УДК 518: 517.949.12

ГЛАДКАЯ ОКРУЖНОСТНАЯ ИНТЕРПОЛЯЦИЯ КРИВЫХ

В.А.Леус

В статье рассматриваются вопросы, связанные с приближенным решением задач прикладной геометрии. Основой вычислений в таких задачах является аппроксимация кривых и поверхностей в 3-мерном пространстве. Процесс аппроксимации поверхности может быть сведен к интерполяции функции одной переменной. Предлагается разновидность интерполяции многозначными функциями (см. статью I настоящего сборника). Проводятся оценки погрешности и доказывается сходимость окружностной интерполяции. Коротко рассказывается о практической реализации излагаемой методики.

Интерполяция плоских кривых

Рассмотрим совокупность $\{\mathcal{D}\}$ плоских непрерывных линий конечной длины, имеющих непрерывную касательную и кусочно-непрерывную ограниченную кривизну. Пусть $\mathcal{D}$ -кривая - некоторый элемент этой совокупности. На $\mathcal{D}$ -кривой выбирается конечная упорядоченная по длине дуги система узлов интерполяции таким образом, что выполняются следующие условия:

а) Если S - максимум длин (по дуге кривой) участков между узлами, K - максимум модуля кривизны $\mathcal{D}$ -кривой, то $K \cdot S < 1$.

б) Прямолинейное расположение узлов допускается только в пределах прямолинейных отрезков $\mathcal{D}$ - кривой.

Зададим каждый узел координатами (XY) в некоторой декартовой прямоугольной системе. Введем понятие местных координат для участка и пары смежных участков. Начало местной системы $(x\,y)$ для i -то участка (между двумя последовательными точками i , $i+1$) совпадает с началом системы (XY), ось абсцисс параллельна хорде участка, а положительное направление совпадает с направлением нумерации узлов. Местная система для пары участков определяется аналогично, только её ось абсцисс параллельна хорде, стягивающей начало первого участка с концом второго. Интерполяционная кривая (И-кривая) строится на каждом участке в отдельности. Для построения используются две точки, предшествующие интерполируемому участку, его концы и две последующие точки. Этим точкам присваиваются номера от единицы до шести в направлении общей нумерации (рис.I). Если точки 2,3,4 или 3,4,5 расположены прямолинейно, то кривая на участке (3,4) интерполируется хордой этого участка. Аналитическая запись $y = y_{34}$, где y_{34} - местная ордината хорды. Если прямолинейности среди узлов 2,3,4 или 3,4,5 нет, то из проходящих через концы участка (3,4) окружностей выбираются стягиваемые хордой (3,4) интерполирующие дуги - первая и вторая, которые могут быть представлены однозначными функциями $y = f_1(x)$, $y = f_2(x)$ в местной системе среднего участка. В случае прямолинейного расположения узлов I,2,3 первая интерполирующая дуга касательна хорде

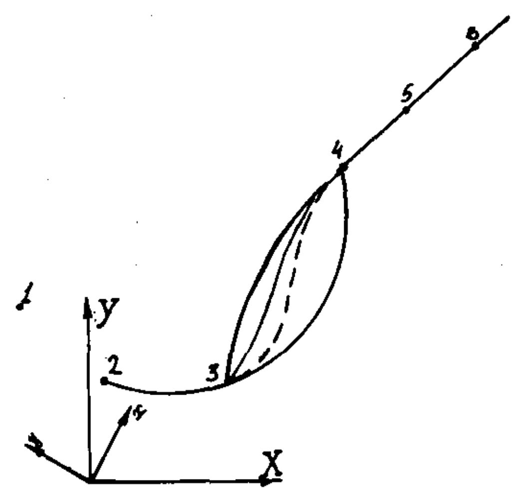

Рис. I

(2,3). Если узлы I,2,3 не лежат на одной прямой, то эта дуга проходит через точку 2. В случае прямолинейного расположения узлов 4, 5, 6 вторая интерполирующая дуга касается хорды (4, 5). Дуга проходит через точку 5, если прямолинейность отсутствует.

Введем две неотрицательные весовые функции $\varphi_1(x)$ и $\varphi_2(x)$, так что $\varphi_1(x)+\varphi_2(x)\equiv 1$. Аналитическое выражение для интерполяционной кривой (сокращенно И-кривая) на участке(3, 4) получается как весовая сумма

$$y(x)=\varphi_1(x)\cdot f_1(x)+\varphi_2(x)\cdot f_2(x). \qquad (I)$$

Практически применялись симметричные линейные веса, меняющиеся от 0 до I. Интерполяционная формула с такими весами имеет вид

$$y(x)=\left(\frac{x_4-x}{\ell}\right)\cdot\left[b_1+\operatorname{sign}(y_{34}-b_1)\cdot\sqrt{r_1^2-(x-a)^2}\right]+ \\ +\left(\frac{x-x_3}{\ell}\right)\cdot\left[b_2+\operatorname{sign}(y_{34}-b_2)\cdot\sqrt{r_2^2-(x-a)^2}\right]. \qquad (2)$$

Здесь a – местная абсцисса центров, b_j – местные ординаты центров, r_j – радиусы интерполирующих окружностей (j = I,2), sign – функция знака, ℓ – длина хорды.

Повторив процесс на всех участках, получим непрерывную кривую, проходящую через узлы. Окружностная интерполяция существует для всякой последовательности узлов, в которой угол поворота хорды от участка к участку меньше $\frac{\pi}{2}$. Исследуя первую производную функции, заданной выражением (2), можно убедиться в том, что касательная к И-кривой в концах участка совпадает с касательной к соответствующей окружности. Следовательно, при переходе через узлы касательная к И- кривой непрерывна.

Кривизна И-кривой имеет разрыв первого рода в каждом узле. Если обозначить f_1, $f_{2,i-1}$, $f_{2,i}$, f_3 функции, графики которых есть интерполирующие окружности i-1-го и i- го участков, то скачок кривизны на i -м узле запишется так:

$$\Delta k=\frac{2\left[f_3'(x_i^+)-f_{2,i}'(x_i^+)\right]+f_{2,i}''(x_i^+)}{\left\{1+\left[f_{2,i}'(x_i^+)\right]^2\right\}^{3/2}}-\frac{2\left[f_{2,i-1}'(x_i^-)-f_1'(x_i^-)\right]+f_{2,i-1}''(x_i^-)}{\left\{1+\left[f_{2,i-1}'(x_i^-)\right]^2\right\}^{3/2}}. \qquad (3)$$

Подчеркнем, что $f_{2,i-1}$ и $f_{2,i}$ соответствуют дугам одной окружности, а значения x_i^- и x_i^+ относятся к разным системам координат. В случае, когда i -й узел отделяет прямолинейный участок от криволинейного, одна из слагаемых дробей в формуле (3) равна нулю.

Чтобы получить И-кривые с непрерывной кривизной, нужно в качестве интерполирующих дуг, касающихся хорд соседних участков, использовать линии непрерывной кривизны,принимающей нулевое значение в точке касания. Кроме того, используемые весовые функции должны иметь непрерывные первые производные, обращающиеся в нуль на концах участка.

Рассматриваемый способ построения И-кривой является примером интерполяции параметризованного одномерного многообразия многозвенными функциями. Введение местных координат участка эквивалентно выбору в качестве параметра длины вдоль вписанной ломаной. Условие (а) гарантирует взаимно однозначное соответствие между значениями параметра и точками кривой.

Прямые и окружности интерполируются с нулевой ошибкой. Погрешность интерполяции произвольной кривой отлична от нуля и должна быть оценена. Далее будут указаны леммы, которые позволят оценить погрешность окружностной интерполяции.

Рассмотрим соотношение, связывающее кривизну $k(x)$ графика функции $y=f(x)$ с производными этой функции в декартовых прямоугольных координатах $(x\ y)$.

$$\frac{\dfrac{d^2f}{dx^2}}{\left[1+\left(\dfrac{df}{dx}\right)^2\right]^{3/2}}=k(x). \qquad (4)$$

При заданной правой части равенство (4) является дифференциальным уравнением относительно неизвестной функции $f(x)$.

Поставим следующую краевую задачу для этого уравнения на отрезке [0 I] оси x.

ЗАДАЧА. На отрезке [0 I] задана кусочно-непрерывная ограниченная функция $k(x)$. Найти на этом отрезке непрерывную, непрерывно дифференцируемую функцию $f(x)$, удовлетворяющую уравнению (4) и принимающую на концах отрезка заданные значения $y_0 = f(0)$, $y_1 = f(1)$.

ЛЕММА I. Если решение граничной задачи существует, то оно единственно.

ДОКАЗАТЕЛЬСТВО. Приводя обычным способом решение уравнения к квадратурам и пользуясь определенным интегралом с переменным верхним пределом как первообразной, получим выражение для производной

$$\frac{df}{dx} = \frac{\mathcal{K}(x) + C}{\sqrt{1-[\mathcal{K}(x)+C]^2}} , \tag{5}$$

где $\mathcal{K}(x) = \int\limits_0^x k(t)dt$, C – постоянная интегрирования. Для непрерывности производной $\frac{df}{dx}$ необходимо, чтобы знаменатель был больше нуля, или $|\mathcal{K}(x)+C| < 1$. Обозначим $\mathcal{K}_{min}$ и $\mathcal{K}_{max}$ – минимальное и максимальное значения функции $\mathcal{K}(x)$ на отрезке $[0,1]$, получим необходимое условие существования решения краевой задачи в виде неравенства

$$\mathcal{K}_{max} - \mathcal{K}_{min} < 2 . \tag{6}$$

Значения константы C можно искать только в пределах:

$$-1-\mathcal{K}_{min} < C < 1-\mathcal{K}_{max} . \tag{7}$$

Второй квадратурой с учетом условия на левом конце отрезка получим:

$$f(x) = \int\limits_0^x \frac{[\mathcal{K}(z)+C]\,dz}{\sqrt{1-[\mathcal{K}(z)+C]^2}} + y_0 = \mathcal{J}(C,x) + y_0 . \tag{8}$$

Вообще говоря, интеграл $\mathcal{J}(C,x)$ не выражается через элементарные функции, поэтому условие на правом конце отрезка

$$\mathcal{J}(C,1) + y_0 - y_1 = \mathcal{J}(C) + y_0 - y_1 = 0 \tag{9}$$

не позволяет явно определить константу C. Ответ на вопрос

о существовании и единственности решения краевой задачи можно получить, исследуя корни уравнения (9) в интервале (7) допустимых значений константы C.

Пусть выполнено необходимое условие (6). Тогда функция $\mathcal{J}(C)$ непрерывна в интервале допустимых значений и её производную можно получить дифференцированием по параметру под знаком интеграла

$$\frac{d\mathcal{J}}{dC} = \int\limits_0^1 \frac{\partial}{\partial C}\left\{\frac{[\mathcal{K}(z)+C]\,dz}{\sqrt{1-[\mathcal{K}(z)+C]^2}}\right\} = \int\limits_0^1 \frac{dz}{\{1-[\mathcal{K}(z)+C]^2\}^{3/2}} .$$

Здесь и в дальнейшем придется ссылаться на свойства функции $g(u) = \frac{u}{\sqrt{1-u^2}}$. Эта функция определена, непрерывна и непрерывно дифференцируема в интервале $(-1,1)$. Её производная $\frac{dg}{du} = \frac{1}{(1-u^2)^{3/2}}$ строго положительна, а сама функция монотонно возрастает от $(-\infty)$ до $(+\infty)$. Очевидно, функция $\mathcal{J}(C)$ тоже монотонно растет.

Пусть $\mathcal{K}_{max} - \mathcal{K}_{min} \leq 1 - \delta$, где δ – фиксированное положительное число меньше единицы. Пусть ε_1 и ε_2 произвольно малые положительные числа. Будем исследовать поведение функции $\mathcal{J}(C)$ на отрезке $-\mathcal{K}_{min}-1+\varepsilon_1 \leq C \leq -\mathcal{K}_{max}+1-\varepsilon_2$ из интервала допустимых значений. В силу монотонности функция $\mathcal{J}(C)$ принимает на левом конце этого отрезка минимальное значение $\mathcal{J}_{min}$, на правом – максимальное $\mathcal{J}_{max}$. Поскольку $\mathcal{K}(z) - \mathcal{K}_{min} \leq \mathcal{K}_{max} - \mathcal{K}_{min} \leq 1-\delta$, $\mathcal{K}(z) - \mathcal{K}_{max} \geq \mathcal{K}_{min} - \mathcal{K}_{max} \geq \delta - 1$,

то, учитывая свойства функции $g(u)$, получаем

$$\mathcal{J}_{min} = \int\limits_0^1 \frac{[\mathcal{K}(z) - \mathcal{K}_{min} - 1 + \varepsilon_1]\,dz}{\sqrt{1-[\mathcal{K}(z)-\mathcal{K}_{min}-1+\varepsilon_1]^2}} \leq \int\limits_0^1 \frac{(1-\delta-1+\varepsilon_1)\,dz}{\sqrt{1-(1-\delta-1+\varepsilon_1)^2}} =$$

$$= \frac{\varepsilon_1 - \delta}{\sqrt{1-(\varepsilon_1-\delta)^2}} \tag{10}$$

$$\mathcal{I}_{max} = \int_0^1 \frac{[\mathcal{K}(z) - \mathcal{K}_{max} + 1 - \varepsilon_2]\,dz}{\sqrt{1 - [\mathcal{K}(z) - \mathcal{K}_{max} + 1 - \varepsilon_2]^2}} \geqslant \int_0^1 \frac{(\delta - 1 + 1 - \varepsilon_2)\,dz}{\sqrt{1 - (\delta - 1 + 1 - \varepsilon_2)^2}} =$$

$$= \frac{\delta - \varepsilon_2}{\sqrt{1 - (\delta - \varepsilon_2)^2}} .$$

Левая часть уравнения (9) монотонно возрастает на отрезке $[-\mathcal{K}_{min} - 1 + \varepsilon_1, -\mathcal{K}_{max} + 1 - \varepsilon_2]$, пробегая значения от $(\mathcal{I}_{min} + y_0 - y_1)$ до $(\mathcal{I}_{max} + y_0 - y_1)$.

Пусть $\delta > \frac{|y_0 - y_1|}{\sqrt{1 + (y_0 - y_1)^2}}$, то есть $|y_0 - y_1| < \frac{\delta}{\sqrt{1 - \delta^2}}$.

При достаточно малых ε_1 и ε_2 правые части неравенств (10) могут быть сколь угодно близки по абсолютной величине к $\frac{\delta}{\sqrt{1 - \delta^2}}$. Следовательно, всегда найдется принадлежащий интервалу (7) отрезок, на концах которого левая часть уравнения (6) принимает значения разных знаков. В силу монотонного возрастания непрерывной функции $\mathcal{Y}(C)$ в допустимом интервале найдется единственный корень уравнения (9), который и дает решение граничной задачи. Достаточным условием существования решения является неравенство

$$\mathcal{K}_{max} - \mathcal{K}_{min} < 1 - \frac{|y_0 - y_1|}{\sqrt{1 + (y_0 - y_1)^2}} . \qquad (11)$$

Лемма доказана.

Сформулируем и докажем лемму 2 о сравнении функций через кривизны их графиков.

Рассмотрим две краевые задачи для уравнения (4) с различными правыми частями $k_1(x)$ и $k_2(x)$ при одинаковых краевых условиях $f_1(0) = f_2(0) = y_0$, $f_1(1) = f_2(1) = y_1$. Иногда, не решая задачи, можно сделать определенные заключения о функциях $f_1(x)$ и $f_2(x)$ только на основе знания кривизны их графиков.

ЛЕММА 2. Если $k_1(x) \geqslant k_2(x)$ на отрезке [0 1], и решения соответствующих краевых задач существуют, то

1) в интервале (0 1) эти решения удовлетворяют неравенству $f_1(x) < f_2(x)$,

2) на концах отрезка [0 1] значения производных удовлетворяют неравенствам $f_1'(0) < f_2'(0)$, $f_1'(1) > f_2'(1)$.

ДОКАЗАТЕЛЬСТВО. Представим функции $f_1(x)$ и $f_2(x)$ в виде (8)

$$f_1(x) = y_0 + \int_0^x \frac{[\mathcal{K}_1(z) + C_1]\,dz}{\sqrt{1 - [\mathcal{K}_1(z) + C_1]^2}} ,$$

$$f_2(x) = y_0 + \int_0^x \frac{[\mathcal{K}_2(z) + C_2]\,dz}{\sqrt{1 - [\mathcal{K}_2(z) + C_2]^2}} .$$

Поскольку $k_1(x) \geqslant k_2(x)$, то

$$\mathcal{K}_1(x) = \int_0^x k_1(t)\,dt \geqslant \int_0^x k_2(t)\,dt = \mathcal{K}_2(x).$$

На правом конце отрезка [0 1] должно выполняться неравенство $\mathcal{K}_1(1) > \mathcal{K}_2(1)$, так как $k_1(x) \not\equiv k_2(x)$. Значения функций $f_1(x)$ и $f_2(x)$ здесь совпадают

$$\int_0^1 \frac{[\mathcal{K}_1(z) + C_1]\,dz}{\sqrt{1 - [\mathcal{K}_1(z) + C_1]^2}} = \int_0^1 \frac{[\mathcal{K}_2(z) + C_2]\,dz}{\sqrt{1 - [\mathcal{K}_2(z) + C_2]^2}} . \qquad (12)$$

В равенстве (12) можно рассматривать подынтегральные выражения как сложные функции $G_1(z) = g(u_1(z))$ и $G_2(z) = g(u_2(z))$, где $u_1(z) = \mathcal{K}_1(z) + C_1$, $u_2(z) = \mathcal{K}_2(z) + C_2$.

Допустим $C_1 \geqslant C_2$, тогда $u_1(z) \geqslant u_2(z)$, причем $u_1(z) \not\equiv u_2(z)$. По свойству монотонности функции $g(u)$ имеем $G_1(z) = g(u_1) \geqslant g(u_2) = G_2(z)$, кроме того $G_1(z) \not\equiv G_2(z)$. Но в таком случае равенство (12) невозможно. Следовательно, допущение неверно и остается принять строгое неравенство $C_1 < C_2$. Учитывая, что $\mathcal{K}_1(0) = \mathcal{K}_2(0) = 0$, получим из соотношения (5) значения производных функций $f_1(x)$ и $f_2(x)$ в нуле

$$f_1'(0)=\frac{C_1}{\sqrt{1-C_1^2}}=g(C_1),\quad f_2'(0)=\frac{C_2}{\sqrt{1-C_2^2}}=g(C_2).$$

В силу монотонного возрастания функции $g(u)$ имеем

$$f_1'(0)<f_2'(0). \qquad (13)$$

Изменив положительное направление оси абсцисс на противоположное и перенося начало координат в правый конец отрезка 0,1, получим при помощи тех же рассуждений неравенство аналогичное (13). В первоначальных осях оно эквивалентно неравенству

$$f_1'(1)>f_2'(1).$$

Второе утверждение леммы доказано.

Поскольку $f_1(x)$ и $f_2(x)$ непрерывны и непрерывно дифференцируемы, то из (13) следует существование интервала $(0\,x_A)$, в котором $f_1(x)<f_2(x)$, причем $f_1(x_A)=f_2(x_A)$. Графики функций $f_1(x)$ и $f_2(x)$ соединяют точки M и N плоскости $(x\,y)$. Точка A является общей для графиков, поэтому функции $f_1(x)$ и $f_2(x)$ можно рассматривать на отрезке $[0\,x_A]$ как решения краевых задач для уравнения (4) с правыми частями $k_1(x)$ и $k_2(x)$ (рис.2). Функции $f_1(x)$ и $f_2(x)$ не совпадают тождественно на этом отрезке. В силу единственности решения краевых задач можно утверждать, что $k_1(x)\not\equiv k_2(x)$ на отрезке $[0\,x_A]$. Но тогда по доказанному

$$f_1'(x_A)>f_2'(x_A). \qquad (14)$$

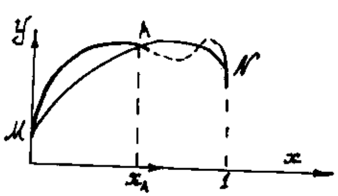

Рис.2

Допустим, что точка A отлична от N. Проводя для отрезка $[x_A\,1]$ те же рассуждения, что и для отрезка $[0\,x_A]$, получим в точке x_A неравенство

$$f_1'(x_A)<f_2'(x_A). \qquad (15)$$

Противоречивость неравенств (14) и (15) говорит о непременном совпадении точек A и N. Интервал $(0\,x_A)$, в котором

$f_1(x)<f_2(x)$, совпадает с интервалом (0, 1), что и утверждается в первом пункте леммы о сравнении.

Отметим, что интегральные представления (5), (8), а следовательно все основанные на них рассуждения, остаются в силе для более широкого (чем кусочно-непрерывные) класса интегрируемых функций $k(x)$.

Оценка погрешности и сходимость

Будем различать погрешность-удаленность полученных точек от истинных и угловую погрешность – разницу в направлениях истинной касательной (нормали) и приближенной.

Пусть на некоторой D-кривой выбрана система узлов, удовлетворяющая условиям (α,δ) и получена окружностная И-кривая. Обозначим h_i максимум модуля разности местных ординат точек кривых U и D на i-м участке, т.е. $h_i=\max_x |f_D(x)-f_U(x)|$. Наибольшую из этих величин по всем участкам будем называть погрешностью интерполяции

$$h(D,U)=\max_i h_i. \qquad (16)$$

Рассмотрим тройку смежных участков между узлами 2,3,4,5 (рис.1). Обозначим l длину хорды среднего участка, k_{min} и k_{max} – минимальное и максимальное значения кривизны Д-кривой на протяжении тройки участков и оценим отклонение h на среднем участке. Пусть k_1 и k_2 – кривизны первой и второй интерполирующих дуг. Первая дуга и кусок Д-кривой между точками 2 и 4 можно рассматривать как отдельные кривые в местной системе координат первой пары участков. Кривизна k_1 удовлетворяет неравенству

$$k_{min}\le k_1\le k_{max}. \qquad (17)$$

Действительно, первая интерполирующая дуга либо имеет с Д-кривой третью общую точку внутри отрезка $[x_2,x_4]$, либо касается Д-кривой в точке 3. Опираясь соответственно на пункт первый или второй леммы о сравнении, заключаем, что k_1 не может быть больше (или меньше) кривизны Д-кривой всюду на протяжении рассматриваемой тройки участков. Аналогично доказывается неравенство для кривизны k_2 второй интерполирующей дуги

$$k_{min} \leqslant k_2 \leqslant k_{max} . \qquad (18)$$

Дуги окружностей с кривизнами k_{min} и k_{max}, проходящие через концы участка и имеющие уравнения $y = f_{min}(x)$, $y = f_{max}(x)$ назовем минимальной и максимальной. Для кривизны Д-кривой на участке (3,4) имеем очевидное неравенство $k_{min} \leqslant k \leqslant k_{max}$. Применяя лемму о сравнении к минимальной дуге среднего участка и Д-кривой на нем, а затем к максимальной и Д-кривой, получим неравенство:

$$f_{max}(x) < f(x) < f_{min}(x)$$

справедливое всюду в интервале $(x_3\, x_4)$. Иными словами, Д-кривая на среднем участке лежит между минимальной и максимальной дугами. С другой стороны, исходя из неравенств (17), (18) на основании леммы о сравнении можно утверждать, что обе интерполирующие окружности и, следовательно, И-кривая заключены между минимальной и максимальной дугами. Отсюда вытекает оценка погрешности окружностной интерполяции на участке (3 4): отклонение И-кривой от Д-кривой по местной ординате не превышает H_1 - толщины "линзы", образованной минимальной и максимальной дугами.

$$H_1(k_{max}, k_{min}, \ell) = \frac{1-\sqrt{1-(\frac{\ell k_{max}}{2})^2}}{k_{max}} - \frac{1-\sqrt{1-(\frac{\ell k_{min}}{2})^2}}{k_{min}} \qquad (19)$$

Оценка погрешности на i -м участке имеет вид:

$$h_i < H_1(k_{i\,max}, k_{i\,min}, \ell_i). \qquad (20)$$

и справедлива при любых весовых функциях, используемых в интерполяционной формуле. Интерполяция сходится, если при стремлении к нулю максимума длин участков между узлами погрешность стремится к нулю.

ТЕОРЕМА О СХОДИМОСТИ. Для всякой Д-кривой при любом $\varepsilon > 0$ найдется $\delta(\varepsilon) > 0$ такое, что окружностная И-кривая, построенная по системе узлов с максимальной длиной дуги участка $s < \delta$, лежит в ε-окрестности Д-кривой.

Доказательство тривиально, если все системы узлов удовлетворяют условиям (α, δ). Действительно, в этом случае погрешность оценивается выражением типа (19), которое стремится к нулю не медленнее ℓ^2, поскольку кривизна Д-кривой ограничена. Если не требовать выполнения условия (б), то на участках, где оно нарушается, и на соседних с ними, нужно проводить оценку погрешности особо. Если s - длина дуги участка Д-кривой, ℓ - длина хорды, k - максимум модуля кривизны Д-кривой на трех смежных участках., то погрешность оценивается величинами соответственно: $H_2(k,\ell) =$

$$= \frac{1-\sqrt{1-(\frac{\ell k}{2})^2}}{k}, \qquad H_3(k,s) = \frac{1-\sqrt{1-(\frac{3k}{2})^2}}{k} .$$

Используя большие буквы для обозначения величин в целом по всей Д-кривой, запишем оценки ε-окрестности при выполнении условия (б)

$$\varepsilon < H_1(K_{max}, K_{min}, L), \qquad (21)$$

при нарушении

$$\varepsilon < H_3(K, S). \qquad (22)$$

В последнем случае также имеет место сходимость второго порядка относительно S, но количественно оценка здесь не лучше, чем для кусочно-линейной интерполяции.

Переходим к угловой погрешности. Предварительно докажем две леммы, на основе которых затем получим оценку угловой погрешности.

Пусть на отрезке [0 1] оси x задана непрерывная, непрерывно-дифференцируемая функция $f(x)$ с кусочно-непрерывной ограниченной второй производной; $f(x)$ принимает нулевые значения на концах отрезка. Её график, очевидно, имеет кусочно-непрерывную ограниченную кривизну $k(x)$. Пусть функции $r(x)$ и $R(x)$, принимающие нулевые значения на концах отрезка, задают дуги минимальной и максимальной окружностей с кривизнами k_{min} и k_{max}.

ЛЕММА 3. Производная функции $f(x)$ по абсолютной величине не превосходит максимума модулей производных функций $r(x)$, $R(x)$

$$\left|\frac{df}{dx}\right| \le max\left\{\max_x\left|\frac{dr}{dx}\right|, \max_x\left|\frac{dR}{dx}\right|\right\} = T. \qquad (23)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай, когда $r(x)$ и $R(x)$ неотрицательны и, следовательно, $T = \left.\frac{dr}{dx}\right|_{x=0}$. В остальных случаях рассуждения аналогичны.

По лемме о сравнении график функции $f(x)$ заключен между дугами минимальной и максимальной окружностей, а её производная $\frac{df}{dx}$ на концах отрезка [0 1] удовлетворяет строгому неравенству (23). Допустим, что в интервале (0 1) есть точка x_A, в которой $\frac{df}{dx} = -T$ (случай с величиной $+T$ можно отдельно не рассматривать, так как он приводится к первому заменой $x' = 1 - x$). Касательную к графику функции $f(x)$ в точке A продолжим до пересечения с осью абсцисс в точке B, принадлежащей интервалу (0 1) (рис.3). На отрезке $[0\ x_B]$ рассмотрим графики функций $\bar{r}(x)$ и $\bar{f}(x)$. Первый есть дуга OB с кривизной окружности $r(x)$, второй составлен из дуги OA графика $f(x)$ и прямолинейного отрезка AB. График функции $\bar{f}(x)$ имеет кривизну не меньшую кривизны окружности $\bar{r}(x)$. На основании леммы о сравнении заключаем

$$\left.\frac{d\bar{r}}{dx}\right|_{x=x_B} < \left.\frac{df}{dx}\right|_{x=x_B} = \left.\frac{dr}{dx}\right|_{x=1},$$

что невозможно, поскольку дуга OB окружности $\bar{r}(x)$ меньше дуги окружности $r(x)$ того же радиуса. Следовательно, точка A не существует, т.е. неравенство (23) выполнено всюду на отрезке [0 1]. Лемма доказана.

Перейдем к следующей лемме. На отрезке $[-\frac{1}{2}\ \frac{1}{2}]$ оси x рассмотрим линейно-весовую сумму функций $f_1(x)$ и $f_2(x)$

$$f(x) = P_1(x) f_1(x) + P_2(x) f_2(x),$$

где

$$P_1(x) = \frac{1}{2} - x, \quad P_2(x) = \frac{1}{2} + x,$$

а функции $f_1(x)$, $f_2(x)$ соответствуют дугам окружностей, проходящих через концы отрезка.

ЛЕММА 4. Производная $\frac{df}{dx}$ линейно-весовой суммы на отрезке $[-\frac{1}{2}\ \frac{1}{2}]$ не превышает по абсолютной величине максимума модулей складываемых окружностей

$$\left|\frac{df}{dx}\right| \le max\left\{\max_x|f_1'|, \max_x|f_2'|\right\} = V. \qquad (24)$$

ДОКАЗАТЕЛЬСТВО. Обозначим r_j - радиусы, b_j - ординаты центров окружностей, проходящих через концы отрезка $[-\frac{1}{2}\ \frac{1}{2}]$. Доказательство проведем только для случая, когда b_1 и b_2 разных знаков. Пусть для определенности $b_1 > 0$ и $b_1 \le |b_2|$ (рис.4). Складываемые функции имеют вид:

$$f_1(x) = b_1 - \sqrt{r_1^2 - x^2} = b_1 - \sqrt{b_1^2 - (\tfrac{1}{4} - x^2)},$$

$$f_2(x) = b_2 - \sqrt{b_2^2 - (\tfrac{1}{4} - x^2)}.$$

С учетом того, что $P_1' \equiv -1$, $P_2' \equiv 1$, имеем:

$$\frac{df}{dx} = (f_2 - f_1) + (P_1 f_1' + P_2 f_2') = d(x) + e(x). \qquad (25)$$

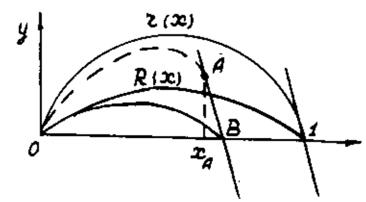

Рис. 3.

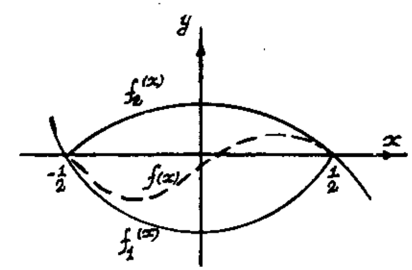

Рис. 4.

Оценим первое слагаемое (25)

$$d(x) \leqslant f_2(0) - f_1(0) \leqslant -2f_1(0) < 2(z_1 - b_1) \cdot$$
$$\cdot \left(1 + \frac{z_1}{b_1}\right) = \frac{2(z_1^2 - b_1^2)}{b_1} = \frac{1}{2b_1}$$

Но в рассматриваемом случае $V = \frac{1}{2b_1}$, следовательно, оценка первого слагаемого: $0 \leqslant d(x) < V$. Второе слагаемое (25) по модулю не превосходит V, так как является весовой суммой величин, каждая из которых по модулю не больше V. Очевидно, всюду на отрезке $[-\frac{1}{2}\ \frac{1}{2}]$ выполняется неравенство $|f_2'(x)| \leqslant |f_1'(x)|$. На отрезке $[-\frac{1}{2}\ 0]$ производная $f_1'(x)$ неположительна и входит в $e(x)$ с весом не меньшим $\frac{1}{2}$, поэтому сама функция $e(x)$ неположительна на отрезке $[-\frac{1}{2}\ 0]$. Поскольку здесь слагаемые в выражении производной $\frac{df}{dx}$ имеют разные знаки, и каждое по модулю не превосходит величины V, то модуль их суммы тоже не больше V. На левой половине отрезка $[-\frac{1}{2}\ \frac{1}{2}]$ неравенство (24) выполняется.

На правой половине этого отрезка исследуем производную:

$$\frac{df}{dx} = (b_2 + \sqrt{z_2^2 - x^2}) - (b_1 - \sqrt{z_1^2 - x^2}) +$$
$$+ \frac{x(\frac{1}{2} - x)}{\sqrt{z_1^2 - x^2}} - \frac{x(\frac{1}{2} + x)}{\sqrt{z_2^2 - x^2}}$$

Пусть $b_1 \leqslant \frac{1}{2}$. Учитывая соотношение $b^2 = z^2 - \frac{1}{4}$, преобразуем выражение производной к виду:

$$\frac{df}{dx} = \frac{x(\frac{1}{2} - x)}{z_1\sqrt{1 - (\frac{x}{z_1})^2}} - \frac{x(\frac{1}{2} + x)}{z_2\sqrt{1 - (\frac{x}{z_2})^2}} + \left[z_2\sqrt{1 - (\frac{x}{z_2})^2} - z_2\sqrt{1 - \frac{1}{4z_2^2}}\right] +$$
$$+ \left[z_1\sqrt{1 - (\frac{x}{z_1})^2} - z_1\sqrt{1 - \frac{1}{4z_1^2}}\right].$$

Оценим первую дробь:

$$0 \leqslant \frac{x(\frac{1}{2} - x)}{z_1\sqrt{1 - (\frac{x}{z_1})^2}} \leqslant \frac{x(\frac{1}{2} - x)}{z_1[1 - (\frac{x}{z_1})^2]} \leqslant \frac{x(\frac{1}{2} - x)}{z_1(1 - \frac{x}{z_1})} =$$
$$= \frac{x(\frac{1}{2} - x)}{z_1 - x} \leqslant \frac{x}{2z_1}$$

Для второй дроби очевидно неравенство:

$$\frac{x(\frac{1}{2} + x)}{z_2\sqrt{1 - (\frac{x}{z_2})^2}} \geqslant \frac{x(\frac{1}{2} + x)}{z_2} \geqslant 0$$

Представим радикал $\sqrt{1 - (\frac{x}{z})^2}$ в виде ряда

$$\sqrt{1 - (\frac{x}{z})^2} = 1 - \frac{1}{2}\left(\frac{x}{z}\right)^2 - \frac{1}{2\cdot 4}\left(\frac{x}{z}\right)^4 - \frac{1\cdot 3}{2\cdot 4\cdot 6}\left(\frac{x}{z}\right)^6 - \ldots,$$

сходящегося при всех $x \in [0\ \frac{1}{2}]$. Любая частичная сумма дает предельную функцию с избытком, поэтому $\sqrt{1 - (\frac{x}{z})^2} < 1 - \frac{1}{2}(\frac{x}{z})^2$. Теперь оценим сверху производную $\frac{df}{dx}$ на отрезке $[0\ \frac{1}{2}]$.

$$\frac{df}{dx} < \frac{x}{2z_1} - \frac{x(\frac{1}{2} + x)}{z_2} + \left\{z_2\left[1 - \frac{1}{2}\left(\frac{x}{z_2}\right)^2\right] - z_2\left(1 - \frac{1}{4z_2^2}\right)\right\} + \left\{z_1\left[1 - \frac{1}{2}\left(\frac{x}{z_1}\right)^2\right] -\right.$$
$$\left. - z_1\left(1 - \frac{1}{4z_1^2}\right)\right\} = \frac{1 + 2x - 2x^2}{4z_1} + \frac{1 - 2x - 6x^2}{4z_2} < \frac{1 + 2x - 2x^2}{4z_1} + \frac{1}{4z_1} =$$
$$= \frac{1 + x - x^2}{2z_1} \leqslant \frac{5}{8z_1} = \frac{5}{4\sqrt{1 + \frac{1}{4b_1^2}}} \cdot \frac{1}{2b_1} < \frac{5}{4\sqrt{2}} \cdot \frac{1}{2b_1} < \frac{1}{2b_1} = V$$

Пусть теперь $b_1 > \frac{1}{2}$. Имеем:

$$\frac{df}{dx} \leqslant \frac{x(\frac{1}{2} - x)}{\sqrt{z_1^2 - x^2}} - 2(b_1 - \sqrt{z_1^2 - x^2}) =$$
$$= \frac{x(\frac{1}{2} - x)}{b_1\sqrt{1 + \frac{\frac{1}{4} - x^2}{b_1^2}}} - 2b_1\left(1 - \sqrt{1 + \frac{\frac{1}{4} - x^2}{b_1^2}}\right).$$

Поскольку на отрезке $[0\ \frac{1}{2}]$ при $b_1 > \frac{1}{2}$ справедливо неравенство $0 \leqslant \frac{\frac{1}{4} - x^2}{b_1^2} < 1$, то можно предложить оценку производной:

$$\frac{df}{dx} \leqslant \frac{x(\frac{1}{2} - x)}{b_1} + \frac{\frac{1}{4} - x^2}{b_1} = \frac{\frac{1}{2} + x - 4x^2}{2b_1}.$$

Квадратичная функция $\frac{1}{2}+x-4x^2$ достигает максимального значения $\frac{9}{16}$ при $x=\frac{1}{8}$, следовательно,

$$\frac{df}{dx} \leqslant \frac{9}{16}\cdot\frac{1}{2b_1} < \frac{1}{2b_1} = V$$

Чтобы получить оценку снизу, вернемся к выражению (25). Достаточно оценить величину $\mathcal{P}_2 f_2'$, так как только она является неположительной на отрезке $[0\ \frac{1}{2}]$. Эта величина минимальна в правом конце отрезка $\mathcal{P}_2 f_2' \geqslant f_2'(\frac{1}{2}) = \frac{1}{2b_2}$.

По условию $b_1 \leqslant |b_2|$, т.е. $\frac{1}{2b_2} \geqslant -\frac{1}{2b_1} = -V$. Следовательно, на отрезке $[0\ \frac{1}{2}]$ выполняется неравенство $\frac{df}{dx} \geqslant \mathcal{P}_2 f_2' \geqslant -V$. Лемма 4 доказана.

При окружностной интерполяции плоской кривой интерполирующие окружности "зажаты" между минимальной и максимальной дугами на каждом участке. Следовательно, в местных осях имеет место неравенство $V<T$. На основании лемм 3 и 4 заключаем, что производные функций, графиками которых являются дуги И-кривой и Д-кривой, отличаются меньше чем на T:

$$\left|\frac{df_Д}{dx} - \frac{df_И}{dx}\right| < T = \frac{\ell k}{2\sqrt{1-(\frac{\ell k}{2})^2}} \quad (26)$$

Здесь ℓ - длина хорды, k - максимум модуля кривизны на трех участках. Там, где условие (б) нарушено, оценка содержит величину s - длину дуги участка Д-кривой, вместо величины ℓ.

Если S -максимум длин участков, K - максимум модуля кривизны Д-кривой в целом, то разности производных по всей кривой оцениваются неравенством:

$$\left|\frac{df_Д}{dx} - \frac{df_И}{dx}\right| < \frac{SK}{2\sqrt{1-(\frac{SK}{2})^2}} = T(S,K). \quad (27)$$

В случае пространственной кривой можно интерполировать её проекции на некоторые координатные плоскости. Второй путь состоит в следующем: начало местной системы $(x\,y\,z)$ помещается в точку 3, ось x направляется вдоль хорды (3 4) ортогонально осям y и z. Плоскость (xy) определяется точками 2,3,4, плоскость (xz) - точками 3,4,5. Пусть через точки 3,4 проходят окружности $y=f_1(x)$ и $z=f_2(x)$.

Радиус-вектор И-кривой является функцией координаты x, а именно $\vec{r}(x)=(x, \mathcal{P}_1(x)f_1(x), \mathcal{P}_2(x)f_2(x))$.

З а д а ч а о п е р е с е ч е н и и. Для практических приложений важно уметь решать задачу об отыскании точки пересечения двух плоских кривых, имеющих уравнения типа $y=f(x)$ в различных системах координат $(x\,y)^{(1)}$ и $(x\,y)^{(2)}$. Пусть система (2) повернута на угол φ относительно системы (I), причем имеет место неравенство $0<\varphi<\pi$. Последнее всегда может быть удовлетворено взаимным изменением нумерации систем.

Случаи $\varphi=0$, $\varphi=\pi$ не рассматриваются, так как при этом задача ограничивается решением одного уравнения $f_1(x)-f_2(x)=0$.

Задачу будем решать последовательными приближениями. Для некоторого значения x_1 абсциссы любой координатной системы, например первой, вычисляется значение ординаты $y_1^{(1)}=f_1(x_1^{(1)})$ точки на соответствующей кривой. Координаты $(x_1 y_1)^{(1)}$ переводятся в другую систему, получаются координаты $(x_1\ y_1)^{(2)}$. Для значения $x_2^{(2)}=x_1^{(2)}$ абсциссы второй системы вычисляется соответствующее значение ординаты $y_2^{(2)}=f_2(x_2^{(2)})$ точки на второй кривой. Координаты $(x_2\ y_2)^{(2)}$ переводятся в первую систему и т.д. Процесс повторяется до тех пор, пока расстояние между точками $(x_n\ y_n)$ и $(x_{n+1}\ y_{n+1})$ не станет меньше заданного ε (рис.5а).

Из наглядного примера (рис.5б) ясно, что характер процесса может меняться: сходимость по "зигзагу", по "спирали", ненаправленный "цикл", вращающийся вокруг точки пересечения. Наконец, процесс может оказаться расходящимся (рис.5в). Исследуем причины такого разнообразия и установим достаточное условие сходимости.

Начнем с простейшего случая. Две прямые пересекаются в точке $\mathcal{P}$. Пусть $a^{(1)}\wedge b^{(1)}$, $a^{(2)}\wedge b^{(2)}$ - ортонормальные базисные реперы, задающие координатные системы $(x\,y)^{(1)}$ и

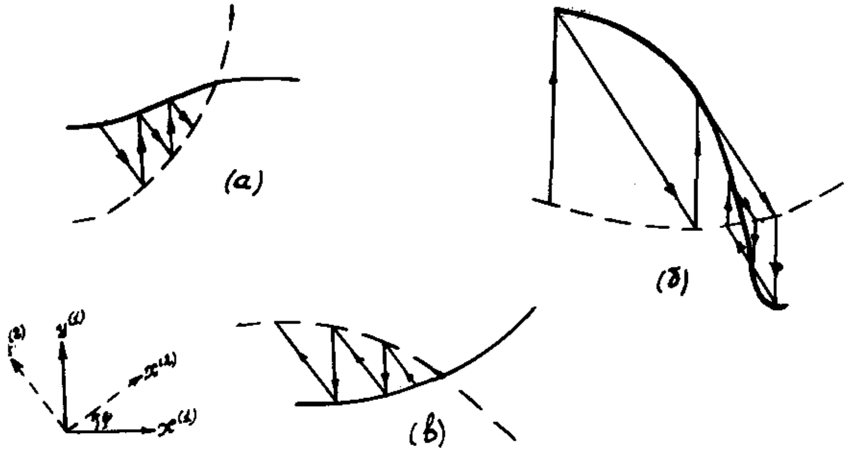

Рис. 5.

$(x\,y)^{(2)}$ с общим началом в точке P. Уравнение первой прямой в первой системе имеет вид $y^{(1)}(x^{(1)}) = tg\,\alpha \cdot x^{(1)}$, второй прямой во второй системе $-y^{(2)}(x^{(2)}) = tg\,\beta \cdot x^{(2)}$, где α – полярный угол единичного вектора „c" первой прямой, β – полярный угол единичного вектора „d" второй прямой. Расположение таково, что $-\frac{\pi}{2} < \alpha, \beta < \frac{\pi}{2}$.

Будем применять к этим прямым процесс последовательных приближений. Введем линейную функцию точки $F(r) = (a^{(1)}, r)$, где r – радиус-вектор произвольной точки относительно полюса в точке P. Линии уровня этой функции суть прямые, параллельные оси ординат первой системы. Один полный "период" последовательных приближений, т.е. переход с первой прямой на вторую и возврат, можно интерпретировать как переход с одного уровня $F(r) = F_1$ на другой $F(r) = F_2$. Точка P лежит на нулевом уровне $F(r) = 0$. Если $|F_2| < |F_1|$, то процесс сходится, в противном случае не сходится или расходится (рис.6).

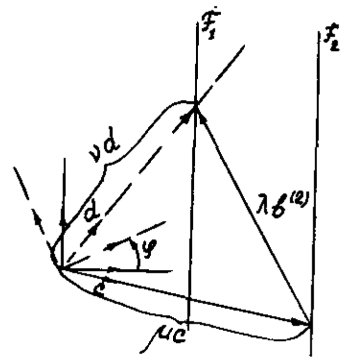

Рис. 6

Пусть исходная точка на первой прямой имеет радиус-вектор $r_1 = \mu c$. Переход в точку νd второй прямой производится вдоль вектора $b^{(2)}$. Имеет место векторное равенство

$$\mu c + \lambda b^{(2)} = \nu d. \qquad (28)$$

Значение исходного уровня функции $F(r)$ есть:

$$F_1 = (a^{(1)}, \mu c) = \mu(a^{(1)}, c) = \mu\cos\alpha,$$

значение последующего уровня:

$$F_2 = (a^{(1)}, \nu d) = \mu(a^{(1)}, c) + \lambda(a^{(1)}, b^{(2)}) = \mu\cos\alpha + \lambda\sin\varphi.$$

Запишем векторное равенство (28) в координатах

$$\mu\cos\alpha - \lambda\sin\varphi = \nu\cos(\varphi+\beta)$$

$$\mu\sin\alpha + \lambda\cos\varphi = \nu\sin(\varphi+\beta).$$

Исключая ν, разрешим равенство относительно λ:

$$\lambda = \frac{\mu\sin(\varphi+\beta-\alpha)}{\cos\beta}.$$

Отношение $\frac{F_2}{F_1}$ обозначим q и назовем коэффициентом направленности процесса последовательных приближений. Учитывая вышенаписанное равенство, имеем:

$$q = \frac{\cos\alpha - \frac{\sin(\varphi+\beta-\alpha)\sin\varphi}{\cos\beta}}{\cos\alpha} = \frac{\cos(\varphi-\alpha)\cos(\varphi+\beta)}{\cos\alpha\cos\beta}.$$

В таблице (рис.7) показана зависимость характера направленности от величины q. При фиксированном φ коэффициент направленности есть функция $q(\alpha,\beta)$. Представим её в следующем виде

$$q(\alpha,\beta) = -\sin^2\varphi\,(tg\,\alpha + ctg\,\varphi)(tg\,\beta - ctg\,\varphi). \qquad (29)$$

$1<\|q\|$	Расходимость	$1<q$	по зигзагу
		$q<-1$	по спирали
$\|q\|=1$	Повторение	$q=1$	отрезка
		$q=-1$	цикла
$\|q\|<1$	Сходимость	$0<q<1$	по зигзагу
		$-1<q<0$	по спирали
		$q=0$	за конечное число шагов

Рис. 7.

Будем считать, что φ входит в это выражение как параметр, т.е. (29) задает однопараметрическую совокупность функций двух переменных. Каждая функция $q(\alpha,\beta)$ из этой совокупности определена, непрерывна и непрерывно дифференцируема внутри квадрата $|\alpha|,|\beta|<\frac{\pi}{2}$ плоскости переменных α, β. Формулы

$$\left.\begin{aligned} \xi &= \mathrm{tg}\,\alpha \\ \eta &= \mathrm{tg}\,\beta \end{aligned}\right\} \qquad \left.\begin{aligned} \alpha &= \mathrm{arc\,tg}\,\xi \\ \beta &= \mathrm{arc\,tg}\,\eta \end{aligned}\right\} \tag{30}$$

задают взаимно однозначное, взаимно непрерывное соответствие (гомеоморфизм) между внутренними точками этого квадрата и точками всей плоскости переменных ξ, η. Из (29) получим коэффициент направленности как функцию от ξ и η:

$$q(\xi,\eta) = -\sin^2\varphi\cdot(\xi+\mathrm{ctg}\,\varphi)(\eta-\mathrm{ctg}\,\varphi).$$

Функцию $q(\xi,\eta)$ проще исследовать, чем $q(\alpha,\beta)$. Основные выводы относительно последней можно получить, учитывая свойства гомеоморфизма (30). Множество линий уровней функции $q(\xi,\eta)$ есть семейство гипербол с общими асимптотами $\xi=-\mathrm{ctg}\,\varphi$ и $\eta=\mathrm{ctg}\,\varphi$. Положительные уровни расположены во второй и четвертой четвертях, отрицательные - в первой и третьей. Гиперболы $q(\xi,\eta)=1$ и $q(\xi,\eta)=-1$ очерчивают область сходимости, содержащую всегда начало координат (рис.8а). В квадрате $|\alpha|,|\beta|<\frac{\pi}{2}$ плоскости α, β асимптотам соответствует "асимптотический крест" из от-

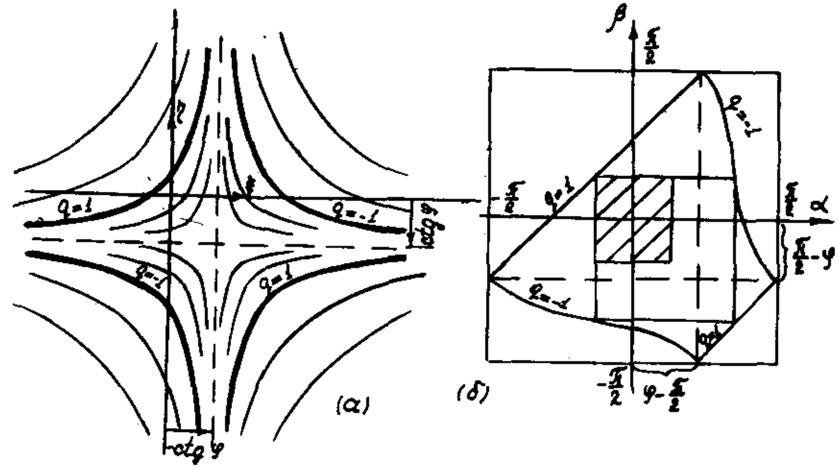

Рис. 8

резков прямых $\alpha=\varphi-\frac{\pi}{2}$ и $\beta=\frac{\pi}{2}-\varphi$, который является нулевым уровнем функции $q(\alpha,\beta)$. Его точкам соответствуют такие направления, для которых процесс последовательных приближений сходится за конечное число шагов. Уровни $q(\alpha,\beta)=1$ и $q(\alpha,\beta)=-1$ ограничивают колоколовидную область сходимости (рис,8б). Начало координат всегда лежит внутри этой области, т.е. процесс последовательных приближений всегда сходится для направлений, параллельных осям абсцисс координатных систем.

Два направления, задаваемые векторами c и d будем называть взаимно сходящимися, если процесс последовательных приближений для соответствующих прямых сходится. Каждой точке (α,β), принадлежащей области сходимости, соответствуют два взаимно сходящихся направления и наоборот. Отрезку $[\alpha_1,\alpha_2]$ оси α соответствует некоторый пучок направлений $\{c\}$, отрезку $[\beta_1,\beta_2]$ - пучок $\{d\}$. Два пучка будем называть взаимно сходящимися, если каждое направление одного из них взаимно сходится с каждым направлением другого. Этим пучкам соответствует координатный (со сторонами параллельными координатным осям) прямоугольник, целиком лежащий внутри области сходимости. Имея в виду указанное соответствие, будем называть такие прямоугольники сходящимися.Если прямоугольник не пересекается с "асимптотическим крестом", то характер сходимости во всех его точках одинаков.

Для любой точки на границе области сходимости можно указать единственный сходящийся прямоугольник, одной из вершин которого является выбранная точка, и который вписан в эту область. Он определяет два открытых взаимно сходящихся пучка, сумма углов которых постоянна и равна π.

Существует единственный сходящийся квадрат, вписанный в область сходимости, который из всех вписанных сходящихся прямоугольников имеет наибольшую площадь. Ему соответствуют два открытых взаимно сходящихся пучка $\{c: \frac{\varphi}{2} - \frac{\pi}{2} < \alpha < \frac{\varphi}{2}\}$ и $\{d: -\frac{\varphi}{2} < \beta < \frac{\pi}{2} - \frac{\varphi}{2}\}$, оптимальных в том смысле, что в них углы α и β равноправны. Оказывается, оптимальные сходящиеся пучки заключены между биссектрисами вертикальных углов, образованных осями абсцисс координатных систем $(xy)^{(1)}$ и $(xy)^{(2)}$ (рис. 9).

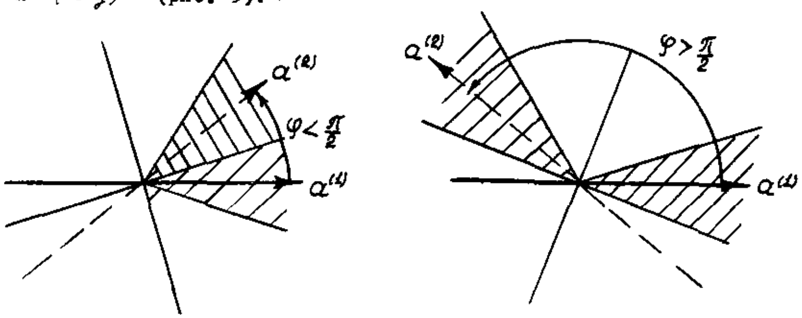

Рис. 9

Введем угол $\Phi = \min\{\frac{\varphi}{2}, \frac{\pi-\varphi}{2}\}$ и рассмотрим пучки направлений $\{c: -\Phi < \alpha < \Phi\}$ и $\{d: -\Phi < \beta < \Phi\}$. Каждый из них симметричен относительно своей оси абсцисс. Из всех симметричных пучков эти являются максимальными взаимно сходящимися. Сходимость в них носит однородный характер – "зигзаг". На рисунках 8 и 9 максимальные симметричные пучки и соответствующий им квадрат отмечены штриховкой.

Рассмотрим теперь более общий случай пересечения в начале координат (точка $\mathcal{P}$) двух точечных множеств, представимых однозначными функциями $y^{(1)} = f_1(x^{(1)})$ и $y^2 = f_2(x^{(2)})$.
Пусть выполнено следующее условие пересения: для произвольного, сколь угодно малого ε найдется δ такое, что любая

точка множества с абсциссой по модулю меньшей δ имеет ординату по модулю меньшую ε, то есть из $|x^{(i)}| < \delta$ следует $|y^{(i)}| < \varepsilon$ $(i=1,2)$. Для множества $\{\mathcal{M}_i\}$ рассмотрим совокупность направлений, задаваемых векторами с координатами $(x^{(i)} \cdot \operatorname{sign} x^{(i)}, f(x^{(i)}) \cdot \operatorname{sign} x^{(i)})$. Пучок направлений, являющийся выпуклой оболочкой указанной совокупности, назовем сопровождающим пучком множества. Процесс последовательных приближений, примененный к множествам $\{\mathcal{M}_1\}$ и $\{\mathcal{M}_2\}$, сходится, если сопровождающие пучки этих множеств – взаимно сходящиеся. Действительно, каждый переход из точки $(x_n^{(1)} y_n^{(1)})$ множества $\{\mathcal{M}_1\}$ в точку $(x_{n+1}^{(2)} y_{n+1}^{(2)})$ множества $\{\mathcal{M}_2\}$ вдоль вектора b_2 можно интерпретировать как переход с некоторой прямой первого сопровождающего пучка на некоторую прямую второго сопровождающего пучка. Относительно функции $\mathcal{F}(r) = (a^{(1)}, r)$ это есть переход с уровня $\mathcal{F}(r) = x_n^{(1)}$ на уровень $\mathcal{F}(r) = x_{n+1}^{(1)}$. Пусть коэффициент направленности для этого перехода есть q_n, тогда $x_{n+1}^{(1)} = x_n^{(1)} \cdot q_n$. В свою очередь, $x_n^{(1)} = x_{n-1}^{(1)} q_{n-1}$ и так далее вплоть до исходного уровня $x_1^{(1)}$ в первом шаге приближений. Используя эти рекурентные соотношения, выразим значения последующего уровня в n-м переходе через $x_1^{(1)}$ и произведение значений коэффициента направленности в каждом из переходов. Получим:

$$x_{n+1}^{(1)} = x_1^{(1)} \cdot \prod_{j=1}^{n} q_j .$$

Каждой паре направлений сопровождающих пучков соответствует определенное значение коэффициента направленности. Максимум модулей этих значений обозначим $\bar{q}$. Тогда имеем неравенство:

$$|x_{n+1}^{(1)}| = |x_1^{(1)} \cdot \prod_{j=1}^{n} q_j| < |x_1^{(1)}| \cdot (\bar{q})^n . \qquad (31)$$

Если сопровождающие пучки являются взаимно сходящимися, то $\bar{q} < 1$ и в силу неравенства (31) $\lim_{n\to\infty} |x_{n+1}^{(1)}| = 0$. По условию пересечения множеств $\{\mathcal{M}_1\}$ и $\{\mathcal{M}_2\}$ процесс последовательных приближений сходится к точке $\mathcal{P}$.

Достаточность доказана, однако сформулированное условие не является необходимым. Действительно, процесс последовательных приближений затрагивает в $\{\mathcal{U}_1\}$ и $\{\mathcal{U}_2\}$ лишь некоторые счетные подмножества, для которых рассматриваемое условие необходимо. Остальным точкам могут соответствовать любые сопровождающие направления. Достаточное условие можно конкретизировать , если рассмотреть более узкий класс пересекающихся множеств. Пусть в точке $\mathcal{P}$ пересекаются непрерывные гладкие кривые, т.е. функции $f_1(x^{(1)})$ и $f_2(x^{(2)})$ имеют первые производные $f_1'(x^{(1)})$ и $f_2'(x^{(2)})$.

Рассмотрим пучок направлений параллельных касательным во всех точках кривой. Очевидно, сопровождающий пучок содержится в пучке касательных направлений, поэтому если последние для обеих кривых являются взаимно сходящимися, то процесс последовательных приближений сходится. Это имеет место,когда пучки касательных направлений содержатся в максимальных симметричных, чему соответствуют неравенства:

$$|f_1'(x^{(1)})| < tg\,\Phi\,; \quad |f_2'(x^{(2)})| < tg\,\Phi .$$

При решении прикладных геометрических задач приходится искать пересечение двух И-кривых. Пусть пересекаются i -й и j -й участки кривых. Используя оценку модуля производной (23), можно записать достаточное условие сходимости в виде двух неравенств

$$T_i < \Phi\;;\quad T_j = \Phi\;. \tag{32}$$

Из (32) ясно, что наивыгодным в смысле сходимости последовательных приближений является тот случай, когда хорды пересекающихся участков взаимно ортогональны. Если при данной системе узлов процесс не сходится, а кривые пересекаются без касания, то, уменьшая длины участков добавлением узлов, можно добиться сходимости.

В принципе, всякий расходящийся процесс можно превратить в сходящийся, изменив направление переходов на противоположное, но это сопряжено с решением трансцендентного уравнения на каждом переходе. Практически окружностная интерполяция осуществляется в ЦВМ стандартным блоком. По координатам шести последовательных точек блок вычисляет параметры интер-

поляции среднего участка, причем допускаются изломы прямых. Угловые точки на криволинейных участках должны быть отмечены неким логическим признаком.

Окружностная интерполяция имеет следующее применение:

I) Управление механизмами,осуществляющими раскрой (или разметку) плоских заготовок.

2) Геометрические расчеты некоторых элементов (например, балок с кручением) пространственных конструкций.

3) Вывод из ЦВМ графической информации на внешние чертежные устройства. В настоящее время создана программа, расчитывающая перспективное изображение кусочно-гладкой замкнутой поверхности.

4) Численное решение некоторых задач математической физики. Имеются в виду первая краевая задача для двумерного уравнения Лапласа в неодносвязной области с криволинейной границей и задача о распределении зарядов на электродах сложной формы в осесимметричной системе.

Поступила в редакцию
20/I - 1970 г.