



ГОСУДАРСТВЕННЫЙ НАУЧНЫЙ ЦЕНТР РОССИЙСКОЙ ФЕДЕРАЦИИ
ИНСТИТУТ ФИЗИКИ ВЫСОКИХ ЭНЕРГИЙ

ИФВЭ 98-9
ОАПиЭС

А. И. Курносенко

**ИНТЕРПОЛЯЦИОННЫЕ СВОЙСТВА
ПЛОСКИХ СПИРАЛЬНЫХ КРИВЫХ**

Направлено в журнал
“Фундаментальная и прикладная математика”

Протвино 1998

Аннотация

Курносенко А.И. Интерполяционные свойства плоских спиральных кривых: Препринт ИФВЭ 98-9. – Протвино, 1998. – 24 с., 11+10 рис., библиогр.: 5.

В работе получены неравенства, описывающие поведение кривых с монотонной функцией кривизны (названных спиральными). Для спирали, представленной последовательностью узлов интерполяции, строится область, внутри которой гарантированно находилась исходная кривая. Ширина области дает оценку детерминированности кривой данным дискретным представлением, безотносительно к какому-либо алгоритму интерполяции. Обсуждается возможность получения такой оценки для произвольной гладкой кривой, исходя из условия “достаточно редкого” расположения вершин.

Источник задачи и ее приложения — практика допускового контроля криволинейных профилей в машиностроении.

Abstract

Kurnosenko A.I. Interpolational Properties of Planar Spiral Curves: IHEP Preprint 98-9. – Protvino, 1998. – p. 24, figs. 11+10, refs.: 5.

Some inequalities on the planar curves termed spirals (due to monotonous curvature function) are considered. For a spiral represented as a set of interpolation nodes a region covering the parent curve is constructed. The width of the region provides an estimate of curve determinism by the given discrete representation, this estimate being independent of any interpolation method. A similar problem is set up for any smooth curve, assuming that vertices are distanced “well away” from one another.

The problem originates from and can be applied to practice of tolerance control in industry.

E-mail: kurnosenko@mx.ihep.su

1. Происхождение и постановка задачи

Настоящей работе предшествовала статья [1], посвященная общим вопросам конструирования и производственной метрологии. Одна из рассмотренных в ней проблем — ошибки в чертежах, которые проявляются как неполная детерминированность геометрии при описании конструкций. Чертеж включает определение идеальной поверхности детали и допустимых отклонений. Если не считать грубых просчетов, неполная определенность геометрии объекта проявляется обычно именно в описании полей допуска. Основные инструменты эксперта-метролога при анализе чертежа — элементарная геометрия и теория ошибок.

Однако при появлении в чертеже *криволинейных профилей*, заданных дискретно-точечным представлением, недетерминированность проявляется уже на уровне описания номинального профиля: она изначально присуща такому заданию кривой. При этом традиционные методы анализа качества чертежа перестают работать, а другие — отсутствуют.

Допуская дискретное представление кривой, теория и практика конструирования оставляют открытым вопрос о его корректности: с какой точностью может быть восстановлена исходная кривая? Естественно, такая постановка вопроса подразумевает определенную дисциплину расстановки точек на кривой. Некоторые алгоритмы интерполяции включают методику разбиения кривой, гарантирующую ее последующее восстановление *данным алгоритмом* с наперед заданной точностью [2].

Мы намерены предложить оценку детерминированности и дисциплину разбиения кривой безотносительно к какому-либо конкретному методу интерполяции, основываясь лишь на характерных свойствах самой кривой, полностью представленной ее натуральным уравнением $k(s)$. Обращение к нему вполне обосновано, поскольку криволинейный элемент есть решение некой задачи, реализующей замысел конструктора.

С этих позиций восстановление кривой мы рассматриваем как попытку “разгадать” характер зависимости $k(s)$, а под дисциплиной расстановки узлов интерполяции подразумеваем предоставление такой возможности в процессе разбиения кривой. Эффективность такого подхода мы продемонстрируем на примере весьма узкого класса кривых, а именно спиралей.

Интерполяционное свойство спиралей. Понятие “спирали” обычно нуждается в дополнительном определении — гиперболическая, архимедова и пр. Однако у большинства кривых, называемых спиралью, есть общее свойство, которое можно возвести в ранг определяющего: это *монотонное изменение кривизны вдоль кривой*.

Спиральные кривые широко используются в различных механизмах, в частности в кулачках. При проектировании альпинистских карабинов эксплуатируется свойство логарифмической спирали сохранять постоянный угол между полярным лучом и нормалью к кривой.

Под дискретно-точечным заданием плоской C^1 -гладкой кривой будем понимать конечный упорядоченный набор n точек $\{P_j=(x_j, y_j)_{j=1}^n\}$ и граничные условия одного из двух типов:

- 1) углы τ_1 и τ_n наклонов касательных в крайних узлах, $n \geq 2$;
- 2) замкнутость кривой; выбор начальной точки в этом случае произволен, $n \geq 3$.

Наиболее общая дисциплина разбиения кривой обычно формулируется в виде ограничений кривизны или поворота в пределах хорды, что дает возможность считать ее функцией в локальной системе координат хорды. Это условие мы будем предполагать всегда выполненным.

Пусть $\{P_j; \tau_1, \tau_n\}$ — дискретно-точечное представление спиральной кривой с граничными условиями первого типа (замкнутость спиральных кривых невозможна). В каждом внутреннем j -том узле построим дугу окружности $K_j(P_{j-1}, P_j, P_{j+1})$, соединяющую его с соседними узлами. Граничные углы позволяют построить дуги K_1 и K_n в крайних узлах по двум точкам. После этого на каждую j -тую хорду, $j \leq n-1$, опирается пара дуг K_j и K_{j+1} с кривизнами q_j и q_{j+1} . Они образуют “линзу” — область, заключенную между этими дугами.

Такое построение (как и сам термин “линза”) используется при окружностной интерполяции плоской кривой ([2, стр. 102], [3]). Его иллюстрирует рисунок 1¹. Мы обращаем внимание на замечательные свойства этого построения в случае спиральных кривых, доказательство которых приведем в разделе 5:

- *исходная кривая лежит внутри области, ограниченной системой “линз” и построенной только по заданному дискретному представлению кривой;*
- *последовательность расчетных кривизн $\{q_j, j=1, n\}$ наследует характер монотонности кривизны исходной кривой.*

Экстремальное расстояние от дуги кривизны q до хорды длины $2c$ равно

$$h(q, c) = \frac{-q c^2}{\sqrt{1-(q c)^2} + 1} \simeq \frac{-q c^2}{2} \left(1 + \frac{q^2 c^2}{4}\right).$$

Локальная ширина δ_j каждой линзы с хордой длины $2c_j$ и глобальная оценка детерминированности кривой Δ имеют вид

$$\delta_j = |h(q_j, c_j) - h(q_{j+1}, c_j)| \simeq \frac{c_j^2}{2} |q_j - q_{j+1}|, \quad \Delta = \max(\delta_j, j=1, n-1). \quad (1)$$

¹Римскими цифрами пронумерованы цветные рисунки, помещенные на стр. 12–13.

Высокая эффективность оценки связана с тем, что при удвоении числа узлов значения δ_j уменьшаются примерно в 8 раз (каждый сомножитель уменьшается вдвое). На рисунке I показан эффект, вызванный увеличением n с 10 до 13.

Общая постановка задачи. Было бы заманчиво получить оценки типа (1) для произвольной гладкой кривой. Ее можно считать кусочно-спиральной, т.е. состоящей из участков с различным характером монотонности кривизны. Вершины кривой (точки экстремальной кривизны) служат границами участков.

Предположим, что точки P_i и граничные углы τ_1, τ_n допускают построение кривой постоянной кривизны (рис. 1). Очевидно, интерполяционное свойство даст в этом случае область нулевой ширины, т.е. однозначно определит интерполирующую кривую как прямую или окружность. Если она не совпадет с исходной кривой, значит, при разбиении информация о вершинах была потеряна. Из этих соображений следует, что каждый участок монотонной кривизны должен быть представлен как минимум двумя узлами. Это исключает наличие двух и более вершин в пределах хорды и обосновывает выбор классов исследуемых функций, описанных в следующем разделе.



Рис. 1.

Наиболее общая формулировка задачи представляется в следующем виде: *среди всех C^1 -гладких кривых, допускающих заданное дискретно-точечное представление, выделить подмножество кривых, имеющих наименьшее количество вершин; оценить ширину покрываемой ими области.*

2. Обозначения и определения

Мы намерены исследовать участки C^1 -гладкой кривой в пределах хорды и изучить связь между характером монотонности кривизны и параметрами кривой в узлах интерполяции. Кривую при этом будем рассматривать как функцию $y=f(x)$ в локальной системе координат, в которой хорда длины $2c$ является отрезком $[-c, c]$ оси абсцисс.

Исследование удобно свести к отрезку $[-1, 1]$, понимая при этом под координатами безразмерные отношения x/c и y/c , а под кривизной — *нормированную кривизну* $k_{norm} = kc$, также безразмерную. Поскольку речь почти всегда будет идти именно о нормированной кривизне, мы будем опускать ее специальное обозначение. Кривизна дуги окружности, выходящей под углом α и опирающейся на хорду $[-1, 1]$, равна $-\sin \alpha$. Уравнение такой дуги, обычно записываемое с радикалами в числителе, путем перенесения иррациональности в знаменатель приводится к виду, который “не боится” нулевой кривизны:

$$\mathbf{A}(x; \alpha) := \frac{\sqrt{1-x^2} \sin^2 \alpha - \cos \alpha}{\sin \alpha} = \frac{\sin \alpha (1-x^2)}{\cos \alpha + \sqrt{1-x^2} \sin^2 \alpha}, \quad \begin{aligned} \cos \alpha &= \sqrt{1-k^2}, \\ \sin \alpha &= -k. \end{aligned} \quad (2)$$

Классы исследуемых функций. К классу \mathcal{M} исследуемых функций отнесем все функции $y=f(x)$, обладающие следующими свойствами:

- функция задана на отрезке $x \in [-1, 1]$, непрерывна и обращается в нуль на его концах: $f(\pm 1) = 0$;
- производная функции всюду на отрезке непрерывна и ограничена; следовательно, существует функция $\tau(x) := \operatorname{arctg} y'(x)$ — угол наклона касательной к графику $y=f(x)$;
- функция кривизны $k(x) = y''/(1 + y'^2)^{3/2}$ нестрого монотонна на $[-1, 1]$; она может иметь конечные разрывы и обращаться в бесконечность, оставаясь интегрируемой на $[-1, 1]$.

С помощью верхнего индекса будем уточнять тип монотонности кривизны (возрастание \mathcal{M}^+ или убывание \mathcal{M}^-). К классу \mathcal{M}^0 отнесем функции семейства (2).

Класс функций \mathcal{V} определим, заменив условие монотонности требованием наличия *единственной* вершины — максимума (\mathcal{V}^+) или минимума (\mathcal{V}^-) кривизны. Помимо традиционной трактовки, подразумевающей экстремум кривизны в точке, мы допускаем *протяженную вершину*, т.е. отрезок $[x_1, x_2] \subset (-1, 1)$, в пределах которого кривизна постоянна, а слева и справа характер ее монотонности противоположен (рис. 2). Абсциссой вершины будем считать любое значение $x \in [x_1, x_2]$.

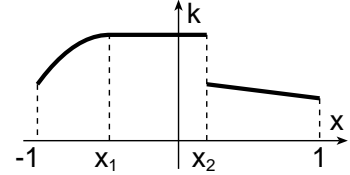


Рис. 2.

В роли “посредника” между натуральным уравнением кривой и ее координатным представлением обычно выступает функция $\tau(s)$ — угол наклона касательной к ориентированной кривой: $\tau'(s) = k(s)$. Аналогичную роль отводим функции $z(x)$:

$$z(x) := \sin \tau(x) = \frac{y'}{\sqrt{1+y'^2}} \quad \Rightarrow \quad \underline{z'(x) = k(x)}. \quad (3)$$

Последнее равенство справедливо и в точках разрыва кривизны, где $z(x)$ имеет изломы с односторонними производными. По известной функции $z(x)$ функция $y=f(x)$ получается интегрированием

$$f(x) = \int_{-1}^x \operatorname{tg} \tau(u) du = \int_{-1}^x \frac{z(u)}{\sqrt{1-z(u)^2}} du. \quad (4)$$

Чтобы на правом конце отрезка выполнялось условие $f(1) = 0$, функция $z(x)$ должна обращать в нуль функционал

$$\mathcal{Y}(z) := \int_{-1}^1 \frac{z(u)}{\sqrt{1-z(u)^2}} du = 0. \quad (5)$$

Из монотонности функции $t \mapsto t/\sqrt{1-t^2}$ следует неравенство

$$\mathcal{Y}(u) > \mathcal{Y}(v) \quad \text{при} \quad u(x) \not\geq v(x), \quad (6)$$

где символ “ $\not\geq$ ” обозначает отношение, исключающее тождественное равенство функций в области определения, т.е. $\{u(x) \geq v(x) \cap u(x) \neq v(x)\}$.

При рассмотрении кривой $y=f(x)$ на хорде $[-1, 1]$ под граничными условиями помимо $f(\pm 1) = 0$ мы будем понимать углы наклона касательных на концах отрезка $\alpha = \tau(-1)$ и $\beta = \tau(1)$ ($|\tau(x)| < \pi/2$) и кривизны $k_1 = k(-1)$ и $k_2 = k(1)$. Для подклассов функций с заданными граничными условиями будем использовать обозначения $\mathcal{M}(\alpha, \beta)$ или $\mathcal{M}(\alpha, \beta, k_1, k_2)$.

Определение линзы. Пусть задан отрезок AB длины $|AB| = 2c \neq 0$. Построим две дуги окружностей, соединяющие концы отрезка и выходящие из точки A под углами ξ и η ($|\xi|, |\eta| < \pi$) к направлению \vec{AB} . Область, ограниченную этими дугами, назовем *линзой* и обозначим $\mathcal{L}(\xi, \eta; AB)$, опуская обозначение отрезка, если оно ясно из контекста. Кривизны дуг связаны с углами соотношением $k_\xi c = -\sin \xi$.

Дугу, выходящую под углом $(\xi + \eta)/2$, назовем *биссектрисой* линзы. Обозначением $\mathcal{L}^*(\xi, \eta)$ мы будем подчеркивать упорядоченность параметров: $\xi \leq \eta$. Очевидно,

$$\mathcal{L}^*(\xi_2, \eta_2) \subseteq \mathcal{L}^*(\xi_1, \eta_1) \iff \xi_1 \leq \xi_2 \leq \eta_2 \leq \eta_1. \quad (7)$$

С граничными углами α и β мы будем связывать линзу $\mathcal{L}(\alpha, -\beta)$ (рис. 3). Ее угловая полуширина ω (со знаком) и угол выхода биссектрисы γ равны

$$\omega = \frac{\alpha + \beta}{2}, \quad \gamma = \frac{\alpha - \beta}{2}. \quad (8)$$

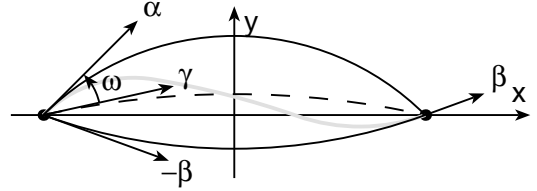


Рис. 3.

3. Задача о сопряжении двух окружностей

Построим и исследуем функции $y=f(x)$, графики которых составлены из двух круговых дуг. Это построение используется в задачах интерполяции как построение кривой, но без стремления построить функцию на хорде [4]. Если такая функция существует, то она принадлежит к классу $\mathcal{M}^\pm(\alpha, \beta, k_1, k_2)$.

Лемма 1. Если функция $y=f(x)$ принадлежит к классу $\mathcal{M}(\alpha, \beta, k_1, k_2)$, то для граничных углов α, β , угла $\omega = (\alpha + \beta)/2$ и кривизн k_1, k_2 справедливы соотношения

$$\begin{aligned} \mathcal{M}^+ : & \quad \omega > 0 \quad \text{и} \quad k_1 < -\sin \alpha < \sin \beta < k_2; \\ \mathcal{M}^- : & \quad \omega < 0 \quad \text{и} \quad k_1 > -\sin \alpha > \sin \beta > k_2; \\ \mathcal{M}^0 : & \quad \omega = 0 \quad \text{и} \quad k_1 = -\sin \alpha = \sin \beta = k_2. \end{aligned} \quad (9)$$

Доказательство. Утверждение для класса \mathcal{M}^0 очевидно. Пусть кривая $y = f(x) \in \mathcal{M}^+(\alpha, \beta, k_1, k_2)$. Обратимся к функции $z(x)$. Согласно ее определению и свойству (3), ее график соединяет точки $A(-1, \sin \alpha)$ и $B(1, \sin \beta)$ (рис. 4). Кроме того, это функция с возрастающей производной, поэтому график представляет собой выпуклую вниз кривую, лежащую ниже хорды AB . Введем функции, представленные прямой AB

$$l(x) = k_0 x + \frac{\sin \alpha + \sin \beta}{2}, \quad k_0 = \frac{\sin \beta - \sin \alpha}{2} \quad (10)$$

и проходящими через начало координат прямыми AA_1 и B_1B :

$$AA_1: z_1(x) = -x \sin \alpha, \quad B_1B: z_2(x) = x \sin \beta. \quad (11)$$

Из неравенства $z(x) < l(x)$ следует $\mathcal{Y}(z) < \mathcal{Y}(l)$ (6). Но $\mathcal{Y}(z) = 0$ (5), а

$$\mathcal{Y}(l) = \int_{-1}^1 \frac{l(u)}{\sqrt{1-l(u)^2}} du = 2 \operatorname{tg} \frac{\alpha+\beta}{2} = 2 \operatorname{tg} \omega.$$

Отсюда и получаем неравенство $\omega > 0$ при $y = f(x) \in \mathcal{M}^+$. Из него следует среднее звено цепочки неравенств (9): $-\sin \alpha < \sin \beta$.

Согласно (3), наклоны левой AP правой PB касательных к графику $z(x)$ равны k_1 и k_2 . Докажем неравенство $k_1 < -\sin \alpha$. Его нарушение означало бы расположение касательной AP к графику и, следовательно, всего графика не ниже отрезка AA_1 : $z(x) \not\geq z_1(x)$. Тождественное равенство исключено, так как $-\sin \alpha = z_1(1) < z(1) = \sin \beta$. Но тогда $\mathcal{Y}(z) > \mathcal{Y}(z_1)$ (6). Очевидно, $\mathcal{Y}(z_1) = 0$, и мы получаем $\mathcal{Y}(z) > 0$, что противоречит граничным условиям (5) и доказывает левую часть цепочки (9). Ее правая часть доказывается аналогично: касательная PB и график $z(x)$ не могут находиться выше отрезка B_1B .

Если кривая $y = f(x) \in \mathcal{M}^-(\alpha, \beta, k_1, k_2)$, то $y_1 = -f(x) \in \mathcal{M}^+(-\alpha, -\beta, -k_1, -k_2)$ и, по доказанному, $-k_1 < \sin \alpha < -\sin \beta < -k_2$, что дает неравенства для класса \mathcal{M}^- . ♦

Следствие 1. Зафиксируем углы α, β , понизив тем самым размерность пространства граничных условий $(\alpha, \beta, k_1, k_2)$ с четырех до двух. Пусть угол $\omega \neq 0$. Тогда в плоскости (k_1, k_2) можно указать одну и только одну из областей \mathcal{K}^+ или \mathcal{K}^- , допускающую принадлежность кривой к одному из двух классов \mathcal{M}^\pm :

$$\begin{aligned} \mathcal{K}^+ &:= \{ (k_1, k_2) : -\infty \leq k_1 < -\sin \alpha \cap \sin \beta < k_2 \leq +\infty, \quad \omega > 0 \}, \\ \mathcal{K}^- &:= \{ (k_1, k_2) : +\infty \geq k_1 > -\sin \alpha \cap \sin \beta > k_2 \geq -\infty, \quad \omega < 0 \}. \end{aligned} \quad (12)$$

Если $\omega > 0$, то область \mathcal{K}^+ существует и расположена выше прямой $k_1 = k_2$; в противном случае имеется область \mathcal{K}^- , расположенная ниже этой прямой. ♦

Первый фрагмент рисунка II иллюстрирует область \mathcal{K}^+ . На плоскости (k_1, k_2) , разрезанной прямыми $k_1 = -\sin \alpha$ и $k_2 = \sin \beta$ на 4 квадранта, она представлена левым верхним квадрантом. Положение его вершины $C = (-\sin \alpha, \sin \beta)$ определяется зафиксированными значениями углов².

Следствие 2. Кривой $y = f(x) \in \mathcal{M}^\pm(\alpha, \beta, k_1, k_2)$ соответствует выпуклая функция $z(x)$. Она вписана в ломаную APB в том смысле, что два звена ломаной являются

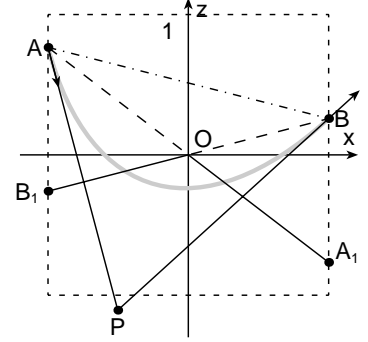


Рис. 4.

²Для иллюстрации убывания понадобился бы другой рисунок с точкой C , лежащей ниже прямой $k_1 = k_2$. На нем был бы выделен правый нижний квадрант. В случае постоянства кривизны области (12) вырождаются в единственную точку C на прямой $k_1 = k_2$ (рис. 10 на стр. 22).

опорными прямыми графика $z(x)$. Поэтому справедливы неравенства

$$\begin{aligned} \mathcal{M}^+ : \quad z(x) &\geq V(x; P), \\ \mathcal{M}^- : \quad z(x) &\leq V(x; P), \end{aligned} \quad V(x; P) := \begin{cases} k_1(x+1) + \sin \alpha, & -1 \leq x \leq x_0, \\ k_2(x-1) + \sin \beta, & x_0 \leq x \leq 1, \end{cases} \quad (13)$$

в которых ломаная APB рассматривается как график некой функции $V(x; P)$, а координаты точки P пересечения касательных равны

$$x_0(k_1, k_2) = \frac{k_1 + k_2 + \sin \alpha - \sin \beta}{k_2 - k_1}, \quad z_0(k_1, k_2) = \frac{2k_1 k_2 + k_2 \sin \alpha - k_1 \sin \beta}{k_2 - k_1}. \quad (14)$$

Следствие 3. Каждой точке $K(k_1, k_2) \in \mathcal{K}^\pm$ соответствует одна и только одна точка $P(x_0, z_0)$. Она расположена в полосе $|x| \leq 1$ строго ниже (\mathcal{M}^+) или выше (\mathcal{M}^-) ломаной B_1OA_1 , т.е. в одной из областей

$$\mathcal{P}^\pm := \{(x, z) : |x| \leq 1 \cap z \leq -x \sin \alpha \cap z \leq x \sin \beta\}. \quad (15)$$

Равенство $|x| = 1$ соответствует бесконечному значению одной из кривизн.

Обратно, каждая точка $P(x, z) \in \mathcal{P}$ однозначно определяет точку $K(k_1, k_2) \in \mathcal{K}$:

$$k_1(x_0, z_0) = \frac{z_0 - \sin \alpha}{x_0 + 1}, \quad k_2(x_0, z_0) = \frac{z_0 - \sin \beta}{x_0 - 1}. \quad (16)$$

Во взаимной однозначности отображений $\mathcal{K} \rightarrow \mathcal{P}$ (14) и $\mathcal{P} \rightarrow \mathcal{K}$ (16) можно убедиться, рассмотрев их якобианы:

$$J_1(k_1, k_2) = 4 \frac{(k_1 - k_0)(k_2 - k_0)}{(k_2 - k_1)^3}, \quad J_2(x_0, z_0) = 2 \frac{z_0 - l(x_0)}{(x_0^2 - 1)^2},$$

где k_0 и $l(x)$ определены в (10). Первый из них не равен нулю в \mathcal{K} , что становится очевидным, если в неравенства (9) ввести k_0 — середину отрезка $[-\sin \alpha, \sin \beta]$:

$$\omega \geq 0 : \quad k_1 \leq -\sin \alpha \leq k_0 \leq \sin \beta \leq k_2. \quad (17)$$

На рисунке II этот факт иллюстрируют точка $N(k_0, k_0)$ и прямые $k_1 = k_0$, $k_2 = k_0$. Якобиан J_2 обращается в нуль на прямой AB — вне области \mathcal{P} . \blacklozenge

Фиксация углов на правом фрагменте рисунка II, иллюстрирующем область \mathcal{P}^+ , отражена положением точек $A(-1, \sin \alpha)$ и $B(1, \sin \beta)$. Признаком возрастания кривизны служит расположение начала координат O ниже отрезка AB (10): $l(0) = (\sin \alpha + \sin \beta)/2 > 0$. Точка O является образом точки C , а исходящие из C лучи — границы области \mathcal{K} — преобразуются в конечные отрезки OB_1 и OA_1 , принадлежащие границе области \mathcal{P} . Лучи, ограничивающие область \mathcal{P} по боковым сторонам, являются образами бесконечно удаленных точек области \mathcal{K} .

В неравенствах (13) интересен случай равенства, т.е. совпадения $z(x)$ со своей “опорной ломаной” $V(x; P)$. Условие (5) принимает вид

$$\left(\frac{\cos \alpha}{k_1} - \frac{\cos \beta}{k_2} \right) - \left(\frac{1}{k_1} - \frac{1}{k_2} \right) \sqrt{1 - z_0^2} = 0, \quad (18)$$

где z_0 определяется в (14). Поскольку функция $z(x)$ кусочно-линейна, кривизна $k(x)$ кусочно-постоянна, и график $y=f(x)$ состоит из двух дуг окружностей. Первая дуга имеет кривизну k_1 , выходит из точки $(-1, 0)$ под углом α , в точке (x_0, y_0) при

наклоне τ_0 гладко переходит во вторую с кривизной $k_2 \neq k_1$, достигающую точки $(1, 0)$ при наклоне β .

При $k_{1,2} \rightarrow 0$ левая часть (18) имеет конечные пределы,³ что позволяет рассматривать прямую как частный случай круговой дуги, не требующий особых оговорок. Освобождение в (18) от радикала дает уравнение связи

$$\begin{aligned} Q(\alpha, \beta, k_1, k_2) &:= k_1 k_2 + k_2 \sin \alpha - k_1 \sin \beta + \sin^2 \gamma = \\ &= (k_1 + \sin \alpha)(k_2 - \sin \beta) + \sin^2 \omega = 0 \end{aligned} \quad (19)$$

(углы γ и ω определены в (8)). В плоскости (k_1, k_2) это гипербола с центром в точке C (рисунок II). Ее асимптоты совпадают с границами области \mathcal{K} . Только одна ветвь лежит в \mathcal{K}^\pm , и именно на ней переход от (18) к (19) является эквивалентным преобразованием без риска получить “лишние” решения⁴.

Выбор кривизны на одном конце отрезка определяет ее значение на другом конце. Оставляя обозначение (k_1, k_2) для несвязанных между собой граничных условий, введем пары (k_1, g_2) и (g_1, k_2) , определяемые уравнениями

$$Q(\alpha, \beta, k_1, g_2) = 0, \quad Q(\alpha, \beta, g_1, k_2) = 0 \quad (20)$$

и допускающие искомое сопряжение. На рисунке II это выглядит как проектирование произвольной точки $K(k_1, k_2) \in \mathcal{K}^\pm$ в точки гиперболы $K_1(k_1, g_2)$ и $K_2(g_1, k_2)$. Кривизны g_2 и g_1 назовем *дополнительными* к заданным кривизнам k_1 и k_2 .

Подстановка (16) в уравнение гиперболы дает ее образ в плоскости (x, z) :

$$\tilde{Q}(x, z) := \frac{E(x, z)}{x^2 - 1} = 0, \quad \text{где} \quad E(x, z) := x^2 \sin^2 \gamma + z^2 + 2xz \sin \gamma \cos \omega - \sin^2 \omega. \quad (21)$$

Кривая $E(x, z) = 0$ — эллипс с центром в начале координат (рисунок II). Он касается горизонталей $z = \pm 1$ и проходит через точки A, B и им симметричные $A_1 = (1, -\sin \alpha)$ и $B_1 = (-1, -\sin \beta)$. Ветвь гиперболы, принадлежащая области \mathcal{K} , отображается в дугу $B_1 A_1$ эллипса, заключенную в области \mathcal{P} :

$$Z(x) = -\sin \omega \sqrt{1 - x^2 \sin^2 \gamma} - x \sin \gamma \cos \omega = -\sin[\omega + \arcsin(x \sin \gamma)], \quad |x| < 1. \quad (22)$$

Образами точек $K_1(k_1, g_2)$ и $K_2(g_1, k_2)$ являются точки $P_1(x_1, z_1)$ и $P_2(x_2, z_2)$. Наклоны звеньев ломаной $V(x; P_1)$ равны k_1 и g_2 , ломаной $V(x; P_2)$ — g_1 и k_2 .

Рисунок III иллюстрирует полученные сопряжения. При фиксированных углах α и β построенные кривые являются членами однопараметрического семейства. В

³Например, слагаемое, критичное к $k_1 \rightarrow 0$, преобразуется так (при этом $\sqrt{1 - z_0^2} \rightarrow \cos \alpha$):

$$\frac{\cos \alpha - \sqrt{1 - z_0^2}}{k_1} = \frac{1}{k_1} \cdot \frac{\cos^2 \alpha - (1 - z_0^2)}{\cos \alpha + \sqrt{1 - z_0^2}} = \frac{(2k_2 + \sin \alpha - \sin \beta)(2k_1 k_2 + 2k_2 \sin \alpha - k_1 \sin \alpha - k_1 \sin \beta)}{(k_2 - k_1)^2 (\cos \alpha + \sqrt{1 - z_0^2})}.$$

⁴Чтобы убедиться в этом, представим (18) в виде

$$\frac{(-k_1 - \sin \alpha) \cos \beta + (k_2 - \sin \beta) \cos \alpha + \sin 2\omega}{k_1 k_2} = \frac{(k_2 - k_1) \sqrt{1 - z_0^2}}{k_1 k_2}.$$

Величины в скобках, согласно (9), *одновременно* с ω положительны или отрицательны. Достаточно того, что левая и правая части имеют одинаковые знаки.

качестве параметра семейства может выступать одна из граничных кривизн или координата x_0 точки сопряжения. Для членов семейства введем обозначения

$$\mathbf{B}_1(x; \alpha, \beta, k_1) \equiv \mathbf{B}_2(x; \alpha, \beta, g_2) \equiv \mathbf{B}_0(x; \alpha, \beta, x_0), \quad (23)$$

отражающие способ параметризации. Параметры связаны соотношениями (20) и

$$\begin{aligned} k_1 &= \frac{\sin \omega}{1 + x_0} (x_0 \cos \gamma - r) - \sin \alpha, \\ g_2 &= \frac{\sin \omega}{1 - x_0} (x_0 \cos \gamma + r) + \sin \beta, \end{aligned} \quad r = +\sqrt{1 - x_0^2 \sin^2 \gamma}, \quad |x_0| < 1.$$

“Лишние” решения (пунктир на рисунке III) получаются заменой знака перед радикалом и/или при $|x_0| > 1$. Точки сопряжения лежат на биссектрисе линзы $\mathcal{L}(\alpha, -\beta)$, направление τ_0 в точке сопряжения составляет с биссектрисой угол ω .

Заметим, что при $k_1 \rightarrow -\infty$ точка P стремится к B_1 по дуге эллипса, ломаная $V(x; P)$ вырождается в отрезок B_1B (11), порождая при интегрировании (4) функцию $\mathbf{A}(x; -\beta)$. При $\omega \rightarrow 0$ в один отрезок сливаются дуга эллипса B_1A_1 , ломаная APB , отрезки AA_1 и B_1B . Функции семейства \mathbf{B} вырождаются в при этом в функцию $\mathbf{A}(x; \gamma)$. Это позволяет дать следующие определения:

$$\begin{aligned} \mathcal{M}^{\pm, 0} : \quad \mathbf{B}_1(x; \alpha, \beta, \mp \infty) &= \mathbf{B}_0(x; \alpha, \beta, -1) = \mathbf{A}(x; -\beta), \\ \mathbf{B}_2(x; \alpha, \beta, \pm \infty) &= \mathbf{B}_0(x; \alpha, \beta, +1) = \mathbf{A}(x; \alpha), \\ \mathbf{B}_i(x; \gamma, -\gamma, *) &= \mathbf{A}(x; \gamma), \quad i = 0, 1, 2. \end{aligned} \quad (24)$$

Рисунок IV иллюстрирует семейство сопряженных дуг с фиксированными кривизнами $k_1 = -\sin 60^\circ$ и $k_2 = \sin 30^\circ$ и свободным параметром α или β . В пределе при $\alpha \rightarrow -\arcsin k_1$ или $\beta \rightarrow \arcsin k_2$ кривые этого семейства также переходят в окружности.

4. Локальные свойства спиральных кривых

В этом разделе мы изучим свойства функций классов \mathcal{M}^\pm . Локальные свойства представляются собой неравенства для граничных условий и координатных представлений кривой в пределах хорды. На их основе затем будут доказаны глобальные свойства, или, заимствуя известный термин, свойства “в целом”.

Теорема 1. *Граничные условия кривой класса $\mathcal{M}^\pm(\alpha, \beta; k_1, k_2)$ удовлетворяют системе неравенств*

$$Q(\alpha, \beta, k_1, k_2) \leq 0 \quad \text{и} \quad \begin{cases} \omega > 0, & \text{если кривизна возрастает } (k_1 < k_2); \\ \omega < 0, & \text{если кривизна убывает } (k_1 > k_2). \end{cases} \quad (25)$$

Доказательство. Рассмотрим случай возрастающей кривизны. Неравенство $\omega > 0$ доказано в лемме 1. Утверждение теоремы усиливает лемму, так как сужает область значений $(k_1, k_2) \in \mathcal{K}^+$, необходимых для принадлежности кривой к классу \mathcal{M}^+ , до выпуклой подобласти

$$\mathcal{H}^+ := \{ (k_1, k_2) : Q(k_1, k_2) \leq 0, \quad k_1 < k_2, \quad \omega > 0 \}.$$

Она ограничена той ветвью гиперболы (19), которая лежит в \mathcal{K}^+ (рисунок II). Образом области \mathcal{H}^+ в (x, z) является, согласно (21), подобласть $\mathcal{E}^+ \subset \mathcal{P}^+$, в которой выполняется неравенство $\tilde{Q}(x, z) \leq 0$. В \mathcal{P}^+ оно принимает вид $E(x, z) \geq 0$ и означает исключение из \mathcal{P}^+ внутренней области эллипса:

$$\mathcal{E}^+ := \{(x, z) : |x| \leq 1, z \leq Z(x), (x, z) \neq A_1, (x, z) \neq B_1\}.$$

Область \mathcal{E}^+ представляет собой часть полосы $|x| \leq 1$, лежащую ниже дуги B_1A_1 , и саму эту дугу с выколотыми концами (точки A_1 и B_1 не принадлежали \mathcal{P}). Таким образом, утверждение теоремы требует исключить расположение точки P пересечения касательных внутри выпуклой области OB_1A_1O (рис. 5).

Предположим, что точка P лежит внутри области OB_1A_1O . Согласно (13), $z(x) \geq V(x; P)$. Спроектируем P в точку P_1 , лежащую на дуге B_1A_1 . Тогда $\mathcal{Y}(V(x; P_1)) = 0$ и $V(x; P) > V(x; P_1)$, что влечет неравенство $\mathcal{Y}(z) > 0$ и означает нарушение граничных условий на правом конце отрезка: $f(1) \neq 0$. Таким образом, точка P может лежать либо на дуге B_1A_1 , либо ниже ее, откуда и следует доказываемое неравенство $Q \leq 0$.

Если кривая $y = f(x) \in \mathcal{M}^-(\alpha, \beta, k_1, k_2)$, то $y_1 = -f(x) \in \mathcal{M}^+(-\alpha, -\beta, -k_1, -k_2)$. Согласно определению (19), $Q(-\alpha, -\beta, -k_1, -k_2) = Q(\alpha, \beta, k_1, k_2)$. Следовательно, это преобразование изменяет знак угла ω , но не изменяет значения Q , что доказывает неравенства для случая убывающей кривизны. \blacklozenge

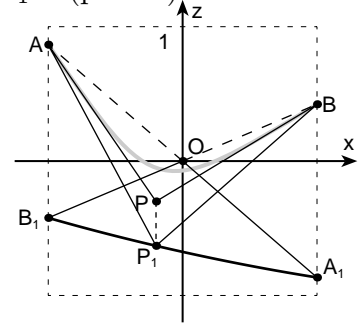


Рис. 5.

Следствие 1. Из доказательства следует, что если в (25) выполнено равенство $Q = 0$, т.е. $P \in B_1A_1$, ломаная $z(x) = V(x; P)$ является единственной функцией, удовлетворяющей условию $\mathcal{Y}(z) = 0$. Порождаемая кривая $y = f(x)$ принадлежит к семейству \mathbf{B} (23), и это построение является единственным в классе \mathcal{M} . \blacklozenge

Некую свободу построения, пределы которой было бы любопытно оценить, мы получаем только при $Q < 0$. Соответствующую оценку даст теорема 3.

Следствие 2. Неравенства (9) могут быть усилены введением дополнительных кривизн g_1 и g_2 . Согласно (19) и (20),

$$k_1 = \frac{Q - \sin^2 \omega}{k_2 - \sin \beta} - \sin \alpha, \quad g_1 = \frac{-\sin^2 \omega}{k_2 - \sin \beta} - \sin \alpha.$$

Учитывая, что $Q \leq 0$, $\omega \neq 0$, а знаменатель $k_2 - \sin \beta$ в случае \mathcal{M}^+ положителен (9), делаем вывод: $k_1 \leq g_1 < -\sin \alpha$. Аналогично, $\sin \beta < g_2 \leq k_2$. Привлекая (17), получаем цепочку неравенств, эквивалентную утверждению теоремы 1

$$\mathcal{M}^+ : \quad k_1 \leq g_1 < -\sin \alpha < k_0 < \sin \beta < g_2 \leq k_2 \quad (26)$$

$$\quad \quad \quad AP_1 \quad AP_2 \quad AA_1 \quad AB \quad B_1B \quad P_1B \quad P_2B$$

и связывающую наклоны семи отрезков, изображенных на рисунке II. Равенства в (26) наступают при $Q = 0$. \blacklozenge

Теорема 2. Если углы α, β и кривизны k_1, k_2 удовлетворяют условиям (25) для одного из классов \mathcal{M}^\pm , то в этом классе существует кривая с заданными граничными условиями.

Доказательство. Пусть $\omega > 0, Q \leq 0$ и $k_1 < k_2$. Построим функцию $z(x)$, порождающую кривую $y = f(x) \in \mathcal{M}^+(\alpha, \beta, k_1, k_2)$. Для случая $Q = 0$ такое построение уже известно (23). При $Q < 0$ будем искать его среди функций, график которых составлен из трех дуг окружностей с последовательно возрастающими кривизнами.

Пусть $Q < 0$, т.е. точка P пересечения касательных к $z(x)$ лежит ниже дуги B_1A_1 (рис. 6). Левая (AP) и правая (PB) касательные к графику искомой функции $z(x)$ пересекают дугу B_1A_1 в точках P_1 и P_2 . Кусочно-линейные функции $V_1(x) = V(x; P_1)$ и $V_2(x) = V(x; P_2)$ обращают в нуль функционал (5). Наклоны отрезков AP_2 и P_1B равны дополнительным кривизнам g_1 и g_2 .

Кандидата на роль функции $z(x)$ будем искать среди семейства функций $W(x; M, N)$, представляющих собой трехзвенные ломаные $AMNB$, где $M \in AP_1, N \in P_2B$. Наклон отрезка MN , равный кривизне промежуточной дуги, обозначим k_{12} . Очевидно, $k_1 < k_{12} < k_2$. Зафиксируем произвольную точку M внутри отрезка AP_1 , а точку N будем передвигать по P_2B , отслеживая значение функционала $\mathcal{Y}(W)$. Оно отрицательно при $N = P_2$, так как $W(x; M, P_2) \leq W(x; A, P_2) \equiv V_2(x)$ (6). На другом конце отрезка P_2B , где трехзвенная ломаная $AMNB$ вырождается в двухзвенную AMB , функционал принимает положительное значение, поскольку $W(x; M, B) \geq W(x; P_1, B) \equiv V_1(x)$. Из соображений непрерывности следует, что внутри отрезка P_2B найдется точка N такая, что $W(x; M, N)$ обращает функционал $\mathcal{Y}(W)$ в нуль. Построенная в итоге гладкая кривая $y = f(x)$ состоит из трех дуг окружностей с кривизнами k_1, k_{12}, k_2 , образующими возрастающую последовательность, и удовлетворяет всем граничным условиям. Аналогично для любой точки $N \in P_2B$ можно найти соответствующую точку $M \in AP_1$. Рисунок VI иллюстрирует семейство построенных кривых.

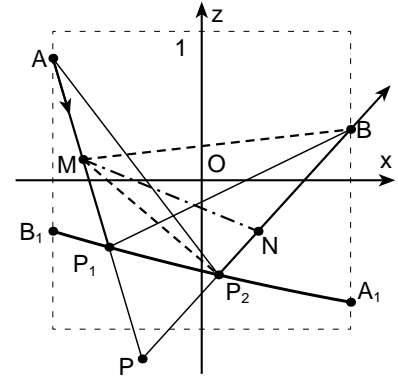


Рис. 6.

Доказательство для случая убывания дает обращение к функции $y = -f(x)$. \blacklozenge

Теорема 3. Если кривая $y = f(x) \in \mathcal{M}(\alpha, \beta)$, то она находится внутри линзы $\mathcal{L}(\alpha, -\beta)$, а именно, при $-1 < x < 1$

$$\mathcal{M}^+ : \{x, f(x)\} \subset \mathcal{L}^*(-\beta, \alpha) \quad \text{или} \quad \mathbf{A}(x; -\beta) < f(x) < \mathbf{A}(x; \alpha); \quad (27)$$

$$\mathcal{M}^- : \{x, f(x)\} \subset \mathcal{L}^*(\alpha, -\beta) \quad \text{или} \quad \mathbf{A}(x; -\beta) > f(x) > \mathbf{A}(x; \alpha).$$

Если известны также граничные кривизны k_1, k_2 или их оценки k', k'' , то кривая находится внутри области, ограниченной парой кривых семейства (23):

$$\mathcal{M}^+ : k' \leq k_1 < k_2 \leq k'' \implies \mathbf{B}_1(x; \alpha, \beta, k') \leq f(x) \leq \mathbf{B}_2(x; \alpha, \beta, k''); \quad (28)$$

$$\mathcal{M}^- : k' \geq k_1 > k_2 \geq k'' \implies \mathbf{B}_1(x; \alpha, \beta, k') \geq f(x) \geq \mathbf{B}_2(x; \alpha, \beta, k'').$$

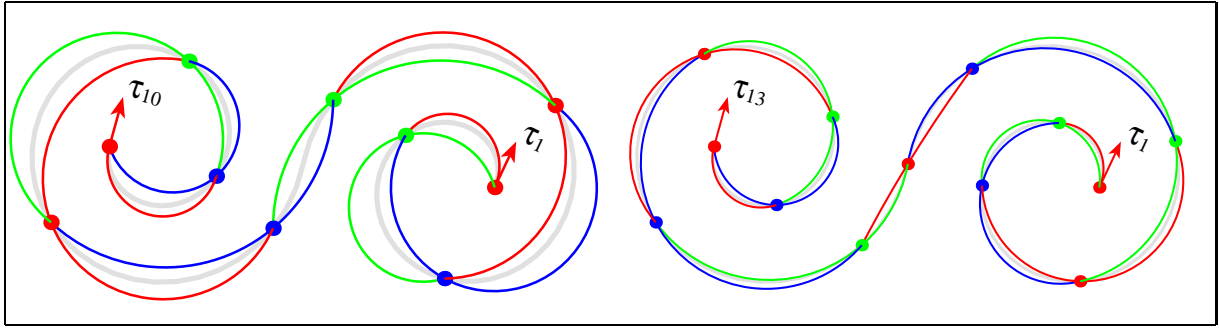


Рис. I (к стр. 2).

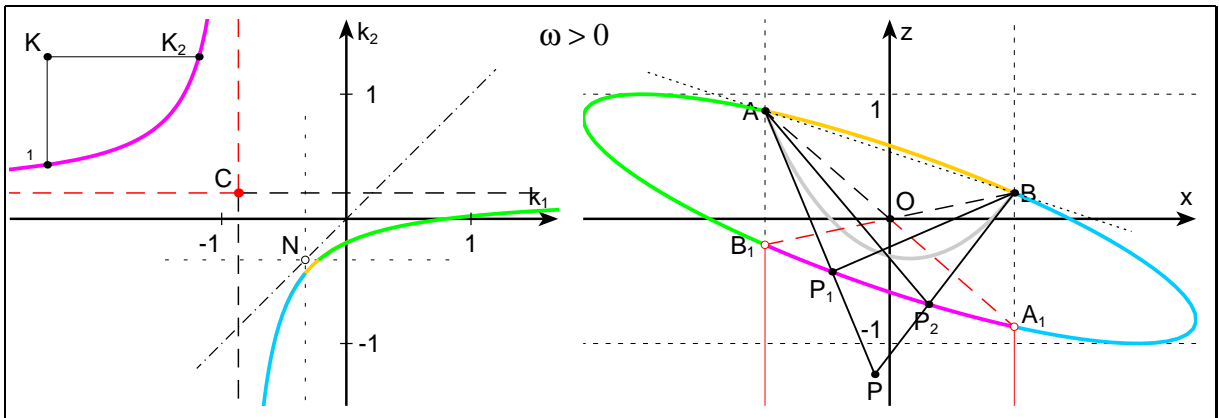


Рис. II (к стр. 6–10).

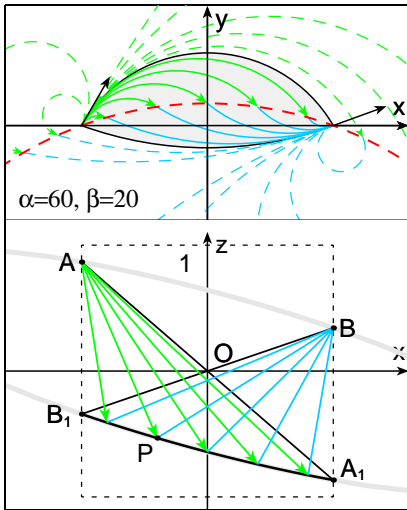


Рис. III (к стр. 8).

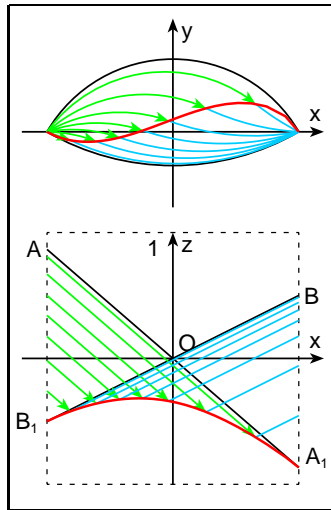


Рис. IV (к стр. 9).

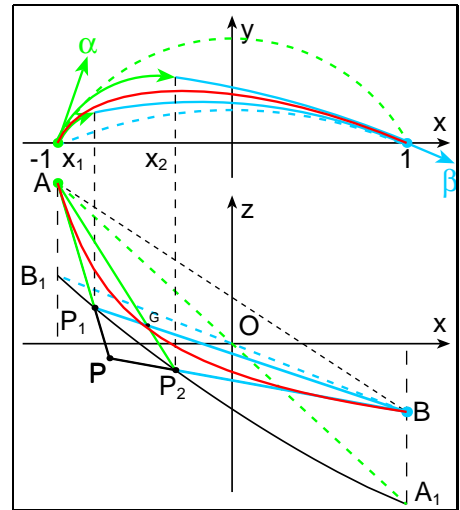


Рис. V (к стр. 14).

Рис. VI
(к стр. 11).

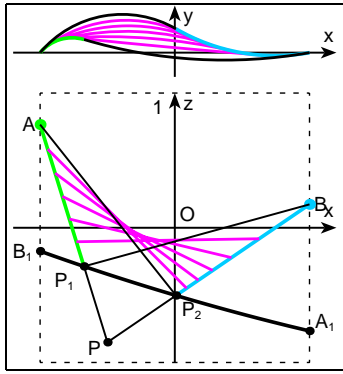


Рис. VII (к стр. 15).

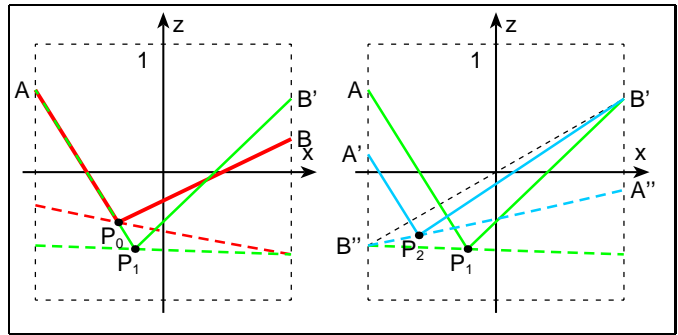


Рис. VIII
(к стр. 16).

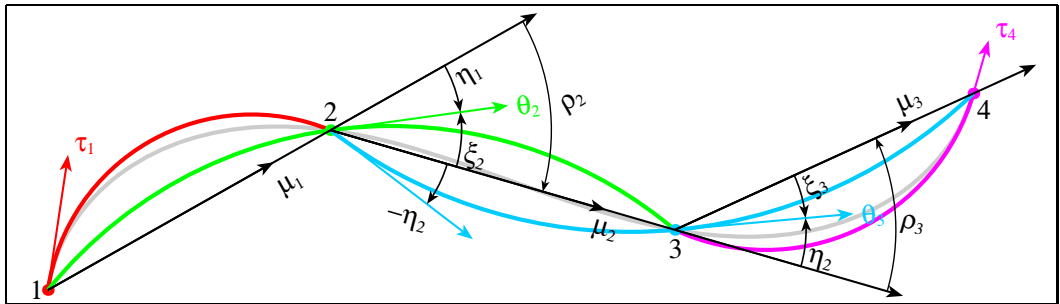


Рис. IX
(к стр. 20).

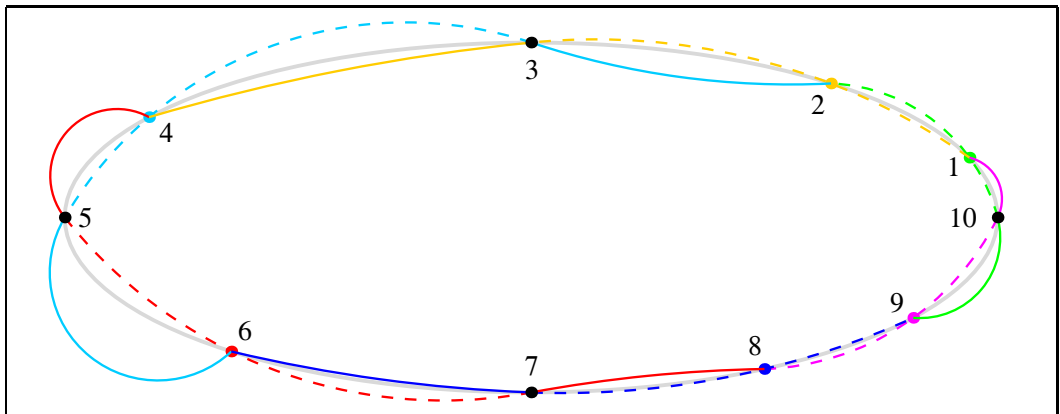
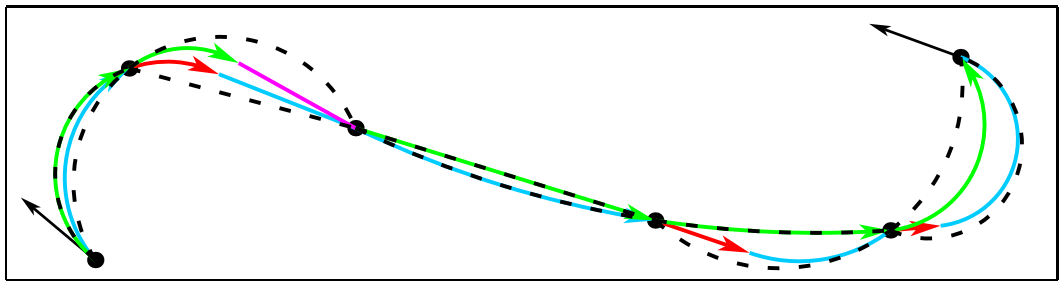


Рис. X
(к стр. 20).



Доказательство. Для функции $y = f(x) \in \mathcal{M}^+(\alpha, \beta, k_1, k_2)$ докажем (28), введя промежуточные неравенства с ограничивающими функциями $b_1(x)$ и $b_2(x)$:

$$b_1(x) := \mathbf{B}_1(x; \alpha, \beta, k_1) \leq f(x) \leq \mathbf{B}_2(x; \alpha, \beta, k_2) =: b_2(x), \quad (29)$$

$$\mathbf{B}_1(x; \alpha, \beta, k') \leq \mathbf{B}_1(x; \alpha, \beta, k_1), \quad \mathbf{B}_2(x; \alpha, \beta, k_2) \leq \mathbf{B}_2(x; \alpha, \beta, k''). \quad (30)$$

Если $Q(\alpha, \beta, k_1, k_2) = 0$, то кривизны k_1, k_2 взаимно дополнительные. Тогда, согласно следствию 1 теоремы 1, функции $b_1(x)$ и $b_2(x)$ совпадают между собой и с $f(x)$ — единственной функцией в классе \mathcal{M}^+ с такими граничными условиями.

Пусть $Q < 0$. Тогда точка P пересечения касательных (рисунок V) лежит ниже дуги B_1A_1 (22), а ломаная APB пересекает ее в точках P_1 и P_2 с абсциссами $x_1 < x_2$. Функции $V_1(x) = V(x; P_1)$ и $V_2(x) = V(x; P_2)$ соответствуют паре кривых $b_1(x)$ и $b_2(x)$. Неравенство (29) означает, что неотрицательна разность

$$\delta_1(x) = f(x) - b_1(x) \stackrel{(4)}{=} \int_{-1}^x \frac{z(u)}{\sqrt{1-z(u)^2}} du - \int_{-1}^x \frac{V_1(u)}{\sqrt{1-V_1(u)^2}} du. \quad (31)$$

Условие экстремума $\delta_1(x)$ принимает вид

$$\delta_1'(x) = \frac{z(x)}{\sqrt{1-z(x)^2}} - \frac{V_1(x)}{\sqrt{1-V_1(x)^2}} = 0, \quad (32)$$

что ввиду монотонности функции $t \mapsto t/\sqrt{1-t^2}$ возможно только при $z(x) = V_1(x)$. Звено AP_1 ломаной $V_1(x)$, будучи опорной прямой графика $z(x)$, не пересекает его, но может совпадать с ним на отрезке $[-1, p]$, $-1 \leq p \leq x_1$. Это влечет равенство $b_1(x) = f(x)$ при $x \in [-1, p]$. Звено P_1B расположено между хордой AB и опорной прямой P_2B (26), и ввиду выпуклости графика пересекает его в единственной точке в интервале $(p, 1)$. Следовательно, в этом интервале уравнение (32) имеет единственное решение, а функция $\delta_1(x)$ — единственный экстремум. На правом конце $\delta_1(x)$ строго убывает (26): $z'(1) = k_2 > g_2 = V_1'(1)$. Поэтому этот экстремум — максимум, и так как $\delta_1(p) = \delta_1(1) = 0$, $\delta_1(x) > 0$ внутри $(p, 1)$.

Для доказательства неравенства $f(x) \leq b_2(x)$ в (29) следует рассмотреть функцию $\delta_2(x) = f(x) - b_2(x)$, которая имеет единственный (отрицательный) минимум в точке, определяемой уравнением $z(x) = V_2(x)$. Отметим, что равенства в (29) могут быть только у концов хорды при наличии участков постоянной кривизны $[-1, p]$ или $[q, 1]$. Точнее, $b_1(x) \leq f(x) < b_2(x)$ при $x \in (-1, p]$, $b_1(x) < f(x) < b_2(x)$ при $x \in (p, q)$, $b_1(x) < f(x) \leq b_2(x)$ при $x \in [q, 1)$.

Прием (31)–(32) делает доказательство неравенств (30) тривиальным: речь идет о сравнении двух функций типа $V(x; P)$, у которых точка излома P смещается по дуге эллипса от B_1 к A_1 . Единственность корня и тип экстремума очевидны⁵. Равенства возможны только при $k' = k_1$ и $k_2 = k''$.

Утверждение (28) для класса \mathcal{M}^+ доказано. Утверждение (27) получаем из него как предел при $k_1 \rightarrow -\infty$, $k_2 \rightarrow +\infty$ (24). Поскольку участки “постоянной бесконечной

⁵В частности, максимум ширины $\delta(x) = |b_2(x) - b_1(x)|$ достигается в точке $x_0(g_1, g_2)$ (14), определяемой как решение уравнения $V_1(x) = V_2(x)$ (рис. 7).

кривизны” невозможны, неравенства (27) строгие. Обращение к функции $y = -f(x)$ дает доказательство для класса \mathcal{M}^- . ♦

Напомним, что каждая точка $K(k_1, k_2) \in \mathcal{K}^\pm(\alpha, \beta)$ порождает подкласс кривых $y = f(x) \in \mathcal{M}^\pm(\alpha, \beta, k_1, k_2)$. Неравенства (28) приводят к определению *билинзы*

$$\mathcal{B}^\pm(\alpha, \beta, k_1, k_2) := \{(x, y) : |x| \leq 1, \mathbf{B}_1(x; \alpha, \beta, k_1) \leq y \leq \mathbf{B}_2(x; \alpha, \beta, k_2)\}, \quad (33)$$

ограничивающей все кривые из области $KK_1DK_2K \in \mathcal{K}$ (рис. 7). Расширение ее в бесконечность означает превращение билинзы в линзу $\mathcal{L}(\alpha, -\beta)$.

Отметим некоторые свойства билинзы. Соединим прямой точку K с центром C гиперболы, пересекая ее в точке $D(d_1, d_2)$. Из геометрии гиперболы следует, что на этой прямой лежит и точка $G(g_1, g_2)$, и что $|CD|^2 = |CG| \cdot |CK|$. Кривую $\mathbf{B}_1(x; \alpha, \beta, d_1) \equiv \mathbf{B}_2(x; \alpha, \beta, d_2)$, порождаемую точкой D , можно считать средней линией билинзы. Максимальный диаметр вписанной в билинзу окружности равен $2|r|$,

$$\text{где } r = \frac{2Q \sin \omega}{N + \sqrt{N^2 + 4QQ'g_1g_2}}, \quad Q = Q(\alpha, \beta, k_1, k_2), \quad Q' = Q - \sin^2 \omega, \\ N = \sin \omega(k_2 - k_1 - \sin \alpha - \sin \beta) - 2Q' \cos \gamma.$$

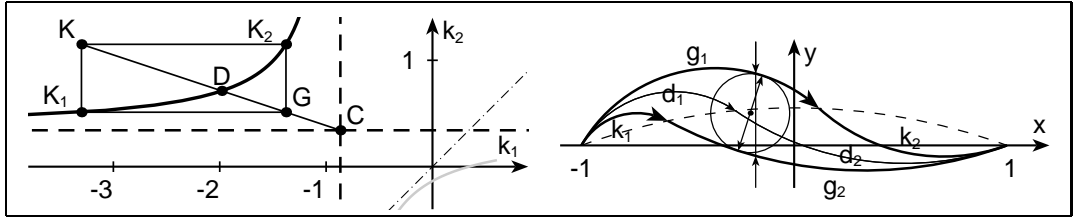


Рис. 7.

Теорема 4. Если кривая $y = f(x) \in \mathcal{M}^+(\alpha, \beta, k_1, k_2)$, и для граничных условий известны оценки

$$k' \leq k_1 < k_2 \leq k'', \quad \alpha' \leq \alpha \leq \alpha'', \quad \text{такие, что} \quad \alpha' + \beta'' \geq 0, \\ \beta' \leq \beta \leq \beta'', \quad \alpha'' + \beta' \geq 0, \quad (34)$$

то кривая находится внутри области, ограниченной парой кривых семейства (23):

$$\mathbf{B}_1(x; \alpha', \beta'', k') \leq f(x) \leq \mathbf{B}_2(x; \alpha'', \beta', k''). \quad (35)$$

Доказательство. Для функции $y = f(x) \in \mathcal{M}^+(\alpha, \beta, k_1, k_2)$ докажем неравенства

$$\mathbf{B}_1(x; \alpha', \beta'', k') \leq \mathbf{B}_1(x; \alpha, \beta'', k') \leq \mathbf{B}_1(x; \alpha, \beta, k') \leq f(x); \\ f(x) \leq \mathbf{B}_2(x; \alpha, \beta, k'') \leq \mathbf{B}_2(x; \alpha'', \beta, k'') \leq \mathbf{B}_2(x; \alpha'', \beta', k''),$$

крайние звенья которых составляют утверждение теоремы. Левое от $f(x)$ неравенство уже доказано (28). Следующий шаг влево означает замену угла β его верхней оценкой β'' при фиксированных значениях α и k' . На рисунке VII показано вызванное этим изменение функции $z(x)$ от $V_0(x)$ к $V_1(x)$ (ломаные AP_0B и AP_1B'). Сравнение функций $V_0(x)$ и $V_1(x)$ по аналогии с (31)–(32) показывает, что уравнение $V_0(x) = V_1(x)$ имеет на участке их несовпадения единственный корень, и разность порождаемых ими функций $\mathbf{B}_1(x; \alpha, \beta, k') - \mathbf{B}_1(x; \alpha, \beta'', k')$ неотрицательна.

Последний шаг влево — замену угла α его нижней оценкой α' — иллюстрирует второй фрагмент рисунка VII. Неравенство было бы строгим, но в оценках мы допускаем $\alpha = \alpha'$. Смещение точки A' вниз имеет предел — точку $B''(-1, \sin \beta'')$, выше которой функция $\mathbf{B}_1(x; \alpha, \beta'', k') \in \mathcal{M}^+$ еще существует. Отсюда в утверждении теоремы условие $\alpha' + \beta'' \geq 0$. В этом условии мы допускаем равенство, имея в виду определения (24) и ограничивающую функцию $y = \mathbf{A}(x; -\beta)$.

Неравенства справа от $f(x)$ доказываются аналогично. \blacklozenge

5. Глобальные оценки детерминированности кривой

Пусть задано дискретно-точечное представление $\{P_i, i=1, n; \tau_1, \tau_n\}$ плоской кривой (рисунок VIII). Для последовательности соседних узлов и исходящих из них хорд введем индексы $i, j=i+1, k=i+2, l=i+3$. В каждом j -том узле, $j \leq n-1$, определим длину $2c_j$ и направление μ_j исходящей хорды:

$$|\vec{P_j P_k}| = 2c_j = \sqrt{(x_k - x_j)^2 + (y_k - y_j)^2}, \quad \cos \mu_j = \frac{x_k - x_j}{2c_j}, \quad \sin \mu_j = \frac{y_k - y_j}{2c_j}.$$

Рассматривая граничные условия как бесконечно малые “псевдохорды” $P_0 P_1$ и $P_n P_{n+1}$, направленные под углами τ_1 и τ_n , полагаем для них

$$c_0 = c_n = 0, \quad \mu_0 = \tau_1, \quad \mu_n = \tau_n.$$

Поворот хорды в j -том узле ρ_j и кривизна q_j ориентированной окружности, соединяющей узлы i, j, k , определяются формулами

$$\rho_j = \mu_j - \mu_i, \quad q_j = \frac{\sin \rho_j}{d_j}, \quad \text{где } d_j = \sqrt{c_i^2 + c_j^2 + 2c_i c_j \cos \rho_j} = \frac{1}{2} |\vec{P_i P_k}|,$$

справедливыми и при расчете кривизн q_1 и q_n . Углы ρ_j предполагаются приведенными к интервалу $(-\pi, \pi)$.

Неизвестные кривизны и углы наклона касательной к кривой в узле j обозначим p_j и τ_j . Преобразованием координат, включающим гомотетию, каждый участок кривой (jk) может быть приведен к виду $f_j(x) \in \mathcal{M}(\alpha_j, \beta_j, k'_j, k''_j)$, где

$$\alpha_j = \tau_j - \mu_j, \quad \beta_j = \tau_k - \mu_j, \quad k'_j = p_j c_j, \quad k''_j = p_k c_j.$$

k' и k'' — нормированные кривизны, входящие в неравенства (9) как k_1 и k_2 .

В качестве оценок величин p_j и τ_j тем лучших, чем мельче разбиение кривой, могут выступать кривизна “трехточечной” дуги q_j и угол ϑ_j — наклон касательной к этой дуге в среднем узле. В узлах P_1 и P_n $\vartheta_1 = \tau_1, \vartheta_n = \tau_n$.

На каждую хорду $P_j P_k$ опираются две дуги — (ijk) и (jkl) с кривизнами q_j и q_k . Они образуют линзу $\mathcal{L}(\xi_j, -\eta_j)$, аналог линзы $\mathcal{L}(\alpha_j, -\beta_j)$:

$$\xi_j = \vartheta_j - \mu_j, \quad \eta_j = \vartheta_k - \mu_j.$$

Углы на смежных хордах $P_i P_j$ и $P_j P_k$ связаны соотношениями

$$\tau_j = \mu_i + \beta_i = \mu_j + \alpha_j \quad \Longrightarrow \quad \beta_i = \rho_j + \alpha_j, \quad (36)$$

$$\vartheta_j = \mu_i + \eta_i = \mu_j + \xi_j \quad \Longrightarrow \quad \eta_i = \rho_j + \xi_j. \quad (37)$$

В отличие от углов α_i, β_i , ограниченных значениями $\pm\pi/2$, углы ξ_i, η_i лежат в интервале $\pm\pi$. Расчетные формулы для них имеют вид

$$\begin{aligned}\sin \xi_j &= \frac{-c_j \sin \rho_j}{d_j} = -q_j c_j, & \sin \eta_j &= \frac{c_j \sin \rho_k}{d_k} = q_k c_j, \\ \cos \xi_j &= \frac{c_i + c_j \cos \rho_j}{d_j}, & \cos \eta_j &= \frac{c_k + c_j \cos \rho_k}{d_k}.\end{aligned}\quad (38)$$

Лемма 2. *Если на участке (ijk) дискретно-точечного представления некой кривой кривизна изменялась нестрого монотонно, то параметры кривой и ее представления на этом участке удовлетворяют неравенствам*

$$\nearrow_{ijk} : \begin{aligned} -\eta_i &\leq -\beta_i \leq \alpha_i \\ -\beta_j &\leq \alpha_j \leq \xi_j \end{aligned} \quad \searrow_{ijk} : \begin{aligned} -\eta_i &\geq -\beta_i \geq \alpha_i \\ -\beta_j &\geq \alpha_j \geq \xi_j, \end{aligned}\quad (39)$$

в которых равенства возможны только в случае постоянства кривизны на всем участке (стрелки в формулах отражают характер монотонности кривизны).

Случай постоянства кривизны очевиден. Исключив его, докажем строгие неравенства для возрастающей кривизны. Соотношения между углами $-\beta_{i,j} < \alpha_{i,j}$ дублируют неравенства (9). Из (9) также следует

$$\begin{aligned} p_j c_i = k_i'' &\geq \sin \beta_i, & \implies & \frac{\sin \beta_i}{c_i} \leq p_j \leq \frac{\sin(\rho_j - \beta_i)}{c_j}. \\ p_j c_j = k_j' &\leq -\sin \alpha_j \stackrel{(36)}{=} \sin(\rho_j - \beta_i) \end{aligned}\quad (40)$$

В каждой паре неравенств хотя бы одно обязательно строгое, так как постоянство кривизны возможно только на одной из хорд⁶. Раскрывая синус разности, получаем

$$\sin \beta_i (c_j + c_i \cos \rho_j) < c_i \sin \rho_j \cos \beta_i. \quad (41)$$

По условиям разбиения кривой, $\cos \beta_i > 0$. Сомножитель $(c_j + c_i \cos \rho_j)$, зависящий от расположения узлов i, j и k , отрицателен только в том случае, если точка P_k находится внутри круга Θ_{ij} , построенного на хорде $P_i P_j$ как на диаметре. Рассмотрим следующие варианты расположения точки P_k (рис. 8).

1. Точка P_k находится вне круга Θ_{ij} , т.е. $0 < c_j + c_i \cos \rho_j \stackrel{(38)}{=} d_j \cos \eta_i$. Следовательно, $\cos \eta_i > 0$, и $|\eta_i| < \pi/2$. Тогда (41) легко преобразуется к желанному соотношению между углами:

$$\operatorname{tg} \beta_i < \frac{c_i \sin \rho_j}{c_j + c_i \cos \rho_j} = \operatorname{tg} \eta_i \implies -\eta_i < -\beta_i.$$

Этим случаем можно было бы и ограничиться: в алгоритмах расстановки узлов интерполяции часто формулируется требование не допускать поворота хорды на угол, превышающий $\pi/2$, а при $|\rho_j| < \pi/2$ неравенство $c_j + c_i \cos \rho_j > 0$ заведомо выполнено. Стремясь к необходимому минимуму ограничений, мы не накладываем этого требования и рассматриваем остальные ситуации.

⁶Либо на обеих, но со скачком кривизны в узле, что не мешает последующему выводу.

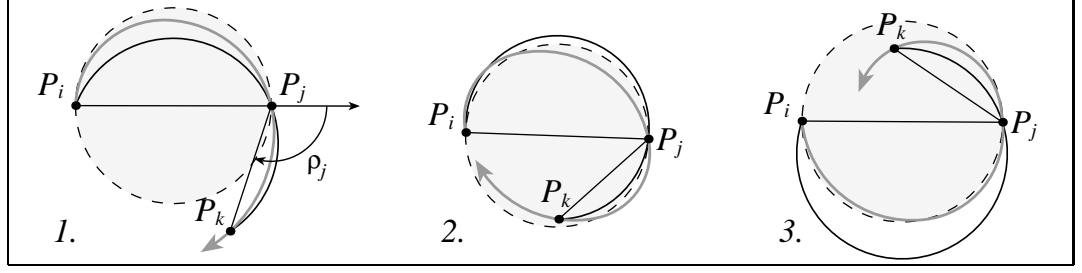


Рис. 8.

2. Точка P_k находится внутри правого (относительно направления $\vec{P_i P_j}$) полукруга Θ_{ij} или на его границе, т.е. $c_j + c_i \cos \rho_j \leq 0$, $-\pi < \rho_j < -\pi/2$. Тогда $\beta_i = \alpha_j + \rho_j < 0$. При таких знаках сомножителей неравенство (41) не может быть выполнено:

$$(\sin \beta_i)^- (c_j + c_i \cos \rho_j)^{-,0} \not\leq c_i^+ (\sin \rho_j)^- (\cos \beta_i)^+.$$

Следовательно, исключено и такое расположение точки P_k : оно несовместимо либо с возрастанием кривизны, либо с представлением участка кривой как функции на хорде.

3. Точка P_k находится внутри или на границе левого полукруга Θ_{ij} , т.е. $c_j + c_i \cos \rho_j \leq 0$, $\pi/2 < \rho_j < \pi$, $\sin \rho_j > 0$. Следовательно (38),

$$\cos \eta_i = \frac{c_j + c_i \cos \rho_j}{d_j} \leq 0, \quad \sin \eta_i = \frac{c_i \sin \rho_j}{d_j} > 0 \quad \implies \quad \eta_i \geq \pi/2. \quad (42)$$

Поскольку $|\beta_i| < \pi/2$, целевое неравенство $-\eta_i < -\beta_i$ выполнено. Особенность этой ситуации состоит в том, что одна из ограничивающих дуг будет больше 180° . Хотя при этом ее нельзя будет рассматривать как функцию на хорде, свою ограничивающую роль она исполнит.

Неравенство в первой строке (39) доказано. Из него получается неравенство второй строки:

$$\eta_i > \beta_i \quad \implies \quad \xi_j \stackrel{(37)}{=} \eta_i - \rho_j > \beta_i - \rho_j \stackrel{(36)}{=} \alpha_j.$$

В случае невозрастающей кривизны лемма доказывается аналогично. \blacklozenge

Теорема 5 (интерполяционное свойство спирали). Если $\{P_j, j=1, n; \tau_1, \tau_n\}$ — дискретно-точечное представление плоской спиральной кривой Γ , то

$$\Gamma \subseteq \bigcup_{j=1}^{j=n-1} \mathcal{L}(-\eta_j, \xi_j; P_j P_{j+1}). \quad (43)$$

Если при этом для поворотов хорд выполнены условия $|\rho_j| \leq \pi/2$ или более слабые ограничения

$$c_{j-1} + c_j \cos \rho_j > 0 \quad \text{и} \quad c_j + c_{j-1} \cos \rho_j > 0, \quad 1 \leq j \leq n, \quad (44)$$

то последовательность расчетных кривизн $\{q_j, j=1, n\}$ наследует характер монотонности кривизны исходной кривой.

Доказательство. Для кривой, представленной n узлами, $n \geq 3$, утверждение (39) охватывает значения $\{\xi_2, \dots, \xi_{n-1}\}$ и $\{\eta_1, \dots, \eta_{n-2}\}$. Значения ξ_1 на первой и η_{n-1} на последней хордах получаются из граничных условий или формул (38): $\xi_1 = \alpha_1 = -\rho_1$, $\eta_{n-1} = \beta_{n-1} = \rho_n$. Это соответствует неравенствам (39) и допускает минимальную конфигурацию ($n = 2$) с единственной хордой. Таким образом, “трехточечное” утверждение (39) может быть распространено на все хорды данного представления спиральной кривой и записано в виде

$$\nearrow : \quad -\eta_i \leq -\beta_i \leq \alpha_i \leq \xi_i, \quad \searrow : \quad -\eta_i \geq -\beta_i \geq \alpha_i \geq \xi_i, \quad 1 \leq i \leq n-1, \quad (45)$$

откуда, согласно (7) и (27), следует первое утверждение теоремы:

$$\begin{aligned} \nearrow : \quad \Gamma_i &\subseteq \mathcal{L}^*(-\beta_i, \alpha_i) \subseteq \mathcal{L}^*(-\eta_i, \xi_i), \\ \searrow : \quad \Gamma_i &\subseteq \mathcal{L}^*(\alpha_i, -\beta_i) \subseteq \mathcal{L}^*(\xi_i, -\eta_i), \end{aligned} \quad 1 \leq i \leq n-1. \quad (46)$$

Здесь Γ_i — множество точек, принадлежащих участку кривой (ij) .

Ограничения (44) означают, что для каждой хорды (ij) соседние узлы $i-1$ и $k = j+1$ расположены вне круга Θ_{ij} . Если они выполнены, то $|\xi_j| < \pi/2$ и $|\eta_j| < \pi/2$. Тогда из равенств $\sin \xi_j = -q_j c_j$, $\sin \eta_j = q_k c_j$ (38) и неравенства $\xi_j \geq -\eta_j$ следует $q_j \leq q_k$ (в случае убывания $\xi_j \leq -\eta_j$ и $q_j \geq q_k$), что составляет второе утверждение теоремы. \blacklozenge

Из неравенств (40) также получаем, что при условиях (44) истинная кривизна в узле ограничена “трехточечными” кривизнами в соседних узлах. Так, в случае неубывающей кривизны $q_i \leq p_j \leq q_k$. На практике несоблюдение условий (44) следует отнести к некорректной (слишком редкой) расстановке точек.

Обобщение для разбиений с выделенной вершиной. Получим ограничения для неспиральных кривых, потребовав разбиения каждого участка монотонности как минимум на две хорды с обязательным наличием узлов в вершинах кривой.

Рассмотрим фрагмент дискретно-точечного представления кривой (1-2-3-4-5) с указанием на наличие вершины-минимума в узле 3. Для участков убывающей (1-2-3) и возрастающей кривизны (3-4-5) неравенства (39) принимают вид

$$\xi_2 < \alpha_2 < -\beta_2 < ?, \quad -\eta_3 < -\beta_3 < \alpha_3 < i.$$

Их левые части дают искомые ограничения для углов α_2 и β_3 . Извлекая отсюда неравенства $-\alpha_3 < \eta_3$ и $\beta_2 < -\xi_2$, получаем недостающие ограничения:

$$-\beta_2 \stackrel{(36)}{=} -\alpha_3 - \rho_3 < \eta_3 - \rho_3, \quad \alpha_3 = \beta_2 - \rho_3 < -\xi_2 - \rho_3.$$

Теперь к каждому из участков (2-3) и (3-4) можно применить теорему 3:

$$\searrow \nearrow : \quad \begin{aligned} \Gamma_2 &\subseteq \mathcal{L}^*(\alpha_2, -\beta_2) \subseteq \mathcal{L}^*(\xi_2, \eta_3 - \rho_3), \\ \Gamma_3 &\subseteq \mathcal{L}^*(-\beta_3, \alpha_3) \subseteq \mathcal{L}^*(-\eta_3, -\xi_2 - \rho_3). \end{aligned} \quad (47)$$

В случае вершины-максимума аналогичные рассуждения дают то же построение:

$$\nearrow \searrow : \quad \begin{aligned} \Gamma_2 &\subseteq \mathcal{L}^*(-\beta_2, \alpha_2) \subseteq \mathcal{L}^*(\eta_3 - \rho_3, \xi_2), \\ \Gamma_3 &\subseteq \mathcal{L}^*(\alpha_3, -\beta_3) \subseteq \mathcal{L}^*(-\xi_2 - \rho_3, -\eta_3). \end{aligned}$$

На рисунке IX это построение выполнено для вершин эллипса, представленного десятью точками. Узлы на большой оси — вершины-максимумы. Большие отклонения у пятого узла связаны со слишком редкой расстановкой точек: длина хорды превышает “диаметр кривизны” в вершине, что в терминах нормированной кривизны означает $\max |k(x)| \gtrsim 2$. На практике обычно соблюдается условие $|k(x)| \ll 1$.

Из рисунка видно, что новые ограничивающие дуги являются гладкими дополнениями к “трехточечным” дугам — приходящей в вершину и исходящей из нее (показаны пунктиром). Описания смежных линз (47) позволяют легко в этом убедиться. Подчеркнем, что данное построение предполагает не только наличие узла в вершине, но и наличие признака “это — вершина” в представлении кривой.

Усиление интерполяционного свойства спирали. В области вершины четыре ограничивающие дуги образуют пару гладких кривых, что на спиральных участках может произойти лишь случайно. Основываясь на гладкости исходной кривой, попытаемся усилить ограничения (45),

Касательная в j -том узле определяет углы β_i и α_j на предшествующей и исходящей хордах и обеспечивает жесткую связь между ними: $\beta_i = \alpha_j + \rho_j$ (36). Рассмотрим случай неубывающей кривизны при соблюдении условий (44). Неравенства (45) преобразуются следующим образом:

$$\begin{aligned} -\xi_i \leq \beta_i \leq \eta_i & \implies -\xi_i - \rho_j \leq \beta_i - \rho_j \leq \eta_i - \rho_j & \implies -\xi_i - \rho_j \leq \alpha_j \leq \xi_j, \\ -\eta_j \leq \alpha_j \leq \xi_j & \implies -\eta_j + \rho_j \leq \alpha_j + \rho_j \leq \xi_j + \rho_j & \implies -\eta_j + \rho_j \leq \beta_i \leq \eta_i. \end{aligned}$$

Отсюда получаются ограничения типа (34)

$$\begin{aligned} \beta'_i &:= \max(-\xi_i, \rho_j - \eta_j) \leq \beta_i \leq \eta_i =: \beta''_i, \\ \alpha'_j &:= \max(-\rho_j - \xi_i, -\eta_j) \leq \alpha_j \leq \xi_j =: \alpha''_j, \\ \beta'_j &:= \max(-\xi_j, \rho_k - \eta_k) \leq \beta_j \leq \eta_j =: \beta''_j \end{aligned}$$

(последнее получено из первого приращением индексов). Убедимся в выполнении условий (34). Например, для суммы $\alpha'_j + \beta''_j$ имеем

$$\alpha'_j + \beta''_j = \max(-\eta_j, -\rho_j - \xi_i) + \eta_j = \max(0, \dots) \geq 0.$$

Ограничения кривизны дают неравенства (40):

$$\frac{\sin \beta'_i}{c_i} \leq \frac{\sin \beta_i}{c_i} \leq p_j \leq \frac{-\sin \alpha_j}{c_j} \leq \frac{-\sin \alpha'_j}{c_j}, \implies \frac{\sin \beta'_i}{c_i} \leq p_j, \quad p_k \leq \frac{-\sin \alpha'_k}{c_k}.$$

Получаем границы возможных значений нормированной кривизны на j -той хорде:

$$k'_1 = -\infty, \quad k'_{j, j>1} = \frac{c_j \sin \beta'_i}{c_i}; \quad k''_{j, j<n-1} = \frac{-c_j \sin \alpha'_k}{c_k}, \quad k''_{n-1} = +\infty.$$

Теорема 4 дает границы участка кривой

$$\mathbf{B}_1(x; \alpha'_j, \beta'_j, k'_j) \leq f_j(x) \leq \mathbf{B}_2(x; \alpha''_j, \beta''_j, k''_j). \quad (48)$$

Соответствующее построение изображено на рисунке X. Пунктиром показана совокупность линз (43).

При приблизительно равномерной расстановке точек на кривой эффект усиления невелик, и для практических целей мы предлагаем более простой вариант (43). Усиление (48) мы рассматриваем как шаг к построению точной границы (каковой оно, возможно, и является).

6. Кривые с единственной вершиной

Требование монотонности кривизны накладывает довольно жесткие рамки на граничные условия кривой, заданной на хорде. С другой стороны, допуская на кривой две и более вершин, можно удовлетворить любые граничные условия. Рассмотрим промежуточный случай с однократной сменой характера монотонности в интервале $(-1, 1)$ ⁷.

Теорема 6. Если кривая принадлежит к одному из классов $\mathcal{V}^\pm(\alpha, \beta, k_1, k_2)$, то

$$\mathcal{V}^+ : \begin{array}{ll} k_1 < -\sin \alpha & \text{при } \omega \geq 0, \\ k_2 < \sin \beta & \text{при } \omega \leq 0; \end{array} \quad \mathcal{V}^- : \begin{array}{ll} k_2 > \sin \beta & \text{при } \omega \geq 0, \\ k_1 > -\sin \alpha & \text{при } \omega \leq 0. \end{array} \quad (49)$$

Доказательство. Пусть $y=f(x) \in \mathcal{V}^+(\alpha, \beta, k_1, k_2)$, и $\omega \geq 0$. Докажем невозможность существования такой кривой при $k_1 \geq -\sin \alpha$ (рис. 9).

Неравенство $k_1 \geq -\sin \alpha$ означало бы, что отрезок AT прямой $t_1(x)$ — левой касательной к графику $z(x)$ — находится не ниже отрезка AA_1 : $t_1(x) \geq z_1(x)$.

Вершина разделяет кривую на два участка монотонности $k(x)$ и порождает точку перегиба S графика $z(x)$ (или “линию перегиба”, если вершина протяженная). На участке AS $k(x) = z'(x)$ возрастает, и выпуклый вниз график уходит вверх от касательной AT . Следовательно, $z(x) \geq t_1(x) \geq z_1(x)$ на отрезке $[-1, x_S]$. Выше отрезка AA_1 лежит и точка перегиба S . На участке SB кривизна убывает, график $z(x)$ имеет выпуклость вверх и лежит выше хорды SB . Если $\omega \geq 0$, то $\sin \beta \geq -\sin \alpha$, и точка B находится не ниже точки A_1 . Значит, вся хорда SB вместе с соответствующим участком графика лежит выше AA_1 , и неравенство $z(x) \geq z_1(x)$ выполнено на всем отрезке $[-1, 1]$. Следовательно, $\mathcal{Y}(z) > \mathcal{Y}(z_1) = 0$, и построение функции $y=f(x)$ невозможно.

Тем самым доказана “четверть теоремы”. Пусть, по-прежнему, кривая $y=f(x)$ принадлежит к классу $\mathcal{V}^+(\alpha, \beta, k_1, k_2)$, но $\omega \leq 0$. Тогда $y_1=f(-x) \in \mathcal{V}^+(\alpha', \beta', k'_1, k'_2)$, где $\alpha'=-\beta$, $\beta'=-\alpha$, $k'_1=k_2$, $k'_2=k_1$ и $\omega'=-\omega \geq 0$. По доказанному существование такой кривой возможно лишь при $k'_1 < -\sin \alpha'$, т.е. при $k_2 < \sin \beta$. Для класса \mathcal{V}^+ утверждение теоремы доказано полностью. Неравенства для кривых $y=f(x)$ с

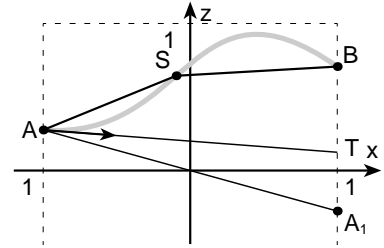


Рис. 9.

⁷Утверждения для этого класса кривых распространяются, по-видимому, и на кривые типа $y=|x|-1$, у которых излом можно рассматривать как вершину бесконечной кривизны. Пример гладкой кривой с вершиной бесконечной кривизны дает функция $y=\sqrt{|x|^3}-1 \in \mathcal{V}^+$.

вершиной-минимумом получаются умножением функции на -1 , что переводит ее из класса $\mathcal{V}^-(\alpha, \beta, k_1, k_2)$ в уже исследованный класс $\mathcal{V}^+(-\alpha, -\beta, -k_1, -k_2)$. \blacklozenge

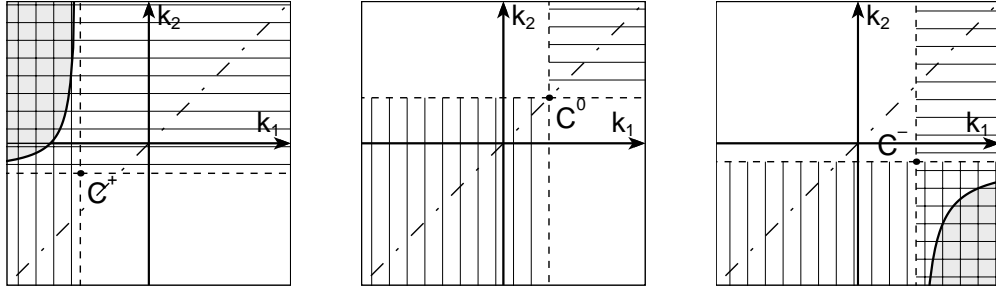


Рис. 10.

Рис. 10 иллюстрирует области (49). Они ограничены асимптотами гиперболы (19) и не включают границу. Области \mathcal{V}^+ показаны вертикальной штриховкой, \mathcal{V}^- — горизонтальной. При $\omega > 0$ и $\omega < 0$ (левый и правый фрагменты) это полуплоскости, целиком включающие область монотонности. На среднем фрагменте ($\omega = 0$) области \mathcal{M}^\pm исчезают; появляется точка, в которой возможно постоянство кривизны. Области \mathcal{V}^\pm при этом ограничены и по k_1 , и по k_2 ⁸.

Теорема 7. Вершина кривой $y=f(x) \in \mathcal{V}^\pm(\alpha, \beta)$, лежит строго ниже (в случае \mathcal{V}^+) или строго выше (в случае \mathcal{V}^-) биссектрисы линзы $\mathcal{L}(\alpha, -\beta)$ — дуги $\mathbf{A}(x; \gamma)$.

Доказательство. Пусть точка $V(x_1, y_1)$ — вершина кривой $y=f(x) \in \mathcal{V}^+(\alpha, \beta)$. Наклон касательной в вершине равен $\tau(x_1) = \tau_1$. Построим хорды, соединяющие вершину с концами отрезка $[-1, 1]$ (рис. 11). Кривизна возрастает на первой хорде с углом наклона μ_1 и убывает на второй хорде с углом наклона μ_2 . Поворот хорды в вершине равен $\mu_2 - \mu_1 = \rho$.

Вершине соответствует точка перегиба $S(x_1, z_1)$ графика $z(x)$, $z_1 = \sin \tau_1$. Рассматривая прямые AS и SB как функции $l_1(x)$ и $l_2(x)$ и учитывая характер выпуклости $z(x)$, получаем

$$y_1 \stackrel{(4)}{=} \int_{-1}^{x_1} \frac{z(u)}{\sqrt{1-z(u)^2}} du < \int_{-1}^{x_1} \frac{l_1(u)}{\sqrt{1-l_1(u)^2}} du = (1+x_1) \operatorname{tg} \frac{\alpha+\tau_1}{2},$$

$$\implies \frac{y_1}{1+x_1} = \operatorname{tg} \mu_1 < \operatorname{tg} \frac{\alpha+\tau_1}{2},$$

т.е. $\tau_1 > 2\mu_1 - \alpha$. Аналогично на участке SB получим $\tau_1 < 2\mu_2 - \beta$, из чего следует

$$2\mu_1 - \alpha < \tau_1 < 2\mu_2 - \beta \implies \rho > \frac{\beta - \alpha}{2} = -\gamma.$$

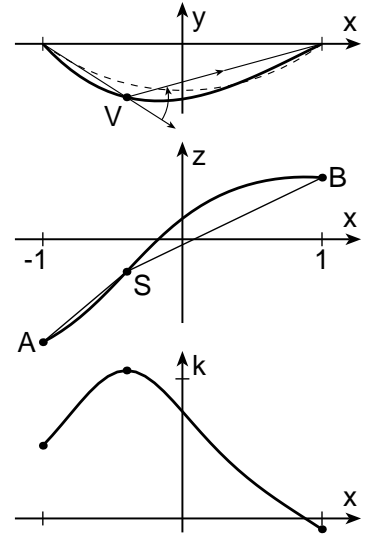


Рис. 11.

⁸Доказательство достаточности условий (49) для построения кривой класса \mathcal{V}^\pm мы опускаем.

Геометрическое место точек, для которых $\rho = -\gamma$, есть дуга $\mathbf{A}(x; \gamma)$, а условие $\rho > -\gamma$ удовлетворяется для всех точек, лежащих ниже этой дуги. \blacklozenge

Из неравенств (9) и (49) следует интересный вывод, касающийся некоторых методов интерполяции кривых. Для оценки углов наклона касательных и кривизн в j -том узле дискретно заданной кривой иногда используются параметры окружности, проходящей через узлы $j-1$, j и $j+1$. Тогда граничные углы α , β и нормированные кривизны k_1 , k_2 интерполирующей кривой на участке (jk) равны

$$\alpha = \xi_j, \quad \beta = \eta_j, \quad k_1 = -\sin \alpha = q_j c_j, \quad k_2 = \sin \beta = q_{j+1} c_j. \quad (50)$$

Вершины — явление достаточно редкое, и большинство хорд исходной кривой заключают участки с монотонной кривизной. Но — *при граничных условиях (50) невозможно построение ни кривой класса \mathcal{M} , ни кривой класса \mathcal{V} ; у интерполирующей кривой неизбежно появление двух вершин на каждой хорде.*

7. Обсуждение и выводы

Основные цели данной работы — сама постановка задачи, привлечение внимания к ее инженерным аспектам и обоснование заявленного во введении геометрического подхода к ее решению. Этот подход апеллирует к характеру натурального уравнения кривой. Оценка точности представления кривой не связывается с каким-либо конкретным алгоритмом, *что дает способ тестирования самих алгоритмов интерполяции.*

Что касается конкретных решений задачи, то в этом смысле мы не считаем исследование завершенным. Полностью сформулированы два возможных решения — (43) и (47). При этом в случае произвольной гладкой кривой требуется ее разбиение на спиральные участки с заданием в них граничных условий. Ослабление этого требования может выглядеть как:

- обязательное наличие и маркировка узлов в вершинах кривой без задания в них граничных условий (47);
- наличие на каждом участке монотонности кривизны $N \geq 3$ узлов без обязательного узла в вершине;
- снижение минимального значения N до 2, что осложнит идентификацию участков кривой в плане их принадлежности к классам \mathcal{M} или \mathcal{V} , но, видимо, оставит возможность получения эффективных оценок.

Несоблюдение условий последнего варианта означает некорректное задание кривой.

Специалисты в области компьютерной графики очень ценят аффинно инвариантные алгоритмы интерполяции, в связи с чем в обсуждениях этой статьи нередко звучал вопрос: является ли предложенная методика аффинно инвариантной? Ответ *“нет, не является”* не достаточен. Более существенно то, что этот вопрос нам *неинтересен*: для рассматриваемых задач, привязанных к инженерии и быту обычного евклидова пространства, все, выходящее за рамки поворота, переноса и

гомометрии является *искажением*. Наиболее веским доводом может служить поездка на велосипеде, колесо которого подверглось аффинно-ударному преобразованию. При этом оно приобрело как минимум 4 вершины [5, стр. 191].

Обсудим в заключение некоторые производственные аспекты проблемы. Недостатком предложенной методики является жесткая дисциплина описания кривой. Однако, по мнению автора, основанному на реальной практике допускового контроля, этот недостаток компенсируется сильным и полезным качеством — возможностью получения оценок (1). Они являются неизбежными составляющими погрешности измерения отклонений реального профиля от номинального и тем самым отражают корректность данного представления кривой.

Стандартизация предложенной методики дала бы участникам производства критерии для разрешения спорных ситуаций и поиска вызвавших их причин. Пока рассматриваемая проблема остается белым пятном в технологии машиностроения.

* * *

Благодарности. Зарождению изложенных здесь идей способствовали работы по окружностной интерполяции В.А. Леуса (Институт математики Сибирского отделения РАН). В поисках формулировок и доказательств автор обратился к теории выпуклых множеств. Почти все путеводные книги оказались написанными, переведенными, либо отредактированными В.А. Залгаллером (Математический институт им. Стеклова, Санкт-Петербург), личное знакомство и переписка с которым вдохновили дальнейшие исследования.

Список литературы

- [1] Курносенко А.И. Метролог, конструктор и метрологичность конструкций: Препринт ИФВЭ 44-96, Протвино, 1996 (Измерительная техника, №4, 1998, стр. 14–19).
- [2] Завьялов Ю.С., Леус В.А., Скороспелов В.А. Сплайны в инженерной геометрии. — М.: Машиностроение, 1985.
- [3] Леус В.А. Гладкая окружностная интерполяция кривых. — В сб.: Вычислительные системы, №38, Новосибирск, 1970.
- [4] Hoschek J. Circular splines. Computer-aided Design, Vol. 24, №11, 1992.
- [5] Бляшке В. Круг и шар. — М.: Наука, 1967.

Рукопись поступила 10 февраля 1998 г.

Содержание

1. Происхождение и постановка задачи	1
2. Обозначения и определения	3
3. Задача о сопряжении двух окружностей	5
4. Локальные свойства спиральных кривых	9
5. Глобальные оценки детерминированности кривой	16
6. Кривые с единственной вершиной	21
7. Обсуждение и выводы	23
Список литературы	24

А.И. Курносенко.

Интерполяционные свойства плоских спиральных кривых.

Оригинал-макет подготовлен с помощью системы \LaTeX .

Редактор Н.В. Ежела.

Подписано к печати 11.02.98. Формат $60 \times 84/8$.

Офсетная печать. Печ.л. 3,0. Уч.-изд.л. 2,3. Тираж 150. Заказ 97.

Индекс 3649. ЛР №020498 17.04.97.

ГНЦ РФ Институт физики высоких энергий
142284, Протвино Московской обл.

