

- цилиндрическая СК со сферической СК:

$$\begin{cases} \rho = r \sin(\psi), & y = r \cos(\psi), \\ r = \sqrt{\rho^2 + y^2}, & \operatorname{tg}(\psi) = \frac{\rho}{y}. \end{cases} \quad (1.3, \text{в})$$

Для единообразия терминов и алгоритмов мы будем работать в правой декартовой СК, оси которой ориентированы в пространстве так, что вращения ортов $x^0 \rightarrow y^0$, $y^0 \rightarrow z^0$ и $z^0 \rightarrow x^0$ происходят в положительном направлении, т. е. против часовой стрелки с точки зрения наблюдателя, находящегося на конце третьего свободного орта.

Взаимосвязь координат графического элемента может быть описана *неявным уравнением $f(p) = 0$ и параметрической функцией $p(t)$* . Частным случаем неявного описания является явная форма, в которой одна из координат выражена через другие. Это возможно лишь для некоторых простейших линий и поверхностей. Не все графические элементы могут быть описаны неявными функциями. Наиболее универсальным является параметрическое описание каждой из координат графического элемента соответствующей функцией времени, угла поворота или других параметров.

Целью настоящего краткого экскурса в вычислительную геометрию, не претендующего на исчерпывающую полноту и строгость, является систематизация формул, необходимых для работы с графическими объектами, приведение их к *алгоритмическому* виду. Это отличает компьютерную графику от обычной, в которой свойства и взаимосвязь графических объектов визуально воспринимаются глазами человека и интеллектуально интерпретируются его мозгом. С методами решения задач классической геометрии можно познакомиться, например, в [6, 7].

Компьютерная или вычислительная геометрия оперирует не с рисунками, а с числами (координатами точек, векторов, матрицами преобразований, списками данных) и отношениями между ними, выраженными также в цифровой форме. Исходная информация, обработанная специальными алгоритмами, преобразуется в другие цифровые данные, интерпретируемые как искомый результат — модель объекта, его экранное изображение, отношения между объектами сцены и т. п. Подробное изложение основ вычислительной геометрии приведено в [24, 29].

Рассмотрим основные соотношения между графическими элементами на плоскости R^2 и в пространстве RA . Все основные формулы будут записаны в координатной и векторной формах. Векторная форма компактнее координатной и предпочтительнее для понимания и программирования алгоритмов, хотя и требует несколько больших машинных ресурсов — объема памяти и времени счета. Увеличение ресурсов объясняется наличием в векторах и матрицах фиксированных элементов (единиц и нулей), с которыми иногда выполняются бесполезные арифметические операции.

Выше приведены основные обозначения и соотношения векторной алгебры. Самые трудоемкие матричные операции — обращение квадратной $n \times n$ -матрицы

M и расчет ее определителя — рекомендуется выполнять с максимально возможной точностью простым в программировании рекуррентным алгоритмом Леве-рье—Фаддеева [М]:

$$\begin{cases} L_1 = E_n, & \alpha_i = \frac{\text{tr}(ML_i)}{i}, & L_{i+1} = ML_i - \alpha_i E_n \quad \forall i = \overline{1, n}, \\ M^{-1} = \frac{L_n}{\alpha_n}, & |M| = (-1)^{n-1} \alpha_n. \end{cases} \quad (1.4)$$

1.1. Графические элементы на плоскости

Точка на плоскости имеет две степени свободы. Расстояние между двумя точками $p_1 = [x_1 \ y_1] \equiv [r_1 \ \varphi_1]$ и $p_2 = [x_2 \ y_2] \equiv [r_2 \ \varphi_2]$ равно:

$$d = \begin{cases} |p_1 - p_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} & \text{— в декартовой СК,} \\ \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\varphi_1 - \varphi_2)} & \text{— в полярной СК.} \end{cases} \quad (1.5)$$

Линия на плоскости имеет одну степень свободы. Уравнения линии в неявной (НФ) и параметрической (ПФ) формах имеют следующий вид:

$$\text{НФ:} \quad f(x, y) = 0 \text{ или } f(p) = 0; \quad (1.6, a)$$

$$\text{ПФ:} \quad p(t) = [x(t) \ y(t)]. \quad (1.6, б)$$

В любой регулярной (гладкой и некротной) точке на линии $p_0 = [x_0 \ y_0] = p(t_0)$ возможна *линеаризация* кривой (рис. 1.2), т. е. проведение к ней *касательной прямой*, уравнения которой имеют вид:

$$\text{НФ:} \quad \begin{cases} N_x(x - x_0) + N_y(y - y_0) = 0 \\ N_x = \left. \frac{df(x, y)}{dx} \right|_{x_0, y_0}, \quad N_y = \left. \frac{df(x, y)}{dy} \right|_{x_0, y_0} \end{cases} \text{ или } \begin{cases} N \circ (p - p_0) = 0; \\ N = \nabla f(p) \Big|_{p_0}; \end{cases} \quad (1.7, a)$$

$$\text{ПФ:} \quad \begin{cases} x(t) = x_0 + V_x t, \quad y(t) = y_0 + V_y t \\ V_x = \left. \frac{dx(t)}{dt} \right|_{x_0, y_0}, \quad V_y = \left. \frac{dy(t)}{dt} \right|_{x_0, y_0} \end{cases} \text{ или } \begin{cases} p(t) = p_0 + Vt; \\ V = \left. \frac{dp(t)}{dt} \right|_{p_0}. \end{cases} \quad (1.7, б)$$

Вектор нормали $N = [N_x \ N_y]$ ортогонален к линии и по свойству градиента направлен в ту сторону, где $f(p) > 0$. При замене уравнения линии на $-f(p) = 0$ направление нормали меняется на противоположное.

Направляющий вектор линии $V = [V_x \ V_y]$ начинается в точке p_0 и направлен по касательной к $p(t)$ в сторону *увеличения* t . Для перехода от одной формы (1.7) к

другой необходимо выбрать одно из двух ортогональных направлений нормали относительно направляющего вектора. Из условия ортогональности

$$N \circ V = 0 \Rightarrow N_x V_x + N_y V_y = 0$$

$$N \circ V = 0 \Rightarrow N_x V_x \pm N_y V_y = 0$$

получаем способ быстрого преобразования $V \rightarrow N$ или $N \rightarrow V$:

- координаты исходного вектора переставляются местами;
- у одной из координат полученного вектора меняется знак.

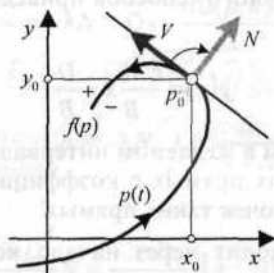


Рис. 1.2

Например, если дан вектор $V = [1 \ 2]$, то после перестановки координат он становится равным $[2 \ 1]$, а после изменения знака, например, у первой координаты получим вектор $N = [-2 \ 1]$, направленный влево от V . Проверка показывает, что векторы V и N ортогональны:

$$N \circ V = (-2) \cdot 1 + 1 \cdot 2 = 0.$$

Для единообразия изложения материала примем *правую* ориентацию нормали, означающую, что при движении точки по линии вектор N направлен перпендикулярно вправо от вектора V . Таким образом, связь векторов нормали и направления прямой будет следующей:

$$N = [V_y \ -V_x], \quad V = [-N_y \ N_x]. \quad (1.8)$$

1.1.1. Модели прямой линии на плоскости

Рассмотрим способы задания прямой на плоскости в различных формах.

1.1.1.1. Неявное уравнение прямой

Неявное уравнение прямой задается тремя коэффициентами A , B и D , составляющими вектор $F = [A \ B \ D]$.

$$\text{НФ:} \quad Ax + By + D = 0 \text{ или } [p \ 1] \circ F = \tilde{p} \circ F = 0. \quad (1.9)$$

Хотя бы одно из чисел A или B должно быть ненулевым. Если оба коэффициента $A \neq 0$ и $B \neq 0$, то прямая проходит наклонно к осям координат и пересекается с ними в следующих точках (рис. 1.3, а):

$$\left[\frac{-D}{A} \quad 0 \right], \left[0 \quad \frac{-D}{B} \right].$$

При $A = 0$ и $B \neq 0$ уравнение $Bu + D = 0$ описывает горизонтальную прямую, а уравнению $Ax + D = 0$ при $A \neq 0$ и $B = 0$ соответствует вертикальная прямая. Отметим ограниченность "школьного" способа приведения неявного координатного уравнения к явному виду

$$y(x) = -\frac{A}{B}x - \frac{D}{B}$$

и построения линии по точкам в желаемом интервале $x \in [x_{\min}, x_{\max}]$. Этот способ неприменим для вертикальных прямых с коэффициентом $B = 0$ и единственной координатой $x = -D/A$ всех точек таких прямых.

Прямая $f(p) \equiv \vec{p} \circ F = 0$ проходит через начало координат при $f(O_2) = D = 0$. Благодаря ее свойству разделять пространство плоскости на два подпространства с противоположными знаками $f(p)$, неявное уравнение позволяет с минимальными вычислительными затратами (за два умножения и два сложения) определить по условию $f(q) = 0$ принадлежность произвольной точки q прямой линии, а также ориентацию двух точек a и b относительно прямой (рис. 1.3, а):

- при $f(a)f(b) > 0$ точки лежат по одну сторону от прямой;
- при $f(a)f(b) < 0$ — по разные стороны от прямой.

Для построения прямой по уравнению $Ax + By + D = 0$ необходимо и достаточно иметь либо две несовпадающие точки p_0 и p_1 , через которые она проходит, либо точку p_0 и направляющий вектор V , с помощью которого вторая точка вычисляется как $p_1 = p_0 + V$. Сравнивая формулы (1.9) и (1.7, а), заключаем, что координатами вектора нормали можно принять коэффициенты неявного уравнения A и B :

$$N = [A \quad B]$$

Тогда по (1.8) получаем направляющий вектор неявной прямой

$$V = [-BA]$$

Рассмотрим два способа выбора точки $p_0 = [x_0 \quad y_0]$.

1. Первое, что приходит на ум — выбрать ее в том месте, где прямая пересекается с осью x при $A \neq 0$ или с осью y при $A = 0$ (только для горизонтальной прямой):

$$p_0 = \text{if} \left(A: \left[\frac{-D}{A} \quad 0 \right], \left[0 \quad \frac{-D}{B} \right] \right). \quad (1.10, a)$$

Несмотря на легкость понимания, этот способ не самый лучший для программирования из-за наличия альтернативы выбора точки в зависимости от коэффициента A .

2. Вычислим ближайшую к началу координат точку p_0 (рис. 1.3, а) методом наименьших квадратов (МНК) путем решения экстремальной задачи

$$|p|^2 = p \circ p \rightarrow \min$$

сограничением $p \circ N + D = 0$. Для этого составим функцию Лагранжа

$$\mathfrak{Z}(p, \lambda) = 0.5(p \circ p) + \lambda(p \circ N + D).$$

Из условий ее экстремума

$$\frac{\partial \mathfrak{Z}(p, \lambda)}{\partial p} = p + \lambda N = O_2, \quad \frac{\partial \mathfrak{Z}(p, \lambda)}{\partial \lambda} = p \circ N + D = 0$$

получим

$$p = -\lambda N, \quad \lambda = \frac{D}{N \circ N}$$

и оптимальное решение

$$p_0 = \frac{-D}{N \circ N} N = \frac{-D}{A^2 + B^2} [A \ B]. \quad (1.10, б)$$

Данное решение безальтернативно по сравнению с (1.10, а), более удобно в программировании и эффективнее в вычислении. Поэтому далее будет использоваться преимущественно оно.

Таким образом, координатная и векторная параметрические формы прямой, заданной вектором $F = [A \ B \ D]^T$, принимают вид

$$\begin{cases} x(t) = \frac{-AD}{A^2 + B^2} - Bt \\ y(t) = \frac{-BD}{A^2 + B^2} + At \end{cases} \quad \text{или} \quad p(t) = \frac{-D}{A^2 + B^2} [A \ B] + [-B \ A]t. \quad (1.11)$$

1.1.1.2. Нормальное уравнение прямой

Нормальное уравнение прямой (p_0, N), заданной точкой p_0 и вектором нормали N , выводится из условия ортогональности векторов N и $p - p_0$ для всех точек p , принадлежащих прямой (рис. 1.3, б):

НФ: $N_x(x - x_0) + N_y(y - y_0) = 0$ или $(p - p_0) \circ N = 0;$ (1.12, а)

ПФ: $x(t) = x_0 - N_y t, y(t) = y_0 + N_x t$ или $p(t) = p_0 + Vt,$ (1.12, б)

$$V = \begin{bmatrix} -N_y \\ N_x \end{bmatrix}.$$

Сравнивая (1.12, а) с (1.9), получим зависимость вектора неявного уравнения от параметров нормального уравнения:

$$F = [N \ -p_0 \circ N]. \quad (1.13)$$

Неявная функция

позволяет оценить положение точки p относительно прямой не только по знаку возвращаемого ей значения, но также и относительно направления нормали:

- при $f(a) > 0$ точка a лежит в том полупространстве, куда направлена нормаль, а угол $\langle a - p_0, N \rangle$ острый;
- при $f(b) < 0$ угол $\langle b - p_0, N \rangle$ тупой, а точка b и нормаль находятся по разные стороны от прямой (рис. 1.3, б).

1.1.1.3. Параметрическая функция прямой

Параметрическая функция прямой $\{p_0, V\}$, проходящей через точку p_0 в направлении вектора V (см. рис. 1.3, б), и эквивалентное ей уравнение в НФ будут иметь вид:

$$\text{ПФ: } x(t) = x_0 + V_x t, y(t) = y_0 + V_y t \text{ или } p(t) = p_0 + Vt; \quad (1.14, a)$$

$$\text{НФ: } V_y x - V_x y + D = 0, D = -V_y x_0 + V_x y_0 \Rightarrow F = [N - p_0 \circ N], \quad (1.14, б)$$

где $N = [V_y, -V_x]$. Параметрическая форма удобна для задания и построения частей прямой — отрезков и лучей. Для этого в (1.14, а) необходимо указать пределы изменения параметра t . Например:

- бесконечный интервал $-\infty < t < \infty$ не ограничивает протяженность бесконечной прямой;

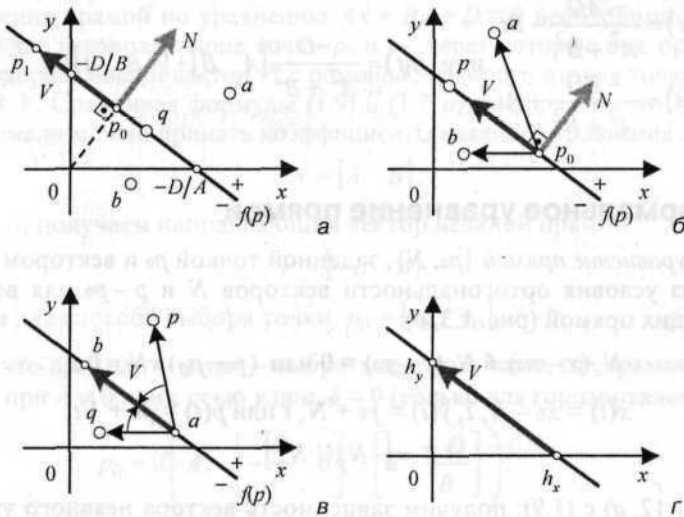


Рис. 1.3

- задание $t \geq 0$ дает луч, выходящий из точки p_0 в бесконечность в направлении вектора V ;
- конечный интервал $t_0 \leq t \leq t_1$ определяет отрезок прямой между точками $p_0 + Vt_0$ и $p_0 + Vt_1$.

Благодаря левой ориентации направляющего вектора V относительно вектора нормали N , эквивалентная нормальной форме функция

$$f(p) = (p - p_0) \cdot [V_y \quad -V_x] = \left| \begin{array}{c} p - p_0 \\ V \end{array} \right|$$

позволяет определить положение точки относительно направления движения по прямой:

- при $f(a) > 0$ точка a лежит справа, так что угол $\angle(a - p_0, V)$ положительный;
- при $f(b) < 0$ угол $\angle(b - p_0, V)$ отрицательный, а точка b лежит слева от прямой.

1.1.1.4. Уравнения прямой, проходящей через две точки

Неявная форма уравнения прямой, проходящей через две точки $a = [a_x \quad a_y]$ и $b = [b_x \quad b_y]$ (рис. 1.3, в), выводится из условия принадлежности прямой этих точек и точки $p = [x \quad y]$:

$$\text{НФ: } \begin{bmatrix} \tilde{p} \\ \tilde{b} \\ \tilde{a} \end{bmatrix} F^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{vmatrix} x & y & 1 \\ b_x & b_y & 1 \\ a_x & a_y & 1 \end{vmatrix} = 0 \Rightarrow \underbrace{\begin{vmatrix} b_y & 1 \\ a_y & 1 \end{vmatrix}}_{N_x} x + \underbrace{\begin{vmatrix} 1 & b_x \\ 1 & a_x \end{vmatrix}}_{N_y} y + \underbrace{\begin{vmatrix} b_x & b_y \\ a_x & a_y \end{vmatrix}}_D = 0. \quad (1.15, a)$$

Выбрав направление движения по прямой от точки a к точке b , получим направляющий вектор $V = b - a$ и параметрическую модель линии:

$$\text{ПФ: } x(t) = a_x + (b_x - a_x)t, y(t) = a_y + (b_y - a_y)t \text{ или } p(t) = a + (b - a)t. \quad (1.15, б)$$

Условие существования прямой очевидно: $|V| \neq 0$, т. е. $a \neq b$.

При изменении параметра от $t = 0$ до $t = 1$ движение точки происходит внутри отрезка ab от точки a до точки b . Ниже рассматриваются свойства плоских многоугольников (полигонов), определяемые ориентацией векторов нормалей к их сторонам. С учетом принятого соглашения (1.8) внешняя ориентация нормалей к сторонам полигона обеспечивается при его обходе против часовой стрелки.

Соответствующая данному способу задания прямой функция

$$f(p) = \left| \begin{array}{c} p - a \\ b - a \end{array} \right|$$

позволяет определить нахождение точки p относительно движения по прямой от a до b . При $f(p) > 0$ точка p лежит справа от прямой, а при $f(q) < 0$ точка q находится от нее слева.

1.1.1.5. Уравнения прямой в отрезках

Неявное уравнение прямой в отрезках h_x и h_y , одновременно не равных нулю и отсекаемых прямой на осях x и y (рис. 1.3, ε), получим из уравнения (1.9) при $D = -1$ с учетом соотношений $h_x = -D/A$ и $h_y = -D/B$:

$$\text{НФ: } \frac{x}{h_x} + \frac{y}{h_y} - 1 = 0 \Rightarrow h_y x + h_x y - h_x h_y = 0 \Rightarrow F = [h_y \quad h_x \quad -h_x h_y]. \quad (1.16, a)$$

Выбрав на прямой две точки $a = [h_x \quad 0]$, $b = [0 \quad h_y]$, получим ее направляющий вектор

$$V = b - a = [-h_x \quad h_y]$$

и координатные параметрические функции прямой в отрезках:

$$\text{ПФ: } x(t) = h_x(1 - t), \quad y(t) = h_y t. \quad (1.16, б)$$

1.1.2. Взаимное положение графических элементов на плоскости

Рассмотрим некоторые полезные формулы, тесты и алгоритмы, описывающие свойства и взаимное расположение точек, прямых и полигонов на плоскости. Для получения необходимых результатов активно используются тесты ортогональности и параллельности векторов, приведенные в перечне основных обозначений и соотношений.

1.1.2.1. Коллинеарность точек

$$\text{Три точки } P_1, P_2, P_3: \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 \text{ или } \begin{vmatrix} p_2 - p_1 \\ p_3 - p_1 \end{vmatrix} = 0. \quad (1.17, a)$$

$$\text{Обобщение (1.1)} \quad \text{rang} \begin{bmatrix} \tilde{p}_1 \\ \tilde{p}_2 \\ \dots \\ \tilde{p}_n \end{bmatrix} = 2 \text{ или } \text{rang} \begin{bmatrix} p_2 - p_1 \\ p_3 - p_1 \\ \dots \\ p_n - p_1 \end{bmatrix} = 1. \quad (1.17, б)$$

Точка p лежит на отрезке ab при нулевом угле между векторами $p - a$ и $b - p$:

$$(p - a) \circ (b - p) = |p - a| \cdot |b - p|. \quad (1.18)$$

Значение параметра t , соответствующее положению точки $p = a + (b - a)t$ на прямой относительно ее отрезка ab , вычисляется с помощью скалярного произведения:

$$(p - a) \circ (b - a) = (b - a)t \circ (b - a) = |b - a|^2 \cdot t,$$

откуда

$$t = \frac{(p - a) \circ (b - a)}{(b - a) \circ (b - a)}. \quad (1.19)$$

1.1.2.2. Взаимное расположение прямых

Две прямые совпадают (рис. 1.4, б), если

$$\text{НФ: } F_1 \times F_2 = 0; \quad \text{ПФ: } \left\{ \begin{array}{l} V_1 \\ V_2 \end{array} \right\} = 0 \cap \left\{ \begin{array}{l} V_1 \\ p_{20} - p_{10} \end{array} \right\} = 0. \quad (1.20)$$

Две прямые параллельны (рис. 1.4, в), если

$$\text{НФ: } \left\{ \begin{array}{l} N_1 \\ N_2 \end{array} \right\} = 0; \quad \text{ПФ: } \left\{ \begin{array}{l} V_1 \\ V_2 \end{array} \right\} = 0. \quad (1.21)$$

Расстояние между параллельными прямыми равно:

$$\text{ПФ: } d = \left| \text{pr}_{N_1}(p_{20} - p_{10}) \right| = \left| (p_{20} - p_{10}) \circ \bar{N}_1 \right|; \quad (1.22, a)$$

$$\begin{aligned} \text{ПФ: } d &= \left| \frac{p_{20} - p_{10}}{V_1} \right| = \left| p_{20} - p_{10} - \text{pr}_{V_1}(p_{20} - p_{10}) \right| = \\ &= \left| p_{20} - p_{10} - ((p_{20} - p_{10}) \circ \bar{V}_1) \bar{V}_1 \right|. \end{aligned} \quad (1.22, б)$$

Две прямые ортогональны (рис. 1.4, з), если

$$\text{НФ: } N_1 \circ N_2 = 0; \quad \text{ПФ: } V_1 \circ V_2 = 0. \quad (1.23)$$

1.1.2.3. Взаимное расположение точки и прямой

Уравнения перпендикуляра, опущенного из точки $q = [q_x \quad q_y]$ на прямую, заданную в НФ или ПФ (рис. 1.4, д), выглядят следующим образом:

$$\text{НФ: } N_y(x - q_x) - N_x(y - q_y) = 0 \Rightarrow N_{\perp} = [N_y \quad -N_x], \quad l_{\perp} = [N_{\perp} \quad -N_{\perp} \circ q]; \quad (1.24, a)$$

$$\text{ПФ: } p_{\perp}(t) = q + N t \text{ или } p_{\perp}(t) = q + V_{\perp} t, \text{ где } V_{\perp} = [V_y \quad -V_x]. \quad (1.24, б)$$

Расстояние от точки q до прямой равно:

$$\text{НФ:} \quad d = \frac{|\tilde{q} \circ F|}{|N|} = |\text{pr}_N(q - p_0)| = |(q - p_0) \circ \bar{N}|; \quad (1.25, a)$$

$$\text{ПФ:} \quad d = \left| \left| \frac{q - p_0}{V} \right| \right| = |q - p_0 - \text{pr}_V(q - p_0)| = |q - p_0 - ((q - p_0) \circ \bar{V})\bar{V}|. \quad (1.25, б)$$

Зеркальное отражение точки q от прямой лежит на перпендикуляре к прямой на расстоянии $2d$ от q в сторону, противоположную проекции вектора $q - p_0$ на нормаль N (рис. 1.4, д):

$$\text{НФ:} \quad q' = q - 2\text{pr}_N(q - p_0) = q - 2((q - p_0) \circ \bar{N})\bar{N}; \quad (1.26, a)$$

$$\text{ПФ:} \quad q' = q - 2(q - p_0 - \text{pr}_V(q - p_0)) = 2p_0 - q - 2((q - p_0) \circ \bar{V})\bar{V}. \quad (1.26, б)$$

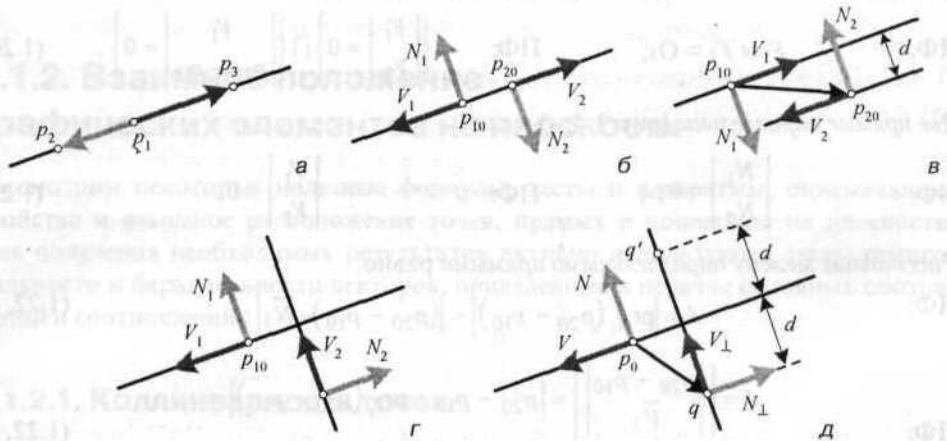


Рис. 1.4

Пример 1.1. Отобразить треугольник $p_1 p_2 p_3$ с вершинами $p_1 = [0 \ 0]$, $p_2 = [0 \ 1]$ и $p_3 = [2 \ 0]$ от прямой линии, проходящей через точки $a = [3 \ 0]$ и $b = [2 \ 2]$ (рис. 1.5).

Решение. Находим направляющий вектор прямой $V = b - a = [-1 \ 2]$, по формуле (1.8) нормаль $N = [2 \ 1]$ и квадрат ее длины $N \circ N = 5$. Пусть $p_0 = a$, тогда по (1.26, а) последовательно для точек p_1, p_2 и p_3 вычисляем отраженные точки

$$p'_i = p_i - \frac{2(p_i - a) \circ N}{N \circ N} N;$$

$$p'_1 = [0 \ 0] - 0.4([-3 \ 0] \circ [2 \ 1])[2 \ 1] = [4.8 \ 2.4],$$

$$p'_2 = [0 \ 1] - 0.4([-3 \ 1] \circ [2 \ 1])[2 \ 1] = [4 \ 3],$$

$$p'_3 = [2 \ 0] - 0.4([-1 \ 0] \circ [2 \ 1])[2 \ 1] = [3.6 \ 0.8]. \square$$

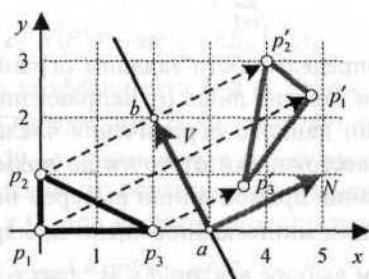


Рис. 1.5

1.1.2.4. Построение прямой, наименее удаленной от совокупности точек

Рассмотрим интересную задачу, связанную с понятием расстояния от точки до прямой линии. Допустим, в R^n (здесь при решении на плоскости xy будем полагать $n = 2$) задано $m \geq 2$ произвольных точек $p_1 \div p_m$, образуем из них матрицу

$$P = \begin{bmatrix} p_1 \\ \dots \\ p_m \end{bmatrix} \in R^{m \times n}. \quad (1.27)$$

Так как в общем случае точки не коллинеарны, то не существует ни одной прямой $f(p) = 0$, для которой все числа $d_i = f(p_i)$ обращаются в ноль. Другими словами, расстояния $|d_i|/N$ от точек p_i до прямой с нормалью N , вычисленные по (1.25, а) и пропорциональные d_i не могут быть одновременно все нулевыми (рис. 1.6).

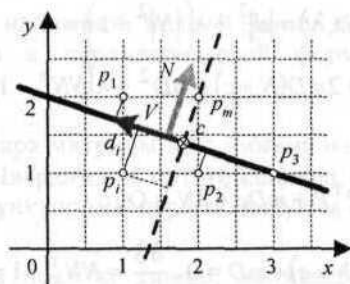


Рис. 1.6

Построим *аппроксимирующую прямую* $\{p_0, N\}$, *наименее удаленную от совокупности точек* p_i *путем решения экстремальной задачи*

$$\sum_{i=1}^m d_i^2 \rightarrow \min.$$

С целью уменьшения неопределенности зададим ограничение $|N|=1$, поскольку ориентация прямой линии зависит лишь от направления, а не от длины вектора нормали. При выполнении данного ограничения числа d_i , являются *алгебраическими*, т. е. со знаками, расстояниями от точек до прямой. Из пяти рассмотренных выше способов описания прямой линии выберем неявную форму (1.9) с вектором $F = [N \ D]$, имеющим минимальное число скалярных параметров, равное трем при $n = 2$. При таком выборе вектор $d \in R^m$ расстояний от точек до прямой принимает вид

$$d = F\tilde{P}^T = NP^T + D \underbrace{[1 \ \dots \ 1]}_m,$$

а квадрат его длины равен

$$\begin{aligned} |d|^2 &= dd^T = \left(NP^T + D [1 \ \dots \ 1] \right) \left(PN^T + D \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix} \right) = \\ &= NP^T PN^T + 2D \left(\sum_{i=1}^m p_i \right) N^T + mD^2 = NP^T PN^T + 2mD(N \circ c) + mD^2, \end{aligned}$$

где точка

$$c = \frac{1}{m} \sum_{i=1}^m p_i \quad (1.28)$$

есть *геометрический центр* массива точек p_i . Задача аппроксимации решается методом наименьших квадратов путем минимизации функции Лагранжа

$$\begin{aligned} \mathfrak{S}(N, D, \lambda) &= |d|^2 + \lambda \left(|N|^2 - 1 \right) = \\ &= NP^T PN^T + 2mD(N \circ c) + mD^2 + \lambda (NN^T - 1). \end{aligned} \quad (1.29, a)$$

Из условий ее экстремума

$$\begin{cases} \frac{1}{2} \frac{\partial \mathfrak{S}}{\partial N} = NP^T P + mDc + \lambda N = 0_n; \\ \frac{1}{2} \frac{\partial \mathfrak{S}}{\partial D} = m(N \circ c) + mD = 0, \quad \frac{\partial \mathfrak{S}}{\partial \lambda} = NN^T - 1 = 0 \end{cases} \quad (1.29, б)$$

и свойства произведения скалярного произведения двух векторов на третий вектор

$$(a \circ b) \cdot c = a \cdot (b^T c)$$

следует, что

$$D = -N \circ c, N(P^T P - mc^T c + \lambda E_n) = O_n. \quad (1.29, в)$$

Сравнение выражений для коэффициента D в (1.29, в) и (1.13) убеждает, что все прямые, доставляющие экстремум $\mathfrak{Z}(\bar{N}, \bar{D}, \bar{\lambda})$, проходят через точку $p_0 = c$.

Второе соотношение в (1.29, в) представляет собой однородную систему линейных алгебраических уравнений относительно координат вектора N . Она имеет нетривиальное решение $N \neq O_n$ при нулевом определителе матрицы $XE_n - M$, где

$$M = mc^T c - P^T P \in \mathbb{R}^{n \times n}. \quad (1.29, г)$$

Следовательно, множитель Лагранжа λ есть одно из n собственных чисел (значений), а N — соответствующий ему нормированный собственный вектор матрицы M . Количество собственных значений квадратной матрицы равно ее размерности. В случае $n = 2$ симметричная матрица M имеет две действительные собственные пары $\{\lambda_1, N_1\}$ и $\{\lambda_2, N_2\}$, причем $N_1 \perp N_2$ при $\lambda_1 \neq \lambda_2$ [см. 9, 13, 32]. Это означает, что прямые, доставляющие экстремум функции $\mathfrak{Z}(N, D, \lambda)$, ортогональны. Вычисление собственных чисел и векторов матрицы — хорошо отработанная процедура вычислительной математики, например, в [10, 27, 32]. Следует обратить внимание на нормирование собственных векторов в том случае, если алгоритм их вычисления не делает эту операцию сам.

Возникает закономерный вопрос: какой из полученных векторов N_1 или N_2 следует выбрать в качестве нормали искомой прямой? Ответ очевиден: тот, чья прямая доставляет минимум функции Лагранжа (1.29, я). Ведь условия (1.29, б) дают лишь экстремальные значения, которые при $D = -N \circ c$ и $|N| = 1$ равны

$$\mathfrak{Z} = NP^T P N^T - m(N \circ c)^2 = -NMN^T = -\lambda NN^T = -\lambda \geq 0.$$

Отсюда следует, во-первых, что все собственные числа матрицы M не положительны, во-вторых, что минимум функции Лагранжа доставляет собственная пара $\{\lambda, N\}$ с максимальным значением $\lambda = \max\{\lambda_1, \lambda_2\}$ и соответствующим ему вектором N — нормалью искомой прямой $\{c, M\}$.

Для удобства построения найденной прямой линии можно перейти от ее нормальной формы (1.12, а) к параметрической форме (1.14, а) с вектором $V = [-N_x, N_y]$.

Оставшаяся собственная пара матрицы M с минимальным собственным числом дает максимум функции Лагранжа и ортогональную (при $\lambda_1 \neq \lambda_2$) прямую (на рис. 1.6 она изображена пунктиром), пересекающуюся с оптимальной прямой в точке c .

☞ *Пример 1.2.* Построить прямую линию, наименее отклоненную от точек $p_1 = [1 \ 2]$, $p_2 = [2 \ 1]$, $p_3 = [3 \ 1]$, $p_4 = [1 \ 1]$ и $p_5 = [2 \ 2]$ (см. рис. 1.6).

Решение. Из заданных точек составим матрицу (1.27) и найдем их центр (1.28):

$$P = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix}, \quad c = \frac{p_1 + p_2 + p_3 + p_4 + p_5}{5} = [1.8 \quad 1.4].$$

Вычислим по (1.29, з) матрицу M и ее собственные параметры:

$$M = 5 \begin{bmatrix} 1.8 \\ 1.4 \end{bmatrix} [1.8 \quad 1.4] - \begin{bmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 1 & 1 & 1 & 2 \end{bmatrix}^T = \begin{bmatrix} -2.8 & 0.6 \\ 0.6 & -1.2 \end{bmatrix},$$

$$\lambda_1 = -3, \lambda_2 = -1, N_1 = [-0.949 \quad 0.316], N_2 = [0.316 \quad 0.949].$$

Отделив максимальное значение $\lambda = \lambda_2$, получим параметры неявной формы искомой прямой:

$$N = [0.316 \quad 0.949], D = -N \circ c = -1.897.$$

Таким образом, аппроксимирующая прямая описывается неявным уравнением

$$0.316x + 0.949y - 1.897 = 0$$

или параметрическими функциями

$$x(t) = 1.8 - 0.949t, \quad y(t) = 1.4 + 0.316t.$$

Алгебраические расстояния от точек до этой прямой равны $d_1 = 0.316$, $d_2 = -0.316$, $d_3 = 0$, $d_4 = -0.632$ и $d_5 = 0.632$, а минимум функции Лагранжа составляет $\mathfrak{I}_{\min} = 1$.

Для сравнения на рис. 1.6 построена пунктирная прямая, соответствующая собственной паре $\{\lambda_1, N_1\}$. Расстояния от нее до заданных точек равны $d_1 = 0.949$, $d_2 = -0.316$, $d_3 = -1.265$, $d_4 = 0.632$ и $d_5 = 0$, а значение функции Лагранжа $\mathfrak{I}_{\max} = 3 > \mathfrak{I}_{\min}$, что подтверждает правильность решения задачи. \square

1.1.2.5. Пересечение двух прямых

Важнейшей задачей геометрии на плоскости является расчет точки q пересечения двух прямых (рис. 1.7, а), который в зависимости от формы их задания выполняется следующим образом:

$$\text{НФ:} \quad q \begin{bmatrix} N_1^T & N_2^T \end{bmatrix} = [p_{10} \circ N_1 \quad p_{20} \circ N_2] \Rightarrow$$

$$\Rightarrow q = [p_{10} \circ N_1 \quad p_{20} \circ N_2] \begin{bmatrix} N_1^T & N_2^T \end{bmatrix}^{-1}; \quad (1.30, a)$$

НФ: $p_{10} + V_1 t_1 = p_{20} + V_2 t_2 \Rightarrow [t_1 \ t_2] = (p_{20} - p_{10}) \begin{bmatrix} V_1 \\ -V_2 \end{bmatrix}^{-1} \Rightarrow$ (1.30, б)

ПФ: $\Rightarrow q = p_{10} + V_1 t_1.$

Условия существования решений в (1.30)

$$|N_1^T \ N_2^T| \neq 0 \text{ или } \begin{vmatrix} V_1 \\ -V_2 \end{vmatrix} \neq 0$$

равносильны *непараллельности* прямых.

Введем функцию вычисления угла между векторами V и W

$$ang(V, W) = \text{if} \left(\left(d = \begin{vmatrix} V \\ W \end{vmatrix} \right) \neq 0 : \text{sgn}(d), 1 \right) \cdot \arccos \left(\frac{V \circ W}{|V| \cdot |W|} \right) \quad (1.31)$$

как угла *кратчайшего* вращения от V к W . По свойству скалярного произведения векторов модуль угла равен $\arccos \left(\frac{V \circ W}{|V| \cdot |W|} \right) \in [0, \pi]$, а правильный знак определяет-ся по *ненулевому* знаку определителя d матрицы, построенной на этих векторах.

При $d=0$ векторы V и W линейно зависимы, а знак $\text{sgn}(d) = 0$ в (1.31) игнорируется. Это может быть в двух случаях:

- при *параллельности* векторов угол между ними нулевой, а $\arccos(1) = 0$;
- при *антипараллельности* векторов угол между ними развернутый, а по (1.31) получим $\arccos(-1) = \pi$.

Таким образом, *угол пересечения* двух прямых находится с помощью функции (1.31) как

$$\gamma = ang(N_1, N_2) \text{ или } \gamma = ang(V_1, V_2). \quad (1.32)$$

1.1.2.6. Уравнения пучка прямых и биссектриса угла

Уравнения пучка прямых, заключенных между двумя прямыми (рис. 1.7, б), выводятся следующим образом:

НФ: $\begin{cases} f_1(p) = \tilde{p} \circ F_1 \\ f_2(p) = \tilde{p} \circ F_2 \end{cases} \Rightarrow f_{\lambda+1}(p) = (1-\lambda)f_1(p) + \lambda f_2(p); \quad (1.33, а)$

ПФ: $\begin{cases} p_1(t) = p_{10} + V_1 t \\ p_2(t) = p_{20} + V_2 t \end{cases} \stackrel{(1.30, б)}{\Rightarrow} q \Rightarrow \begin{cases} p_{\lambda+1}(t) = q + V_{\lambda+1} t; \\ V_{\lambda+1} = (1-\lambda)V_1 + \lambda V_2. \end{cases} \quad (1.33, б)$

Изменение параметра пучка в интервале $0 \leq \lambda \leq 1$ дает такие промежуточные прямые, что вращение векторов $N_{\lambda+1}$ и $V_{\lambda+1}$ происходит по *кратчайшим* углам от начальных положений N_i и V_i до конечных положений N_2 и V_2 . Уравнение (1.33, б) можно также рассматривать как параметрическую (с параметром λ) мо-

дель прямой линии, соединяющей точки исходных прямых $p_1 = q + V_1 t$ и $p_2 = q + V_2 t$ при фиксированном значении параметра t .

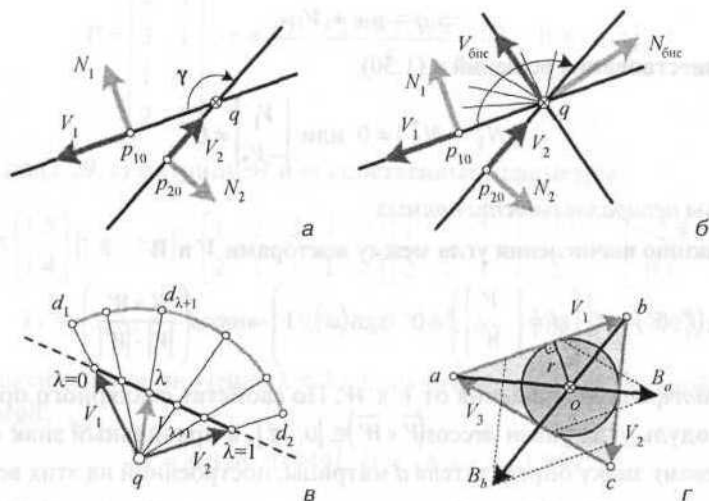


Рис. 1.7

На основе этой трактовки несложно и без вычисления углов решается задача построения дуги окружности радиуса r , заключенной в створе между векторами V_1 и V_2 , выходящими из точки q (рис. 1.7, в). Точки дуги $d_{\lambda+1}$, соответствующие параметру $\lambda \in [0, 1]$, вычисляются по формуле

$$d_{\lambda+1} = q + \overline{V}_{\lambda+1} r,$$

в которой нормируемый вектор $\overline{V}_{\lambda+1}$ находится по (1.33,6). Для изображения дуги в виде ломаной линии, состоящей из N отрезков между точками d_1 и d_2 , можно разбить диапазон изменения параметра λ на N равных интервалов (рис. 1.7, в построен при $N = 5$), сформировать набор значений $\lambda = i / N \forall i = 0, N$, рассчитать векторы $\overline{V}_{\lambda+1}$ и точки дуги $d_{\lambda+1}$, после чего соединить эти точки отрезками прямых.

Уравнения биссектрисы угла между двумя прямыми получим по уравнениям (1.33) с параметром $\lambda = 0.5$ только при $|N_1| = |N_2|$ или $|V_1| = |V_2|$, т. е. когда треугольник с вершиной в точке q и сторонами V_1 и V_2 является равнобедренным, у которого биссектриса совпадает с медианой. В противном случае можно искусственно создать равнобедренный треугольник со сторонами $|V_2|V_1$ и $|V_1|V_2$ равной длины $|V_1| \cdot |V_2|$. В результате параметры биссектрисы можно найти по формулам

$$F_{\text{бис}} = |N_2|F_1 + |N_1|F_2, \quad p_{\text{бис}}(t) = q + V_{\text{бис}} t, \quad V_{\text{бис}} = |V_2|V_1 + |V_1|V_2. \quad (1.34)$$

Расчет биссектрис бывает необходим, например, при построении окружности, вписанной в треугольник (рис. 1.7, г). Как известно, ее центр лежит в точке пересечения любых двух биссектрис *внутренних* углов этого треугольника. При построении биссектрисы внутреннего угла следует учитывать направления подставляемых в (1.34) векторов сторон треугольника: они должны либо оба выходить из вершины, либо оба в нее входить. При несоблюдении этого правила по (1.34) будет построена биссектриса дополнительного угла треугольника, а окружность окажется вневписанной.

Итак, допустим, что вершины треугольника находятся в точках a , b и c . Вычислив векторы сторон

$$V_1 = b - a, \quad V_2 = c - b, \quad V_3 = a - c,$$

найдем по (1.34) направляющие векторы биссектрис двух внутренних углов

$$B_a = |V_3|V_1 - |V_1|V_3, \quad B_b = |V_1|V_2 - |V_2|V_1.$$

Тогда центр вписанной окружности o лежит на пересечении параметрически заданных прямых $\{a, B_j\}$ и $\{b, B_i\}$ и может быть вычислен методом (1.30,6), а ее радиус r есть расстояние от точки o до прямой $\{a, V_j\}$, вычисляемое по (1.25, б).

1.1.2.7. Тесты свойств графических элементов на плоскости

Рассмотрим вначале тесты, основанные на *свойстве сепарабельности* прямой $f(p) = 0$ разделять плоскость на два подпространства с противоположными знаками функции $f(p)$.

- **Тест ориентации точки $p = [x \ y]$ относительно прямой** заключается в проверке знака числа $f(p)$, для вычисления которого в зависимости от способа задания прямой *в разд. 1.1.1* используется соответствующая функция nf^* , где "*" означает варьируемый третий символ "F", "N", "V или "2" в имени функции:

$$nfF(F, p) = \tilde{p} \circ F; \quad (1.35, a)$$

$$nfN(p_0, N, p) = (p - p_0) \circ N; \quad (1.35, б)$$

$$nfV(p_0, V, p) = \left| \frac{p - p_0}{V} \right|; \quad (1.35, в)$$

$$nf2(a, b, p) = \left| \frac{p - a}{b - a} \right|. \quad (1.35, г)$$

Все точки q справа от прямой, т. е. со стороны ее нормали, дают значения каждой из этих функций $nf^* > 0$, а слева — $nf^* < 0$. Для доказательства данного факта подставим, например, в функцию (1.35,б) точку $q = p_0 + Vt + \lambda N$, правую при $\lambda > 0$ и векторе $V \perp N$ (рис. 1.8):

$$nfN(p_0, N, q) = (Vt + \lambda N) \circ N = (V \circ N)t + \lambda(N \circ N) = 0 \cdot t + \lambda|N|^2 > 0.$$

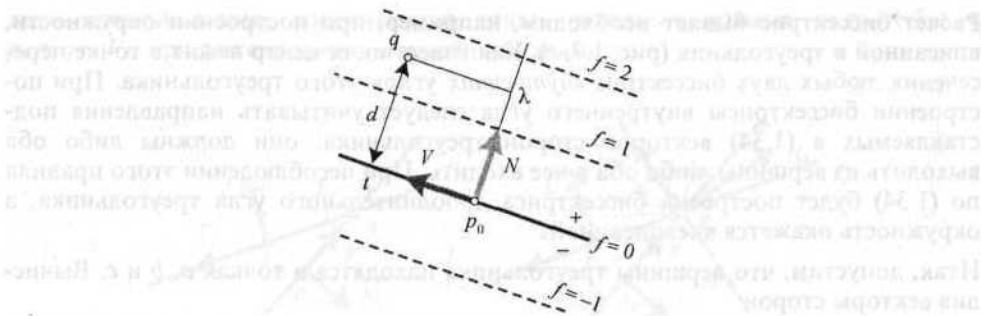


Рис. 1.8

Так как значение $f(q)$ не зависит от коэффициента t , отсчитываемого вдоль направляющего вектора V , то линии постоянных уровней $f(p)$ проходят параллельно прямой и также являются прямыми. Как следствие, получим расстояние от точки q до прямой, равное

$$d = |\lambda|N| = \frac{|f(q)|}{|N|}$$

и совпадающее с (1.25, а).

Ввиду простоты вычисления функция $f(p)$, реализуемая на практике одной из функций ny^* в (1.35), идеально подходит для сортировки точек относительно прямой линии на правые, левые и лежащие на ней. Тест ориентации точки относительно прямой применяется в алгоритмах, связанных с обработкой точек, вершин полигонов, а также во многих других задачах вычислительной геометрии.

- Тест пересечения прямой линии $f(p) = 0$ с полигоном $P = \{p_1, p_2, \dots, p_n, p_1\}$, имеющим произвольное направление обхода вершин, выполняется с помощью следующей функции:

$$cross2(f, P) = \{\exists 1 \leq j < i \leq n : f(p_i) \cdot f(p_j) < 0\}. \quad (1.36)$$

Прямая пересекает полигон, если существует хотя бы одна пара вершин, лежащих от нее по разные стороны (рис. 1.9, а).

Обсудим некоторые подходы к реализации теста в виде конкретного алгоритма, стремясь получить его наиболее эффективным в смысле минимальной трудоемкости.

Следует ясно понимать, что нужно сравнивать положение относительно, прямой линии не только смежных вершин полигона p_i и p_{i+1} , а всех различных пар вершин $\{p_i, p_j\}$, т. к. прямая может, пересекая полигон только в вершинах, не пересечь ни одного ребра внутри (рис. 1.9, б). Свойство сепарабельности позволяет по разным знакам чисел $f(p_i)$ и $f(p_j)$ обнаружить вершины p_i и p_j лежащие по разные стороны от пересекающей полигон прямой. Если же не

найдется ни одной пары разносторонних вершин, то прямая и полигон не пересекаются.

Число всех пар из n вершин полигона равно $n(n + 1) / 2$. Если решать задачу алгоритмизации теста (1.36) "в лоб" (с двумя вычислениями $f(p)$ для каждой пары вершин), то при больших значениях n (например, когда некоторая замкнутая кривая аппроксимируется полигоном с числом сторон порядка нескольких тысяч) трудоемкость теста будет пропорциональной n^2 , т. е. недопустимо большой. Уменьшения этой зависимости до линейной можно достичь с помощью специальных приемов. Рассмотрим некоторые из них:

- во внешнем цикле $i = \overline{1, n}$ вычисляем значения $s_i = f(p_i)$. Во внутреннем цикле $j = \overline{i+1, n} \forall i > 1$ анализируем знак произведения $s_i s_j$ (заметим, что число $s_j = f(p_j)$ уже вычислено). При $s_i s_j < 0$ тест немедленно завершается, возвращая 1 — признак пересечения прямой с полигоном. После нормального окончания внешнего цикла алгоритм возвращает 0 — признак отсутствия пересечения либо касания прямой с полигоном;

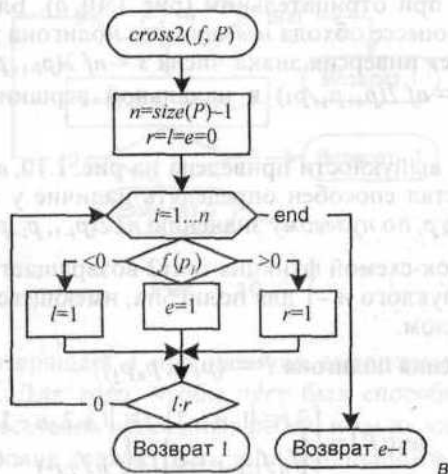
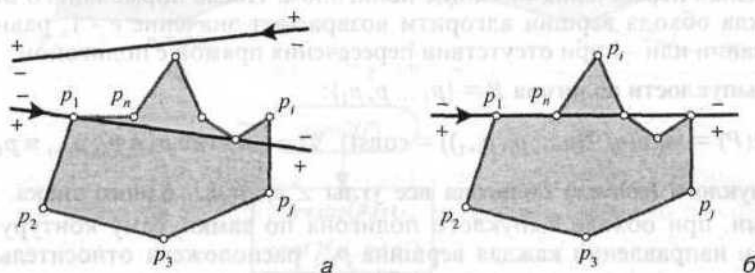


Рис. 1.9

- в цикле $i = \overline{1, n}$ вычислим значения $S_j = f(p_i)$, а затем сделаем сортировку вектора s . Результат тестирования возвращается числом $res = \text{sgn}(s_1, s_n)$. При $res = -1$ крайние элементы s_1 и s_n имеют разные знаки, что означает пересечение прямой с полигоном. По сравнению с предыдущим алгоритмом данный вариант теста способен различить касание ($res = 0$) и отсутствие пересечения прямой с полигоном ($res = 1$). В качестве недостатка отметим наличие операции сортировки, трудоемкость лучших методов которой оценивается зависимостью $n \cdot \log(n) < n^2$ [13, 24];
- наилучший, по мнению автора, алгоритм теста пересечения представлен блок-схемой на рис. 1.9, в и использует три флажка, сигнализирующих о наличии вершин полигона слева (флажок l), справа (флажок r) и точно на прямой (флажок e). Вначале все флажки сбрасываем в ноль. В цикле $i = \overline{1, n}$ вычисляем значение $s = f(p_i)$ и устанавливаем в единицу один из флажков: l при $s < 0$, e при $s = 0$ или r при $s > 0$. Тут же по условию $l \cdot r \neq 0$ проверяем наличие *любой* пары вершин полигона по разные стороны от прямой. Если это условие выполняется, то тест немедленно завершается, возвращая 1 — признак пересечения прямой с полигоном. После нормального окончания цикла обхода вершин алгоритм возвращает значение $e - 1$, равное 0 при касании или -1 при отсутствии пересечения прямой с полигоном.

- Тест выпуклости полигона $P = \{p_1 \dots p_n p_1\}$

$$\text{conv2}(P) = \{\text{sgn}(nf2(p_{i-1}, p_i, p_{i+1})) = \text{const}\} \quad \forall i = \overline{1, n}, \text{ где } p_0 \equiv p_n, p_{n+1} \equiv p_1. \quad (1.37)$$

У **выпуклого** (*convex*) полигона все углы $\angle p_{i-1} p_i p_{i+1}$ одного знака. Другими словами, при обходе выпуклого полигона по замкнутому контуру в *произвольном* направлении каждая вершина p_{i+1} расположена относительно ребра $p_{i-1} p_i$ одинаково для всех значений i : слева при положительном направлении обхода и справа при отрицательном (рис. 1.10, о). Благодаря свойству сепарабельности в процессе обхода *невыпуклого* полигона рано или поздно обязательно произойдет инверсия знака числа $s = nf2(p_{i-1}, p_i, p_{i+1})$ по сравнению со знаком числа $c = nf2(p_0, p_1, p_2)$ в начальной вершине p_1 , как показано на рис. 1.10, б.

Блок-схема теста выпуклости приведена на рис. 1.10, в. По сравнению с (1.37) алгоритм *conv2* стал способен определять наличие у полигона развернутого угла при вершине p_i по *нулевому* значению $nf2(p_{i-1}, p_i, p_{i+1}) = 0$.

Определенная блок-схемой функция *conv2* возвращает 1 для выпуклого полигона, 0 для невыпуклого и -1 для полигона, имеющего хотя бы одну вершину с развернутым углом.

- Тест самопересечения полигона $P = \{p_1, \dots, p_n p_1\}$

$$\text{self_test}(P) = \left\{ \begin{array}{l} \exists i \in [1, n-2], j \in [i+2, n-1 + \{i > 1\}] : \\ \{f_{ij} f_{i, j+1} < 0\} \cap \{f_{ji} f_{j, i+1} < 0\} \end{array} \right\},$$

где $n > 3$, а $f_{ij} = nf2(p_i, p_{i+1}, p_j)$ — значение НФ (1.35, г) прямой отрезка $p_i p_{i+1}$ в точке p_j . Полигон является самопересекающейся замкнутой ломаной линией,

если у него существует хотя бы одна пара пересекающихся отрезков. Два отрезка $P_i P_{i+1}$ и $P_j P_{j+1}$ пересекаются друг с другом, если концы одного находятся по разные стороны от прямой другого и наоборот. Интервалы изменения индексов i и j выбраны так, чтобы были протестированы все пары не смежных ребер полигона:

- индекс i перебирается в интервале от 1 до $n - 2$;
- при $i = 1$ индексу перебирается в интервале от 3 до $n - 1$;
- при $i > 1$ индексу перебирается в интервале от $i + 2$ до n .

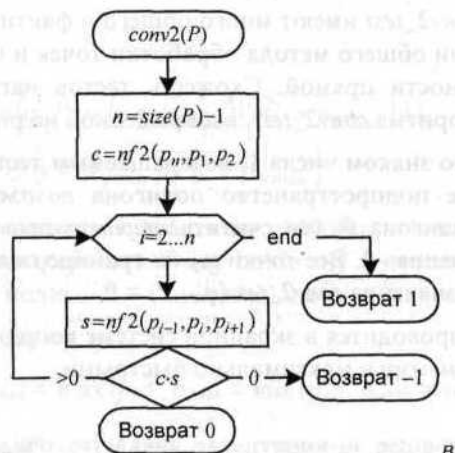
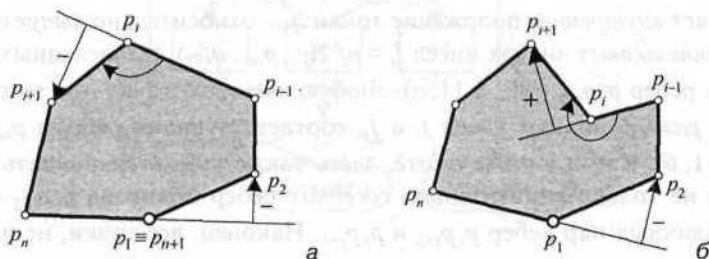


Рис. 1.10

Функция *self_test* возвращает 1 при наличии самопересечения полигона и 0 при его отсутствии. Для того чтобы тест был способен обнаруживать не только "чистые" пересечения несмежных ребер, но и их касания, используемое в нем логическое условие должно быть записано немного сложнее:

$$\{f_{ij} f_{i,j+1} < 0\} \cap \{f_{ji} f_{j,j+1} \leq 0\} \cup \{f_{ij} f_{i,j+1} \leq 0\} \cap \{f_{ji} f_{j,j+1} < 0\}.$$

1.1.2.8. Тесты ориентации точки относительно полигона

Рассмотрим тесты ориентации точки $q = [x \ y]$ относительно полигона $P = \{P_1, \dots, P_m, P_1\}$.

II Выпуклый тест

$$\text{conv2_test}(q, P) = \begin{cases} -1, & \text{если } \text{sgn}(f_i) = \text{const} \quad \forall i = \overline{1, n}; \\ 1, & \text{если } \exists 1 \leq j < i \leq n: f_j f_i < 0; \\ 0, & \text{если не } \pm 1 \end{cases} \quad (1.38)$$

определяет внутреннее положение точки $q_{\text{вну}}$ относительно выпуклого полигона при одинаковых знаках чисел $f_i = \text{nf2}(p_i, p_{i+1}, q_{\text{вну}})$, вычисленных по (1.35, з) для всех ребер p_i, p_{i+1} (рис. 1.11, а). Любую внешнюю точку $q_{\text{вне}}$ тест обнаруживает по разным знакам чисел f_i и f_j соответствующих ребрам p_i, p_{i+1} и p_j, p_{j+1} (рис. 1.11, б). Как и в тесте *cross2*, здесь также нужно сравнивать положение точки q не только относительно соседних ребер полигона p_i, p_{i+1} и p_{i+1}, p_{i+2} , а всех различных пар ребер p_i, p_{i+1} и p_j, p_{j+1} . Наконец, все точки, не идентифицированные как внутренние или внешние, считаются граничными точками $q_{\text{гп}}$.

Тесты *cross2* и *conv2_test* имеют много общего и фактически являются двойственными формами общего метода обработки точек и прямых на основе свойства сепарабельности прямой. Схожесть тестов наглядно иллюстрируется блок-схемой алгоритма *conv2_test*, изображенной на рис. 1.11, в.

В соответствии со знаком числа 1, возвращаемым тестом для внешних точек, назовем внешнее подпространство полигона *положительным*. Внутреннее пространство полигона будем считать *отрицательным* согласно знаку возвращаемого значения -1. Все точки $q_{\text{гп}}$ на границе полигона (в вершинах или на ребрах) дают значение $\text{conv2_test}(q_{\text{гп}}, P) = 0$.

Если тест (1.38) проводится в экранной системе координат, то все расчеты являются *целочисленными* и максимально быстрыми.

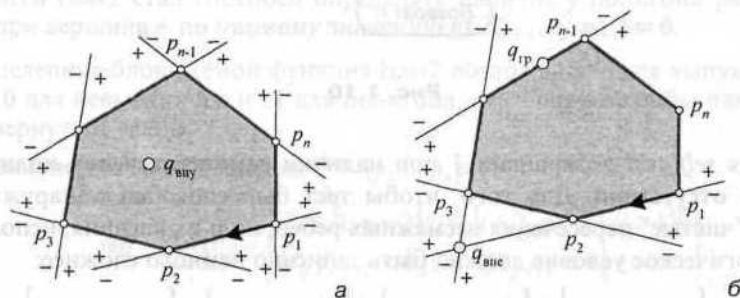


Рис. 1.11, а, б

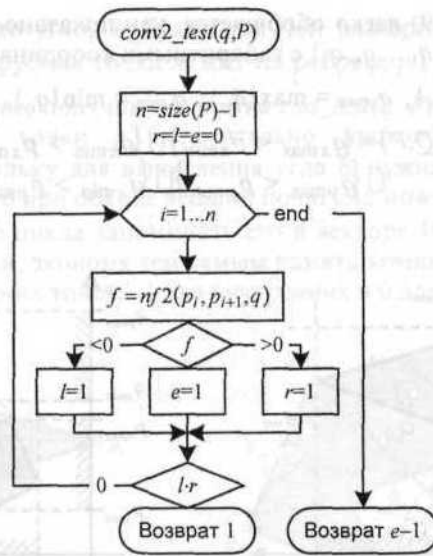


Рис. 1.11, в

• Габаритный тест

$$\begin{aligned}
 gab2_test(q, P) = & \{x < p_{x\min}\} \cup \{x > p_{x\max}\} \cup \\
 & \cup \{y < p_{y\min}\} \cup \{y > p_{y\max}\}
 \end{aligned}
 \quad (1.39, a)$$

определяет по возвращаемому значению $gab2_test = 1$ гарантированную непринадлежность точки q произвольному полигону P путем сравнения ее координат с габаритами полигона — минимальными и максимальными координатами его вершин $p_i \forall i = \overline{1, n}$:

$$p_{x\min} = \min\{p_{ix}\}, p_{x\max} = \max\{p_{ix}\}, p_{y\min} = \min\{p_{iy}\}, p_{y\max} = \max\{p_{iy}\}. \quad (1.39, б)$$

Полностью габаритный тест задачу ориентации не решает, возвращая значение 1 только для тех внешних точек ($q_{\text{вне}}$, которые расположены вне габаритного прямоугольника (рис. 1.12, а). Результат $gab2_test = 0$ означает, что точка находится внутри габаритного прямоугольника, но не может быть однозначно идентифицирована как внешняя, внутренняя либо граничная. Тем не менее, благодаря своей простоте габаритный тест применяется во многих алгоритмах для быстрого обнаружения заведомо непересекающихся геометрических объектов, что освобождает от необходимости решения трудоемких задач внешнего и внутреннего отсека.

Точечный тест (1.39) легко обобщается, как показано на рис. 1.12,6, на случай полигона $Q = \{q_1, \dots, q_n, q_1\}$ с габаритными координатами

$$q_{x,\min} = \min\{q_{ix}\}, \quad q_{x,\max} = \max\{q_{ix}\}, \quad q_{y,\min} = \min\{q_{iy}\}, \quad q_{y,\max} = \max\{q_{iy}\};$$

$$gab2_poly(Q, P) = \{q_{x,\max} < p_{x,\min}\} \cup \{q_{x,\min} > p_{x,\max}\} \cup \\ \cup \{q_{y,\max} < p_{y,\min}\} \cup \{q_{y,\min} > p_{y,\max}\}. \quad (1.40)$$

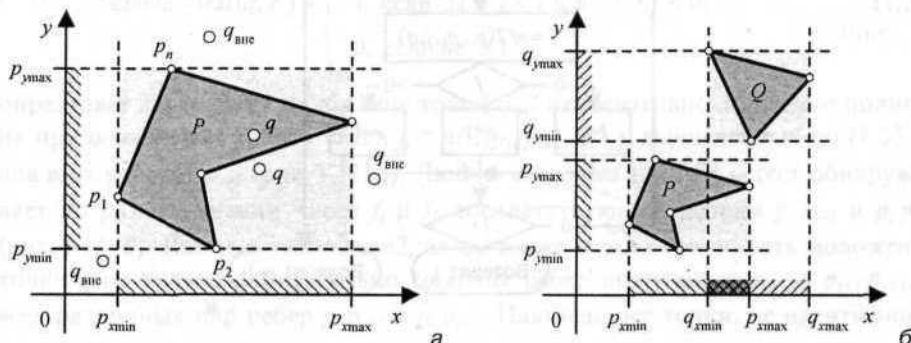


Рис. 1.12

- **Угловой тест** основан на вычислении и анализе алгебраической суммы углов $\delta_i = \langle V_i, V_{i+1} \rangle$ между смежными векторами $V_i = p_i - q$, соединяющими точку q с вершинами p_i , при обходе произвольного полигона P по замкнутому контуру в произвольном направлении. Тест основан на следующем опыте: наблюдатель, просматривающий вершины полигона из внутренней точки $q_{\text{вне}}$, совершает вокруг себя полный оборот (рис. 1.13, а), а из любой внешней точки $q_{\text{вне}}$ — ни одного оборота (рис. 1.13, б).

Вспользуемся для вычисления угла между двумя векторами функцией ang , определенной в (1.31) и возвращающей радианную меру угла. Тогда при всех ненулевых векторах V_i радианный вариант углового теста примет вид

$$\left| \sum_{i=1}^n ang(V_i, V_{i+1}) \right| = \begin{cases} 0 < \pi \Rightarrow q_{\text{вне}} \notin P; \\ 2\pi > \pi \Rightarrow q_{\text{вне}} \in P. \end{cases} \quad (1.41)$$

В связи с неизбежными погрешностями, возникающими при суммировании углов, вычисляемых с помощью приближенной функции ang , в (1.41) нельзя сравнивать результат ни с нулем, ни с 2π : он всегда будет хоть на немного, но отличаться от этих чисел. Предлагается взять пороговое значение, равное n : оно ни при каких погрешностях не даст неверного результата тестирования.

Граничная точка $q_{\text{г}}$ обнаруживается по следующим признакам:

- если при расчете векторов будет получен нулевой вектор длиной $V_i = 0$, то тестируемая точка совпадает с вершиной p_i ;

• если при расчете углов δ_i будет получен развернутый угол с модулем $|\delta_i| = \pi$, то тестируемая точка лежит на ребре $P_i P_{i+1}$ (рис. 1.13, б).

Резюмируя сказанное, получим алгоритм *rad_test*(q, P) радианного углового теста ориентации точки q относительно произвольного полигона P (рис. 1.13, в). Поскольку для вычисления угла δ_i нужна лишь пара смежных векторов V_i и V_{i+1} , то при обходе вершин полигона можно вычислять текущий вектор V , а в конце цикла запоминать его в векторе W для использования в следующей итерации, экономя тем самым память компьютера. Алгоритм возвращает 1 для внешних точек, -1 для внутренних и 0 для граничных.

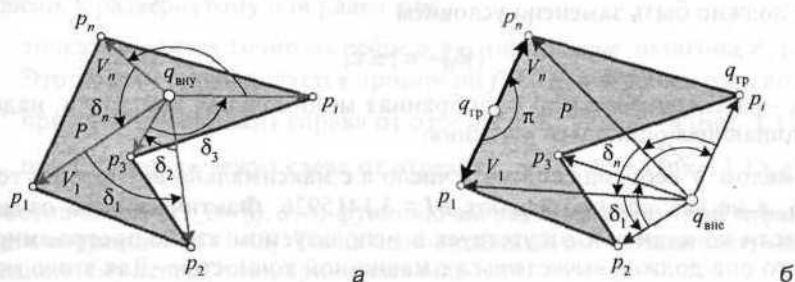


Рис. 1.13

Ввиду важности места, занимаемого угловым тестом в задачах компьютерной графики, обсудим как явные, так и скрытые *недостатки* алгоритма *rad_test*:

- необходимо нормировать все ненулевые векторы V_i , что связано с операциями вычисления квадратного корня и деления, которые относятся к не-

целочисленным операциям с *плавающей точкой*, выполняющимся гораздо медленнее операций с целыми числами;

- вычисление угла в (1.31) тригонометрической функцией \arccos производится путем суммирования бесконечного степенного ряда до сходимости по точности, предусмотренной компилятором используемого языка программирования. Этот итерационный процесс с плавающей точкой также не отличается высоким быстродействием;
- вычисления с плавающей точкой приводят к накоплению погрешностей при суммировании степенных рядов и углов в (1.41). В связи с этим обнаружение граничной точки $q \in p_i p_{i+1}$ с помощью сравнения $|\delta_i| = \pi$ на практике должно быть заменено условием

$$||\delta_i| - \pi| \leq \varepsilon,$$

где ε — экспериментально подобранная минимальная константа, надежно поглощающая ошибки округления;

- при малом ε необходимо иметь число с максимально возможной точностью, а не как обычно задают $PI = 3.1415926$. Фактически это означает, что если константа π отсутствует в используемом языке программирования, то она должна *вычисляться* с машинной точностью. Для этого можно воспользоваться абсолютно, но медленно сходящимся степенным рядом

$$4 \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \pi.$$

Модифицируем радианный угловой тест так, чтобы расчет углов δ_i с высокой точностью стал не обязателен. Будем считать углы в целочисленных октантах. Руководствуясь неравенствами на рис. 1.14, сопоставим вектору $V = [x \ y]$ *октантный угол* $1 \leq v \leq 8$ и функцию вычисления октантного угла $oct(V)$, возвращающую 0 при $x = y = 0$ и целое число $v \in \{1, 2, \dots, 8\}$ при $|V| \neq 0$. Как и в радианном тесте, получение нулевого значения $v_i = oct(p_i - q)$ означает принадлежность точки q *границе* полигона в его вершине p_i и немедленное завершение теста с возвращением нуля.

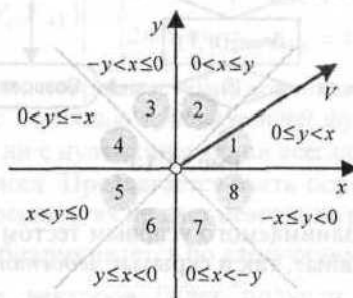


Рис. 1.14

При $q \neq p_i \forall i=1, n$ вычислим приращения октантных углов $\Delta_i = v_{i+1} - v_i$ на переходах $p_i \rightarrow p_{i+1}$. Для получения кратчайших угловых приращений необходима коррекция величин и знаков октантных углов с $|\Delta_i| \geq 4$:

• если $\Delta_i > 4$, то $\Delta_i = \Delta_i - 8$; (1.42, а)

• если $\Delta_i < -4$, то $\Delta_i = \Delta_i + 8$; (1.42, б)

• если $\Delta_i = \pm 4$ и $f = nf 2(p_i, p_{i+1}, q) \neq 0$, то $\Delta_i = -4 \text{sgn}(f)$. (1.42, в)

Поясним коррекцию в (1.42, в). Значение $|\Delta_i| = 4$ возможно в трех случаях расположения точки q в окрестности ребра $p_i p_{i+1}$, при которых угол $\angle p_i q p_{i+1}$ близок к развернутому или равен ему:

- точка находится точно на ребре $p_i p_{i+1}$ (на границе полигона P , рис. 1.15, а). Этот факт устанавливается проверкой $f = 0$ и завершает тест, возвращая 0;
- при $f > 0$ точка лежит справа от отрезка $p_i p_{i+1}$, а $\Delta_i = -4$ (рис. 1.15, б);
- при $f < 0$ точка лежит слева от отрезка $p_i p_{i+1}$, а $\Delta_i = 4$ (рис. 1.15, в).

Обозначим за $oct_cor(q, a, b)$ функцию вычисления октантной меры угла $\angle aqb$ и его коррекции по (1.42). Тогда целочисленный октантный тест, аналогичный радианному тесту (1.41), принимает вид

$$\sum_{i=1}^n oct_cor(q, p_i, p_{i+1}) = \begin{cases} 0 \Rightarrow q \notin P; \\ \pm 8 \neq 0 \Rightarrow q \in P. \end{cases} \quad (1.43)$$

Точка q является внешней относительно полигона P , если суммарный октантный угол поворота вектора V_i при обходе вершин p_i по замкнутому контуру равен нулю. Точка лежит внутри полигона, если этот угол ненулевой. Наконец, точка находится на границе полигона, если либо при $oct(p_i, q) = 0$ она совпадает с вершиной p_i , либо при $|\Delta_i| = 4$ и $nf 2(p_i, p_{i+1}, q) = 0$ она лежит на ребре $p_i p_{i+1}$.

Обозначим за $oct_test(q, P)$ функцию октантного углового теста ориентации точки q относительно плоского полигона $P = \{p_1 \dots p_n p_1\}$, возвращающую -1 для точки внутри полигона, 1 — вне полигона и 0 — на его границе. Блок-схема октантного теста построена на рис. 1.15, в. Пунктиром выделена часть алгоритма, в которой происходит вычисление и коррекция приращения октантного угла по (1.42).

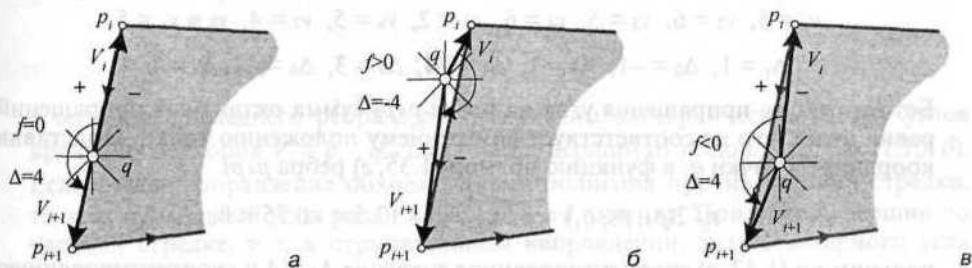


Рис. 1.15, а, б, в

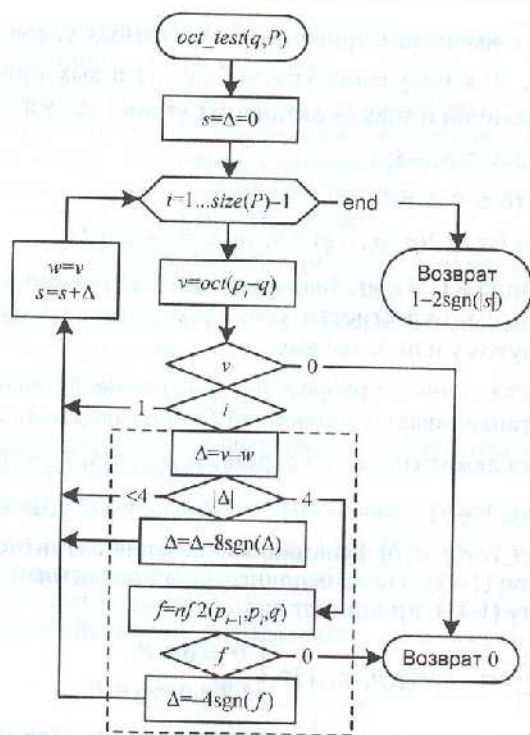


Рис. 1.15, г

Пример 1.3. Определить ориентацию точек $q_1 = [4.5 \ 3]$, $q_2 = [2.5 \ 1]$ и $q_3 = [1.5 \ 1.5]$ относительно невыпуклого полигона $P = \{p_1 p_2 \dots p_7\}$ с вершинами $p_1 = [1 \ 2]$, $p_2 = [2.5 \ 0.5]$, $p_3 = [2 \ 2.5]$, $p_4 = [4 \ 0.5]$, $p_5 = [5 \ 4]$, $p_6 = [3 \ 3]$ и $p_7 = [1.5 \ 4.5]$ (рис. 1.16).

Решение. Для первой тестируемой точки $q_1 = [x_1 \ y_1]$ вычислим октантные углы векторов $V_i = p_i - q_1$ и их приращения:

$$v_1 = 5, \ v_2 = 6, \ v_3 = 5, \ v_4 = 6, \ v_5 = 2, \ v_6 = 5, \ v_7 = 4, \ v_8 \equiv v_1 = 5;$$

$$\Delta_1 = 1, \ \Delta_2 = -1, \ \Delta_3 = 1, \ \Delta_4 = -4, \ \Delta_5 = 3, \ \Delta_6 = -1, \ \Delta_7 = 1.$$

Без коррекции приращения угла на ребре $p_4 p_5$ сумма октантных приращений равна нулю, что не соответствует внутреннему положению точки. Подставив координаты точки q_1 в функцию прямой (1.35, з) ребра $p_4 p_5$

$$nf\ 2(p_4, p_5, q_1) = 3.5x_1 - y_1 - 13.5 = -0.75 < 0,$$

получим по (1.42, в) скорректированное значение $\Delta_4 = 4$ и скорректированную сумму $s = 8$. Согласно (1.43) точка q_1 лежит внутри полигона.

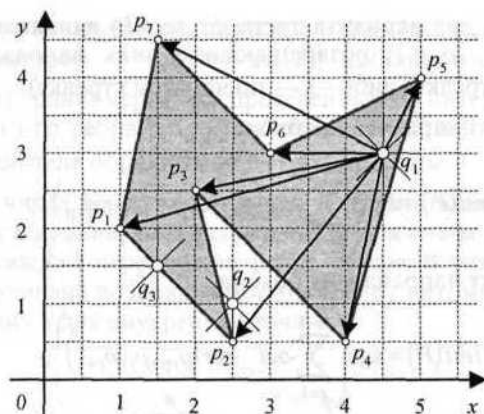


Рис. 1.16

Для второй точки q_2 октантные углы и их приращения равны:

$$v_1 = 4, v_2 = 7, v_3 = 3, v_4 = 8, v_5 = 2, v_6 = 2, v_7 = 3, v_8 \equiv v_1 = 4;$$

$$\Delta_1 = 3, \Delta_2 = -4, \Delta_3 = 5 \equiv -3, \Delta_4 = -6 \equiv 2, \Delta_5 = 0, \Delta_6 = 1, \Delta_7 = 1.$$

Значение функции прямой для ребра p_2p_3 в точке q_2

$$nf 2(p_2, p_3, q_2) = 2x_2 + 0.5y_2 - 5.25 = 7.75 > 0$$

положительно, поэтому $\Delta_2 = -4$, а сумма приращений октантных углов $5 = 0$. Из (1.43) следует, что точка q_2 является внешней.

Наконец, для точки q_3 на первом же ребре p_1, P_2 получим приращение октантного угла $\Delta_1 = 8 - 4 = 4$. Подставив координаты этой точки в функцию ребра $P_1 P_2$

$$nf 2(p_1, p_2, q_3) = 1.5x_3 + 1.5y_3 - 4.5 = 0,$$

закключаем, что точка q_3 принадлежит границе полигона. \square

С помощью суммирования углов можно алгоритмически определить направление обхода вершин полигона. Установим *границную* точку

$$q_{гр} = \frac{p_1 + p_n}{2}$$

в середине последнего ребра p_n, p_1 и вычислим алгебраическую сумму углов вращения векторов $V_i = p_i - q_{гр}$ при обходе вершин от p_1 , до p_n (см. рис. 1.13, б). Если общее направление обхода вершин полигона против часовой стрелки, то этот суммарный угол равен π [рад] или $+4$ октанта. При обходе вершин по часовой стрелке, т. е. в отрицательном направлении, знак суммарного угла отрицателен.

Отсюда получаем два варианта теста $dir_test(P)$ направления обхода вершин полигона $P = \{p_1, \dots, p_n, p_1\}$, возвращающего знак направления обхода (1 – против часовой стрелки, либо -1 – по часовой стрелке):

- радианный тест направления обхода

$$dir_test(P) = \operatorname{sgn} \left(\sum_{i=1}^{n-1} \operatorname{ang}(p_i - q_{\text{тр}}, p_{i+1} - q_{\text{тр}}) \right); \quad (1.44, a)$$

- октантный тест направления обхода

$$dir_test(P) = \operatorname{sgn} \left(\sum_{i=1}^{n-1} \operatorname{oct_cor}(q_{\text{тр}}, p_i, p_{i+1}) \right), \quad (1.44, б)$$

гораздо более эффективный, чем радианный тест, благодаря целочисленным и не итерационным вычислениям.

- Лучевой тест ориентации точки q относительно произвольного полигона P заключается в выпускании из этой точки в произвольном направлении V луча $p(t) = q + Vt \forall t \geq 0$ и подсчете числа его пересечений с ребрами полигона. Параметры пересечения луча с отрезками $p_i + (p_{i+1} - p_i)\tau \forall 0 \leq \tau \leq 1$, непараллельными вектору V , находятся по следующей из (1.30, б) формуле

$$[t_i \quad \tau_i] = (p_i - q) \begin{bmatrix} V \\ p_i - p_{i+1} \end{bmatrix}^{-1} \forall i = \overline{1, n}. \quad (1.45)$$

Анализ параметрических пар $\{t_i, \tau_i\}$ дает следующие критерии ориентации точки относительно полигона:

- точка является *внутренней*, если число пар $\{t_i > 0, 0 < \tau_i < 1\}$ нечетно (рис. 1.17, а);
- точка является *внешней*, если число таких пар четно, в том числе и равно нулю (рис. 1.17, б);
- точка лежит на *границе* полигона, если найдется хотя бы одна пара $\{t_i = 0, 0 \leq \tau_i \leq 1\}$ (рис. 1.17, в).

Особенности лучевого теста:

- неопределенность числа пересечений при прохождении луча точно через вершину p_i , при $\tau_i = 1$ или вершину p_{i+1} , при $\tau_i = 0$ (рис. 1.17, а, б). Необходимо повторить тест заново с другим направлением луча V ;
- требуется расчет по (1.45) (с плавающей точкой и вычислением обратной матрицы) параметров пересечений луча со всеми ребрами полигона.

Обозначим за $ray2_test(q, P)$ функцию лучевого теста ориентации точки q относительно плоского полигона P , возвращающую результат:

- -1, если точка находится внутри полигона;
- 0, если точка на границе полигона;
- 1, если точка вне полигона.