5.1 Типы геометрических моделей

Что такое геометрическая модель в машинной графике?

В системах машинной графики информация о различных физических процессах или объектах отображается на экране дисплея в виде синтезированного изображения их математических моделей. Если математическая модель сцены и обрабатывающая программа точно передают трехмерную геометрическую форму объектов, их взаимное положение, освещенность, тени и т.д., то и синтезируемое изображение будет максимально приближаться к реалистичному. Для отображения *геометрических* свойств *объектов* и создания их виртуального представления во всех трех измерениях применяются *геометрические модели объектов.* Процесс создания геометрической модели связан с получением понятного машине *математического описания* геометрических свойств объекта, в результате чего формируется *набор данных*, представляющих свойства объекта, и совокупность отношений между этими данными. Эта информация хранится, наряду с другой, в базе данных и используется для вывода изображения на графические устройства вывода: графопостроитель или экран дисплея. Данные, которые хранятся в модели, различаются и сильно зависят от требуемого качества моделирования. Действительно, моделирование всегда несовершенно, и получить отображение предмета в ЭВМ, соответствующее реалистичному, можно лишь частично. Поэтому, в машинной графике для достижения требуемого качества изображения предмета используются различные типы моделей.

Геометрические модели, формируемые методом *машинного* моделирования, могут быть *двумерными и трехмерными* и, по мере возрастания сложности, распределяются в следующем порядке:

-двумерная проволочная (каркасная),
-двумерная полигональная,
-трехмерная проволочная (каркасная),
-трехмерная полигональная (поверхностная),
-трехмерная объемная.

Геометрические модели, как двух- так и трехмерные, имеют многоуровневую иерархическую структуру (рис.5.1.1.), нулевой уровень которой представлен *элементарными* геометрическими объектами (ЭГО). На основе ЭГО можно сформировать *базовые* геометрические объекты (БГО), которые, в свою очередь, могут быть использованы для создания новых, более сложных *составных* геометрических объектов (СГО).

Двумерные модели (2D) обычно представлены совокупностью таких ЭГО, как точки, прямые и отрезки прямых, окружности и дуги окружностей, плоские кривые, многоугольники и т.п. В *трехмерных моделях* (3D), кроме перечисленных выше ЭГО, используются также плоскости, линейчатые поверхности, поверхности вращения, криволинейные поверхности. Для обоих типов моделей (2D и 3D) следует отметить понятие БГО. Это такие СГО, которые являются неделимыми на данном уровне формирования модели и зависят от ее сложности.

Например, для 2D-модели это может быть плоский контур сложной конфигурации, а для 3D-объемной модели - тела (параллелепипед, цилиндр, конус, сфера и др.).

На примере рис.5.1.2. показана модель пирамиды SABC в виде графа, которая состоит из ряда точек и линий, определяющих взаимосвязи между точками, ребрами и плоскостями геометрических элементов пирамиды.

Данные о геометрических объектах хранятся в базе данных в структурированной форме, воспроизводящей *структуру* отображаемых объектов, и представлены в памяти компьютера *массивами чисел*, которые могут задаваться параметрами как исходные данные или определяться в ходе построения самой модели как результат вычислений.

В каркасных и полигональных моделях используются данные следующих типов:

-геометрические (координаты вершин, уравнения ребер, плоскостей, поверхностей);

-топологические (данные, определяющие связи между вершинами и ребрами, ребрами и гранями, гранями и телами);

-вспомогательные (данные, передающие цвет, фактуру, освещенность объекта, прозрачность граней и т.п., которые используются для визуализации модели).

Рассмотрим типы трехмерных моделей более подробно.

Каркасные модели (рис.5.1.3.) появились одними из первых.

Информацией, хранящейся в этих моделях, являются координаты *(x, y, z) вершин* и соединяющие их *ребра*. С помощью этой модели можно представить такие объекты, в которых аппроксимирующими поверхностями являются, как правило, плоскости. Достоинством такой модели является ее простота, а недостатком – невозможность автоматически анализировать и осуществлять операции удаления невидимых линий, а также строить различные сечения.

Полигональные модели (рис.5.1.4.) используются чаще всего при описании сложных форм, например, в авиа- или автопромышленности. Конструктивными элементами *полигональной* модели объекта являются *точки (вершины), ребра и поверхности*. При этом вершины определяются как точки пересечения трех поверхностей, а ребра - как линии пересечения или касания поверхностей. Поверхности, используемые в этих моделях, могут задаваться плоскостями, поверхностями вращения, линейчатыми поверхностями и др.

Сложная поверхность объекта может быть аппроксимирована многогранником, каждая грань которого является простейшим плоским многоугольником (треугольник, четырехугольник). Использование метода аппроксимации поверхностей общего вида плоскими гранями имеет свои достоинства и недостатки. К достоинствам можно отнести простоту математических методов обработки плоскостей. К недостаткам – прямая зависимость достоверности вида поверхностей объекта от количества граней, используемых для аппроксимации каждой из поверхностей модели. Вид и формы объекта приближаются к действительным при увеличении числа граней, однако, это влечет к увеличению объема и времени вычислений и размера памяти.

При использовании полигональной модели возможно применение автоматической операции удаления невидимых линий и поверхностей, а также визуальная проверка пересечений поверхностей.

Объемная модель (рис.5.1.5.) строится из тех же конструктивных элементов, что и предыдущие модели, но поверхности при этом всегда принадлежат объемам. Объемы могут быть представлены в виде различных структур, для формирования которых применяются любые способы, приведенные ниже:

А) объем определяется совокупностью *ограничивающих его поверхностей;*

Б) объем определяется комбинацией элементарных объемов (параллелепипеда, конуса, цилиндра, сферы и др.) с использованием булевых операций объединения, пересечения и вычитания.

В общем случае объемную модель можно представить как совокупность каркасной и полигональной модели, в ней можно выполнить операции удаления невидимых линий и поверхностей, а также построение сечений и разрезов (со штриховкой). В объемной модели хранится также информация, позволяющая отличать материал от пустоты (пустота может рассматриваться как особый вид материала).

5.2. Геометрические объекты, применяемые для формирования моделей

Описание внешних геометрических форм 3D объекта - сложный процесс. В нем просматриваются следующие *подходы* к решению этой проблемы:

-использование методов *точного аналитического описания* ограничивающих контуров или поверхностей в *полярной* или *декартовой прямоугольной системе координат*;

-использование *приближенных методов интерполяции и аппроксимации*. Кривые и поверхности, получаемые таким образом, называются *аналитически неописываемыми* геометрическими объектами.

Во втором случае чаще всего используется параметрическое представление поверхностей и кривых, так как, во-первых, параметрическое представление дает возможность изучать неявные функции, когда переход к их явному (аналитическому) заданию без посредства параметров затруднителен, и, во-вторых, удается выражать многозначные функции посредством однозначных.

Но в обоих случаях любая задача машинной графики (двух- или трехмерной) сводится к определению точек в двумерном пространстве с последующим выделением некоторым способом отдельных точек для попарного соединения их прямыми линиями или для заполнения областей, очерченных такими линиями. То есть, задача фактически сводится к тому, как определить и использовать эти точки при формировании геометрической модели.

5.2.1. Геометрические объекты, описываемые аналитически.

Точка. Из аналитической геометрии известно, что любая точка пространства задается тройкой чисел *(x, y, z)* или радиус–вектором.

В однородных координатах она может быть представлена *вектором* следующим образом: $[P] = [x \quad y \quad z \quad 1]$.

Точка как результат пересечения трех плоскостей. Рассмотрим систему

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= h_1, \\ a_2 x + b_2 y + c_2 z &= h_2, . \\ a_3 x + b_3 y + c_3 z &= h_3 \end{aligned}$$

Введем обозначения

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \text{ (определитель системы),}$$

$$\Delta_x = \begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix}, \quad \Delta_y = \begin{vmatrix} a_1 & h_1 & c_1 \\ a_2 & h_2 & c_2 \\ a_3 & h_3 & c_3 \end{vmatrix}, \quad \Delta_z = \begin{vmatrix} a_1 & b_1 & h_1 \\ a_2 & b_2 & h_2 \\ a_3 & b_3 & h_3 \end{vmatrix}.$$

Если определитель системы не равен нулю $\Delta \neq 0$, то система имеет единственное решение:

$$x = \frac{\Delta_x}{\Delta}, \qquad y = \frac{\Delta_y}{\Delta}, \qquad z = \frac{\Delta_z}{\Delta}.$$

То есть три плоскости пересекаются в одной точке.

Прямая в пространстве. Зададим на прямой L точку $M_0(x_0, y_0, z_0)$ и приложенный к ней *направляющий* ненулевой *вектор* $v=(a, b, c)\neq 0$, $(\overrightarrow{M_0M_1} = v)$ (рис. 5.2.1.). Тогда произвольная точка $M(x, y, z)$ будет лежать на заданной прямой только в том случае, если векторы $\overrightarrow{M_0M}$ и v коллинеарны, то есть когда существует такое число t, что

$$\overrightarrow{M_0M} = vt.$$

Если $\overrightarrow{OM_0} = r_0$ и $\overrightarrow{OM} = r$ - радиус-векторы точек M_0 и M (соответственно), то вектор $\overrightarrow{M_0M} = r - r_0$ можно записать в виде:

$$r - r_0 = tv,$$

где t –некоторое число (скаляр). Если действительная переменная t пробегает интервал $(-\infty,\ \infty)$, то конец вектора $r = r_0 + tv$ пробегает всю прямую L.

Полученное уравнение определяет прямую, а каждая точка прямой определяется значением переменной t. Эта переменная называется *параметром,* а уравнение называется *векторным уравнением прямой в параметрической форме.*

Его можно переписать в виде трех уравнений для каждой из координат точки на прямой, если a, b, c – координаты вектора v:

$$x = x_0 + ta,$$
$$y = y_0 + tb,$$
$$z = z_0 + tc.$$

Исключив параметр t, получим уравнение прямой в каноническом виде:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c},$$

где a, b, c одновременно не равны нулю.

Прямая, проходящая через две заданные и точки M_1 и M_0 с радиус-векторами r_1, r_0 имеет параметрическое уравнение вида

$$r = (r_1 - r_0)t + r_0 = (1 - t)r_0 + tr_1.$$

Если выразить координаты вектора $\overrightarrow{M_0M_1}$ через координаты точек, то получим:

$$a = x_1 - x_0, \quad b = y_1 - y_0, \quad c = z_1 - z_0.$$

И тогда прямая может быть представлена уравнением:

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0}=\frac{z-z_0}{z_1-z_0}.$$

Прямая, как линия пересечения двух плоскостей (рис.5.2.2.). Прямая, представляющая собой пересечение двух плоскостей задается системой двух уравнений, задающих эти плоскости. Плоскости пересекаются по прямой, если их нормали неколлинеарны. В векторной форме их уравнениями будут

$$n(r-r_0)=0,\ \ n_1(r-r_1)=0.$$

Направляющий вектор линии пересечения ортогонален нормалям пересекающихся плоскостей. В качестве его можно взять векторное произведение $n\times n_1$.

Поэтому, если r_0 – радиус-вектор какой-нибудь общей точки плоскостей $\alpha,\ \ \alpha_1$, то прямую их пересечения можно задать параметрическим уравнением

$$r=(n\times n_1)t+r_0.$$

Плоскость. Как известно из аналитической геометрии любая плоскость общего положения, проходящая *через три различные точки* $M_1(x_1,y_1,z_1),\ M_2(x_2,y_2,z_2),\ M_3(x_3,y_3,z_3)$, не лежащие на одной прямой, описывается уравнениями первой степени:

$$\left\{\begin{array}{c} a(x-x_1)+b(y-y_1)+c(z-z_1)=0 \\ a(x_2-x_1)+b(y_2-y_1)+c(z_2-z_1)=0 \\ a(x_3-x_1)+b(y_3-y_1)+c(z_3-z_1)=0 \end{array}\right\}.$$

Здесь *(x, y, z)* – произвольная точка лежит в одной плоскости с точками M_1, M_2, M_3 тогда и только тогда, когда определитель этой системы равен нулю:

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0.$$

Это же уравнение может быть представлено и в следующем виде:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix}=0.$$

Если задать плоскость α двумя неколлинеарными векторами *m* и *n,* приложенными к некоторой точке *A* с радиус-вектором r_0, лежащей в этой плоскости, и выразить произвольную точку плоскости *X* через ее радиус-вектор *r*, то можно получить *параметрическое* уравнение плоскости (рис.5.2.3.):

$$r=r_0+tm+sn.$$

Из этого уравнения положение любой точки *X* плоскости определяется заданием упорядоченной пары действительных чисел *(t, s),* причем каждой такой паре соответствует некоторая точка плоскости α. Например, точка *A* отвечает паре (0,0).

В *матричной форме* уравнение $ax+by+cz+d=0$ произвольной плоскости в пространстве имеет вид:

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0,$$

или

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} [P]^T = 0,$$

где $[P]^T$ представляет собой плоскость.

Поэтому любой выпуклый многогранник можно представить в виде матрицы, состоящей из коэффициентов уравнений плоскостей

$$[V] = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \\ c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{bmatrix},$$

где каждый столбец содержит коэффициенты плоскости (грани).

Другие фигуры, ограниченные криволинейными поверхностями, или имеющие невыпуклые формы, можно привести к этому случаю путем аппроксимации или разбиением на выпуклые части.

Поверхности второго порядка. *Поверхность второго порядка* определяется общим уравнением второй степени (с тремя неизвестными (x,y,z):

$$f(x,y,z) = Ax^2 + By^2 + Cz^2 + 2Dxy + 2Eyz + 2Fzx + 2Gx + 2Hy + 2Jz + K = 0,$$

где, по крайней мере, один из коэффициентов A, B, C, D, E, F не равен нулю.

Или в матричном виде:

$$f(x,y,z) = \begin{bmatrix} x & y & z & 1 \end{bmatrix} P \begin{bmatrix} x & y & z & 1 \end{bmatrix}^T,$$

где

$$P = \begin{bmatrix} A & D & F & G \\ D & B & E & H \\ F & E & C & J \\ G & H & J & K \end{bmatrix}.$$

В зависимости от значений коэффициентов этим уравнением могут быть описаны различные типы поверхностей второго порядка: *эллипсоиды, гиперболоиды, параболоиды,* и в их числе – поверхности вращения: *сфера, конус, цилиндр*. За исключением эллипсоида и сферы все поверхности второго порядка не замкнуты. Поэтому при образовании объемной модели эти поверхности следует ограничивать линиями или другими поверхностями.

Каждая поверхность второго порядка описывается своей системой уравнений. Приведем некоторые из них.

Эллипсоид, заданный каноническим уравнением:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

где a,b,c – полуоси эллипсоида, может быть также представлен параметрически в виде:

$$x = a \quad \cos\theta \sin\varphi, \quad 0 \le \theta \le 2\pi,$$
$$y = b \quad \sin\theta \quad \sin\varphi, \quad 0 \le \varphi \le 2\pi,$$
$$z = c \quad \cos\varphi,$$

где θ – долгота, φ – широта.

Нормаль к поверхности *эллипсоида* определяется по формуле:

$$N = i \quad bc \quad \cos\theta \quad \cos\varphi + j \quad ca \quad \sin\theta \quad \cos\varphi + k \quad ab \quad \sin\varphi,$$

где *i, j, k* – орты, направленные соответственно по осям *Ox, Oy, Oz*.

В случае, когда какие-нибудь две оси одинаковы, то эллипсоид будет являться поверхностью вращения. Например, при *a=b* осью вращения будет *Oz*. Если *a=b=c,* то эллипсоид обращается в *сферу*.

Если по оси поверхности вращения направить ось *Oz* прямоугольной системы координат *Oxyz,* то *параметрические уравнения поверхности вращения* можно записать в следующем виде:

$$x = f(u)\cos v, \quad y = f(u)\sin v, \quad z = u,$$

где *f(u)* – функция, определяющая форму меридиана, а *v* – угол поворота плоскости меридиана.

Общее уравнение второй степени

$$f(x, y, z) = Ax^2 + By^2 + Cz^2 + 2Dxy + 2Eyz + 2Fzx + 2Gx + 2Hy + 2Jz + K = 0$$

представляет *сферу* только при следующих условиях:

$$A=B=C, \qquad D=0, \qquad E=0, F=0,$$
$$G^2 + H^2 + J^2 - 4AK > 0.$$

При этих условиях

$$a = -\frac{G}{2A}, \quad b = -\frac{H}{2A}, \quad c = -\frac{J}{2A}, \quad R^2 = \frac{G^2 + H^2 + J^2 - 4AK}{4A^2}.$$

Сфера радиуса *R* и с центром в начале координат описывается уравнением второй степени:

$$x^2 + y^2 + z^2 = R^2.$$

Если центр сферы не совпадает с началом, то сфера может быть представлена следующим уравнением:

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = R^2,$$

где *a, b, c* – координаты центра сферы в некоторой точке *C(a, b, c).*

Круговой конус (рис.5.2.4.) с вершиной в начале координат определяется уравнением:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

В векторной форме уравнение кругового конуса имеет вид:

$$x \cdot a_0 = |x| \quad \cos\alpha, \quad |\mathrm{a}_0| = 1.$$

Уравнение *цилиндрической* (рис.5.2.5.) поверхности в векторной форме имеет вид:

$$|x|^2 - (x \cdot a_0)^2 = R^2, \quad |a_0| = 1,$$

где *R* – радиус цилиндра*; x* – радиус-вектор любой точки на боковой поверхности цилиндра; a_0 - единичный вектор в направлении оси вращения.

Поверхности вращения общего вида (рис.5.2.6.). Поверхность, которая получается путем вращения плоской кривой (образующей) вокруг оси, называется поверхностью вращения.

Поверхность вращения можно задать параметрическими уравнениями

$$x = f(u)\ \cos v,\quad y = f(u)\ \sin v,\quad z = u\,,$$

где *f(u)* – функция, определяющая форму образующей *C*, *v* – угол поворота.

Или в общем виде:

$$r(u,v) = r_0 + ua_0 + f(u)(\cos v\cdot e_1 + \sin v\cdot e_2)\,,$$

где a_0 – единичный вектор в направлении оси вращения, r_0 – радиус-вектор некоторой точки, лежащей на оси вращения.

Линейчатые поверхности (рис.5.2.7.). Поверхность, образуемая перемещением по определенному закону прямолинейной образующей по пространственной кривой (*направляющей*), называется линейчатой поверхностью.

Поверхность строится следующим образом. В каждой точке пространственной кривой *p(u)* (*направляющей*) задается непрерывная вектор-функция *g(u)* *(образующая*). Тогда линейчатая поверхность описывается уравнением

$$r(u,v)=p(u)+vg(u),$$

где *v* – параметр вдоль образующей. В частном случае, если *g(u)* – единичный вектор, то *v* приобретает смысл расстояния вдоль образующей.

5.3. Общие принципы конструирования пространственных кривых и поверхностей сложных форм

В практике проектирования кривых и поверхностей сложной формы приходится решать разнообразные задачи построения кривых и поверхностей с заданными свойствами, которые чаще всего определяются конкретной предметной областью. Например, в автомобильной промышленности при проектировании новых изделий разрабатываются сложные поверхности, которые должны удовлетворять, в частности, критерию непрерывности или определенным требованиям дизайнера. Кроме этого, есть целый ряд особенно важных требований, выполнение которых необходимо предусмотреть в процессе генерирования кривых и поверхностей. Это следующие требования:

-глобальные или локальные формы кривых или поверхностей должны контролироваться с помощью контрольных точек (это – точки, через которые *проходит кривая* или те точки, которые *влияют на форму* кривой глобально или локально);

-кривая или поверхность должны быть *гладкими и непрерывными,* то есть в точках контакта кривые или поверхности должны касаться друг друга (непрерывность первой производной) или иметь одинаковую кривизну (непрерывность второй производной).

Для выполнения перечисленных требований с целью получения кривой или поверхности с заданными характеристиками используется интерактивное конструирование, которое носит итерационный характер. Модель, полученную на некотором шаге итерации, модифицируют и улучшают до тех пор, пока не будет достигнута желаемая форма кривой или поверхности. Для этих целей на разных шагах итерационного процесса применяются различные *методы аппроксимации, интерполяции и сглаживания.*

В зависимости от конкретных исходных данных задачи геометрического моделирования кривых и поверхностей по их содержанию можно отнести к одному из следующих видов:

-*приближенное описание* этих геометрических объектов и

-*восстановление* их по некоторым данным.

5.3.1. Приближенное описание кривых и поверхностей

В этих задачах используются данные, полученные в результате расчетов или эксперимента. Например, множество точек, принадлежащих кривой или поверхности, может быть определено путем оцифровки существующей физической модели или чертежа проектируемого объекта (например, в авиа-, судостроении и др.). В этих случаях *форма кривой или поверхности приблизительно известна,* и задача состоит в *подборе* простого *математического выражения* для этих форм с выполнением всех требуемых ограничений. В зависимости от конкретных требований для решения подобных задач применяются различные математические методы.

Ниже приводятся описания *классических решений некоторых задач* из числа перечисленных выше.

Линейная интерполяция. Простейшим видом интерполяции является *линейная* интерполяция, в основе которой лежит аппроксимация кривой на участке между точками (x_k, y_k) и (x_{k+1}, y_{k+1}) прямой, проходящей через те же точки (рис.5.3.1.).

Уравнение прямой можно представить в виде:

$$\frac{y - y_k}{x - x_k} = \frac{y_{k+1} - y_k}{x_{k+1} - x_k}$$

или в виде:

$$y = \frac{y_k(x - x_{k+1}) - y_{k+1}(x - x_k)}{x_k - x_{k+1}}.$$

Таким образом, можно найти значение функции *y* при любом значении *x* в интервале $[x_k, x_{k+1}]$, если известны значения табличных данных y_k и y_{k+1}, соответствующих x_k и x_{k+1}.

Полиномиальная интерполяция. При использовании большего числа соседних точек и аппроксимировании истинной кривой более *сложной линией* полученный результат можно уточнить *методами полиномиальной интерполяции*.

Задача. Пусть на сегменте *[a,b]* заданы *n+1* узловых точек (узлов) $a \le x_0 < x_1 < x_2 \ldots < x_n \le b$, а также значения y_j, *(j=0,1,...,n)* в узлах такие, что *f(x)* = y_j. Найти многочлен $P_n(x)$ степени *n* такой, что $P_n(x_j) = y_j$ для $0 \le j \le n$.

Существует всегда только *один интерполяционный многочлен*, который может быть определен в различных формах – в форме Лагранжа, в форме Ньютона и т. д.

Форма Лагранжа. При этой интерполяции задается *n+1* табличное значение (x_i, y_i), где *i=0,1,...,n*. Считается, что точки (x_i, y_i) принадлежат кривой *y=f(x)* в интервале $x_0 \le x \le x_n$. *Интерполяционный многочлен* для этого метода имеет вид:

$$P_n(x) = y_0 b_0(x) + y_1 b_1(x) + \ldots + y_n b_n(x),$$

где все $b_j(x)$ – многочлены степени *n*, коэффициенты которых можно найти с помощью *n+1* уравнения

$$P_n(x_i) = y_i,$$

где *i=0,1,...,n.*

В результате получим систему уравнений

$$y_0 b_0(x_0) + y_1 b_1(x_0) + ... + y_n b_n(x_0) = y_0, \quad y_0 b_0(x_n) + y_1 b_1(x_n) + ... + y_n b_n(x_n) = y_n$$

……………………………………………………

Если значения $b_j(x_i)$ выбраны так, что

$$b_j(x) = \begin{cases} 1 \quad при \quad i = j, \\ 0 \quad при \quad i \neq j \end{cases},$$

то уравнения, указанные выше, будут удовлетворены. Это означает, что любой многочлен $b_j(x)$ равен нулю при каждом x_i, кроме x_j.

Следовательно, в общем случае многочлен $b_j(x)$ имеет вид

$$b_j(x) = C_j(x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n).$$

Так как $b_j(x_j) = 1$, то коэффициент C_j определяется выражением

$$C_j = 1/(x_j - x_0)(x_j - x_1)...(x_j - x_{j-1})(x_j - x_{j+1})...(x_j - x_n).$$

Наконец, для искомого многочлена получаем

$$P_n(x) = \sum_{j=0}^{n} y_j \frac{(x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n)}{(x_j - x_0)(x_j - x_1)...(x_j - x_{j-1})(x_j - x_{j+1})...(x_j - x_n)}.$$

Введя обозначения

$$L_j(x) = (x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n),$$

можно записать полученный многочлен в следующем виде:

$$P_n(x) = \sum_{j=0}^{n} y_j \frac{L_j(x)}{L_j(x_j)}.$$

Форма Ньютона. Интерполяционный многочлен имеет вид

$$P_n(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + ... + c_n(x - x_0)(x - x_1)...(x - x_{n-1}).$$

Коэффициенты c_j находятся из уравнений

$$P_n(x) = y_i, \quad i = 0,1,...,n,$$

позволяющих записать систему

$$c = y_{0,}$$

$$c_0 + c_1(x_1 - x_0) = y_1,$$

$$c_0 + c_1(x_2 - x_0) + c_2(x_2 - x_0)(x_2 - x_1) = y_2,$$

……………………………………………………………

$$c_0 + ... + c_n(x_n - x_0)(x_n - x_1)...(x_n - x_{n-1}) = y_n.$$

Если значения *x* заданы через *равные промежутки*

$$x_{i+1} - x_i = h,$$

то в общем случае

$$x_i = x_0 + ih, \ где \ i=1,2,...,n.$$

Тогда решаемые уравнения можно привести к виду

$$y_0 = c_0,$$

$$y_1 = c_0 + c_1 h,$$

$$y_2 = c_0 + c_1(2h) + 2h^2 c_2,$$

……………………………………………………………

$$y_i = c_0 + c_1 ih + c_2 ih[(i-1)h + ... + c_i(i!)h^i,$$

откуда для коэффициентов получаем

$$c = y_{0,}$$

$$c_1 = \frac{y_1 - c_0}{h} = \frac{y_1 - y_0}{h} = \frac{\Delta y_0}{h}$$

$$c_2 = \frac{1}{2h^2}(y_2 - c_0 - 2hc_1) = \frac{1}{2h^2}[(y_2 - y_1) - (y_1 - y_0)] =$$

$$= \frac{1}{2h^2}[\Delta(\Delta y)] = \frac{\Delta^2 y_0}{2h^2},$$

где $\Delta^2 y_0$ – *разность разностей*.

Коэффициент c_j можно представить в виде

$$c_j = \frac{\Delta^j y_0}{(j!)h^j}.$$

Замечание. Полиномиальную интерполяцию имеет смысл применять лишь для небольшого числа точек (не более пятнадцати) из-за того, что вместе с числом точек растет степень полинома и имеют место большие осцилляции в промежутках между точками. Если количество точек небольшое (не больше пятнадцати), то используется *интерполяционный полином Лагранжа*.

Из двух рассмотренных выше форм чаще отдается предпочтение *полиномам Ньютона* с нахождением коэффициентов с помощью *разделенных разностей*. На рис.5.3.2. показана интерполяция с помощью полиномов Ньютона на 3, 6 и 7 точках. Добавление одной точки приводит к появлению значительных осцилляций.

В тех случаях, когда задана геометрическая информация о концевых точках, а также первые производные в этих точках, применяются *интерполяционные полиномы Эрмита.*

Интерполяция на больших отрезках, т.е. с относительно большим количеством узловых точек, имеет дополнительные трудности. С одной стороны, при больших расстояниях между узловыми точками точность очень мала, а с другой стороны, интерполяционные многочлены высокого порядка на концах отрезка значительно колеблются, что существенно искажает поведение функции. Поэтому при задании большего числа точек применяют *кусочно-полиномиальную интерполяцию*. Идея состоит в том, чтобы строить независимо друг от друга полиномы на каждом интервале $[t_i,\ t_{i+1}]$ отрезка $[a,\ b]$ с последующей *сшивкой* их между собой. Это возможно осуществить только при выполнении требования о непрерывности производных в точках сшивки t_i.

Для вычисления интерполяционного полинома на большом числе точек следует использовать *интерполяционные сплайны*.

Сплайн – это гибкая линейка, которую деформируют так, чтобы по ней можно было провести кривую через заданные точки $(x_i,\ y_i)$. Математически сплайн представляет собой кусочно-полиномиальную функцию $S(x)$ степени m с узлами $x_0, x_1, \cdots, x_n$, определенную на всей вещественной прямой, заданной последовательностью вещественных чисел $x_0 < x_1 < \cdots < x_n$ и обладающую следующими свойствами:

1) На каждом интервале (x_i, x_{i+1}) для $i=0,1,\ldots, n$ функция $S(x)$ задается некоторым многочленом степени m или ниже (здесь $x_0 = -\infty$, а $x_{n+1} = +\infty$).

2) Сплайн-функция $S(x)$ и ее производные порядков 1,2,…, m всюду непрерывны.

Для $m=0$ условие 2 не выполняется, так как сплайн степени 0 – это ступенчатая функция. Сплайн степени 1 – это ломаная, которая проходит через заданные точки, и между соседними узлами реализуется линейная интерполяция. Сплайн степени m для m >0 можно определить также как функцию из C^{m-1}, построенную в результате m-кратного неопределенного интегрирования ступенчатой функции.

Используя теорию изгиба бруса при малых деформациях, можно показать, что сплайн – это группа *сопряженных кубических полиномов,* в местах сопряжения которых *первая и вторая производные непрерывны*. Такие функции называются *кубическими сплайнами.*

Чтобы построить *кубический сплайн,* необходимо задать *коэффициенты,* которые единственным образом определяют кубический полином в промежутке *между заданными точками*. Например, в случае рис. 5.3.3. необходимо задать все кубические функции $q_1(x), q_2(x), \ldots, q_m(x)$.

Таким образом, форма кубического сплайна задается *кубическим полиномом*, который в наиболее общем случае имеет вид:

$$q_i(x) = \kappa_{1i}x^3 + \kappa_{2i}x^2 + \kappa_{3i}x + \kappa_{4i}, \quad i=1,\ 2,, \ldots,\ m.$$

Здесь κ_{ji} – коэффициенты, которые определяются следующими условиями:

1) первые 2m условий требуют, чтобы сплайны соприкасались в заданных точках

$$q_i(x_i) = y_i,\ i=1,\ldots,\ m$$

$$q_{i+1}(x_i) = y_i,\ i=0,\ldots,\ m\text{-}1;$$

2) следующие 2m-2 условия требуют, чтобы в местах соприкосновения сплайнов были равны первые и вторые производные

$$q'_{i+1}(x_i) = q'_i(x_i), \quad i=1,\ldots,\ m\text{-}1,$$

$$q''_{i+1}(x_i) = q''_i(x_i), \quad i=0,\ldots,\ m\text{-}1.$$

Чтобы система алгебраических уравнений имела решение, необходимо, чтобы число уравнений точно равнялось числу неизвестных. На данном этапе мы имеем *4m* неизвестных и *4m-2* уравнения. Следовательно, необходимо найти еще два уравнения. В таких случаях обычно используются следующие уравнения:

$$q''_1(x_0) = 0 \ \text{ и } \ q''_m(x_m) = 0.$$

Определив коэффициенты сплайна, можно использовать эту кусочно-гладкую полиномиальную функцию для представления данных при интерполяции, подгонке кривых или поверхностей.

Кубические сплайны обладают следующими достоинствами:

-вследствие простого задания кривой, удобны в использовании, так как для построения кривой необходимы только значения сплайн-функции в узлах (опорные точки) и значения первых производных в концевых точках;

-на каждом интервале кривая определяется кубическим полиномом;

-так как кривая на всем интервале является дважды непрерывно дифференцируемой функцией, то у кривой нет точек перегиба.

К недостаткам можно отнести необходимость задания *(n-1)* значений производных.

Следует отметить, что существуют и другие сплайны, получающиеся при других граничных условиях или использовании полиномов более высоких степеней. Для выбора других методов решения задач рекомендуется обратиться к специальной литературе.

5.3.2. Восстановление кривых и поверхностей

В этих задачах требуется *восстановить* кривую или поверхность по некоторым данным о них, причем, контрольные точки определяют только направление изгиба в то время как *сведения* об *изначальной форме* кривой или поверхности *отсутствуют*. В этих случаях математическое описание кривых генерируется путем итераций в интерактивном режиме с использованием методов *Бернштейна, Безье и В-сплайнов*. Эти методы позволяют создавать кривые и поверхности любой формы.

Метод аппроксимации Бернштейна. Этот метод составляет *основу метода Безье*. Он состоит в следующем. Пусть $f(t)$ – произвольная непрерывная на интервале *[0,1]* вещественная функция. Аппроксимирующий многочлен Бернштейна степени *n* для функции $f(t)$ определяется выражением:

$$B_n(f(t)) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\varphi_{k,n}(t),$$

где

$$\varphi_{k,n}(t) = \frac{n!}{k!(n-k)!}t^k(1-t)^{n-k}, \quad k = 0,1,\dots,n.$$

Функции $\varphi_{k,n}(t)$ образуют *базис* для множества всех полиномов степени, меньшей или равной *n*: любой полином степени *n* можно единственным образом представить линейной комбинацией этих функций.

Многочлены Бернштейна редко используются для решения задач аппроксимации из-за того, что они очень медленно сходятся в равномерной норме. Но эти многочлены, как показал Безье, удобно использовать в задачах интерактивного проектирования гладких свободных кривых и поверхностей.

Кривая Безье. Пусть в пространстве или на плоскости дано упорядоченное множество *n+1* точек *P,* в результате соединения которых получается *n*-звенная незамкнутая ломаная, называемая *характеристической ломаной* или характеристическим *многоугольником.* Вершины ломаной определяются радиус-векторами $r_0,\ r_1,\dots,r_n$. Кривой Безье называется кривая, которая аппроксимирует эту характеристическую ломаную и определяется относительно ее вершин следующим параметрическим векторным уравнением:

$$r(u) = \sum_{i=0}^{n} B_{i,n}(u)r_i, \quad 0 \le u \le 1,$$

где

$$B_{i,n}(u) = C_n^i u^i(1-u)^{n-i}, \quad C_n^i = \frac{n!}{(n-1)!i!}$$

– базис Безье или Бернштейна, или функция аппроксимации.

Кривая Безье обладает следующими свойствами:

-кривая является гладкой;

-*степень многочлена,* определяющего сегмент (участок) кривой, на единицу меньше количества точек соответствующего характеристического многоугольника;

-*граничные точки кривой* совпадают с *крайними вершинами* характеристической ломаной;

-первое и последнее звенья ломаной являются *касательными* к кривой соответственно в начальной и конечной точке кривой;

-основа формы кривой повторяет очертания многоугольника;

-изменение всей формы кривой связано с изменением положения вершин характеристической ломаной;

-кривая целиком содержится в выпуклой оболочке точек *P*.

Из перечисленных свойств вытекают следующие особенности применения функций Безье:

-изменение положения хотя бы одной вершины приводит к заметному изменению кривой;

-добавление хотя бы еще одной вершины приводит к необходимости решения задачи заново;

-значительное увеличение количества вершин приводит к увеличению степени полинома, что значительно затрудняет вычисление функций Безье.

Пример (рис.5.3.4.). Заданы вершины многоугольника Безье $P_0[1, 1]$, $P_1[2, 3]$, $P_2[4, 3]$, $P_3[3, 1]$. Найти семь точек, принадлежащих кривой Безье для $0 \le u \le 1$.

Для данного случая степень многочлена равна $n=3$.

Из уравнения кривой Безье

$$r(u) = \sum_{i=0}^{n} B_{i,n}(u) r_i,$$

где

$$B_{i,n}(u) = C_n^i u^i (1-u)^{n-i}$$

и

$$C_n^i = \frac{n!}{(n-i)!i!}$$

определяем

$$C_n^i = \frac{3!}{(3-i)!i!} = \frac{6}{(3-i)!i!},$$

а затем вычисляем коэффициенты $B_{i,n}(u)$:

$B_{0,3}(u) = (1)u^0(1-u)^3 = (1-u)^3$ (при $i=0$ $i!=0!=1$, $u^0=1$)

$B_{1,3}(u) = 3u(1-u)^2$,

$B_{2,3}(u) = 3u^2(1-u)$,

$B_{3,3}(u) = u^3$.

Значения коэффициентов для различных параметров *u* сведены в таблицу:

u	$B_{0,3}$	$B_{1,3}$	$B_{2,3}$	$B_{3,3}$
0	1	0	0	0
0.15	0.614	0.325	0.058	0.003
0.35	0.275	0.444	0.239	0.042

0.5	0.125	0.375	0.375	0.125
0.65	0.042	0.239	0.444	0.275
0.85	0.003	0.058	0.325	0.614
1	0	0	0	1

Подставляя найденные значения коэффициентов в уравнение кривой Безье

$$p(u) = P_0 B_{0,3} + P_1 B_{1,3} + P_2 B_{2,3} + P_3 B_{3,3} = P_0 (1-u)^3 + P_1 3u(1-u)^2 + P_2 3u^2 (1-u) + P_3 u^2$$

получаем координаты *x, y* искомых точек кривой:

$p(0) = P_0 = [\text{[illegible]}\;\; 1]$;

$p(0.15) = 0.614 P_0 + 0.325 P_1 + 0.058 P_2 + 0.003 P_3 = [\text{[illegible]}.5\;\; 1.765]$;

$p(0.35) = [\text{[illegible]}.248\;\; 2.367]$;

$p(0.5) = [\text{[illegible]}.75\;\; 2.5]$;

$p(0.65) = [\text{[illegible]}.122\;\; 2.367]$;

$p(0.85) = [\text{[illegible]}.248\;\; 1.765]$;

$p(1) = P_3 = [\text{[illegible]}\;\; 1]$.

Кривая В-сплайна. Под сплайнами понимаются гибкие деревянные или металлические линейки, которые, изгибаясь, могут проходить через две заданные точки. В настоящее время разработаны математические методы описания таких кривых, и они применяются в машинной графике для моделирования криволинейных форм.

Кривая В-сплайна определяется полиномом:

$$S(u) = \sum_{i=0}^{n} N_{i,k}(u) r_i,$$

где

$$N_{i,1}(u) = \begin{cases} 1, & при \;\; a_i \le u \le a_{i+1} \\ 0 & в \;\; других \;\; случаях \end{cases},$$

$$N_{i,k}(u) = \frac{(u - a_i) N_{i,k-1}(u)}{a_{i+k-1} - a_i} - \frac{(a_{i+k} - u) N_{i+1,k-1}(u)}{a_{i+k} - a_{i+1}}.$$

С помощью задания a_i можно выбирать соотношение между u и контрольными точками.

Кривая, построенная с помощью В-сплайнов, обладает следующими свойствами:

-кривая определяется линейной комбинацией сплайн-функций, коэффициентами которой являются координаты опорных точек $P_i = (P_x, P_y, P_z)$,

-кривая является кусочно-полиномиальной;

-локальное изменение опорных точек не приводит к изменению всей формы кривой;

-кривая находится внутри выпуклой оболочки;

-кривая может содержать прямолинейные отрезки;

-возможно задание координат вершин характеристической ломаной;

-возможно распознание нежелательных экстремумов и петель по виду характеристической ломаной.

В общем случае кривая не проходит точно через заданные точки. Они используются только для формирования гладкой кривой. Аппроксимация будет тем лучше, чем ближе расположены соседние точки. Если полученная кривая не удовлетворительна, то ее можно улучшить, задавая большее количество точек.

Пример. Из заданных точек m первая и последняя не аппроксимируются, хотя и оказывают влияние на формирование кривой; точки последовательно соединяются и образуют $m-1$ отрезков или интервалов, которым соответствуют $m-3$ сегментов сглаживающей кривой. Поэтому для В-сплайна необходимо задавать не менее четырех точек.

Коэффициенты В-сплайна для каждого сегмента кривой вычисляются по формулам:

$$a_3 = (-x_{i-1} + 3x_i - 3x_{i+1} + x_{i+2})/6$$
$$a_2 = (x_{i-1} - 2x_i + x_{i+1})/2$$
$$a_1 = (-x_{i-1} + x_{i+1})/2$$
$$a_0 = (x_{i-1} + 4x_i + x_{i+1})/6$$
$$b_3 = (-y_{i-1} + 3y_i - 3y_{i+1} + y_{i+2})/6$$
$$b_2 = (y_{i-1} - 2y_i + y_{i+1})/2$$
$$b_1 = (-y_{i-1} + y_{i+1})/2$$
$$b_0 = (y_{i-1} + 4y_i + y_{i+1})/6$$
$$c_3 = (-z_{i-1} + 3z_i - 3z_{i+1} + z_{i+2})/6$$
$$c_2 = (z_{i-1} - 2z_i + z_{i+1})/2$$
$$c_1 = (-z_{i-1} + z_{i+1})/2$$
$$c_0 = (z_{i-1} + 4z_i + z_{i+1})/6.$$

Каждый сегмент кривой, лежащий между точками i и $i+1$, описывается полиномами третьей степени в параметрической форме

$$x = a_3u^3 + a_2u^2 + a_1u + a_0,$$
$$y = b_3u^3 + b_2u^2 + b_1u + b_0,$$
$$z = z_3u^3 + z_2u^2 + z_1u + z_0.$$

Зная коэффициенты полиномов, можно вычислить координаты точек кривой, задавая различные значения параметра в пределах $0<u<1$. Причем, при $u=0$ конечная точка кривой будет вблизи точки i, а при $u=1$ производится аппроксимация точки $i+1$.

Как и в предыдущем примере с кривой Безье, результаты можно свести в таблицу.

Основное и существенное *различие между кривыми Безье и В-сплайнами* (рис.5.3.5.) состоит в следующем:

-общая форма кривой В-сплайна повторяет форму определяющей ее характеристической ломаной,

-кривая В-сплайна изменяется только вблизи изменяемой вершины (в связи с локальным характером аппроксимации на одном сегменте), в то время как изменение только одной контрольной точки кривой Безье требует полного пересчета всей кривой.

5.3.3. Бикубические поверхности

Сложная криволинейная поверхность произвольной формы обычно составляется из отдельных кусков (участков). Участок такой поверхности строится в виде гладкого изогнутого четырехугольника по заданным координатам четырех угловых точек и частным производным в этих точках.

Границы кусков могут быть представлены в виде произвольных кривых. Эти куски затем между собой сшиваются так, чтобы в местах сшивки сохранялась гладкость создаваемой поверхности. Этого можно достигнуть за счет совпадения в местах сшивки координат угловых точек, а также совпадения первых производных. Участок такой поверхности $x = x(u,v), \quad y = y(u,v), \quad z = z(u,v)$ описывается, как *бикубическая поверхность* и может быть представлен в виде кубических полиномов в параметрической форме. Например, для $x = x(u,v)$:

$$\begin{aligned} x(u,v) = {} & a_{11}u^3v^3 + a_{12}u^3v^2 + a_{13}u^3v + a_{14}u^3 + \\ & + a_{21}u^2v^3 + a_{22}u^2v^2 + a_{23}u^2v + a_{24}u^2 + \\ & + a_{31}uv^3 + a_{32}uv^2 + a_{33}uv + a_{34}u + \\ & + a_{41}v^3 + a_{42}v^2 + a_{43}v + a_{44} \ , \end{aligned}$$

где u,v – параметры, изменяющиеся в заданном диапазоне $v \in [0,1]$, $u \in [0,1]$; $a_{11},...,a_{44}$- постоянные коэффициенты в пределах данной поверхности. Коэффициенты могут быть записаны в виде матрицы:

$$C_x = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Аналогичные выражения существуют для

$$y = y(u,v), \quad z = z(u,v) .$$

Если обозначить через C_y, C_z матрицы коэффициентов при переменных в этих выражениях и ввести векторы-строки

$$U = [u^3 \quad u^2 \quad u \quad 1], \; V = [v^3 \quad v^2 \quad v \quad 1],$$

то *уравнение для бикубического участка* поверхности будет иметь следующий вид:

$$x(u,v) = UC_xV^T, \quad y(u,v) = UC_yV^T, \quad z(u,v) = UCzV^T .$$

В зависимости от типа имеющихся исходных данных о геометрии моделируемой поверхности участок бикубической поверхности может быть представлен в виде *форм Фергюсона, Эрмита, Кунса, Безье, В-сплайнов.*

Бикубическая поверхность в форме Безье (рис.5.3.6.). Геометрически поверхность Безье определяется множеством точек, принадлежащих образующей кривой в процессе ее перемещения и деформации. Можно считать, что каждая вершина характеристического многоугольника *образующей кривой* описывает траекторию, которая сама определяется характеристическим многоугольником. Если все образующие имеют одинаковую степень, т.е. их характеристические многоугольники имеют одинаковое число сторон, то *характеристическая сетка поверхности* получается соединением вершин с одинаковыми индексами.

Таким образом, вершины характеристической сетки образуют матрицу размером $(m+1)\xi(n+1)$, в которой все элементы с одинаковым первым индексом образуют *строку.* Через точки поверхности, соответствующие таким элементам, проводится характеристическая ломаная для каждой строки матрицы. В результате образуются *(m+1)* характеристических ломаных и соответствующие им *(m+1)* кривые Безье.

При степени (n=4) образуется характеристическая сетка с общим количеством вершин равным 16, из которых только угловые вершины (точки) принадлежат самой поверхности. Эти 16 точек называют также *управляющими точками*, так как их положение влияет на геометрию куска бикубической поверхности.

Поверхность в форме Безье описывается так же, как и кривая Безье:

$$P(u,v)=\sum_{i=0}^{n}\ \sum_{j=0}^{m}P_{i,j}B_{i,n}(u)B_{j,m}(v).$$

В матричной форме участок поверхности Безье определяется уравнением:

$$P(u.v)=(1\quad u\quad u^2\quad u^3)MBM^T\begin{pmatrix}1\\ v\\ v_2\\ v^3\end{pmatrix},$$

где
$$M=\begin{pmatrix}1&0&0&0\\ -3&3&0&0\\ 3&-6&3&0\\ -1&3&-3&1\end{pmatrix},\quad B=\begin{pmatrix}p_{00}&p_{01}&p_{02}&p_{03}\\ p_{10}&p_{11}&p_{12}&p_{13}\\ p_{20}&p_{21}&p_{22}&p_{23}\\ p_{30}&p_{31}&p_{32}&p_{33}\end{pmatrix}.$$

Здесь B – матрицы 4ξ4 координат x, y, z для 16 управляющих точек, из которых p_{00}, p_{03}, p_{30}, p_{33} - координаты (x,y,z) четырех угловых точек, а остальные точки p_{01}, p_{02}, p_{10}, p_{11}, p_{12}, p_{13}, p_{20}, p_{21}, p_{22}, p_{23}, p_{31}, p_{32} являются концами касательных векторов. Значения вектора p и его производных вычисляется в углах участка поверхности следующим образом:

$$\begin{pmatrix}p(0,0)&p(0,1)&p_v(0,0)&p_v(0,1)\\ p(1,0)&p(1,1)&p_v(1,0)&p_v(1,1)\\ p_u(0,0)&p_u(0,1)&p_{uv}(0,0)&p_{uv}(0,1)\\ p_u(1,0)&p_u(1,1)&p_{uv}(1,0)&p_{uv}(1,1)\end{pmatrix}=$$

$$\begin{pmatrix}1&0&0&0\\ 0&0&0&1\\ -3&3&0&0\\ 0&0&-3&3\end{pmatrix}\begin{pmatrix}p_{00}&p_{01}&p_{02}&p_{03}\\ p_{10}&p_{11}&p_{12}&p_{13}\\ p_{20}&p_{21}&p_{22}&p_{23}\\ p_{30}&p_{31}&p_{32}&p_{33}\end{pmatrix}\begin{pmatrix}1&0&-3&0\\ 0&0&3&0\\ 0&0&0&-3\\ 0&1&0&3\end{pmatrix}.$$

Значения вектора p и его производных в углу с параметрами u=v=0 определяются из соотношений:

$$p(0,0)=p_{00},$$
$$p_u(0,0)=3(p_{i0}-p_{00}),$$
$$p_v(0,0)=3(p_{01}-p_{00}),$$
$$p_{uv}(0,0)=9(p_{00}-p_{01}-p_{10}+p_{11}).$$

Поверхности типа B-сплайна. Определение поверхности *методом B-сплайна* аналогично определению поверхности методом Безье для случая, когда степень В-сплайн-функции соответствует числу опорных точек. В матричной форме участок поверхности типа В-сплайна определяется уравнением:

$$P(u.v) = (1 \quad u \quad u^2 \quad u^3) M_S B M_S{}^T \begin{pmatrix} 1 \\ v \\ v_2 \\ v^3 \end{pmatrix},$$

где

$$M_S = (1/6) \begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{pmatrix}$$

и B – матрицы 4ξ4 координат *x, y, z* для 16 управляющих точек.

Поверхность В-сплайна, связанная с сеткой точек в трехмерном пространстве, может быть также определена следующим образом:

$$P(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} N_{i,n}(u) N_{j,m}(v) P_{ij},$$

где P_{ij} - вершины характеристической сетки.

Пример. Рассмотрим поверхности, точки которых имеют координаты $x(u,v), y(u,v), z(u,v)$ где u и v - два независимых параметра. Будем использовать аппроксимацию с помощью В-сплайна. Зададим набор из nm точек, которые образуют таблицу следующего вида:

$$\begin{matrix} P_{11} & P_{12} & . & . & .P_{1m} \\ P_{21} & P_{22} & . & . & .P_{2m} \\ . & . & . & .. & . \\ . & . & . & . & . \\ . & . & . & . & . \\ P_{n1} & P_{n2} & . & . & .P_{nm} \end{matrix}.$$

То есть каждая точка P_{ij} в таблице имеет номер (индекс) $k = (i-1)m + j$.

Таблицу точек P_{ij} можно также интерпретировать как два набора кривых линий: для каждой горизонтальной кривой (строка) параметр i является константой, а параметр j изменяется от 1 до m, а для каждой вертикальной кривой (столбец) параметр j – константа, а параметр i изменяется от 1 до n. Как известно, в процессе аппроксимации В-сплайном самая первая и самая последняя точки, хотя и участвуют в процессе вычислений, но не аппроксимируются. То же самое относится и к точкам P_{ij} таблицы, для которых $i=1$ или $i=n$, а также для тех точек, для которых $j=1$ или $j=m$. Для вычисления точек сетки на отсеке (четырехугольнике) поверхности потребуются по четыре точки в каждом направлении, что вместе дает следующие 16 точек:

$$\begin{matrix} P_{i-1,j-1} & P_{i-1,j} & P_{i-1,j+1} & P_{i-1,j+2} \\ P_{i,j-1} & P_{i,j} & P_{i,j+1} & P_{i,j+2} \\ P_{i+,j-1} & P_{i+1,j} & P_{i+1,j+1} & P_{i+1,j+2} \\ P_{i+2,j-1} & P_{i+2,j} & P_{i+2,j+1} & P_{i+2,j+2} \end{matrix}.$$

Все эти точки будут использованы для аппроксимации четырехугольника в середине (с точкой P_{ij} в верхнем левом углу и с точкой $P_{i+1,j+1}$ в нижнем правом углу). Этот четырехугольник в середине будет поделен на *NxM* малых элементов поверхности. То есть можно будет пересчитать *N* элементов по вертикальному и *M* элементов по горизонтальному направлениям. Значения *N* и *M* – положительные целые числа.

Приведем формулу для координаты *x* (для *y* и *z* формулы определяются аналогично):

$$
\begin{gathered}
x = F_{-1}(u)F_{-1}(v)x_{i-1,j-1} \quad + \quad F_{-1}(u)F_0(v)x_{i-1,j} \\
F_{-1}(u)F_1(v)x_{i-1,j+1} \quad + \quad F_{-1}(u)F_2(v)x_{i-1,j+2} \\
F_0(u)F_{-1}(v)x_{i,j-1} \quad + \quad F_0(u)F_0(v)x_{i,j} \\
F_0(u)F_1(v)x_{i,j+1} \quad + \quad F_0(u)F_2(v)x_{i,j+2} \\
F_1(u)F_{-1}(v)x_{i+1,j-1} \quad + \quad F_1(u)F_0(v)x_{i+1,j} \\
F_1(u)F_1(v)x_{i+1,j+1} \quad + \quad F_1(u)F_2(v)x_{i+1,j+2} \\
F_2(u)F_{-1}(v)x_{i+2,j-1} \quad + \quad F_2(u)F_0(v)x_{i+2,j} \\
F_2(u)F_1(v)x_{i+2,j+1} \quad + \quad F_2(u)F_2(v)x_{i+2,j+2}\,.
\end{gathered}
$$

Метод Кунса. Принцип метода. Решение, предложенное Кунсом является универсальным и вкратце состоит в следующем.

Сложная поверхность задается набором плоских сечений, ориентированных в двух направлениях, которые образуют сетку кривых, определяющих проектируемую поверхность.

Каркас сечений, составленный из кусков таких параметрических кривых, позволяет определить проектируемую поверхность по методу Кунса (рис. 5.3.7.) когда *координаты точек каждого куска поверхности,* ограниченного парой плоскостей поперечных сечений U_i и U_{i+1} и парой плоскостей продольных сечений V_j и V_{j+1} определяются по формуле:

$$
\begin{aligned}
P(u,v) &= P(u_i,v)\cdot F_0(u) + P(u_{i+1},v)\cdot F_1(u) + P(u,v_j)\cdot F_0(v) + P(u,v_{j+1})\cdot F_1(v) - \\
&- P(u_i,v_i)\cdot F_0(u)\cdot F_0(v) - P(u_i,v_{j+1})\cdot F_0(u)\cdot F_1(v) - P(u_{i+1},v_j)\cdot F_1(u)\cdot F_0(v) - \\
&- P(u_{i+1},v_{j+1})\cdot F_1(u)\cdot F_1(v),
\end{aligned}
$$

где:

-$P(u,v)$ -координата *x, y* или *z* точки поверхности, соответствующей параметрам *u* и *v*;

-$P(u_i,v)$, $P(u_{i+1},v)$, $P(u,v_j)$, $P(u,v_{j+1})$ - координаты *x, y* или *z* точек кривых в плоскостях U_i, U_{i+1} и V_j, V_{j+1};

-$P(u_i,v_j)$, $P(u_{i+1},v_j)$, $P(u_i,v_{j+1})$, $P(u_{i+1},v_{j+1})$- координаты точек в углах рассматриваемого куска поверхности;

-F_0, F_1 - функции сшивки.

Для выполнения сшивки с сохранением непрерывности кривизны в углах сетки, а следовательно, и во всех точках сшивки применяются функции пятой степени:

$$F_0(t) = 6\cdot t^5 - 15\cdot t^4 + 10\cdot t^3;$$
$$F_1(t) = 1 - F_0(t),$$

где

$$t = \frac{u - u_i}{u_{i+1} - u_i}$$

или

$$t = \frac{v - v_j}{v_{j+1} - v_j}.$$

5.3.4. Касательная, нормаль, кривизна поверхности.

К основным характеристикам поверхности относятся: кривизна, нормаль и касательная.

Уравнение касательной плоскости. Касательная плоскость к поверхности S в точке M есть плоскость, проходящая через точку M и характеризующаяся тем свойством, что расстояние от этой плоскости до переменной точки M' поверхности S при стремлении M' к M является бесконечно малым по сравнению с расстоянием MM'. Если поверхность представлена уравнением вида $z = f(x, y)$, то уравнение касательной плоскости к поверхности будет иметь следующий вид:

$$Z - z = p(X - x) + q(Y - y),$$

где (X, *Y, Z)* – текущие координаты точки, (x, y, z) - координаты точки касания, а p и q – соответствующие значения частных производных $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

Если поверхность представляется уравнением вида *F(x,y,z)=0,* то касательная плоскость будет выражена следующим уравнением:

$$F'_x(X - x) + F'_y(Y - y) + F'_z(Z - z) = 0.$$

Пример. Эллипсоид задан уравнением: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$.

Тогда уравнение касательной плоскости в точке *M*(*x, y, z)* при $F'_x = \frac{2x}{a^2}$, $F'_y = \frac{2y}{b^2}$, $F'_z = \frac{2z}{c^2}$ будет представлено в виде:

$$\frac{2x}{a^2}(X - x) + \frac{2y}{b^2}(Y - y) + \frac{2z}{c^2}(Z - z) = 0$$

или окончательно:

$$\frac{xX}{a^2} + \frac{yY}{b^2} + \frac{zZ}{c^2} - 1 = 0.$$

Уравнение касательной плоскости можно также получить путем дифференцирования уравнения данной поверхности и заменяя дифференциалы *dx, dy, dz* разностями *X-x, Y-y, Z-z.*

Уравнения нормали. Нормаль к поверхности в данной ее точке есть прямая, проходящая через эту точку и перпендикулярная к касательной плоскости в этой же точке поверхности. Понятие нормали играет существенную роль, в частности, в геометрической оптике (например, в формулировке основных законов преломления и отражения световых лучей), законы которой используются в машинной графике при синтезе реалистичных изображений моделей сцен.

Нормаль к поверхности F(x,y,z)=0 в точке $M(x, y, z)$ представляется уравнениями:

$$\frac{X-x}{F'_x}=\frac{Y-y}{F'_y}=\frac{Z-z}{F'_z}.$$

Если поверхность задана уравнением *z=f(x,y),* то нормаль к поверхности представляется уравнением:

$$\frac{X-x}{p}=\frac{Y-y}{q}=\frac{Z-z}{-1}.$$

Уравнение нормали к поверхности эллипсоида (см. пример выше):

$$\frac{a^2(X-x)}{x}=\frac{b^2(Y-y)}{y}=\frac{c^2(Z-z)}{z}.$$

Нормальное сечение и кривизна поверхности. Нормальное сечение поверхности *S* в данной ее точке *M* есть линия *L* пересечения поверхности с плоскостью, проведенной в направлении $\frac{du}{dv}$ через нормаль *n* в точке *M.*

С помощью нормального сечения изучается искривление поверхности *S* в различных (касательных) направлениях, выходящих из точки *M.* Величина, характеризующая отклонение поверхности от плоскости, называется *кривизной* поверхности. Кривизны нормальных сечений в точке М называются *нормальными кривизнами* поверхности в этой точке.

Нормальная кривизна *1/R* вычисляется по формуле:

$$\frac{1}{R}=\frac{L\ \ du^2+2M\ \ du\ \ dv+N\,dv^2}{E\,du^2+2F\ \ du\ \ dv+G\ \ dv^2}.$$

Максимальная и минимальная из нормальных кривизн в данной точке *M* именуются *главными кривизнами.* Если k_1 и k_2 - главные кривизны, то величины $K=k_1\,k_2$ *и* $H=1/2(k_1+k_2)$ называются соответственно *гауссовой кривизной* и *средней кривизной* поверхности в точке *M*. Эти кривизны поверхности определяют нормальные кривизны и поэтому могут служить характеристикой отклонения поверхности от плоскости. В частности, если *K=0* и *H=0* во всех точках поверхности, то поверхность представляет собой плоскость.

Гауссова кривизна не меняется при изгибаниях поверхности (деформация поверхности без изменения длин линий на ней). Если, например, гауссова кривизна равна нулю во всех точках поверхности, то каждый достаточно малый ее кусок может быть изогнут на плоскость. Гауссова кривизна на поверхности составляет объект внутренней геометрии поверхности, а средняя кривизна связана с внешней формой поверхности. Для поверхности в трехмерном эклидовом пространстве внешняя кривизна совпадает с ее внутренней кривизной. *Внутренняя* кривизна поверхности в эклидовом пространстве является ее *гауссовой кривизной.* (Внешняя кривизна зависит от кривизны пространства, в которой она располагается, например, Риманово пр-во).

При этом кривизны берутся со знаком + (или –), если направление вогнутости сечения совпадает (или не совпадает) с положительным направлением нормали к поверхности. *Нормальные кривизны* поверхности в произвольных направлениях достаточно просто можно выразить через главные кривизны. Так *кривизна поверхности* k_n *нормального сечения,* проведенного в направлении, составляющем угол φ с направлением наибольшей кривизны, связана с главными кривизнами в данной точке следующим соотношением (формула Эйлера):

$$k_n = k_1 \cos^2 \varphi + k_2 \sin^2 \varphi \ .$$

С помощью кривизн нормального сечения изучаются также кривизны наклонных сечений поверхности. Например, кривизна наклонного сечения плоскостью π, проходящей через данную касательную прямую, выражается формулой Мёнье:

$$k = k_n / \cos\theta,$$

где θ - угол между плоскостью π и нормалью к поверхности,
k_n - нормальная кривизна поверхности в направлении данной касательной.