5 \iiint_V x^2 dV, \\ \iint_S \mathbf{m} \cdot (y^2 \mathbf{r}) dS &= 5 \iiint_V y^2 dV, \\ \iint_S \mathbf{m} \cdot (z^2 \mathbf{r}) dS &= 5 \iiint_V z^2 dV, \\ \iint_S \mathbf{m} \cdot (xy \mathbf{r}) dS &= 5 \iiint_V xy dV, \\ \iint_S \mathbf{m} \cdot (yz \mathbf{r}) dS &= 5 \iiint_V yz dV, \\ \iint_S \mathbf{m} \cdot (zx \mathbf{r}) dS &= 5 \iiint_V zx dV. \end{aligned}$$

Моменты инерции тела вычислим по формулам

$$J_{xx} = \frac{\rho}{5} \iiint_S (y^2 + z^2) \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iiint_{\Omega} (y^2 + z^2) \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv,$$

$$J_{xy} = \frac{\rho}{5} \iint_S (z^2 + x^2) \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iint_{\Omega} (z^2 + x^2) \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv,$$

$$J_{yz} = \frac{\rho}{5} \iint_S (x^2 + y^2) \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iint_{\Omega} (x^2 + y^2) \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv,$$

$$J_{xy} = \frac{\rho}{5} \iint_S xy \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iint_{\Omega} xy \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv,$$

$$J_{yz} = \frac{\rho}{5} \iint_S yz \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iint_{\Omega} yz \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv,$$

$$J_{xz} = \frac{\rho}{5} \iint_S xz \mathbf{m} \cdot \mathbf{r} dS = \frac{\rho}{5} \iint_{\Omega} xz \mathbf{m} \cdot \mathbf{r} \sqrt{g_{11}g_{22} - g_{12}^2} dudv.$$

8.10. Вычисление инерционных характеристик

При численном интегрировании по поверхности разобьем область определения параметров грани на небольшие четырехугольные и треугольные подобласти. От размеров подобластей зависит точность вычисления. В качестве управляющего параметра разбиения области интегрирования будем использовать угловое отклонение $\Delta\alpha$. Размер каждой подобласти определим условием: угловое изменение нормали поверхности в подобласти не должно превышать некоторую заданную величину $\Delta\alpha$. Описанным выше способом заменим двумерную область определения параметров совокупностью стыкующихся между собой четырехугольников и треугольников.

Вычислим определенный интеграл

$$I = \iint_{\Omega} f(u, v) dudv \quad (8.10.1)$$

на выпуклой четырехугольной области с использованием квадратурных формул Гаусса. Пусть вершины четырехугольной области расположены в точках $\mathbf{p}_a = [u_a \ v_a]^T$, $\mathbf{p}_b = [u_b \ v_b]^T$, $\mathbf{p}_c = [u_c \ v_c]^T$, $\mathbf{p}_d = [u_d \ v_d]^T$ на плоскости параметров, рис 8.10.1.

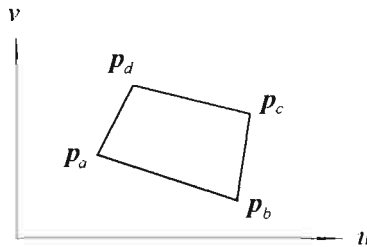


Рис. 8.10.1

Заменим переменные в двойном интеграле (8.10.1). Для произвольной выпуклой четырехугольной области введем новые параметры x и y , связанные с параметрами u и v соотношениями

$$\begin{aligned} u(x,y) &= \\ &= u_a \frac{1-x-y+xy}{4} + u_b \frac{1+x-y-xy}{4} + u_c \frac{1+x+y+xy}{4} + u_d \frac{1-x+y-xy}{4}, \\ v(x,y) &= \\ &= v_a \frac{1-x-y+xy}{4} + v_b \frac{1+x-y-xy}{4} + v_c \frac{1+x+y+xy}{4} + v_d \frac{1-x+y-xy}{4}. \end{aligned} \quad (8.10.2)$$

Точкам p_a, p_b, p_c, p_d соответствуют параметры $x_a = -1, y_a = -1; x_b = 1, y_b = -1; x_c = 1, y_c = 1; x_d = -1, y_d = 1$. Функции (8.10.2) являются непрерывными, имеют непрерывные частные производные первого порядка и отображают квадрат с вершинами в точках $x = \pm 1, y = \pm 1$ на заданную четырехугольную область интегрирования. Это отображение взаимно однозначное и имеет отличный от нуля якобиан

$$J(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}.$$

По каждому параметру поверхности выполним интегрирование с помощью квадратурных формул Гаусса. После замены параметров интеграл (8.10.1) можно вычислить как взвешенную сумму значений интегрируемой функции внутри области интегрирования:

$$\begin{aligned} \iint_D f(u,v) du dv &= \int_{-1}^1 \int_{-1}^1 f(u(x,y), v(x,y)) J(x,y) dx dy \approx \\ &\approx \sum_{i=1}^m \sum_{j=1}^n w_i z_j f(u(x_i, y_j), v(x_i, y_j)) J(x_i, y_j), \end{aligned} \quad (8.10.3)$$

где x_1, x_2, \dots, x_m — корни полинома Лежандра степени m ; w_1, w_2, \dots, w_m — соответствующие им весовые множители; y_1, y_2, \dots, y_n — корни полинома Лежандра степени n ; z_1, z_2, \dots, z_n — соответствующие им весовые множители. Корни $x_i, i = 1, \dots, m$ полинома Лежандра степени m и соответствующие им весовые множители приведены в табл. 8.10.1.

Таблица 8.10.1

Число точек m	Координаты точек x_i	Весовые множители точек w_i
1	0	2
2	$\pm 0,57735026918962576450$	1
3	0 $\pm 0,77459666924148337703$	8/9 5/9
4	$\pm 0,33998104358485626480$ $\pm 0,86113631159405257522$	0,65214515486254614262 0,34785484513745385737
5	0 $\pm 0,53846931010568309103$ $\pm 0,90617984593866399279$	0,5688888888888888889 0,47862867049936646804 0,23692688505618908751
6	$\pm 0,23861918608319690863$ $\pm 0,66120938646626451366$ $\pm 0,93246951420315202781$	0,46791393457269104738 0,36076157304813860756 0,17132449237917034504
7	0 $\pm 0,40584515137739716690$ $\pm 0,74153118559939443986$ $\pm 0,94910791234275852452$	0,41795918367346938775 0,38183005050511894495 0,27970539148927666790 0,12948496616886969327
8	$\pm 0,18343464249564980493$ $\pm 0,52553240991632898581$ $\pm 0,79666647741362673959$ $\pm 0,96028985649753623168$	0,36268378337836198296 0,31370664587788728733 0,22238103445337447054 0,10122853629037625915

Корни полиномов Лежандра и весовые множители определяют-ся из условия, что равенство (8.10.3) является точным, если $f(x,y)J(x,y) = p_{2m-1}(x)p_{2n-1}(y)$, где $p_{2m-1}(x)$ и $p_{2n-1}(y)$ – полиномы степени $2m - 1$ и $2n - 1$, соответственно.

Вычислим определенный интеграл

$$I = \iint_{\Sigma} f(u,v) du dv \quad (8.10.4)$$

на треугольной области параметров поверхности. Пусть вершины треугольной области расположены в точках $\mathbf{p}_a = [u_a \ v_a]^T$, $\mathbf{p}_b = [u_b \ v_b]^T$, $\mathbf{p}_c = [u_c \ v_c]^T$ на плоскости параметров, рис 8.10.2.

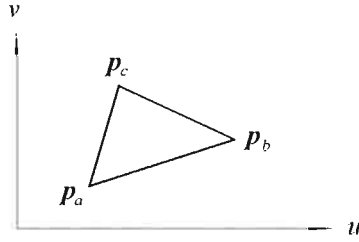


Рис. 8.10.2

Для треугольных областей удобно перейти от координат u и v к трем барицентрическим координатам a, b, c , построенным по точкам P_a, P_b, P_c . Координаты произвольной точки $p = [u \ v]^T$ через барицентрические координаты a, b, c выражаются с помощью формул

$$u = au_a + bu_b + cu_c; \quad v = av_a + bv_b + cv_c. \quad (8.10.5)$$

Барицентрические координаты удовлетворяют равенству: $a + b + c = 1$. Интегрируемую функцию $f(u, v)$ представим в виде функции барицентрических координат $f(a, b, c) \equiv f(u(a, b, c), v(a, b, c))$. Из трех барицентрических координат только две являются независимыми. В качестве независимых параметров, не теряя общности, выберем первые две барицентрические координаты. Используя равенства (8.10.5) в форме

$$u = a(u_a - u_c) + b(u_b - u_c) + u_c, \quad v = a(v_a - v_c) + b(v_b - v_c) + v_c,$$

перепишем интеграл (8.10.4) в виде

$$\begin{aligned} \iint_{\Delta} f(u, v) du dv &= \iint_{\Delta} f(u(a, b, c), v(a, b, c)) J(a, b, c) da db = \\ &= \int_0^1 \left(\int_0^{1-a} f(u(a, b), v(a, b)) J(a, b) db \right) da, \end{aligned}$$

где $J(a, b) = \begin{vmatrix} \frac{\partial u}{\partial a} & \frac{\partial u}{\partial b} \\ \frac{\partial v}{\partial a} & \frac{\partial v}{\partial b} \end{vmatrix}$ – якобиан преобразования координат.

С помощью перехода к барицентрическим координатам и замены параметров интеграл (8.10.4) можно вычислить как взвешенную сумму значений интегрируемой функции внутри области интегрирования:

$$\begin{aligned} \iint_{\Delta} f(u, v) du dv &= \iint_{\Delta} f(u(a, b, c), v(a, b, c)) J(a, b, c) da db = \\ &\approx S \sum_i^m w_i f(u(a_i, b_i, c_i), v(a_i, b_i, c_i)), \end{aligned} \quad (8.10.6)$$

где $S = \frac{1}{2}J(a,b)$ – площадь треугольника $p_p p_c$. Баричесентрические координаты $a_i, b_i, c_i, i = 1, \dots, m$ и весовые множители для кубатурных формул по треугольной области приведены в табл. 8.10.2.

Таблица 8.10.2

m	a_i	b_i	c_i	Весовые множители
1	1/3	1/3	1/3	1
3	2/3	1/6	1/6	1/3
	1/6	2/3	1/6	1/3
	1/6	1/6	2/3	1/3
4	1/3	1/3	1/3	-0,5625
	0,6	0,2	0,2	0,5208333333333333
	0,2	0,6	0,2	0,5208333333333333
	0,2	0,2	0,6	0,5208333333333333
6	x_1	y_1	y_1	0,1099517436553224
	y_1	x_1	y_1	0,1099517436553224
	y_1	y_1	x_1	0,1099517436553224
	x_2	y_2	y_2	0,2233815896780112
	y_2	x_1	y_2	0,2233815896780112
	y_2	y_2	x_2	0,2233815896780112,
				где
				$x_1 = 0,7974269853530880$
				$y_1 = 0,1012865073234560$
				$x_2 = 0,1081030181680702$
				$y_2 = 0,4459484909159648$
7	1/3	1/3	1/3	0,225
	x_1	y_1	y_1	0,1259391805448265
	y_1	x_1	y_1	0,1259391805448265
	y_1	y_1	x_1	0,1323941527885112
	x_2	y_2	y_2	0,1323941527885112
	y_2	x_1	y_2	0,1323941527885112
				где
				$x_1 = 0,7974269853530880$
				$y_1 = 0,1012865073234560$
				$x_2 = 0,0597158717897742$
				$y_2 = 0,4701420641051129$

Баричесентрические координаты a_i, b_i, c_i и весовые множители w_i для кубатурных формул по треугольной области определяются из условия,

что равенство (8.10.6) является точным, если $f(u(a,b,c),v(a,b,c)) = (A_m a^m + B_m b^m + C_m c^m)$ – полином степени m .

8.11. Главные моменты инерции

При параллельном переносе начала декартовой системы координат изменяется тензор инерции тела и его компоненты. Пусть начало декартовой системы координат перемещено в центр масс тела \mathbf{c} , имеющий в исходной системе координаты x_c, y_c, z_c . Система координат с началом в центре масс объекта называется *центральной системой координат* для этого объекта. В центральной системе координат радиус-вектор точки тела равен $\mathbf{r}' = \mathbf{r} - \mathbf{c}$, а компоненты тензора инерции будут равны:

$$\begin{aligned} J_{xx}' &= \iiint_V (y^2 + z^2 - 2yy_c - 2zz_c + y_c^2 + z_c^2) \rho dV = J_{xx} - (y_c^2 + z_c^2)M, \\ J_{yy}' &= \iiint_V (z^2 + x^2 - 2zz_c - 2xx_c + z_c^2 + x_c^2) \rho dV = J_{yy} - (z_c^2 + x_c^2)M, \\ J_{zz}' &= \iiint_V (x^2 + y^2 - 2xx_c - 2yy_c + x_c^2 + y_c^2) \rho dV = J_{zz} - (x_c^2 + y_c^2)M, \\ J_{yc}' &= \iiint_V (yz - yz_c - zy_c + y_c z_c) \rho dV = J_{yc} - y_c z_c M, \\ J_{xc}' &= \iiint_V (zx - zx_c - xz_c + z_c x_c) \rho dV = J_{xc} - z_c x_c M, \\ J_{xy}' &= \iiint_V (xy - xy_c - yx_c + x_c y_c) \rho dV = J_{xy} - x_c y_c M. \end{aligned}$$

Моменты инерции в центральной системе координат называются *центральными моментами инерции*.

В формуле $\mathbf{L} = \mathbf{J}\boldsymbol{\omega}$ вектор количества движения \mathbf{L} в общем случае не совпадает по направлению с вектором угловой скорости $\boldsymbol{\omega}$. Это приводит к тому, что при вращении тела вокруг заданной оси в точках крепления тела возникают зависящие от модуля угловой скорости реакции. Такое вращение является динамически неуравновешенным. Найдем для данного тензора инерции \mathbf{J} такие направления, при которых вектор количества движения \mathbf{L} совпадает по направлению с вектором угловой скорости $\boldsymbol{\omega}$. Пусть это направление определяется единичным вектором $\mathbf{e} = [x \ y \ z]^T$, тогда вектор $\mathbf{J}\mathbf{e}$ должен быть пропорционален вектору \mathbf{e}

$$\mathbf{J} \cdot \mathbf{e} - \lambda \mathbf{e} = 0,$$

где λ – пока неизвестный скаляр. Проецируя это равенство на оси координат, получим три равенства

$$\begin{aligned}
 (J_{xx} - \lambda)x - J_{xy}y - J_{xz}z &= 0, \\
 -J_{xy}x + (J_{yy} - \lambda)y - J_{yz}z &= 0, \\
 -J_{xz}x - J_{yz}y + (J_{zz} - \lambda)z &= 0,
 \end{aligned}
 \tag{8.11.1}$$

которые будем рассматривать как однородную систему линейных алгебраических уравнений относительно x, y, z . Тривиальное решение нам не подходит, так как необходимо найти вектор, для которого

$$x^2 + y^2 + z^2 = 1.$$

Система (8.11.1) имеет нетривиальные решения, когда определитель системы равен нулю:

$$\begin{vmatrix}
 J_{xx} - \lambda & -J_{xy} & -J_{xz} \\
 -J_{xy} & J_{yy} - \lambda & -J_{yz} \\
 -J_{xz} & -J_{yz} & J_{zz} - \lambda
 \end{vmatrix} = 0.
 \tag{8.11.2}$$

Это равенство выполняется, когда скаляр λ удовлетворяет уравнению

$$\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0, \tag{8.11.3}$$

где

$$\begin{aligned}
 I_1 &= J_{xx} + J_{yy} + J_{zz}, \\
 I_2 &= \begin{vmatrix} J_{xx} & -J_{xy} \\ -J_{xy} & J_{yy} \end{vmatrix} + \begin{vmatrix} J_{yy} & -J_{yz} \\ -J_{yz} & J_{zz} \end{vmatrix} + \begin{vmatrix} J_{zz} & -J_{xz} \\ -J_{xz} & J_{xx} \end{vmatrix}, \\
 I_3 &= \begin{vmatrix} J_{xx} & -J_{xy} & -J_{xz} \\ -J_{xy} & J_{yy} & -J_{yz} \\ -J_{xz} & -J_{yz} & J_{zz} \end{vmatrix}.
 \end{aligned}$$

Уравнение (8.11.3) называется *характеристическим уравнением тензора инерции*. Корни $\lambda_1, \lambda_2, \lambda_3$ характеристического уравнения являются действительными, так как матрица тензора инерции является симметричной. Скаляры $\lambda_1, \lambda_2, \lambda_3$ являются *собственными значениями матрицы* моментов инерции. Каждому $\lambda_i, i = 1, 2, 3$ соответствует собственный вектор \mathbf{e}_i .

Предположим, что собственные значения матрицы инерции λ_1 и λ_2 различны, тогда соответствующие им собственные векторы \mathbf{e}_1 и \mathbf{e}_2 ортогональны. Действительно, для данных собственных значений выполняются равенства

$$\mathbf{J} \cdot \mathbf{e}_1 - \lambda_1 \mathbf{e}_1 = 0, \quad \mathbf{J} \cdot \mathbf{e}_2 - \lambda_2 \mathbf{e}_2 = 0.$$

Умножим первое равенство скалярно на \mathbf{e}_2 , а второе — на \mathbf{e}_1 , и вычтем второе из первого. Используя симметрию тензора инерции, получим

$$(\lambda_2 - \lambda_1) \mathbf{e}_1 \cdot \mathbf{e}_2 = 0.$$

Из последнего равенства в силу предположения $\lambda_2 \neq \lambda_1$ следует, что собственные векторы \mathbf{e}_1 и \mathbf{e}_2 ортогональны. Аналогично доказывается ортогональность других собственных векторов. Собственный вектор определяет главное направление тензора инерции в пространстве. Если два из трех собственных значения матрицы инерции равны, то мы сможем определить только одно из главных направлений, а другими главными направлениями могут быть любые направления, ортогональные найденному направлению. Если все три собственных значения матрицы инерции равны между собой, то главными направлениями могут быть любые направления в пространстве. В любом случае можно получить три взаимно ортогональных главных направления.

Три взаимно ортогональных главных направления, соответствующих собственным векторам $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, определяют *главные оси инерции* тела и являются базисными векторами главной системы координат. Если орты новой декартовой прямоугольной системы координат ориентировать вдоль главных направлений, то осевые моменты инерции тела будут равны собственным значениям матрицы инерции, а центробежные моменты инерции будут равны нулю, то есть $J_{11} = \lambda_1, J_{22} = \lambda_2, J_{33} = \lambda_3, J_{12} = 0, J_{23} = 0, J_{13} = 0$. Действительно, для главных направлений выполняются равенства $\mathbf{J}\mathbf{e}_i = \lambda_i\mathbf{e}_i, i = 1, 2, 3$. Тогда $\mathbf{e}_k\mathbf{J}\mathbf{e}_i = \lambda_i\delta_{ik}$, что доказывает высказанное утверждение. Матрица тензора инерции в главной системе координат имеет вид

$$\mathbf{J} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \equiv \begin{bmatrix} J_{11} & 0 & 0 \\ 0 & J_{22} & 0 \\ 0 & 0 & J_{33} \end{bmatrix}.$$

Система координат, в которой все центробежные моменты инерции тела равны нулю, называется *главной системой координат*. Моменты инерции J_{11}, J_{22}, J_{33} в главной системе координат называются *главными моментами инерции*. Главные направления тензора инерции зависят от неподвижной точки тела (от выбора начала координат). Центральная система координат, в которой все центробежные моменты инерции тела равны нулю, называется *главной центральной системой координат*. Моменты инерции в главной центральной системе координат называются *главными центральными моментами инерции*. Направления главных осей инерции тела лежат в плоскостях симметрии тела. Начало главной центральной системы координат находится в центре масс тела.

Найдем направления главных осей инерции. Пусть главному моменту инерции λ_i соответствует главное направление, определяемое ортом $\mathbf{e}_i = [x_i, y_i, z_i]^T$. Возможны три случая соотношений между корнями характеристического уравнения. Первый случай – все корни характеристического уравнения различны, второй случай – два из трех корней характе-

ристического уравнения равны и отличны от третьего корня, третий случай – все три корня характеристического уравнения равны между собой.

Если все корни характеристического уравнения различные ($\lambda_1 \neq \lambda_2 \neq \lambda_3$), то для каждого λ_i составим систему уравнений (8.11.1)

$$\begin{aligned}(J_{11} - \lambda_i) x_i + J_{12} y_i + J_{13} z_i &= 0, \\ J_{12} x_i + (J_{22} - \lambda_i) y_i + J_{23} z_i &= 0, \\ J_{13} x_i + J_{23} y_i + (J_{33} - \lambda_i) z_i &= 0.\end{aligned}$$

Среди данных трех уравнений одно является линейной комбинацией двух других, так как определитель системы (8.11.2) равен нулю. Нам нужно выяснить, какие два из трех уравнений являются линейно независимыми. Для этого вычислим определители:

$$A_1 = \begin{vmatrix} J_{22} - \lambda_i & J_{23} \\ J_{23} & J_{33} - \lambda_i \end{vmatrix}, \quad A_2 = \begin{vmatrix} J_{33} - \lambda_i & J_{13} \\ J_{13} & J_{11} - \lambda_i \end{vmatrix}, \quad A_3 = \begin{vmatrix} J_{11} - \lambda_i & J_{12} \\ J_{12} & J_{22} - \lambda_i \end{vmatrix}.$$

По крайней мере, один из них должен быть отличен от нуля. Пусть $A_2 \neq 0$. Тогда компоненты соответствующего главного вектора найдем из системы уравнений:

$$\begin{aligned}(J_{11} - \lambda_i) x_i + J_{12} y_i + J_{13} z_i &= 0, \\ J_{13} x_i + J_{23} y_i + (J_{33} - \lambda_i) z_i &= 0, \\ x_i^2 + y_i^2 + z_i^2 &= 1.\end{aligned}$$

Решив три подобные системы уравнений, мы определим собственные векторы $\mathbf{e}_i = [x_i, y_i, z_i]^T$ для каждого собственного значения λ_i , $i = 1, 2, 3$.

Если два из трех корней характеристического уравнения равны между собой, например, $\lambda_1 = \lambda_2 \neq \lambda_3$, то базисный вектор $\mathbf{e}_3 = [x_3, y_3, z_3]^T$ может быть найден описанным выше способом. Двумя другими базисными векторами могут служить любые ортогональные друг другу и вектору \mathbf{e}_3 орта.

Если $\lambda_1 = \lambda_2 = \lambda_3$, то собственными векторами могут служить любые три взаимно ортогональные орта, а $\lambda_i = I/3$.

Для определения собственных значений матрицы инерции нужно решить кубическое уравнение (8.11.3). Заменим в уравнении (8.11.3) неизвестную λ новой неизвестной χ , связанной с λ равенством $\lambda = \chi + I/3$, и вместо уравнения (8.11.3) получим неполное кубическое уравнение относительно χ , не содержащее квадрата новой неизвестной:

$$\chi^3 + p\chi + q = 0, \quad (8.11.4)$$

где введены обозначения

$$p = -\frac{1}{3}I_1^2 + I_2, \quad (8.11.5)$$

$$q = -\frac{2}{27}I_1^3 + \frac{1}{3}I_1I_2 - I_3. \quad (8.11.6)$$

Решение неполного кубического уравнения (8.11.4) дает **формула Кардано**:

$$\chi_i = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

В формуле Кардано должны быть суммированы кубические радикалы

$$\gamma_1 = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}, \quad \gamma_2 = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}},$$

удовлетворяющие условию: $\gamma_1\gamma_2 = -p/3$. Таким образом, корнями неполного кубического уравнения (8.11.4) будут

$$\chi_1 = \gamma_1 + \gamma_2, \quad \chi_2 = \epsilon\gamma_1 + \epsilon^2\gamma_2, \quad \chi_3 = \epsilon^2\gamma_1 + \epsilon\gamma_2,$$

где $\epsilon = -1/2 + i\sqrt{3}/2$ – кубический корень из единицы, а $\epsilon^2 = -1/2 - i\sqrt{3}/2$ – квадрат этого кубического корня из единицы.

Формула Кардано позволяет найти три корня, в общем случае комплексных, кубического уравнения (8.11.4). Коэффициенты кубического уравнения в общем случае могут быть комплексными. В случае вычисления главных моментов инерции коэффициенты (8.11.5) и (8.11.6) неполного кубического уравнения, так же как и инварианты I_1, I_2, I_3 , являются действительными числами. В формуле Кардано величина

$$D = -4p^3 - 27q^2$$

называется *дискриминантом* неполного кубического уравнения (8.11.4). Если коэффициенты p и q являются действительными числами, то дискриминант также является действительным числом и может принимать значения $D < 0, D = 0, D > 0$.

Если $D < 0$, то под знаком каждого из кубических радикалов формулы Кардано стоят действительные числа. Кубический корень из действительного числа имеет одно действительное и два комплексных сопряженных значения. Пусть действительными значениями этих кубических радикалов являются γ_1 и γ_2 , тогда кубическое уравнение с действительными коэффициентами имеет один действительный корень и два сопряженных комплексных корня:

$$\chi_1 = \gamma_1 + \gamma_2, \quad \chi_2 = -\frac{\gamma_1 + \gamma_2}{2} + i\sqrt{3}\frac{\gamma_1 - \gamma_2}{2}, \quad \chi_3 = -\frac{\gamma_1 + \gamma_2}{2} - i\sqrt{3}\frac{\gamma_1 - \gamma_2}{2}.$$

Если $D = 0$, то $\gamma_1 = \gamma_2 = \sqrt[3]{-q/2}$, $\varepsilon + \varepsilon^2 = -1$ и кубическое уравнение с действительными коэффициентами имеет три действительных корня, два из которых равны между собой, а третий равен сумме первых двух, взятой с противоположным знаком:

$$\chi_1 = 2\sqrt[3]{-q/2}, \quad \chi_2 = \chi_3 = -\sqrt[3]{-q/2}.$$

Если $D > 0$, то под знаками кубических радикалов формулы Кардано стоят сопряженные комплексные числа, кубические корни которых в общем случае являются комплексными числами. Среди корней кубического уравнения с действительными коэффициентами обязательно один из корней должен быть действительным числом, следовательно, γ_1 и γ_2 являются сопряженными комплексными числами:

$$\gamma_1 = -r(\cos(\gamma/3) + i\sin(\gamma/3)), \quad \gamma_2 = -r(\cos(\gamma/3) - i\sin(\gamma/3)),$$

где $r = \frac{q}{|q|} \sqrt[3]{\frac{|p|}{3}}$ (знак r должен совпадать со знаком q), $\gamma = \arccos\left(\frac{q}{2r^3}\right)$.

Таким образом, если $D > 0$, то кубическое уравнение с действительными коэффициентами имеет три действительных корня:

$$\begin{aligned} \chi_1 &= -2r \cos(\gamma/3), \\ \chi_2 &= r(\cos(\gamma/3) + \sqrt{3} \sin(\gamma/3)) = 2r \cos(\pi/3 - \gamma/3), \\ \chi_3 &= r(\cos(\gamma/3) - \sqrt{3} \sin(\gamma/3)) = 2r \cos(\pi/3 + \gamma/3). \end{aligned}$$

Все корни λ_1 , λ_2 и λ_3 характеристического кубического уравнения (8.11.3) являются действительными числами и определяются формулами

$$\begin{aligned} \lambda_1 &= \frac{I_1}{3} - 2r \cos(\gamma/3), \\ \lambda_2 &= \frac{I_1}{3} + 2r \cos(\pi/3 - \gamma/3), \\ \lambda_3 &= \frac{I_1}{3} + 2r \cos(\pi/3 + \gamma/3), \end{aligned}$$

где $p = -\frac{1}{3}I_1^2 + I_2$, $q = -\frac{2}{27}I_1^3 + \frac{1}{3}I_1I_2 - I_3$, $r = \frac{q}{|q|} \sqrt[3]{\frac{|p|}{3}}$, $\gamma = \arccos\left(\frac{q}{2r^3}\right)$.

Возможны частные случаи: если $p = 0$, то $\lambda_1 = \lambda_2 = \lambda_3 = \frac{I_1}{3}$, если $D = 0$, то

$$\lambda_1 = \frac{I_1}{3} - 2\sqrt[3]{\frac{q}{2}}, \quad \lambda_2 = \lambda_3 = \frac{I_1}{3} + \sqrt[3]{\frac{q}{2}}.$$

Упражнения

1. Где используются геометрические модели?
2. Опишите алгоритм построения векторного изображения тела.
3. Опишите алгоритм построения точечного изображения тела.
4. Сравните различные алгоритмы триангуляции.

Приложение. Криволинейные системы

В определенных случаях для построения геометрической модели удобно использовать криволинейные системы координат, например, цилиндрические или сферические. Криволинейные системы координат используются при постановке численных экспериментов в механике сплошной среды. Мы рассмотрим описание векторов и точек, а также преобразования координат точек и компонент векторов в криволинейных системах координат.

А.1. Криволинейные координаты

В криволинейной системе координат в различных точках пространства базисные векторы системы в общем случае имеют различные длины и направления, что усложняет описание геометрических объектов. Но в определенных случаях криволинейные координаты достаточно удобны.

Пусть мы выбрали в трехмерном пространстве некоторую точку и три неколлинеарных вектора. Построим на выбранной точке и векторах некоторую аффинную систему координат в пространстве. Координаты точек в этой системе обозначим через x^1, x^2, x^3 . Индексы координат расположим сверху, что даст возможность упрощать громоздкие суммы однотипных слагаемых путем введения соглашения о суммировании по повторяющимся индексам. Пусть имеются непрерывные дифференцируемые и однозначные функции

$$\begin{aligned}u^1 &= u^1(x^1, x^2, x^3), \\u^2 &= u^2(x^1, x^2, x^3), \\u^3 &= u^3(x^1, x^2, x^3),\end{aligned}$$

такие, что обратные им функции

$$\begin{aligned}x^1 &= x^1(u^1, u^2, u^3), \\x^2 &= x^2(u^1, u^2, u^3), \\x^3 &= x^3(u^1, u^2, u^3)\end{aligned}\tag{A.1.1}$$

также являются непрерывными дифференцируемыми и однозначными.

Если один из параметров в (A.1.1) зафиксировать, например, $u^i = \text{const}$, то мы получим некоторую поверхность в пространстве, которую назовем поверхностью i -й координаты. Функции (A.1.1) должны быть такими, чтобы поверхности каждой i -й координаты не пересекались друг с другом. В этом случае параметры u^1, u^2, u^3 могут служить ко-

ординатами точек в рассматриваемом пространстве. Для описания точки в пространстве с координатами u^1, u^2, u^3 будем использовать обозначение

$$\mathbf{r} = \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix},$$

аналогичное обозначению точек в аффинной системе координат. Потребуем, чтобы определитель матрицы Якоби системы (A.1.1) был отличен от нуля:

$$\begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \frac{\partial x^1}{\partial u^2} & \frac{\partial x^1}{\partial u^3} \\ \frac{\partial x^2}{\partial u^1} & \frac{\partial x^2}{\partial u^2} & \frac{\partial x^2}{\partial u^3} \\ \frac{\partial x^3}{\partial u^1} & \frac{\partial x^3}{\partial u^2} & \frac{\partial x^3}{\partial u^3} \end{vmatrix} \neq 0. \quad (\text{A.1.2})$$

Кривую в криволинейной системе координат опишем как совокупность функций $u^1 = u^1(t)$, $u^2 = u^2(t)$, $u^3 = u^3(t)$ общего параметра t , то есть в виде векторной функции:

$$\mathbf{r}(t) = \begin{bmatrix} u^1(t) \\ u^2(t) \\ u^3(t) \end{bmatrix}. \quad (\text{A.1.3})$$

Координатными линиями будем называть кривые, у которых изменяется только одна координата, а остальные координаты не меняют свои значения. Кривую $\mathbf{r}(u^k)$, у которой изменяется только k -я координата ($t = u^k$), а остальные координаты являются константами, будем называть k -й координатной линией. В общем случае координатные линии не являются прямыми, поэтому систему координат, полученную описанным способом, называют криволинейной системой координат.

В каждой точке пространства с координатами u^1, u^2, u^3 можно построить локальную систему координат с началом в рассматриваемой точке и базисными функциями $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$, являющимися производными координатных линий в рассматриваемой точке:

$$\begin{aligned} \mathbf{r}_1 &= \frac{d\mathbf{r}(u^1)}{du^1}, \\ \mathbf{r}_2 &= \frac{d\mathbf{r}(u^2)}{du^2}, \\ \mathbf{r}_3 &= \frac{d\mathbf{r}(u^3)}{du^3}. \end{aligned} \quad (\text{A.1.4})$$

В силу (А.1.2) векторы $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ не лежат в одной плоскости и могут быть использованы в качестве базиса для разложения по нему любого другого вектора в этой точке пространства. Совокупность векторов $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ будем называть **касательным базисом**. В общем случае базисные векторы $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ в каждой точке различны, поэтому по ним можно производить разложение только тех векторов, которые вычислены в этой же точке пространства. Базисный вектор \mathbf{r}_k направлен по касательной к k -й координатной линии в данной точке, его длина в общем случае отлична от единицы.

Квадрат длины бесконечно малого отрезка, заданного бесконечно малыми приращениями координат du^1, du^2, du^3 , равен

$$ds^2 = \sum_{i=1}^3 \sum_{j=1}^3 \mathbf{r}_i \cdot \mathbf{r}_j du^i du^j.$$

Коэффициенты

$$g_{ij} = g_{ij}(u^1, u^2, u^3) = \mathbf{r}_i \cdot \mathbf{r}_j \quad i, j = 1, 2, 3 \quad (\text{A.1.5})$$

называются **ковариантными компонентами метрического тензора** в точке пространства с координатами u^1, u^2, u^3 . Матрицу, составленную из ковариантных компонент метрического тензора, будем обозначать через \mathbf{G} .

Соглашение о суммировании по повторяющимся индексам позволяет многим формулам придать компактный вид. Для этого координатам приписан верхний индекс, а векторам касательного базиса и ковариантным компонентам приписаны нижние индексы.

Соглашение о суммировании по повторяющимся индексам гласит: *в выражениях выполняется суммирование по повторяющимся один раз снизу и один раз сверху индексам, при котором индексы пробегают значения от единицы до размерности пространства*. Соглашение позволяет опустить знаки суммы и придать выражениям компактный вид, например:

$$a_i b_k^i \equiv \sum_{i=1}^3 a_i b_k^i, \quad a_i b^{ik} c_k \equiv \sum_{i=1}^3 \sum_{k=1}^3 a_i b^{ik} c_k.$$

С применением соглашения о суммировании по повторяющимся индексам выражение для квадрата длины бесконечно малого отрезка примет вид

$$ds^2 = \mathbf{r}_i \cdot \mathbf{r}_j du^i du^j = g_{ij} du^i du^j.$$

Индексы, по которым производится суммирование, называют *немymi*, а остальные индексы называют *свободными*. Немые индексы отсутствуют в одной из частей равенства. Свободные индексы принимают значение номера одного из измерений пространства. Чтобы не путаться в индексах при большом их количестве, следует помнить о трех правилах.

1. Свободные индексы по обе стороны знака равенства занимают одинаковые позиции.
2. Индексы, по которым производится суммирование, встречаются один раз вверху и один раз внизу и только в одной части равенства.
3. Индекс параметра, по которому выполняется дифференцирование, является нижним.

При переходе к новой криволинейной системе координат u^1, u^2, u^3 ковариантные компоненты метрического тензора g_{kn} в новой системе координат связаны с компонентами g_{ij} соотношениями

$$g_{kn} = A^i_k A^j_n g_{ij},$$

где

$$A^i_k = \frac{\partial u^i}{\partial u^{k'}}.$$

Аналогичными соотношениями связаны и базисные векторы криволинейных систем:

$$\mathbf{r}_k = A^i_k \mathbf{r}_i. \quad (\text{A.1.6})$$

Для касательного базиса $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ в каждой точке можно построить **взаимный базис** $\mathbf{r}^1, \mathbf{r}^2, \mathbf{r}^3$ по правилу

$$\mathbf{r}_i \mathbf{r}^k = \delta_i^k, \quad (\text{A.1.7})$$

где δ_i^k – символы Кронекера, принимающие значения $\delta_i^k = 1$ при $i = k$ и $\delta_i^k = 0$ при $i \neq k$. Для нахождения векторов взаимного базиса представим их в виде разложения по касательному базису с неизвестными коэффициентами. Из приведенных выше условий получим систему линейных алгебраических уравнений относительно искомым коэффициентов. После ее решения получим коэффициенты разложения векторов взаимного базиса по векторам касательного базиса:

$$\mathbf{r}^k = g^{ki} \mathbf{r}_i, \quad (\text{A.1.8})$$

где g^{ki} – коэффициенты матрицы \mathbf{G}^{-1} . Коэффициенты $g^{ki} = \mathbf{r}^k \mathbf{r}^i$ называются **контравариантными компонентами метрического тензора**, заданного в системе координат u^1, u^2, u^3 . При переходе к другой криволинейной системе координат $u^{1'}, u^{2'}, u^{3'}$ контравариантные компоненты метрического тензора $g^{j'i'}$ в новой системе координат связаны с компонентами g^{ki} соотношениями

$$g^{j'i'} = A^{n'}_k A^i'_n g^{ki},$$

где $A^{i'}_i = \frac{\partial u^{i'}}{\partial u^i}$.

Аналогичными соотношениями связаны и векторы взаимных базисов криволинейных систем

$$\mathbf{r}^i = A^i_k \mathbf{r}^k. \quad (\text{A.1.9})$$

Так как \mathbf{G}^{-1} является обратной матрицей матрицы \mathbf{G} , то их произведение равно единичной матрице, откуда следует, что смешанные компоненты метрического тензора равны функции δ_i^k :

$$g_{ij} g^{jk} = g_i^k = \delta_i^k.$$

Векторы взаимного базиса и касательного базиса связаны соотношениями

$$\begin{aligned} \mathbf{r}^1 &= \frac{\mathbf{r}_2 \times \mathbf{r}_3}{\sqrt{g}}, \quad \mathbf{r}^2 = \frac{\mathbf{r}_3 \times \mathbf{r}_1}{\sqrt{g}}, \quad \mathbf{r}^3 = \frac{\mathbf{r}_1 \times \mathbf{r}_2}{\sqrt{g}}, \\ \mathbf{r}_1 &= \mathbf{r}^2 \times \mathbf{r}^3 \sqrt{g}, \quad \mathbf{r}_2 = \mathbf{r}^3 \times \mathbf{r}^1 \sqrt{g}, \quad \mathbf{r}_3 = \mathbf{r}^1 \times \mathbf{r}^2 \sqrt{g}, \end{aligned} \quad (\text{A.1.10})$$

где g – определитель матрицы \mathbf{G} . Последние соотношения можно записать в виде

$$\mathbf{r}^k = e^{jk} \frac{\mathbf{r}_1 \times \mathbf{r}_j}{\sqrt{g}}, \quad \mathbf{r}_k = e_{jk} \mathbf{r}^j \times \mathbf{r}^j \sqrt{g}$$

или

$$\mathbf{r}_i \times \mathbf{r}_j = e_{jk} \mathbf{r}^k \sqrt{g}, \quad \mathbf{r}^i \times \mathbf{r}^j = e^{jk} \frac{\mathbf{r}_k}{\sqrt{g}},$$

где используются символы Леви-Чивита e^{ijk} и e_{ijk} . Величины e^{ijk} и e_{ijk} равны 0, если среди индексов встречаются одинаковые; равны 1 для последовательности индексов 1,2,3 и получающейся из нее круговой перестановкой последовательностей 2,3,1 и 3,1,2; равны -1 при нарушении этого порядка (для последовательности индексов 3,2,1 и получающейся из нее круговой перестановкой последовательностей 2,1,3 и 1,3,2).

В криволинейной системе координат произвольный вектор \mathbf{a} может быть разложен по касательному или взаимному базису и представлен в виде

$$\mathbf{a} = a^j \mathbf{r}_j = a_j \mathbf{r}^j.$$

Результатом операции скалярного произведения векторов \mathbf{a} и \mathbf{b} является число, которое в зависимости от представления векторов равно

$$\mathbf{a} \cdot \mathbf{b} = a^i b^j g_{ij} = a_j b^j = a^i b_i = a_j b^j. \quad (\text{A.1.11})$$

Векторное произведение векторов \mathbf{a} и \mathbf{b} может быть представлено в виде

$$\mathbf{a} \times \mathbf{b} = a^i b^j e_{ijk} \sqrt{g} \mathbf{r}^k = a_i b_j e^{ijk} \frac{\mathbf{r}_k}{\sqrt{g}}, \quad (\text{A.1.12})$$

где использовались равенства

$$e^{ijk} \frac{1}{\sqrt{g}} = (\mathbf{r}^j \times \mathbf{r}^i) \cdot \mathbf{r}^k, \quad e_{ijk} \sqrt{g} = (\mathbf{r}_i \times \mathbf{r}_j) \cdot \mathbf{r}_k.$$

Величины

$$\varepsilon^{ijk} = e^{ijk} \frac{1}{\sqrt{g}} = (\mathbf{r}^j \times \mathbf{r}^i) \cdot \mathbf{r}^k, \quad \varepsilon_{ijk} = e_{ijk} \sqrt{g} = (\mathbf{r}_i \times \mathbf{r}_j) \cdot \mathbf{r}_k$$

являются соответственно контравариантными и ковариантными компонентами тензора Леви-Чивита.

A.2. Дифференцирование в криволинейных координатах

В аффинной системе координат базисные векторы во всех точках пространства одинаковы по величине и направлению, поэтому при дифференцировании они выступают в роли констант. В криволинейной системе координат базисные векторы во всех точках пространства различны, поэтому при дифференцировании следует учитывать изменение базисных векторов при переходе из одной точки пространства в другую.

Найдем изменение векторов касательного базиса при переходе из точки пространства с координатами u^1, u^2, u^3 в бесконечно близкую к ней точку с координатами $u^1 + \delta u^1, u^2 + \delta u^2, u^3 + \delta u^3$. С точностью до линейных слагаемых относительно δu^i , $i = 1, 2, 3$ изменение базисных векторов можно представить в виде

$$\delta \mathbf{r}_i = \frac{\partial \mathbf{r}_i}{\partial u^j} \delta u^j \equiv \mathbf{r}_{ij} \delta u^j = \Gamma_{ij}^k \mathbf{r}_k \delta u^j, \quad (\text{A.2.1})$$

где Γ_{ij}^k – неизвестные пока коэффициенты разложения производных \mathbf{r}_{ij} касательного базиса по базису $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. Во всех выражениях мы используем соглашение о суммировании по повторяющимся индексам. Умножив равенство

$$\frac{\partial \mathbf{r}_i}{\partial u^j} = \frac{\partial \mathbf{r}_i}{\partial u^j} = \mathbf{r}_{ij} = \Gamma_{ij}^k \mathbf{r}_k \quad (\text{A.2.2})$$

скалярно на $\mathbf{r}_1, \mathbf{r}_2$ и \mathbf{r}_3 , получим систему уравнений для Γ_{ij}^k :

$$\begin{aligned}\Gamma_{ij}^1 g_{11} + \Gamma_{ij}^2 g_{12} + \Gamma_{ij}^3 g_{13} &= \Gamma_{ij^{\gamma 1}} \equiv \mathbf{r}_{ij} \cdot \mathbf{r}_1, \\ \Gamma_{ij}^1 g_{21} + \Gamma_{ij}^2 g_{22} + \Gamma_{ij}^3 g_{23} &= \Gamma_{ij^{\gamma 2}} \equiv \mathbf{r}_{ij} \cdot \mathbf{r}_2, \\ \Gamma_{ij}^1 g_{31} + \Gamma_{ij}^2 g_{32} + \Gamma_{ij}^3 g_{33} &= \Gamma_{ij^{\gamma 3}} \equiv \mathbf{r}_{ij} \cdot \mathbf{r}_3.\end{aligned}\tag{A.2.3}$$

Введем обозначение $\Gamma_{ij^{\gamma k}} = \mathbf{r}_{ij} \cdot \mathbf{r}_k$. Коэффициенты $\Gamma_{ij^{\gamma k}}$ могут быть выражены через производные ковариантных компонент метрического тензора. Для этого выпишем известные равенства

$$\mathbf{r}_j \cdot \mathbf{r}_k = g_{jk}, \quad \mathbf{r}_k \cdot \mathbf{r}_i = g_{ki}, \quad \mathbf{r}_i \cdot \mathbf{r}_j = g_{ij},$$

в которых циклически переставляются индексы i, j, k ; принимающие значения от единицы до размерности пространства. Продифференцируем первое равенство по u^i , второе – по u^j , третье – по u^k и получим:

$$\begin{aligned}\mathbf{r}_{ji} \cdot \mathbf{r}_k + \mathbf{r}_j \cdot \mathbf{r}_{ki} &= \frac{\partial g_{jk}}{\partial u^i}, \\ \mathbf{r}_{kj} \cdot \mathbf{r}_i + \mathbf{r}_k \cdot \mathbf{r}_{ij} &= \frac{\partial g_{ki}}{\partial u^j}, \\ \mathbf{r}_{ik} \cdot \mathbf{r}_j + \mathbf{r}_i \cdot \mathbf{r}_{jk} &= \frac{\partial g_{ij}}{\partial u^k}.\end{aligned}$$

Если сложим первые два равенства и вычтем из них третье, то получим формулу для определения коэффициентов $\Gamma_{ij^{\gamma k}}$:

$$\Gamma_{ij^{\gamma k}} = \frac{1}{2} \left(\frac{\partial g_{kj}}{\partial u^i} + \frac{\partial g_{ik}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k} \right).\tag{A.2.4}$$

Теперь решим систему уравнений (A.2.3) относительно Γ_{ij}^k и получим равенства

$$\Gamma_{ij}^n = g^{nk} \Gamma_{ij^{\gamma k}}.\tag{A.2.5}$$

Систему (A.2.3) перепишем в виде

$$\Gamma_{ij^{\gamma n}} = g_{nk} \Gamma_{ij}^k.$$

Подставим в (A.2.5) равенства (A.2.4) и получим окончательное выражение для коэффициентов в разложении (A.2.1):

$$\Gamma_{ij}^n = \frac{1}{2} g^{nk} \left(\frac{\partial g_{kj}}{\partial u^i} + \frac{\partial g_{ik}}{\partial u^j} - \frac{\partial g_{ij}}{\partial u^k} \right).$$

Коэффициенты $\Gamma_{ij^{\gamma k}}$ называются **символами Кристоффеля первого рода**, а коэффициенты Γ_{ij}^k называются **символами Кристоффеля второго рода**. Они выражаются через компоненты метрического тензора и их частные

производные по криволинейным координатам. Символы Кристоффеля обладают симметрией $\Gamma_{im^k} = \Gamma_{mi^k}$, $\Gamma_{im}^k = \Gamma_{mi}^k$. Символы Кристоффеля не являются компонентами тензора.

Вычислим производную векторной функции в криволинейной системе координат u^1, u^2, u^3 . Пусть компоненты векторной функции $\mathbf{a}(t)$ являются непрерывными функциями $a^1 = a^1(t)$, $a^2 = a^2(t)$, $a^3 = a^3(t)$ общего параметра t . Пусть изменению параметра δt векторной функции соответствует приращение координат $\delta u^1, \delta u^2, \delta u^3$. Потребуем, чтобы максимальное приращение координат δu^i стремилось к нулю при стремлении к нулю δt . Приращение векторной функции $\mathbf{a}(t)$ в местном базисе точки пространства с координатами u^1, u^2, u^3 равно

$$\begin{aligned} \delta \mathbf{a} &= \mathbf{a}(t + \delta t) - \mathbf{a}(t) = a^i(t + \delta t) (\mathbf{r}_i + \delta \mathbf{r}_i) - a^i(t) \mathbf{r}_i = \\ &= a^i(t + \delta t) (\mathbf{r}_i + \Gamma_{ij}^k \mathbf{r}_k \delta u^j) - a^i(t) \mathbf{r}_i. \end{aligned}$$

Поделив последнее равенство на δt и устремив δt к нулю, получим формулу для производной векторной функции

$$\begin{aligned} \frac{d\mathbf{a}}{dt} &= \frac{da^i}{dt} \mathbf{r}_i + a^i \Gamma_{ij}^k \mathbf{r}_k \frac{du^j}{dt} = \frac{da^i}{dt} \mathbf{r}_i + a^n \Gamma_{ni}^j \mathbf{r}_j \frac{du^i}{dt} = \\ &= \left(\frac{da^i}{dt} + a^n \Gamma_{ni}^j \frac{du^j}{dt} \right) \mathbf{r}_i. \end{aligned} \quad (\text{A.2.6})$$

Выражение

$$\frac{Da^i}{dt} = \frac{da^i}{dt} + a^n \Gamma_{ni}^j \frac{du^j}{dt}$$

называется *абсолютной* или **ковариантной производной** контравариантных компонент векторной функции в направлении кривой $u^1(t), u^2(t), u^3(t)$.

Касательный базис и взаимный базис в некоторой точке пространства равноправны, поэтому вектор $\mathbf{a}(t)$ кривой может быть разложен также по взаимному базису в точке кривой:

$$\mathbf{a}(t) = a_i \mathbf{r}^i, \quad (\text{A.2.7})$$

где $a_i = a_i(t)$ – ковариантные компоненты вектора \mathbf{a} . Для вывода производной $\mathbf{a}(t)$ в данном представлении нужно знать, как меняются векторы взаимного базиса при переходе из точки с координатами u^1, u^2, u^3 в бесконечно близкую к ней точку с координатами $u^1 + \delta u^1, u^2 + \delta u^2, u^3 + \delta u^3$. Дифференцируя равенства (A.1.7), получим

$$\mathbf{r}_i \cdot \frac{\partial \mathbf{r}^k}{\partial u^i} = - \frac{\partial \mathbf{r}_i}{\partial u^i} \cdot \mathbf{r}^k = - \Gamma_{ij}^n \mathbf{r}_n \cdot \mathbf{r}^k = - \Gamma_{ij}^k, \quad (\text{A.2.8})$$

откуда следует, что $-\Gamma_{ij}^k$ являются коэффициентами разложения производных $\partial \mathbf{r}^k / \partial u^j$ по взаимному базису

$$\mathbf{r}_j^k \equiv \frac{\partial \mathbf{r}^k}{\partial u^j} = -\Gamma_{ij}^k \mathbf{r}^i. \quad (\text{A.2.9})$$

Действительно, если мы скалярно умножим равенство (A.2.9) на \mathbf{r}_i , то получим равенства (A.2.8). Таким образом, с точностью до линейных слагаемых относительно δu^j изменение векторов взаимного базиса равно

$$\delta \mathbf{r}^k = \frac{\partial \mathbf{r}^k}{\partial u^j} \delta u^j = -\Gamma_{ij}^k \mathbf{r}^i \delta u^j.$$

Используя покомпонентное представление (A.2.7), получим формулу для производной векторной функции кривой:

$$\begin{aligned} \frac{d\mathbf{a}}{dt} &= \frac{da_i}{dt} \mathbf{r}^i - a_i \Gamma_{ij}^k \mathbf{r}^k \frac{du^j}{dt} = \frac{da_i}{dt} \mathbf{r}^i - a_n \Gamma_{ij}^n \mathbf{r}^i \frac{du^j}{dt} = \\ &= \left(\frac{da_i}{dt} - a_n \Gamma_{ij}^n \frac{du^j}{dt} \right) \mathbf{r}^i. \end{aligned} \quad (\text{A.2.10})$$

Выражение

$$\frac{Da_i}{dt} = \frac{da_i}{dt} - a_n \Gamma_{ij}^n \frac{du^j}{dt}$$

называют **ковариантной производной** ковариантных компонент векторной функции в направлении кривой $u^1(t)$, $u^2(t)$, $u^3(t)$.

Как можно видеть, производная векторной функции представляет собой также векторную функцию. Дифференцируя представления

$$\frac{d\mathbf{a}}{dt} = \frac{Da_i}{dt} \mathbf{r}^i, \quad \frac{d\mathbf{a}}{dt} = \frac{Da^i}{dt} \mathbf{r}_i$$

и используя формулы (A.2.6) и (A.2.10), можно получить для векторной функции $\mathbf{a}(t)$ производные более высокого порядка.

A.3. Тензоры в криволинейных координатах

Тензоры представляют собой объекты, которые описываются несколькими компонентами и операции над которыми не зависят от выбранной координатной системы. В разных системах координат один и тот же тензор имеет разные компоненты. В одной и той же системе координат тензор можно представить или его ковариантными компонен-

тами, или его контравариантными компонентами, или его смешанными компонентами. Ковариантные компоненты имеют нижние индексы, контравариантные компоненты имеют верхние индексы, а смешанные компоненты имеют как нижние, так и верхние индексы. Одним из тензоров является метрический тензор. Примерами тензоров могут служить тензор инерции твердого тела, тензор кривизны пространства, тензор деформации и тензор напряжений сплошной среды. Тензоры, заданные для каждой точки некоторой области пространства, образуют *тензорное поле*.

Векторная функция представляет собой тензор первого ранга. Представим, что в каждой точке пространства задана векторная функция, компоненты которой зависят от положения в пространстве, то есть являются функциями координат u^1, u^2, u^3 . Такая векторная функция определяется в пространстве *векторное поле*.

Найдем изменение векторного поля $\mathbf{a}(u^1, u^2, u^3)$ при переходе из точки с координатами u^1, u^2, u^3 в бесконечно близкую к ней точку вдоль одной из координатных линий. Пусть этому переходу соответствует изменение координаты u^j на бесконечно малую величину δu^j . Приращение векторного поля в представлении через контравариантные компоненты $\mathbf{a}(u^1, u^2, u^3) = a^i \mathbf{r}_i$ с точностью до линейных по δu^j членов равно

$$\delta \mathbf{a} = \frac{\partial a^i}{\partial u^j} \delta u^j \mathbf{r}_i + a^i \frac{\partial \mathbf{r}_i}{\partial u^j} \delta u^j.$$

Напомним, что в выражениях мы используем соглашение о суммировании по повторяющимся индексам. Поделив обе части этого равенства на δu^j и устремив δu^j к нулю, получим формулу для производной векторного поля в представлении через контравариантные компоненты:

$$\frac{\partial \mathbf{a}}{\partial u^j} = \frac{\partial a^i}{\partial u^j} \mathbf{r}_i + a^i \Gamma_{ij}^k \mathbf{r}_k = \left(\frac{\partial a^i}{\partial u^j} + a^n \Gamma_{nj}^i \right) \mathbf{r}_i.$$

Выражение

$$\nabla_j a^i = \frac{\partial a^i}{\partial u^j} + a^n \Gamma_{nj}^i \quad (\text{A.3.1})$$

называется **ковариантной производной контравариантных компонент** векторного поля. Символ ∇_j будем называть знаком ковариантной производной по координате u^j .

Приращение векторного поля в представлении через ковариантные компоненты $\mathbf{a}(u^1, u^2, u^3) = a_i \mathbf{r}^i$ с точностью до линейных по δu^j членов равно

$$\delta \mathbf{a} = \frac{\partial a_i}{\partial u^j} \delta u^j \mathbf{r}^i + a_i \frac{\partial \mathbf{r}^i}{\partial u^j} \delta u^j.$$

Отсюда получим формулу для производной векторного поля в представлении через ковариантные компоненты:

$$\frac{\partial \mathbf{a}}{\partial u^i} = \frac{\partial a_j}{\partial u^i} \mathbf{r}^j - a_j \Gamma_{ij}^k \mathbf{r}^k = \left(\frac{\partial a_j}{\partial u^i} - a_n \Gamma_{ij}^n \right) \mathbf{r}^j.$$

Выражение

$$\nabla_j a_i = \frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \quad (\text{A.3.2})$$

называют **ковариантной производной ковариантных компонент** векторного поля.

Метрический тензор является тензором второго ранга и образует тензорное поле в пространстве. По аналогии с записью векторов в виде $\mathbf{a} = a_i \mathbf{r}^i = a'_i \mathbf{r}'_i$ можно записать тензор второго ранга, например, метрический тензор, с использованием диад базисных векторов в виде $\mathbf{M} = g_{ik} \mathbf{r}^i \mathbf{r}^k = g'^k_i \mathbf{r}'_i \mathbf{r}'_k$. Дифференцируя тензор в данном представлении по правилу дифференцирования суммы и произведения функций, получим выражение для ковариантных производных компонент тензора второго ранга:

$$\begin{aligned} \frac{\partial (g_{ik} \mathbf{r}^i \mathbf{r}^k)}{\partial u^j} &= \frac{\partial g_{ik}}{\partial u^j} \mathbf{r}^i \mathbf{r}^k + g_{ik} \frac{\partial \mathbf{r}^i}{\partial u^j} \mathbf{r}^k + g_{ik} \mathbf{r}^i \frac{\partial \mathbf{r}^k}{\partial u^j} = \\ &= \frac{\partial g_{ik}}{\partial u^j} \mathbf{r}^i \mathbf{r}^k - g_{ik} \Gamma_{ij}^n \mathbf{r}^n \mathbf{r}^k - g_{ik} \mathbf{r}^i \Gamma_{ij}^n \mathbf{r}^n = \\ &= \left(\frac{\partial g_{ik}}{\partial u^j} - g_{nk} \Gamma_{ij}^n - g_{in} \Gamma_{kj}^n \right) \mathbf{r}^i \mathbf{r}^k = \nabla_j g_{ik} \mathbf{r}^i \mathbf{r}^k, \\ \frac{\partial (g'^k_i \mathbf{r}'_i \mathbf{r}'_k)}{\partial u^j} &= \frac{\partial g'^k_i}{\partial u^j} \mathbf{r}'_i \mathbf{r}'_k + g'^k_i \frac{\partial \mathbf{r}'_i}{\partial u^j} \mathbf{r}'_k + g'^k_i \mathbf{r}'_i \frac{\partial \mathbf{r}'_k}{\partial u^j} = \\ &= \frac{\partial g'^k_i}{\partial u^j} \mathbf{r}'_i \mathbf{r}'_k + g'^k_i \Gamma_{ij}^n \mathbf{r}'_n \mathbf{r}'_k + g'^k_i \mathbf{r}'_i \Gamma_{kj}^n \mathbf{r}'_n = \\ &= \left(\frac{\partial g'^k_i}{\partial u^j} + g'^nk \Gamma_{ij}^n + g'^n \Gamma_{kj}^n \right) \mathbf{r}'_i \mathbf{r}'_k = \nabla_j g'^k_i \mathbf{r}'_i \mathbf{r}'_k, \end{aligned}$$

где

$$\nabla_j g_{ik} = \frac{\partial g_{ik}}{\partial u^j} - g_{nl} \Gamma_{ij}^n - g_{kl} \Gamma_{ij}^n, \quad (\text{A.3.3})$$

$$\nabla_j g'^k_i = \frac{\partial g'^k_i}{\partial u^j} + g'^nk \Gamma_{ij}^n + g'^kn \Gamma_{ij}^n \quad (\text{A.3.4})$$

есть ковариантные производные компонент метрического тензора. Так как при выводе равенств (А.3.3) и (А.3.4) никакие свойства компонент метрического тензора не использовались, то вместо компонент метрического тензора могут быть подставлены компоненты любого другого тензора второго ранга. Равенства (А.3.3) и (А.3.4) представляют собой определения **ковариантных производных ковариантных и контравариантных компонент тензоров второго ранга**.

Аналогично записи векторов в виде $\mathbf{a} = a\mathbf{r}^j = a^i\mathbf{r}_i$ можно записать тензор произвольного ранга с использованием его компонент и групп соответствующих базисных векторов в виде $\mathbf{T} = t^{km}\mathbf{r}_k\mathbf{r}_l\mathbf{r}_m = t_{kn}\mathbf{r}^k\mathbf{r}^l\mathbf{r}^m = t_{im}^k\mathbf{r}_k\mathbf{r}^l\mathbf{r}^m = t_n^{ki}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m$. Каждому верхнему индексу компоненты тензора соответствует вектор касательного базиса \mathbf{r}_i , а каждому нижнему индексу соответствует вектор взаимного базиса \mathbf{r}^i . Используя аналогию с векторами, можно получить формулы для ковариантных производных компонент тензора произвольного ранга. Например, дифференцируя правую часть выражения $\mathbf{T} = t_m^{ki}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m$ по правилам дифференцирования суммы и произведения функций

$$\begin{aligned} & \frac{\partial(t_m^{ki}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m)}{\partial u^j} = \\ & = \frac{\partial t_m^{ki}}{\partial u^j}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\frac{\partial\mathbf{r}_k}{\partial u^j}\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\mathbf{r}_k\frac{\partial\mathbf{r}_l}{\partial u^j} + t_m^{ki}\mathbf{r}_k\mathbf{r}_l\frac{\partial\mathbf{r}^m}{\partial u^j} = \\ & = \frac{\partial t_m^{ki}}{\partial u^j}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\Gamma_{kj}^n\mathbf{r}_n\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\Gamma_{lj}^n\mathbf{r}_k\mathbf{r}_n\mathbf{r}^m - t_m^{ki}\Gamma_{mj}^n\mathbf{r}_k\mathbf{r}_l\mathbf{r}^n = \\ & = \frac{\partial t_m^{ki}}{\partial u^j}\mathbf{r}_k\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\Gamma_{kj}^n\mathbf{r}_n\mathbf{r}_l\mathbf{r}^m + t_m^{ki}\Gamma_{lj}^n\mathbf{r}_k\mathbf{r}_n\mathbf{r}^m - t_m^{ki}\Gamma_{mj}^n\mathbf{r}_k\mathbf{r}_l\mathbf{r}^n = \\ & = \left(\frac{\partial t_m^{ki}}{\partial u^j} + t_m^{ki}\Gamma_{kj}^n + t_m^{ki}\Gamma_{lj}^n - t_m^{ki}\Gamma_{mj}^n \right) \mathbf{r}_k\mathbf{r}_l\mathbf{r}^m = \\ & = \nabla_j t_m^{ki} \mathbf{r}_k\mathbf{r}_l\mathbf{r}^m \end{aligned}$$

и взяв компоненты при одинаковых триадах базисных векторов результата, получим выражения для ковариантных производных смешанных компонент тензора третьего ранга:

$$\nabla_j t_m^{ki} = \frac{\partial t_m^{ki}}{\partial u^j} + t_m^{ki}\Gamma_{kj}^n + t_m^{ki}\Gamma_{lj}^n - t_m^{ki}\Gamma_{mj}^n.$$

Из полученного результата можно сформулировать общее правило вычисления ковариантной производной компонент тензора произвольного ранга. Кроме частной производной данной компоненты тензора $\partial t_{***}^{***} / \partial u$ (индексы заменены точками) в ковариантной производной

присутствуют еще несколько сумм, каждая сумма соответствует одному из индексов. Так, каждому нижнему индексу соответствует сумма вида $-t_{*n*}^{***}\Gamma_{*j}^n$, а каждому верхнему индексу соответствует сумма вида $+t_{***}^{*n*}\Gamma_{nj}^*$, где значение соответствующего индекса (обозначенного точкой) переходит от компоненты к символам Кристоффеля, а по индексу n производится суммирование. В символической записи это выглядит так:

$$\nabla_i t_{***}^{***} = \frac{\partial t_{***}^{***}}{\partial u^j} + \dots + t_{***}^{*n*}\Gamma_{ni}^* + \dots - \dots - t_{*n*}^{***}\Gamma_{*i}^n - \dots$$

Покажем, что компоненты метрического тензора можно выносить из-под знака ковариантной производной ∇_j . Дифференцируя равенства $g_{ki} = \mathbf{r}_k \cdot \mathbf{r}_i$ и $g^{ki} = \mathbf{r}^k \cdot \mathbf{r}^i$ и используя (A.2.2) и (A.2.9), получим выражения для производных компонент метрического тензора:

$$\frac{\partial g_{kn}}{\partial u^j} = g_{kn}\Gamma_{ij}^n + g_{ni}\Gamma_{kj}^n, \quad (\text{A.3.5})$$

$$\frac{\partial g^{ki}}{\partial u^j} = -g^{kn}\Gamma_{nj}^i - g^{ni}\Gamma_{nj}^k. \quad (\text{A.3.6})$$

Подставив в равенство (A.3.1) соотношения $a^i = g^{jk}a_k$ и (A.3.6), получим

$$\begin{aligned} \nabla_j a^i &= \nabla_j(g^{jk}a_k) = g^{jk}\frac{\partial a_k}{\partial u^j} + a_k\frac{\partial g^{jk}}{\partial u^i} + a_k g^{nk}\Gamma_{nj}^i = \\ &= g^{jk}\frac{\partial a_k}{\partial u^j} - a_k g^{ni}\Gamma_{nj}^k = g^{jk}\left(\frac{\partial a_k}{\partial u^j} - a_n\Gamma_{kj}^n\right) = g^{jk}\nabla_j a_k. \end{aligned} \quad (\text{A.3.7})$$

Во время преобразований мы меняли обозначение немых индексов. Аналогично подставив в равенство (A.3.2) соотношения $a_i = g_{ik}a^k$ и (A.3.5), получим

$$\begin{aligned} \nabla_j a_i &= \nabla_j(g_{ik}a^i) = g_{ik}\frac{\partial a^i}{\partial u^j} + a^i\frac{\partial g_{ik}}{\partial u^j} - a^k g_{kn}\Gamma_{ij}^n = \\ &= g_{ik}\frac{\partial a^i}{\partial u^j} + a^k g_{ni}\Gamma_{kj}^n = g_{ik}\left(\frac{\partial a^i}{\partial u^j} + a^n\Gamma_{nj}^k\right) = g_{ik}\nabla_j a^k. \end{aligned} \quad (\text{A.3.8})$$

Из равенств (A.3.7) и (A.3.8) следует вывод о том, что компоненты метрического тензора можно выносить из-под знака ковариантной производной. Как и следовало ожидать, из соотношений (A.3.5) и (A.3.6) следует, что ковариантные производные компонент метрического тензора (A.3.3) и (A.3.4) равны нулю.

Набла-оператор, или **оператор Гамильтона**, обозначают символом ∇ , формальная запись которого имеет вид

$$\nabla = \frac{\partial}{\partial u^j} \mathbf{r}^j = \frac{\partial}{\partial u^1} \mathbf{r}^1 + \frac{\partial}{\partial u^2} \mathbf{r}^2 + \frac{\partial}{\partial u^3} \mathbf{r}^3.$$

Оператору Гамильтона приписывают атрибуты вектора. В результате действия оператора Гамильтона на тензор может быть построен новый тензор. Например, в результате действия оператора Гамильтона на тензор нулевого ранга (скаляр) f получим тензор первого ранга (вектор), который называют градиентом скаляра:

$$\nabla f = \mathbf{r}^j \frac{\partial f}{\partial u^j} = \mathbf{r}^1 \frac{\partial f}{\partial u^1} + \mathbf{r}^2 \frac{\partial f}{\partial u^2} + \mathbf{r}^3 \frac{\partial f}{\partial u^3} = \text{grad } f.$$

Оператор Гамильтона может действовать на тензоры тензорно (результатом является тензор, ранг которого на единицу выше ранга исходного тензора), скалярно (при этом ранг результата на единицу меньше ранга исходного тензора) и векторно (при этом ранг результата равен рангу исходного тензора):

$$\nabla \mathbf{a} = \mathbf{r}^j \frac{\partial \mathbf{a}}{\partial u^j} = \mathbf{r}^j \mathbf{r}^i \left(\frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \right) = \nabla_j a_i \mathbf{r}^j = \text{grad } \mathbf{a},$$

$$\nabla \cdot \mathbf{a} = \mathbf{r}^j \cdot \frac{\partial \mathbf{a}}{\partial u^j} = \mathbf{r}^j \cdot \mathbf{r}^i \left(\frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \right) = \nabla_j a^j = \text{div } \mathbf{a},$$

$$\nabla \times \mathbf{a} = \mathbf{r}^j \times \frac{\partial \mathbf{a}}{\partial u^j} = \mathbf{r}^j \times \mathbf{r}^i \left(\frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \right) = \frac{1}{\sqrt{g}} e^{ijk} \mathbf{r}_k \nabla_j a_i = \text{rot } \mathbf{a}.$$

Для тензора $\mathbf{T} = t_m^{ki} \mathbf{r}_k \mathbf{r}_i \mathbf{r}^m$ результат действия оператора Гамильтона имеет вид

$$\nabla \mathbf{T} = \mathbf{r}^j \frac{\partial \mathbf{T}}{\partial u^j} = \nabla_j t_m^{ki} \mathbf{r}_k \mathbf{r}_i \mathbf{r}^m = \text{grad } \mathbf{T},$$

$$\nabla \cdot \mathbf{T} = \mathbf{r}^j \cdot \frac{\partial \mathbf{T}}{\partial u^j} = \nabla_j t_m^{ji} \mathbf{r}_i \mathbf{r}^m = \text{div } \mathbf{T},$$

$$\nabla \times \mathbf{T} = \mathbf{r}^j \times \frac{\partial \mathbf{T}}{\partial u^j} = \frac{e^{jkm}}{\sqrt{g}} g_{ka} \nabla_j t_{in}^{ki} \mathbf{r}_a \mathbf{r}_i \mathbf{r}^m = \frac{e^{jkn}}{\sqrt{g}} \nabla_j t_{lm}^{li} \mathbf{r}_i \mathbf{r}^m = \text{rot } \mathbf{T}.$$

Оператор Гамильтона может воздействовать несколько раз на один и тот же объект. Так, скалярное произведение операторов Гамильтона называется оператором Лапласа

$$\nabla^2 = \nabla \cdot \nabla = g'' \frac{\partial^2}{\partial u^i \partial u^i} - g'^a \Gamma_{ik}^a \frac{\partial}{\partial u^j}.$$

Действие оператора Лапласа на скаляр записывается в виде

$$\nabla \nabla f = \nabla^2 f = \operatorname{div} \operatorname{grad} f.$$

В декартовой системе координат оператор Лапласа имеет вид

$$\nabla^2 = \frac{\partial^2}{(\partial x^1)^2} + \frac{\partial^2}{(\partial x^2)^2} + \frac{\partial^2}{(\partial x^3)^2}.$$

Оператор Гамильтона используется при описании тензорных (в частом случае скалярных или векторных) полей. Приведенные формулы позволяют записывать уравнения для тензоров, имеющие инвариантный вид относительно системы координат. Например, дифференциал da векторного поля $\mathbf{a}(u^1, u^2, u^3)$ с использованием оператора Гамильтона можно выразить следующим образом:

$$d\mathbf{a} = \frac{\partial \mathbf{a}}{\partial u^i} du^i = du^n \mathbf{r}_n \cdot \mathbf{r}^j \frac{\partial \mathbf{a}}{\partial u^j} = d\mathbf{r} \cdot \nabla \mathbf{a} = (\nabla \mathbf{a})^T \cdot d\mathbf{r},$$

где мы использовали обозначение $d\mathbf{r} = du^n \mathbf{r}_n$.

А.4. Примеры тензоров

Тензор деформации используется в механике сплошной среды для определения изменений расстояний и углов между точками среды. Тензор деформации является симметричным тензором второго ранга. Рассмотрим две бесконечно близкие точки P и Q среды. Пусть до деформирования точка P имела координаты u^1, u^2, u^3 , а точка Q имела координаты $u^1 + du^1, u^2 + du^2, u^3 + du^3$. Бесконечно малое волокно, соединяющее эти точки, можно описать вектором $d\mathbf{r} = du^i \mathbf{r}_i$. Напомним, что в выражениях мы используем соглашение о суммировании по повторяющимся индексам. Квадрат длины волокна, соединяющего точки P и Q среды, до деформирования выражается через компоненты метрического тензора:

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = du^i \mathbf{r}_i \cdot \mathbf{r}_j du^j = du^i g_{ij} du^j = [du^1 \ du^2 \ du^3] \cdot \mathbf{G} \cdot [du^1 \ du^2 \ du^3]^T,$$

где $\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix}$ – матрица метрического тензора. Пусть векторное

поле $\mathbf{a}(u^1, u^2, u^3) = \begin{bmatrix} a^1(u^1, u^2, u^3) \\ a^2(u^1, u^2, u^3) \\ a^3(u^1, u^2, u^3) \end{bmatrix}$ описывает перемещение материальной

точки с координатами u^1, u^2, u^3 . В результате деформирования сплошной

среды рассматриваемые точки сместятся и займут новое положение. После деформирования точка P будет описываться координатами $u^1 + a^1, u^2 + a^2, u^3 + a^3$, а точка Q будет описываться координатами $u^1 + du^1 + a^1 + da^1, u^2 + du^2 + a^2 + da^2, u^3 + du^3 + a^3 + da^3$. Дифференциал векторного поля перемещений $da(u^1, u^2, u^3)$ описывает изменение расстояний между бесконечно близкими точками сплошной среды. После деформирования рассматриваемое бесконечно малое волокно будет описываться вектором

$$d\mathbf{r}' = d\mathbf{r} + da = d\mathbf{r} + (\nabla \mathbf{a})^T \cdot d\mathbf{r} = du^i \mathbf{r}_i + \nabla_j a^n du^j \mathbf{r}_n.$$

Квадрат длины волокна, соединяющего точки P и Q среды, после деформирования выразится следующим образом:

$$\begin{aligned} (ds')^2 &= d\mathbf{r}' \cdot d\mathbf{r}' = (d\mathbf{r} + (\nabla \mathbf{a})^T \cdot d\mathbf{r}) \cdot (d\mathbf{r} + (\nabla \mathbf{a})^T \cdot d\mathbf{r}) = \\ &= (du^i \mathbf{r}_i + \nabla_j a^n du^j \mathbf{r}_n) \cdot (du^i \mathbf{r}_i + \nabla_j a^k du^j \mathbf{r}_k) = \\ &= du^i (g_{ij} + \nabla_j a_i + \nabla_j a_i + \nabla_j a^n \nabla_j a_n) du^i = \\ &= [du^1 \ du^2 \ du^3] \cdot \mathbf{G}' \cdot [du^1 \ du^2 \ du^3]^T, \end{aligned}$$

где \mathbf{G}' – матрица с компонентами $g'_{ij} = g_{ij} + \nabla_j a_i + \nabla_j a_i + \nabla_j a^n \nabla_j a_n$. Тензор деформации определяют равенством

$$\boldsymbol{\varepsilon} = \frac{1}{2}(\mathbf{G}' - \mathbf{G}),$$

где $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix}$ – матрица тензора деформации в ковариантном

представлении. Ковариантные компоненты тензора деформации определяются формулой

$$\varepsilon_{ij} = \frac{1}{2}(\nabla_j a_i + \nabla_j a_i + \nabla_j a^n \nabla_j a_n). \quad (\text{A.4.1})$$

Формула (A.4.1) получена в предположении, что координатная система u^1, u^2, u^3 остается неизменной в процессе деформирования. В механике сплошной среды рассмотрение некоторого процесса в неподвижной системе координат называют *пространственным представлением*, а саму систему – *эйлеровой*. Возможен и другой подход, при котором координатная система u^1, u^2, u^3 считается связанной с материальными точками и деформируется вместе со средой. При таком подходе координаты материальных точек остаются неизменными в процессе деформирования, а меняются компоненты метрического тензора. Бесконечно малое

волокно, соединяющее рассматриваемые точки, после деформирования будет описываться вектором $d\mathbf{r}' = du^i \mathbf{r}'_i$, где базисные векторы \mathbf{r}'_i отличаются от базисных векторов \mathbf{r}_i исходного состояния. Описание некоторого процесса таким способом называют *материальным представлением*, а используемую систему — *лагранжевой*. Ковариантные компоненты тензора деформации в материальном представлении определяются формулой

$$\varepsilon_{ij}' = \frac{1}{2}(\mathbf{g}'_{ij} - \mathbf{g}_{ij}). \quad (\text{A.4.2})$$

Тензор деформации может быть задан в системе координат с базисными векторами \mathbf{r}_i (в исходной конфигурации при пространственном представлении) или может быть задан в системе координат с базисными векторами \mathbf{r}'_i (в конечной конфигурации при материальном представлении). В первом случае тензор деформации носит название тензора деформации Коши—Грина, во втором случае он носит название тензора деформации Альманзи—Гамеля. В общем случае тензор деформации может быть описан в произвольной системе координат.

В качестве примера получим тензор деформации в некоторой системе координат, приближенно описывающей деформированное состояние сплошной среды. Использование формулы (A.4.1) при больших перемещениях осложнено нелинейными относительно компонент вектора перемещения слагаемыми $\nabla_i a^j \nabla_j a_n$, которыми не всегда можно пренебречь. Для использования формулы (A.4.2) необходимо знать метрический тензор для начального состояния и для конечного состояния, которое является неизвестным и подлежит определению. Предположим, что мы пытаемся применить материальное представление, и в процессе решения задачи нам приближенно известно деформированное состояние среды (например, из предыдущей итерации). Так как деформированное состояние мы знаем приближенно, то координаты материальных точек в нашем подходе не остаются неизменными, но их изменение невелико. Основные изменения претерпевают базисные векторы координатной системы, переходя от \mathbf{r}_i в недеформированном состоянии к \mathbf{r}'_i , приближенно описывающем деформированное состояние среды. При таком подходе после деформирования бесконечно малое волокно, соединяющее рассматриваемые точки P и Q , будет описываться вектором

$$d\mathbf{r}' = d\mathbf{r} + d\mathbf{v} = d\mathbf{r} + (\nabla \mathbf{v})^T \cdot d\mathbf{r} = du^i \mathbf{r}'_i + \nabla_i v^j du^i \mathbf{r}'_j,$$

где векторное поле $\mathbf{v}(u^1, u^2, u^3) = \begin{bmatrix} v^1(u^1, u^2, u^3) \\ v^2(u^1, u^2, u^3) \\ v^3(u^1, u^2, u^3) \end{bmatrix}$ описывает перемещение в

конечное состояние точки среды с координатами u^1, u^2, u^3 . Если деформированное состояние среды мы знали бы точно, то выполнялось бы равенство $\mathbf{r}_n'' = \mathbf{r}_n'$ и векторное поле $\mathbf{v}(u^1, u^2, u^3)$ было бы равно нулю. Вектор $\mathbf{v}(u^1, u^2, u^3)$ описывает величину нашей ошибки в знании деформированного состояния среды. Квадрат длины волокна между точками P и Q среды после деформирования в данном случае выразится следующим образом:

$$\begin{aligned} (ds'')^2 &= (du^i \mathbf{r}_i'' + \nabla_j v^i du^j \mathbf{r}_i'') \cdot (du^k \mathbf{r}_k'' + \nabla_l v^k du^l \mathbf{r}_k'') = \\ &= du^i (g_{ij}'' + \nabla_j v_i + \nabla_i v_j + \nabla_i v^k \nabla_j v_k) du^j. \end{aligned}$$

Компоненты тензора деформации в таком подходе определяются равенством

$$\varepsilon_{ij} = \frac{1}{2} (g_{ij}'' - g_{ij}' + \nabla_j v_i + \nabla_i v_j + \nabla_i v^k \nabla_j v_k). \quad (\text{A.4.3})$$

Благодаря тому что \mathbf{r}_n'' близки к \mathbf{r}_n' , компоненты вектора $\mathbf{v}(u^1, u^2, u^3)$ малы, и формулу (A.4.3) можно линеаризовать, представив в виде

$$\varepsilon_{ij} = \frac{1}{2} (g_{ij}'' - g_{ij}' + \nabla_j v_i + \nabla_i v_j). \quad (\text{A.4.4})$$

Формула (A.4.4) может быть использована при больших перемещениях. Ее достоинство в том, что она содержит только линейные слагаемые относительно компонент вектора $\mathbf{v}(u^1, u^2, u^3)$. Определение деформированного состояния должно выполняться итерационными методами. На каждой итерации нам приближенно известно деформированное состояние среды и связанной с ней координатной системы u^1, u^2, u^3 . Для этого состояния нужно вычислять базисные векторы \mathbf{r}_n'' и компоненты метрического тензора g_{ij}'' . Далее, используя (A.4.4), найдем очередную поправку $\mathbf{v}(u^1, u^2, u^3)$ к векторному полю перемещений. В точке решения векторное поле $\mathbf{v}(u^1, u^2, u^3)$ становится равным нулю.

Тензор кривизны используется для определения свойства пространства. В искривленных пространствах последовательность дифференцирования по координатам в общем случае играет роль. При изменении последовательности дифференцирования появляются дополнительные слагаемые. Пусть дана некоторая векторная функция $\mathbf{a} = a_i(u^1, u^2, u^3) \mathbf{r}_i'$. Рассмотрим некоторую поверхность $u^i = u^i(t, w)$, $i = 1, 2, 3$ в пространстве. Всегда можно выбрать поверхность так, чтобы она проходила через лю-

бую заданную точку или линию. Вычислим абсолютный дифференциал $D_a \mathbf{a}$ векторной функции при бесконечно малом смещении из точки с координатами u^1, u^2, u^3 в точку с координатами $u^1 + \frac{\partial u^1}{\partial t} dt, u^2 + \frac{\partial u^2}{\partial t} dt, u^3 + \frac{\partial u^3}{\partial t} dt$ вдоль поверхности в направлении первого ее параметра:

$$D_t \mathbf{a} = \left(\left(\frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \right) \mathbf{r}^i \frac{\partial u^j}{\partial t} \right) dt = \left(\frac{\partial a_i}{\partial t} - a_n \Gamma_{ij}^n \frac{\partial u^j}{\partial t} \right) \mathbf{r}^i dt \equiv (D_t a)_i \mathbf{r}^i dt.$$

Аналогично вычислим абсолютный дифференциал $D_w \mathbf{a}$ векторной функции при бесконечно малом смещении из той же точки в точку с координатами $u^1 + \frac{\partial u^1}{\partial w} dw, u^2 + \frac{\partial u^2}{\partial w} dw, u^3 + \frac{\partial u^3}{\partial w} dw$ вдоль поверхности в направлении второго ее параметра:

$$D_w \mathbf{a} = \left(\left(\frac{\partial a_i}{\partial u^j} - a_n \Gamma_{ij}^n \right) \mathbf{r}^i \frac{\partial u^j}{\partial w} \right) dw = \left(\frac{\partial a_i}{\partial w} - a_n \Gamma_{ij}^n \frac{\partial u^j}{\partial w} \right) \mathbf{r}^i dw \equiv (D_w a)_i \mathbf{r}^i dw.$$

Полученные дифференциалы можно также рассматривать как векторные функции тех же параметров в той же точке. Вычислим дифференциал векторной функции $D_a \mathbf{a}$ вдоль поверхности в направлении второго ее параметра и дифференциал векторной функции $D_w \mathbf{a}$ вдоль поверхности в направлении первого ее параметра

$$\begin{aligned} D_w D_t \mathbf{a} &= \left(\frac{\partial (D_t a)_i}{\partial w} - (D_t a)_n \Gamma_{ij}^n \frac{\partial u^j}{\partial w} \right) \mathbf{r}^i dw = \\ &= \left(\frac{\partial^2 a_i}{\partial t \partial w} - \frac{\partial a_n}{\partial w} \Gamma_{ij}^n \frac{\partial u^j}{\partial t} - a_n \frac{\partial \Gamma_{ij}^n}{\partial w} \frac{\partial u^j}{\partial t} - a_n \Gamma_{ij}^n \frac{\partial^2 u^j}{\partial t \partial w} - \right. \\ &\quad \left. - \frac{\partial a_n}{\partial t} \Gamma_{ir}^n \frac{\partial u^r}{\partial w} + a_n \Gamma_{ir}^m \frac{\partial u^i}{\partial t} \Gamma_{rn}^m \frac{\partial u^r}{\partial w} \right) \mathbf{r}^i dt dw, \\ D_t D_w \mathbf{a} &= \left(\frac{\partial (D_w a)_i}{\partial t} - (D_w a)_n \Gamma_{ij}^n \frac{\partial u^j}{\partial t} \right) \mathbf{r}^i dt = \\ &= \left(\frac{\partial^2 a_i}{\partial w \partial t} - \frac{\partial a_n}{\partial t} \Gamma_{ij}^n \frac{\partial u^j}{\partial w} - a_n \frac{\partial \Gamma_{ij}^n}{\partial t} \frac{\partial u^j}{\partial w} - a_n \Gamma_{ij}^n \frac{\partial^2 u^j}{\partial w \partial t} - \right. \\ &\quad \left. - \frac{\partial a_n}{\partial w} \Gamma_{ir}^n \frac{\partial u^r}{\partial t} + a_n \Gamma_{ir}^m \frac{\partial u^i}{\partial w} \Gamma_{rn}^m \frac{\partial u^r}{\partial t} \right) \mathbf{r}^i dw dt. \end{aligned}$$

Полученные выражения отличаются тем, что в них переставлены местами производные по t и по w . Первые и четвертые слагаемые обоих выра-

жений равны. Второе слагаемое первого выражения равно пятому слагаемому второго выражения (если немой индекс r обозначить через j). Вычтем из первого выражения второе и, меняя обозначение немых индексов, получим

$$\begin{aligned}
 & D_w D_t \mathbf{a} - D_t D_w \mathbf{a} = \\
 & = \left(-a_n \frac{\partial \Gamma_{ij}^n}{\partial w} \frac{\partial u^i}{\partial t} + a_m \Gamma_{ij}^m \frac{\partial u^i}{\partial t} \Gamma_{ir}^n \frac{\partial u^r}{\partial w} \right) \mathbf{r}^i dw dt + \\
 & + \left(a_n \frac{\partial \Gamma_{ij}^n}{\partial t} \frac{\partial u^j}{\partial w} - a_m \Gamma_{ij}^m \frac{\partial u^j}{\partial w} \Gamma_{ir}^n \frac{\partial u^r}{\partial t} \right) \mathbf{r}^i dw dt = \\
 & = \left(\left(-\frac{\partial \Gamma_{ij}^n}{\partial u^r} + \Gamma_{ij}^n \Gamma_{ir}^q \right) \frac{\partial u^i}{\partial t} \frac{\partial u^r}{\partial w} + \left(\frac{\partial \Gamma_{ir}^n}{\partial u^j} - \Gamma_{qr}^n \Gamma_{ij}^q \right) \frac{\partial u^j}{\partial t} \frac{\partial u^r}{\partial w} \right) a_n \mathbf{r}^i dw dt = \\
 & = \left(-\frac{\partial \Gamma_{ij}^n}{\partial u^r} + \Gamma_{ij}^n \Gamma_{ir}^k + \frac{\partial \Gamma_{ir}^n}{\partial u^j} - \Gamma_{kr}^n \Gamma_{ij}^k \right) \frac{\partial u^i}{\partial t} \frac{\partial u^r}{\partial w} a_n \mathbf{r}^i dw dt = \\
 & = R_{ji}^n \frac{\partial u^i}{\partial t} \frac{\partial u^r}{\partial w} a_n dw dt \mathbf{r}^i, \tag{A.4.5}
 \end{aligned}$$

где введено обозначение

$$R_{ji}^n = \frac{\partial \Gamma_{ir}^n}{\partial u^j} - \frac{\partial \Gamma_{ij}^n}{\partial u^r} + \Gamma_{ij}^n \Gamma_{ir}^k - \Gamma_{kr}^n \Gamma_{ij}^k. \tag{A.4.6}$$

Коэффициенты R_{ji}^n являются компонентами *тензора кривизны* пространства. Тензор кривизны называют *тензором Римана-Кристоффеля*. Он зависит лишь от точки пространства, в которой вычислен.

Если в качестве поверхности $u^i = u^i(t, w)$ взять одну из координатных поверхностей $u^1 = u^1(u^2, u^3)$, $u^2 = u^2(u^3, u^1)$, $u^3 = u^3(u^1, u^2)$, проходящую через рассматриваемую точку, то выражение (A.4.5) примет вид

$$D_r D_s \mathbf{a} - D_s D_r \mathbf{a} = R_{ji}^n a_n \mathbf{r}^i du^j du^k, \tag{A.4.7}$$

так как $dw = du^i$ и $dt = du^i$. Левая часть (A.4.7) называется вторым альтернативным дифференциалом векторной функции. Если векторную функцию взять в виде разложения по касательному базису $\mathbf{a} = a^i(u^1, u^2, u^3) \mathbf{r}_i$, то проведя аналогичный вывод, получим следующее выражение для второго альтернативного дифференциала векторной функции:

$$D_r D_s \mathbf{a} - D_s D_r \mathbf{a} = -R_{ji}^n a^i \mathbf{r}_n du^j du^k.$$

Из формулы (A.4.7) видно, что тензор кривизны кососимметричен по первым двум индексам:

$$R_{ji}^n = -R_{jri}^n.$$

Компоненты тензора кривизны (А.4.6) являются трижды ковариантными и один раз контравариантными. С помощью компонент метрического тензора можно получить полностью ковариантные компоненты тензора кривизны:

$$R_{rijk} = g_{nk} R_{rji}^n.$$

Ковариантные компоненты тензора кривизны удовлетворяют условиям симметрии:

$$R_{rijk} = -R_{jrik} = -R_{rjki} = R_{ikrj}.$$

Если над тензором кривизны выполнить операцию свертки по индексам n и r , то мы получим *тензор Риччи*, ковариантные компоненты которого равны

$$R_{ji} = R_{njin} = \frac{\partial \Gamma_{in}^n}{\partial u^j} - \Gamma_{kn}^n \Gamma_{ji}^k - \frac{\partial \Gamma_{ij}^n}{\partial u^n} + \Gamma_{kn}^n \Gamma_{in}^k.$$

Свертыванием тензора Риччи получим инвариант кривизны:

$$R = g^{ji} R_{ji}.$$

Тензор кривизны характеризует искривленность пространства. В евклидовом пространстве все компоненты тензора кривизны равны нулю. Двумерное пространство на некоторой поверхности в общем случае не является евклидовым и имеет ненулевой тензор кривизны. Плоская поверхность представляет собой пример евклидова двумерного пространства. Хотя координатные линии на плоскости могут быть кривыми, все компоненты тензора кривизны плоского пространства равны нулю. Пространства, для которых компоненты тензора кривизны везде равны нулю, по аналогии с поверхностями называются плоскими.

В пространстве любой размерности тензор кривизны выражается через компоненты метрического тензора и его производные первого и второго порядка. Это свойство использовал А. Эйнштейн для получения уравнений гравитационного поля в общей теории относительности. Он искал уравнение, аналогичное уравнению Пуассона, в котором тензор энергии-импульса был бы пропорционален дифференциальному выражению второго порядка, образованному компонентами метрического тензора четырехмерного пространства-времени. Такое уравнение в общем случае имеет вид

$$R_{ik} + c R g_{ik} + \lambda g_{ik} + \mu T_{ik} = 0,$$

где R_{ik} – ковариантные компоненты тензора Риччи; R – инвариант кривизны; g_{ik} – ковариантные компоненты метрического тензора; T_{ik} – ковариантные компоненты тензора энергии-импульса; c, λ, μ – константы.

Исходя из принципов общей теории относительности, А. Эйнштейн пришел к уравнению

$$R_{ik} - \frac{1}{2} R g_{ik} + \lambda g_{ik} = -\mu T_{ik}.$$

В первоначальном варианте уравнения А. Эйнштейн положил $\lambda = 0$, но позже ввел лямбду, которая позволила получить стационарное решение. Затем Э. Хаббл открыл расширение вселенной.

Изучением пространств различной размерности с ненулевым тензором кривизны занимается риманова геометрия.

А.5. Ортогональные криволинейные координаты

Описание геометрии объектов в криволинейных координатах является достаточно сложным. На практике в большинстве случаев используются ортогональные криволинейные координаты. В ортогональной системе координат координатные линии различного семейства взаимно ортогональны. Взаимно ортогональны и векторы $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ касательного базиса, так как они направлены по касательным к соответствующим координатным линиям. В ортогональной криволинейной системе координат векторы $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ касательного базиса и соответствующие им векторы $\mathbf{r}^1, \mathbf{r}^2, \mathbf{r}^3$ взаимного базиса совпадают по направлению, но длина их в общем случае различна; недиагональные компоненты метрического тензора равны нулю, то есть $g_{ij} = 0$ и $g^{ij} = 0$ при $i \neq j$, а $g^{ii} = \frac{1}{g_{ii}}$. Часть символов

Кристоффеля в ортогональной криволинейной системе координат равна нулю и, следовательно, многие формулы упрощаются.

Цилиндрическая система координат является ортогональной системой координат. Параметрами цилиндрической системы координат служат полярный радиус $u^1 = r$, полярный угол $u^2 = \varphi$ и вертикальная ось $u^3 = z$. Цилиндрические координаты r, φ, z связаны с декартовыми координатами x^1, x^2, x^3 равенствами

$$r = \sqrt{x^1 x^1 + x^2 x^2}, \quad \varphi = \arctg\left(\frac{x^2}{x^1}\right), \quad z = x^3.$$

Обратные зависимости имеют вид

$$x^1 = r \cos \varphi, \quad x^2 = r \sin \varphi, \quad x^3 = z.$$

На рис. А.5.1 приведены декартовы и цилиндрические координаты точки \mathbf{p} .

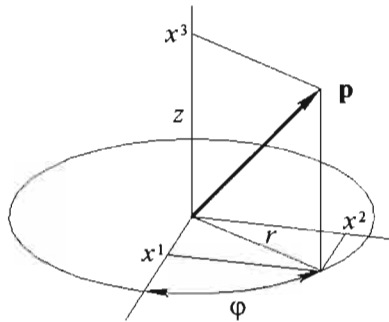


Рис. А.5.1

Будем считать, что параметры цилиндрической системы координат изменяются в следующих пределах: $0 \leq r$, $0 \leq \varphi < 2\pi$.

Матрица Якоби перехода от декартовой системы координат к цилиндрической системе координат равна

$$\begin{bmatrix} \frac{\partial x^1}{\partial u^1} & \frac{\partial x^1}{\partial u^2} & \frac{\partial x^1}{\partial u^3} \\ \frac{\partial x^2}{\partial u^1} & \frac{\partial x^2}{\partial u^2} & \frac{\partial x^2}{\partial u^3} \\ \frac{\partial x^3}{\partial u^1} & \frac{\partial x^3}{\partial u^2} & \frac{\partial x^3}{\partial u^3} \end{bmatrix} = \begin{bmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Компоненты метрического тензора в новой системе координат определяются по формуле

$$g_{\mu\nu} = \frac{\partial x^i}{\partial u^\mu} \frac{\partial x^j}{\partial u^\nu} \delta_j^i,$$

где δ_j^i – символы Кронекера (А.1.7). Компоненты метрического тензора и отличные от нуля символы Кристоффеля в ней равны

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G}^{-1} = \begin{bmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\Gamma_{12^2 2} = \Gamma_{21^2 2} = r, \quad \Gamma_{22^2 1} = -r,$$

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r}, \quad \Gamma_{22}^1 = -r.$$

Как можно заметить, в цилиндрических координатах длины первого и третьего векторов касательного базиса равны единице, а длина второго вектора, соответствующего полярному углу, равна r .

Векторные функции в цилиндрической системе координат будем выражать с помощью векторов касательного и взаимного базисов следующим образом:

$$\mathbf{a} = a^r \mathbf{e}_r + a^\varphi \mathbf{e}_\varphi + a^z \mathbf{e}_z = a_r \mathbf{e}^r + a_\varphi \mathbf{e}^\varphi + a_z \mathbf{e}^z. \quad (\text{A.5.1})$$

Частные производные векторов касательного базиса (A.2.2) по цилиндрическим координатам равны

$$\begin{aligned} \frac{\partial \mathbf{e}_r}{\partial r} &= 0, & \frac{\partial \mathbf{e}_r}{\partial \varphi} &= \frac{1}{r} \mathbf{e}_\varphi, & \frac{\partial \mathbf{e}_r}{\partial z} &= 0, \\ \frac{\partial \mathbf{e}_\varphi}{\partial r} &= \frac{1}{r} \mathbf{e}_\varphi, & \frac{\partial \mathbf{e}_\varphi}{\partial \varphi} &= -r \mathbf{e}_r, & \frac{\partial \mathbf{e}_\varphi}{\partial z} &= 0, \\ \frac{\partial \mathbf{e}_z}{\partial r} &= 0, & \frac{\partial \mathbf{e}_z}{\partial \varphi} &= 0, & \frac{\partial \mathbf{e}_z}{\partial z} &= 0. \end{aligned}$$

Частные производные векторов взаимного базиса (A.2.9) по цилиндрическим координатам равны

$$\begin{aligned} \frac{\partial \mathbf{e}^r}{\partial r} &= 0, & \frac{\partial \mathbf{e}^r}{\partial \varphi} &= r \mathbf{e}^\varphi, & \frac{\partial \mathbf{e}^r}{\partial z} &= 0, \\ \frac{\partial \mathbf{e}^\varphi}{\partial r} &= -\frac{1}{r} \mathbf{e}^\varphi, & \frac{\partial \mathbf{e}^\varphi}{\partial \varphi} &= -\frac{1}{r} \mathbf{e}^r, & \frac{\partial \mathbf{e}^\varphi}{\partial z} &= 0, \\ \frac{\partial \mathbf{e}^z}{\partial r} &= 0, & \frac{\partial \mathbf{e}^z}{\partial \varphi} &= 0, & \frac{\partial \mathbf{e}^z}{\partial z} &= 0. \end{aligned}$$

Мы видим, что некоторые векторы касательного и взаимного базисов цилиндрической системы координат изменяются при переходе от одной точки пространства к другой. Все это в общем случае усложняет описание геометрических объектов, но в некоторых частных случаях использование криволинейных систем является оправданным. Формула (A.2.6) для производной векторной функции в цилиндрической системе координат имеет вид

$$\frac{d\mathbf{a}}{dt} = \left(\frac{da^r}{dt} - r a^\varphi \frac{d\varphi}{dt} \right) \mathbf{e}_r + \left(\frac{da^\varphi}{dt} + \frac{a^\varphi}{r} \frac{dr}{dt} + \frac{a^r}{r} \frac{d\varphi}{dt} \right) \mathbf{e}_\varphi + \frac{da^z}{dt} \mathbf{e}_z. \quad (\text{A.5.2})$$

Сферическая система координат является ортогональной системой координат. Параметрами сферической системы координат служат ради-

ус $u^1 = r$, долгота $u^2 = \varphi$ и широта $u^3 = \theta$. Сферические координаты r, φ, θ связаны с декартовыми координатами x^1, x^2, x^3 равенствами

$$r = \sqrt{x^1 x^1 + x^2 x^2 + x^3 x^3}, \quad \varphi = \operatorname{arctg}\left(\frac{x^2}{x^1}\right), \quad \theta = \operatorname{arctg}\left(\frac{\sqrt{x^1 x^1 + x^2 x^2}}{x^3}\right).$$

Обратные зависимости имеют вид

$$x^1 = r \cos\varphi \sin\theta, \quad x^2 = r \sin\varphi \sin\theta, \quad x^3 = r \cos\theta.$$

На рис. А.5.2 приведены декартовы и сферические координаты точки \mathbf{p} .

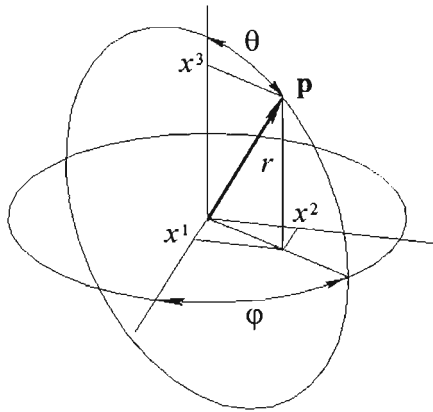


Рис. А.5.2

Будем считать, что параметры сферической системы координат изменяются в следующих пределах: $0 \leq r$, $0 \leq \varphi < 2\pi$, $0 \leq \theta \leq \pi$.

Матрица Якоби перехода от декартовой системы координат к цилиндрической системе координат равна

$$\begin{bmatrix} \frac{\partial x^1}{\partial u^1} & \frac{\partial x^1}{\partial u^2} & \frac{\partial x^1}{\partial u^3} \\ \frac{\partial x^2}{\partial u^1} & \frac{\partial x^2}{\partial u^2} & \frac{\partial x^2}{\partial u^3} \\ \frac{\partial x^3}{\partial u^1} & \frac{\partial x^3}{\partial u^2} & \frac{\partial x^3}{\partial u^3} \end{bmatrix} = \begin{bmatrix} \cos\varphi \sin\theta & -r \sin\varphi \sin\theta & r \cos\varphi \cos\theta \\ \sin\varphi \sin\theta & r \cos\varphi \sin\theta & r \sin\varphi \cos\theta \\ \cos\theta & 0 & -r \sin\theta \end{bmatrix}.$$

Компоненты метрического тензора в сферической системе координат равны

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 \sin^2 \theta & 0 \\ 0 & 0 & r^2 \end{bmatrix},$$

$$\mathbf{G}^{-1} = \begin{bmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2 \sin^2 \theta} & 0 \\ 0 & 0 & \frac{1}{r^2} \end{bmatrix}.$$

Отличны от нуля только следующие производные компонент метрического тензора:

$$\frac{\partial g_{22}}{\partial u^1} = 2r \sin^2 \theta, \quad \frac{\partial g_{22}}{\partial u^3} = 2r^2 \sin \theta \cos \theta, \quad \frac{\partial g_{33}}{\partial u^1} = 2r.$$

В соответствии с (A.2.4) отличные от нуля символы Кристоффеля в сферической системе координат равны

$$\begin{aligned} \Gamma_{12^2 2} &= \Gamma_{21^2 2} = r \sin^2 \theta, & \Gamma_{22^2 1} &= -r \sin^2 \theta, \\ \Gamma_{23^2 2} &= \Gamma_{32^2 2} = r^2 \sin \theta \cos \theta, & \Gamma_{22^2 3} &= -r^2 \sin \theta \cos \theta, \\ \Gamma_{13^2 3} &= \Gamma_{31^2 3} = r, & \Gamma_{33^2 1} &= -r, \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{1}{r}, & \Gamma_{22}^1 &= -r \sin^2 \theta, \\ \Gamma_{23}^2 &= \Gamma_{32}^2 = \text{ctg} \theta, & \Gamma_{22}^3 &= -\sin \theta \cos \theta, \\ \Gamma_{13}^3 &= \Gamma_{31}^3 = \frac{1}{r}, & \Gamma_{33}^1 &= -r. \end{aligned}$$

Векторные функции в сферической системе координат будем выражать с помощью векторов касательного и взаимного базисов следующим образом:

$$\mathbf{a} = a^r \mathbf{e}_r + a^\varphi \mathbf{e}_\varphi + a^\theta \mathbf{e}_\theta = a_r \mathbf{e}^r + a_\varphi \mathbf{e}^\varphi + a_\theta \mathbf{e}^\theta. \quad (\text{A.5.3})$$

Частные производные векторов касательного базиса (A.2.2) по сферическим координатам равны

$$\begin{aligned} \frac{\partial \mathbf{e}_r}{\partial r} &= 0, & \frac{\partial \mathbf{e}_r}{\partial \varphi} &= \frac{1}{r} \mathbf{e}_\varphi, & \frac{\partial \mathbf{e}_r}{\partial \theta} &= \frac{1}{r} \mathbf{e}_\theta, \\ \frac{\partial \mathbf{e}_\varphi}{\partial r} &= \frac{1}{r} \mathbf{e}_\varphi, & \frac{\partial \mathbf{e}_\varphi}{\partial \varphi} &= -r \sin^2 \theta \mathbf{e}_r - \sin \theta \cos \theta \mathbf{e}_\theta, & \frac{\partial \mathbf{e}_\varphi}{\partial \theta} &= \text{ctg} \theta \mathbf{e}_\varphi, \\ \frac{\partial \mathbf{e}_\theta}{\partial r} &= \frac{1}{r} \mathbf{e}_\theta, & \frac{\partial \mathbf{e}_\theta}{\partial \varphi} &= \text{ctg} \theta \mathbf{e}_\varphi, & \frac{\partial \mathbf{e}_\theta}{\partial \theta} &= -r \mathbf{e}_r. \end{aligned}$$

Частные производные векторов взаимного базиса (А.2.9) по сферическим координатам равны

$$\begin{aligned} \frac{\partial \mathbf{e}^r}{\partial r} &= 0, & \frac{\partial \mathbf{e}^r}{\partial \varphi} &= r \sin^2 \theta \mathbf{e}^\varphi, & \frac{\partial \mathbf{e}^r}{\partial \theta} &= r \mathbf{e}^\theta, \\ \frac{\partial \mathbf{e}^\varphi}{\partial r} &= -\frac{1}{r} \mathbf{e}^\varphi, & \frac{\partial \mathbf{e}^\varphi}{\partial \varphi} &= -\frac{1}{r} \mathbf{e}^r - \operatorname{ctg} \theta \mathbf{e}^\theta, & \frac{\partial \mathbf{e}^\varphi}{\partial \theta} &= -\operatorname{ctg} \theta \mathbf{e}^\varphi, \\ \frac{\partial \mathbf{e}^\theta}{\partial r} &= -\frac{1}{r} \mathbf{e}^\theta, & \frac{\partial \mathbf{e}^\theta}{\partial \varphi} &= \sin \theta \cos \theta \mathbf{e}^\varphi, & \frac{\partial \mathbf{e}^\theta}{\partial \theta} &= -\frac{1}{r} \mathbf{e}^r. \end{aligned}$$

Формула (А.2.6) для производной векторной функции в сферической системе координат имеет вид

$$\begin{aligned} \frac{d\mathbf{a}}{dt} &= \left(\frac{da^r}{dt} - a^\varphi r \sin^2 \theta \frac{d\varphi}{dt} - a^\theta r \frac{d\theta}{dt} \right) \mathbf{e}_r + \\ &+ \left(\frac{da^\varphi}{dt} + \frac{a^\varphi}{r} \frac{dr}{dt} + \frac{a^r}{r} \frac{d\varphi}{dt} + a^\varphi \operatorname{ctg} \theta \frac{d\theta}{dt} + a^\theta \operatorname{ctg} \theta \frac{d\varphi}{dt} \right) \mathbf{e}_\varphi + \\ &+ \left(\frac{da^\theta}{dt} + \frac{a^\theta}{r} \frac{dr}{dt} + \frac{a^r}{r} \frac{d\theta}{dt} - a^\varphi \sin \theta \cos \theta \frac{d\varphi}{dt} \right) \mathbf{e}_\theta. \end{aligned} \quad (\text{А.5.4})$$

Геометрическое моделирование обычно выполняется в декартовых системах координат, но входная и выходная информация о моделируемом объекте может быть представлена в криволинейных координатах. Чаще других криволинейных систем используются цилиндрическая и сферическая система координат. Приведенная в приложении информация позволяет построить любую криволинейную систему координат и получить ее связь с декартовой системой.

Пример 1. Окружность в цилиндрической системе координат можно описать векторной функцией

$$\mathbf{r}(t) = \begin{bmatrix} \rho \\ t \\ 0 \end{bmatrix}, \quad 0 \leq t \leq 2\pi.$$

Координаты связаны с параметрами окружности равенствами $r = \rho$, $\varphi(t) = t$, $z = 0$. Тогда производная окружности будет описываться векторной функцией

$$\mathbf{a} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Производную окружности представим в виде (А.5.1), то есть $\mathbf{a} = \mathbf{e}_\varphi$. Производная окружности направлена вдоль локального вектора \mathbf{e}_φ в текущей точке. Вторая производная окружности в соответствии с (А.5.2) равна

$$\frac{d\mathbf{a}}{dt} = -r\dot{\alpha} \frac{d\varphi}{dt} \mathbf{e}_r = -\rho \mathbf{e}_r,$$

и направлена к центру окружности против локального вектора \mathbf{e}_r в текущей точке. На рис. А.5.3 приведены векторы \mathbf{e}_r , \mathbf{e}_φ , \mathbf{a} и производная $d\mathbf{a}/dt$.

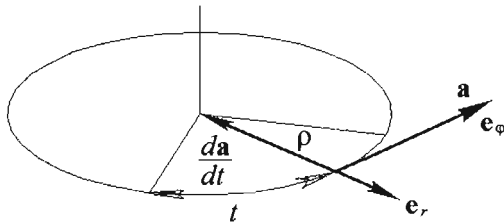


Рис. А.5.3

Пример 2. Поверхность сферы в сферической системе координат можно описать векторной функцией

$$\mathbf{r}(u, v) = \begin{bmatrix} \rho \\ u \\ \frac{\pi}{2} - v \end{bmatrix}, \quad u \in [0, 2\pi], v \in [-\pi/2, \pi/2].$$

Координаты связаны с параметрами поверхности равенствами $r = \rho$, $\varphi(u) = u$, $\theta(v) = \pi/2 - v$. Тогда частная производная поверхности по первому параметру будет описываться векторной функцией

$$\mathbf{a}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

а частная производная поверхности по второму параметру будет описываться векторной функцией

$$\mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}.$$

Частные производные поверхности представим в виде (А.5.3), то есть $\mathbf{a}_1 = \mathbf{e}_\varphi$, $\mathbf{a}_2 = -\mathbf{e}_\theta$. Частные производные поверхности направлены вдоль локальных векторов \mathbf{e}_φ и \mathbf{e}_θ в текущей точке. Частные производные вектора \mathbf{a}_1 в соответствии с (А.5.4) равны

$$\begin{aligned} \frac{\partial \mathbf{a}_1}{\partial u} &= -a^\circ r \sin^2 \theta \frac{\partial \varphi}{\partial u} \mathbf{e}_r - a^\circ \sin \theta \cos \theta \frac{\partial \varphi}{\partial u} \mathbf{e}_\theta = -\rho \cos^2 \nu \mathbf{e}_r - \cos \nu \sin \nu \mathbf{e}_\theta, \\ \frac{\partial \mathbf{a}_1}{\partial \nu} &= a^\circ \operatorname{ctg} \theta \frac{\partial \theta}{\partial \nu} \mathbf{e}_\varphi = -\operatorname{tg} \nu \mathbf{e}_\varphi. \end{aligned}$$

Частные производные вектора \mathbf{a}_2 в соответствии с (А.5.4) равны

$$\begin{aligned} \frac{\partial \mathbf{a}_2}{\partial u} &= a^\circ \operatorname{ctg} \theta \frac{\partial \varphi}{\partial u} \mathbf{e}_\varphi = -\operatorname{tg} \nu \mathbf{e}_\varphi, \\ \frac{\partial \mathbf{a}_2}{\partial \nu} &= -a^\circ r \frac{\partial \theta}{\partial \nu} \mathbf{e}_r = -\rho \mathbf{e}_r. \end{aligned}$$

Производная $\frac{\partial \mathbf{a}_1}{\partial u}$ направлена перпендикулярно оси сферы. Для производных $\frac{\partial \mathbf{a}_1}{\partial \nu}$ и $\frac{\partial \mathbf{a}_2}{\partial u}$ выполняется равенство $\frac{\partial \mathbf{a}_1}{\partial \nu} = \frac{\partial \mathbf{a}_2}{\partial u}$. Производная $\frac{\partial \mathbf{a}_2}{\partial \nu}$ направлена к центру сферы против локального вектора \mathbf{e}_r в текущей точке. На рис. А.5.4 приведены векторы \mathbf{e}_r , \mathbf{e}_φ , \mathbf{a}_1 , \mathbf{a}_2 и их производные.

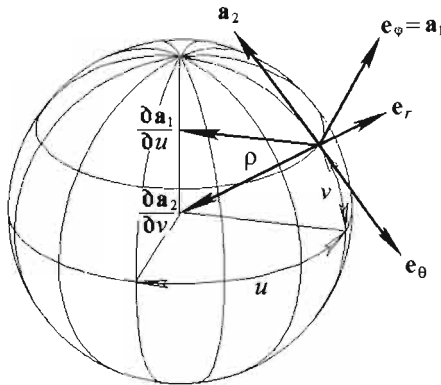


Рис. А.5.4

Список литературы

1. *Гильберт Д., Кон–Фоссен С.* Наглядная геометрия. – М. : Наука, 1981. (Hilbert D., Cohn–Vossen S. Anschauliche Geometrie. – Berlin, 1932).
2. *Голованов Н.Н.* Геометрическое моделирование. – М. : Физматлит, 2002.
3. *Де Бор К.* Практическое руководство по сплайнам. – М. : Радио и связь, 1985. (De Boor C. A Practical Guide to Splines. – Springer ; Berlin ; 1879).
4. Компьютерная геометрия. Голованов Н.Н., Ильютко Д.П., Носовский Г.В., Фоменко А.Т. – М. : Академия. : 2006.
5. *Препарата Ф., Шеймос М.* Вычислительная геометрия. – М. : Мир, 1989. (Preparata F.P., Shamos M. Computational Geometry: An Introduction. – Springer-Verlag. – New York ; Berlin ; Tokyo, 1985).
6. *Рашевский П.К.* Курс дифференциальной геометрии. – М ; Л. : Физматгиз, 1950.
7. *Фокс А., Прайт М.* Вычислительная геометрия. – М. : Мир, 1982. (Faux I.E., Pratt M.J. Computational Geometry for Design and Manufacture. – Ellis Horwood ; Chichester, 1979).
8. *Фоли Дж., Вен Дем А.* Основы интерактивной машинной графики. – Т. 1, Т. 2. – М. : Мир, 1985. (James D. Foley, Andries Van Dam. Fundamentals Interactive Computer Graphics. – Addison-Wesley Publishing Company Reading. – Massachusetts ; Menlo Park ; California ; London ; Amsterdam ; Don Mills ; Ontario ; Sydney, 1982).
9. *Bartels R.H., Beatty J.C., Barsky B.A.* An Introduction to Splines for use in Computer Graphics and Geometric Modeling. – Morgan Kaufmann Publishers Inc. – San Mateo ; California, 1987.
10. *Bohm W.* A Survey of Curve and Surface Methods in CAGD. – Computer Aided Geometric Design. – North-Holland Amsterdam, v.1, n.1. – July 1984.
11. *Cardon D.L.* T-Spline Simplification. – Brigham Young University, 2007.
12. *Christoph M. Hoffman.* Geometric and Solid Modeling. – Morgan Kaufmann Publishers Inc. – San Mateo ; California, 1989.

13. *Kramer C.A.* Solving Geometric Constraint System: A Case Study in Kinematics. – The MIT Press. Cambridge ; Massachusetts ; London, 1992.
14. *Les Piegl, Wayne Tiller.* The NURBS Book. – Springer, 1997.
15. *Neamtu M.* Homogeneous Simplex Splines. – Vanderbilt University, Nashville.
16. *Sokolnikoff I.S.* Tensor Analysis. Theory And Applications To Geometry And Mechanics Of Continua.
17. *Versprille K.J.* Computer-aided design applications of the rational *B*-spline approximation from. – Diss. Syracuse University, 1975.

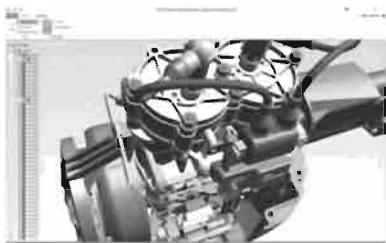
Предметный указатель

Алгоритм		
Де Бура	72	
Де Кастелье	41	
Бариецентрические		
координаты	149	
Булевы операции над телами	262	
Вектор Дарбу	25	
Вершина	236	
Вес точки	20	
Взаимный базис	369	
Вторая квадратичная форма		
поверхности	104	
Гауссова кривизна		
поверхности	108	
Геометрическая модель	320	
Геометрические		
ограничения	302	
Главная нормаль кривой	23	
Главные кривизны	107	
Грань	235	
Дерево построения	320	
Деривационные формулы		
Вейнгартена	104	
Касательный базис	368	
Ковариантная		
производная	373, 374	
Контрольные точки	137	
Контрольный		
многогранник	137, 143, 154	
Контур	96	
Кратные узлы	56	
Кривая	22	
аналитическая	27	
Безье	38	
ломаная	30	
на поверхности	100, 200	
отрезок	31	
очерка	202	
пересечения поверхностей	205	
перехода	96	
плоская	199	
продолженная	92	
проекционная	201	
рациональная	51	
репараметризованная	94	
составная	96	
ссылочная	95	
усеченная	93	
эквилистантная	95	
Кривизна кривой	25, 27	
Кручение кривой	25	
Кубические функции		
смещения	129	
Метод Фонга	336	
Моменты инерции	353	
Моменты инерции		
главные	361	
центральные	360	
Натуральная параметризация	23	
Нормаль поверхности	103	
Нормальная плоскость		
кривой	23	
Нормальное сечение		
поверхности	105	
Оболочка	236	
однородная	242	
Однородные координаты	19	
Оператор		
Гамильтона	378	
сдвига	41	
Ориентируемость оболочки	240	
Первая квадратичная форма		
поверхности	101	
Поверхность	99	
NURBS	139	
аналитическая	111	
Безье	136	
бикубическая	133	
билинейная	126	
вращения	116	
выдавливания	115	
Гордона	130	
деформируемая	177	
кинематическая	117	

коническая	111	Спрямяющая плоскость	23
Кунса кубическая	129	Средняя кривизна	
Кунса линейная	129	поверхности.....	108
Лагранжа	121	Степнная функция	
линейчатая.....	120	двумерных точек.....	161
на базе триангуляции.....	179	Тело.....	243
перехода	123	движения кривой.....	251
продолженная	174	элементарное	246
с произвольной границей	182	Тензор.....	374
сдвига	115	деформации.....	380
скругления	216	инерции.....	353
ссылочная	176	кривизны	383
тора	112	Леви Чивита	371
фаски	224	метрический.....	368, 369
эквидистантная.....	175	Риччи.....	386
Эрмита	122	Теорема Менье	106
Полуоткрытая		Топологический тип	241
выпуклая область	160	Треугольная поверхность	153
Радиус кривизны кривой.....	26	Безье	154
Радиус-вектор	12	Триангуляция	341
Разделенная разность	54	Делоне	341
однородная	159	Уравнение Эйнштейна	387
полинома	56	Усеченная степенная	
Расширенная матрица	19	функция	58
Ребро.....	235	Формула	
Связность оболочки	238	Кардано.....	363
Символы		Кокса—Де Бура	62
Кристоффеля	372	Лейбница	58
Леви-Чивита	370	Пуанкаре.....	242
Симплексный сплайн	166	Формулы	
Синхронное моделирование.....	290	Френе—Серре.....	25, 27
Система координат.....	13	Функции Бернштейна	39
криволинейная	366	трехиндексные	154
сферическая	389	Цикл грани	235
цилиндрическая	387	Эйлерова характеристика	
Соприкасающаяся плоскость... ..	23	оболочки	236
Сплайн.....	31	<i>B</i> -поверхность	139
кубический	33	<i>B</i> -сплайн.....	59
Лагранжа.....	31	ненормированный	61
Ньютона	32	DMS-сплайн.....	174
симплексный	166	NURBS	67
Эрмита	35	<i>T</i> -поверхность	144

ОТ ТЕОРИИ К ПРАКТИКЕ: ГЕОМЕТРИЧЕСКОЕ ЯДРО C3D ДЛЯ ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ

Автор данной книги – Николай Николаевич Голованов – руководитель разработки геометрического ядра C3D российской компании C3D Labs (дочернее предприятие АСКОН). C3D – программный компонент для разработчиков прикладных решений, реализующий описанные в книге математические методы построения и управления геометрией численных моделей. Ядро C3D применяется как компаниями-разработчиками коммерческого программного обеспечения, так и учебными заведениями, специально для которых в C3D Labs разработана образовательная программа.



Использование C3D в научно-образовательной среде открывает возможности для:

- Изучения внутренней структуры типовых CAD/CAM/CAE-систем;
- Знакомства с современным математическим аппаратом при выполнении геометрических расчётов и построении 2D и 3D-моделей;
- Обучения стандартам разработки программных средств с применением геометрических ядер;
- Закрепления теоретического материала по высшей математике, инженерной графике, линейной алгебре и аналитической геометрии;
- Написания методического обеспечения к лабораторным работам и практикумам;
- Выполнения курсовых, выпускных квалификационных работ и диссертаций;
- Создания программных инструментов для осуществления НИиОКР;
- Разработки коммерческих IT-продуктов на базе инновационных предприятий при учебном заведении.

Образовательная лицензия C3D включает, кроме самого программного компонента, бесплатную техническую поддержку, документацию по геометрическому ядру, методические указания и информационную поддержку проектов, выполненных с использованием C3D.

По всем вопросам, связанным с лицензированием и использованием геометрического ядра C3D, обращайтесь в компанию C3D Labs любым удобным для вас способом:



C3D Labs

127106, Москва, Алтуфьевское шоссе, д. 1,
Бета-Центр, офис 112
E-mail: info@c3dlabs.com
Тел.: +7 (495) 783-2560
<http://c3dlabs.com/ru/>

ОТ ТЕОРИИ К ПРАКТИКЕ: ГЕОМЕТРИЧЕСКОЕ ЯДРО C3D ДЛЯ ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ

Автор данной книги – Николай Николаевич Голованов – руководитель разработки геометрического ядра C3D российской компании C3D Labs (дочернее предприятие АСКОН). C3D – программный компонент для разработчиков прикладных решений, реализующий описанные в книге математические методы построения и управления геометрией численных моделей. Ядро C3D применяется как компаниями-разработчиками коммерческого программного обеспечения, так и учебными заведениями, специально для которых в C3D Labs разработана образовательная программа.



Использование C3D в научно-образовательной среде открывает возможности для:

- Изучения внутренней структуры типовых CAD/CAM/CAE-систем;
- Знакомства с современным математическим аппаратом при выполнении геометрических расчётов и построении 2D и 3D-моделей;
- Обучения стандартам разработки программных средств с применением геометрических ядер;
- Закрепления теоретического материала по высшей математике, инженерной графике, линейной алгебре и аналитической геометрии;
- Написания методического обеспечения к лабораторным работам и практикумам;
- Выполнения курсовых, выпускных квалификационных работ и диссертаций;
- Создания программных инструментов для осуществления НИиОКР;
- Разработки коммерческих IT-продуктов на базе инновационных предприятий при учебном заведении.

Образовательная лицензия C3D включает, кроме самого программного компонента, бесплатную техническую поддержку, документацию по геометрическому ядру, методические указания и информационную поддержку проектов, выполненных с использованием C3D.

По всем вопросам, связанным с лицензированием и использованием геометрического ядра C3D, обращайтесь в компанию C3D Labs любым удобным для вас способом:



127106, Москва, Алтуфьевское шоссе, д. 1,
Бета-Центр, офис 112
E-mail: info@c3dlabs.com
Тел.: +7 (495) 783-2560
<http://c3dlabs.com/ru/>

Голованов Николай Николаевич

Геометрическое моделирование

Учебное пособие

Оригинал-макет подготовлен в Издательстве «КУРС»

Подписано в печать 27.07.2015. Гарнитура Newton.
Бумага офсетная. Формат 60×90/16. Усл. печ. л. 25.
Тираж 500 (1-й завод – 300) экз. Заказ №
Цена свободная

ТК 437600 – 520536 – 270715

ООО Издательство «КУРС»
127273, Москва, ул. Олонецкая, д. 17А, офис 104.
Тел.: (499) 709-16-28.
E-mail: kursizdat@gmail.com

ООО «Научно-издательский центр ИНФРА-М»
127282, Москва, ул. Полярная, д. 31В, стр. 1
Тел.: (495) 280-15-96, 280-33-86. Факс: (495) 280-36-29
E-mail: books@infra-m.ru <http://www.infra-m.ru>