

Шаг 1. Рассчитываются все точки q'_i пересечения луча

$$p(t) = q + Vt, t \geq 0$$

с неограниченной поверхностью. Для этого решается скалярное уравнение

$$f(q + Vt) = 0$$

относительно параметра t , если поверхность задана в НФ уравнением $f(p) = 0$. Если она определена параметрической функцией $p(x, \theta)$, то решается векторное уравнение

$$q + Vt = p(\tau, \theta)$$

относительно трех параметров t , τ и θ . Отсутствие действительных решений t означает, что несущая луч (I) прямая не пересекает поверхность.

Шаг 2. Список полученных точек q'_i сортируется по возрастанию значений t_i и из него исключаются элементы с $t_i < 0$, не принадлежащие лучу. Если оставшийся список пуст, то луч (2) не пересекает поверхность.

Шаг 3. Ищется ближайший к началу списка элемент, для которого выполняются условия $\psi(q'_i) \geq 0$ или $\psi(\tau_i, \theta_i) \geq 0$ принадлежности точки q'_i ограниченной поверхности. Невыполнение любого из этих условий означает, что луч (5) не пересекает поверхность внутри ее границ. Найденный элемент является искомой точкой q' на луче (4). \square

Рассмотрим *вычислительные аспекты* расчета пересечения прямой с поверхностью. Точное аналитическое решение уравнений $f(q + Vi) = 0$ или $q + Vi = p(\tau, \theta)$ возможно лишь для некоторых простейших поверхностей первого (плоскость в НФ или ПФ) и второго (в НФ) порядков. Эти решения получены в (1.102) и 3d-варианте (1.54)–(1.56). Пересечение луча $q + Vi$ с более сложными поверхностями находится *численными методами*. Рассмотрим некоторые из них.

Корни скалярного уравнения $f(q + Vi) = 0$ можно найти любым известным численным методом решения нелинейных уравнений [13, 19]: проб, касательных, секущих, итераций и т. п. Простейший для программирования и не требующий вычисления производной df/dt метод проб заключается в трассировке луча от точки q в направлении вектора V с шагом t . В точках траектории

$$p_i = q + Vi = q + iV\Delta t, i = 0, 1, 2, \dots$$

неявная функция поверхности имеет значения $f_i = f(p_i)$. Первый же найденный интервал $[t_i, t_{i+1}]$, на котором выполняется условие $f_i f_{i+1} < 0$, пропорционально дробится точкой

$$\tau^{(0)} = \frac{|f_{i+1}|t_i + |f_i|t_{i+1}}{|f_i| + |f_{i+1}|}$$

на два интервала $[t_i, \tau^{(0)})$ и $(\tau^{(0)}, t_{i+1}]$. Из них выбирается интервал $[t'_i, t'_{i+1}]$ с *разнозначными* концевыми значениями НФ, для которых $f(t'_i)f(t'_{i+1}) < 0$. Дихотомиче-

ский процесс дает последовательность значений $\{\tau^{(k)}\}$ и заканчивается при достижении желаемой точности ϵ ПО аргументу $|\tau^{(k+1)} - \tau^{(k)}| \leq \epsilon$ или (и) функции $|f(\tau^{(k+1)}) - f(\tau^{(k)})| \leq \epsilon$.

Основные проблемы метода:

- компромиссный выбор шага Δt : достаточно малого, на котором не будет четного числа пересечений лучом поверхности, ограничивающей малый объем, и в то же время достаточно большого для получения приемлемого машинного времени расчета;
- необходимость ограничения пространства распространения луча габаритами сцены на случай, если пересечения не существует.

Параметрическая форма поверхности $p(\tau, \theta)$ не имеет такого индикатора пересечения ее лучом, каким в НФ является инверсия знака функции $f(p)$. Будем искать в пространстве параметров t, τ и θ корни уравнения $q + Vt = p(\tau, \theta)$, доставляющие *нулевое* значение целевой функции

$$J(t, \tau, \theta) = |q + Vt - p(\tau, \theta)|^2 \geq 0,$$

т. е. квадрату расстояния между лежащими на одной вертикали точками луча и поверхности.

Алгоритм МНК-решения задачи минимизации $J(t, \tau, \theta)$, стартуя с начальных приближений $t_0 = 0, \tau_0$ и θ_0 , обновляет параметры по итерационной формуле

$$[t_{i+1} \quad \tau_{i+1} \quad \theta_{i+1}] = [t_i \quad \tau_i \quad \theta_i] + [\Delta t_i \quad \Delta \tau_i \quad \Delta \theta_i], \quad i = 0, 1, 2, \dots \quad (1.137)$$

Различные методы оптимизации отличаются способами расчета вектора обновлений $[\Delta t_i \quad \Delta \tau_i \quad \Delta \theta_i]$. Метод **градиентного спуска** полагает

$$[\Delta t_i \quad \Delta \tau_i \quad \Delta \theta_i] = -\gamma_i \nabla J(t_i, \tau_i, \theta_i) \quad (1.138, a)$$

и имеет скорость сходимости, зависящую от выбора шага настройки γ , что само по себе представляет нетривиальную задачу. Квадратичной скоростью сходимости обладает метод **линеаризации**, заключающийся в замене по (1.61, б) поверхности в текущей точке $p(\tau_i, \theta_i)$ касательной плоскостью $p(\tau_i, \theta_i) + V_i \Delta \tau_i + W_i \Delta \theta_i$, расчете по (1.102, б) вектора обновлений и нового приближения точки пересечения

$$[\Delta t_i \quad \Delta \tau_i \quad \Delta \theta_i] = (p(\tau_i, \theta_i) - q - V t_i) \begin{bmatrix} V \\ -V_i \\ -W_i \end{bmatrix}^{-1} \Rightarrow p_{i+1} = p_i + V \Delta t_i \quad (1.138, б)$$

до сходимости по аргументу $||[\Delta t_i \quad \Delta \tau_i \quad \Delta \theta_i]|| \leq \epsilon$ или (и) функции $J(t_i, \tau_i, \theta_i) \leq \epsilon$.

Большое число поверхностей являются *периодическими* по одному или обоим параметрам τ и θ . Например, если период m функции $p(\tau, \theta)$ по параметру τ равен 2π , то

$$p(\tau, \theta) = p(\tau \pm mk, \theta) \quad \forall k = 1, 2, \dots$$

Формула обновления (1.137) может вывести параметр за допустимый интервал $0 \leq \tau_{\min} \leq \tau \leq \tau_{\max} \leq m$. Во избежание потери решения на третьем этапе алгоритма пересечения, где проверяются условия допустимости параметров $\psi(\tau_{n_i}, \theta_{n_i}) \geq 0$, необходимо после каждого обновления сдвигать значение τ_{i+1} на целое число периодов по формуле

$$\tau_{i+1} = \text{mod}(\text{mod}(\tau_{i+1}, m) + m, m) \in [0, m). \quad (1.139)$$

Итерационные методы оптимизации являются *локальными* (сходятся к локальному минимуму $J(t, \tau, \theta)$, причем необязательно нулевому), в связи с чем требуется выполнить несколько расчетов с различных начальных значений t_0 и θ_0 , например, равномерно распределенных в областях их допустимых значений. Полученные решения группируются на прямой $q + Vt$ в окрестностях точек ее пересечения с поверхностью. Далее выполняется сортировка, описанная во втором и третьем шагах алгоритма пересечения.

▣ *Пример 1.7.* Из точки $q = [10 \ 1 \ 0]$ выпущен луч $p(t) = q - Vt$ в направлении, задаваемом вектором $V = [-2 \ -0.1 \ -0.2]$. Рассчитать пересечение луча с поверхностью, описываемой явным уравнением

$$y = f(x, z) = e^{-a(x^2+z^2)} \cos(\omega_x x) \cos(\omega_z z)$$

с параметрами $a = 0.02$, $\omega_x = 1$ и $\omega_z = 0.5$ (рис. 1.53, а).

Решение. Так как уравнение поверхности задано в явной форме, будем решать задачу пересечения как задачу поиска корней функции

$$h(t) = y(t) - f(x(t), z(t)) = 1 - 0.1t - f(10 - 2t, -0.2t)$$

зависимости вертикального расстояния от точки $p(1) = [x(t) \ y(t) \ z(t)]$ до поверхности. Из ее графика, приведенного на рис. 1.53, б, видно, что луч пересекает поверхность в пяти точках. Ввиду некоторой сложности расчета производной dh/dt зададим шаг трассировки луча $\Delta t = 0.1$, допустимую погрешность $E = 10^{-6}$ и выполним поиск первой точки пересечения методом проб.

Реализация метода в среде MathCAD показала, что необходимо выполнить 45 шагов трассировки для достижения интервала $e \in [4.5, 4.6]$ с разнозначными значениями $h(4.5) = 0.0808$ и $h(4.6) = -0.066$. Пропорциональное дробление этого интервала еще за 4 итерации дало параметр пересечения $t = 4.553155$ с абсолютной погрешностью $h(t) = 5.208 \cdot 10^{-9} < 10^{-6}$ и точку

$$q' = [0.8936 \ 0.5447 \ -0.9106].$$

Заметим, что половинное дробление интервала $[4.5, 4.6]$ с той же погрешностью потребовало бы 14 итераций. Задание слишком большого шага трассировки $\Delta t > 1.1$ приводит к пропуску первых двух ближайших точек пересечения луча с поверхностью и получению сразу третьей точки $q'_3 = [-0.9861 \ 0.4507 \ -1.0986]$ с параметром $t_3 = 7.678274$. □

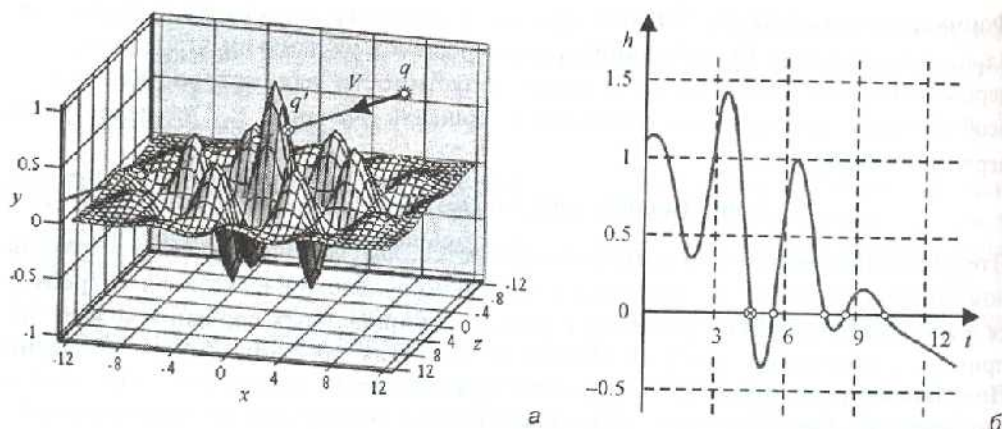


Рис. 1.53

Пример 1.8. Рассчитать пересечение луча, выпущенного из точки $q = [-8 \ 0 \ 8]$ в направлении $V = [4 \ 0.5 \ -4]$, с параметрической поверхностью (рис. 1.54, а)

$$\begin{cases} p(\tau, \theta) = [(1 + \tau - \sin(\tau)) \cos(\theta) \ 1 - \cos(\tau) \ -(\tau - \sin(\tau)) \sin(\theta)] \\ \forall 0.1\pi \leq \tau \leq 1.4\pi, \ 0 \leq \theta \leq 1.5\pi. \end{cases}$$

Решение. Луч $p(t) = [-8 + 4t \ 0.5t \ 8 - 4t]$, $t \geq 0$, пересекает поверхность $p(\tau, \theta)$ в трех точках. Найдем их методом линеаризации (1.137) и (1.138, б), минимизируя целевую функцию

$$J(t, \tau, \theta) = |p(t) - p(\tau, \theta)|^2$$

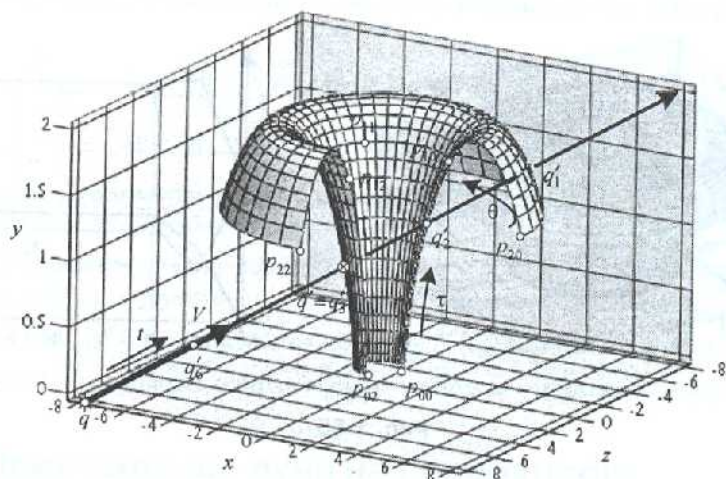
с начального приближения $t_0 = 0$ и девяти сочетаний начальных приближений

$$t_0 \in \{0.1\pi, 0.75\pi, 1.4\pi\}, \theta_0 \in \{0, 0.75\pi, 1.5\pi\}.$$

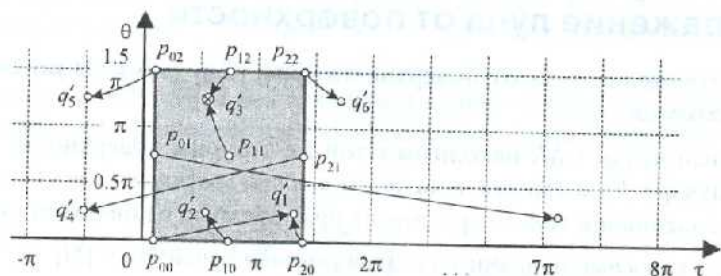
На рис. 1.54, б изображены пространство параметров τ и θ , допустимая область их изменения, начальные приближения $\{t_0, \theta_0\}$ и результаты сходимости алгоритма линеаризации с погрешностью $\varepsilon = 10^{-6}$:

- приближения $\{0.1\pi, 0\}$ и $\{1.4\pi, 0\}$ дают сходимость к точке q'_1 с параметрами $t_1 = 3.061$, $\tau_1 = 1.322\pi$, $\theta_1 = 0.25\pi$;
- приближение $\{0.75\pi, 0\}$ дает сходимость к точке q'_2 с параметрами $t_2 = 2.307$, $\tau_2 = 0.549\pi$, $\theta_2 = 0.25\pi$;
- приближение $\{0.1\pi, 0.75\pi\}$ дает лишь *относительную* сходимость к параметрам $t = 5.911$, $\tau = 7.128\pi$, $\theta = 0.25\pi$ и *ненулевому* установившемуся значению $J = 3.827 \neq 0$, в связи с чем этот вариант в список анализа не включается;

- приближения $\{0.75\pi, 0.75\pi\}$ и $\{0.75\pi, 1.5\pi\}$ дают сходимость к точке q'_3 с параметрами $t_3 = 1.744$, $\tau_3 = 0.459\pi$, $\theta_3 = 1.25\pi$;
- приближение $\{1.4\pi, 0.75\pi\}$ дает сходимость к точке q'_4 с параметрами $t_4 = 2.07$, $\tau_4 = -0.511\pi$, $\theta_4 = 0.25\pi$;
- приближение $\{0.1\pi, 1.5\pi\}$ дает сходимость к точке q'_5 с параметрами $t_5 = 1.917$, $\tau_5 = -0.487\pi$, $\theta_5 = 1.25\pi$;
- приближение $\{1.5\pi, 1.5\pi\}$ дает сходимость к точке q'_6 с параметрами $t_6 = 0.732$, $\tau_6 = 1.719\pi$, $\theta_6 = 1.25\pi$.



а



б

Рис. 1.54

Таким образом, сортируя точки q'_i по возрастанию параметра $t_i > 0$, получим упорядоченный список точек $L = \{q'_6, q'_3, q'_5, q'_4, q'_2, q'_1\}$, расположенных на луче $q + VI$ в порядке удаления от начала луча — точки q .

На третьем этапе исключим из списка L точки, параметры которых не удовлетворяют условиям $0.1\pi \leq \tau \leq 1.4\pi$ и $0 \leq \theta \leq 1.5\pi$. Это, во-первых, точка q'_6 пересечения луча с продолжением поверхности при $\tau_6 > 1.4\pi$, изображенной на рис. 1.55, а для интервалов параметров $0 \leq \tau \leq 3.4\pi$ и $0 \leq \theta \leq 2\pi$. Во-вторых, это точки q'_5 и q'_4 с отрицательными значениями t . Первая в оставшемся списке точка

$$q'_3 = q + 1.744V = [-1.026 \quad 0.872 \quad 1.026]$$

и является точкой q' пересечения луча с заданной поверхностью. На рис. 1.55,6 изображены графики итерационных процессов поиска параметров t_i , τ_i и θ_i . □

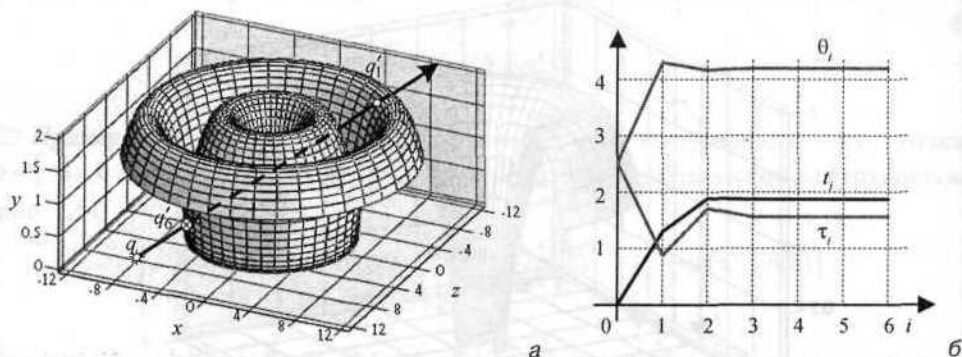


Рис. 1.55

1.3.2. Отражение луча от поверхности

Зеркальное отражение луча от поверхности (рис. 1.56) строится по следующим законам отражения.

- Отраженный луч $q' + Mt$ находится с той же стороны поверхности, что и падающий луч $q + Vt$, а эти два луча лежат в одной плоскости с вектором нормали N , построенным в точке пересечения прямого луча с поверхностью q' .
- Равенство скоростей падающего и отраженного лучей: $V = M$.
- Равенство углов падения α и отражения β .

Найдем нормальную и тангенциальную составляющие падающего луча:

$$V_{\perp} = \text{pr}_N V = |V| \cos(\angle VN) \bar{N} = |V| (\bar{V} \circ \bar{N}) \bar{N} = (V \circ \bar{N}) \bar{N},$$

$$V_{||} = V - V_{\perp} = V - (V \circ \bar{N}) \bar{N}.$$

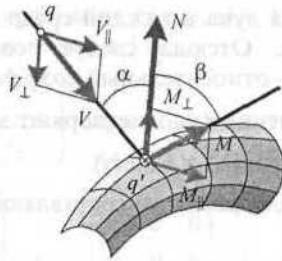


Рис. 1.56

На основе законов отражения вычислим составляющие и сам вектор отраженного луча:

$$M_{\perp} = -V_{\perp} = -(V \circ \bar{N})\bar{N}, \quad M_{\parallel} = V_{\parallel} = V - (V \circ \bar{N})\bar{N}, \quad (1.140)$$

$$M = M_{\perp} + M_{\parallel} = V - 2(V \circ \bar{N})\bar{N}.$$

Таким образом, алгоритм расчета отражения луча от зеркальной ограниченной поверхности имеет следующий вид.

Шаг 1. Решается задача пересечения луча $q + Vt$ с ограниченной поверхностью. Если точка q' отсутствует, то луч проходит мимо поверхности.

Шаг 2. По (1.61), (1.62) вычисляется нормаль N к поверхности в точке q' .

Шаг 3. По (1.140) находится направляющий вектор M отраженного луча $q' + Mt$. \square

1.3.3. Преломление луча на поверхности

Идеальное преломление луча на поверхности раздела двух сред с показателями преломления n_1 и n_2 (рис. 1.57) строится по следующим законам преломления.

- Преломленный луч $q' + Rt$ находится с другой стороны поверхности, чем падающий луч $q + Vt$, а эти два луча лежат в одной плоскости с вектором нормали N , построенным в точке пересечения прямого луча с поверхностью q' .

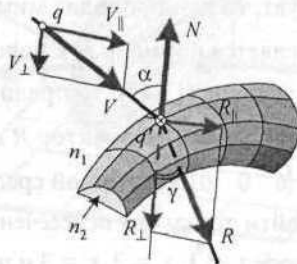


Рис. 1.57

• Скорость распространения луча в каждой среде *обратно пропорциональна* ее показателю преломления. Отсюда следует соотношение длин векторов $|R| = n_n |V|$, где $n_n = n_1 / n_2$ — относительный коэффициент преломления.

• Углы падения α и преломления γ удовлетворяют *закону Снеллиуса—Декарта*

$$n_1 \sin(\alpha) = n_2 \sin(\gamma). \quad (1.141)$$

Вычислим *нормальную* и *тангенциальную* составляющие вектора преломленного луча:

$$\begin{aligned} R_{\perp} &= |R| \cos(\gamma) \bar{V}_{\perp} = n_n |V| \sqrt{1 - n_n^2 \sin^2(\alpha)} \bar{V}_{\perp} = n_n |V| \sqrt{1 - n_n^2 |\bar{V} \times \bar{N}|^2} \frac{V \circ \bar{N}}{|V \circ \bar{N}|} \bar{N} = \\ &= \operatorname{sgn}(V \circ \bar{N}) n_n \sqrt{|V|^2 - n_n^2 |V \times \bar{N}|^2} \bar{N}, \\ R_{\parallel} &= |R| \sin(\gamma) \bar{V}_{\parallel} = n_n^2 |V| \sin(\alpha) \bar{V}_{\parallel} = n_n^2 V_{\parallel} = n_n^2 (V - (V \circ \bar{N}) \bar{N}), \end{aligned}$$

откуда следует вектор

$$R = R_{\perp} + R_{\parallel} = n_n^2 V + n_n \left(\operatorname{sgn}(V \circ \bar{N}) \sqrt{|V|^2 - n_n^2 |V \times \bar{N}|^2} - n_n V \circ \bar{N} \right) \bar{N}. \quad (1.142)$$

Существование преломленного луча в среде с меньшим показателем преломления $n_2 < n_1$ (например, при переходе из воды в воздух) возможно при положительности подкоренного выражения в (1.142), что совпадает с известным в оптике неравенством [16]:

$$|V|^2 - n_n^2 |V \times \bar{N}|^2 > 0 \Rightarrow |V| > n_n |V \times \bar{N}| \Rightarrow 1 > n_n \sin(\alpha). \quad (1.143)$$

При падении луча на оптически менее плотную среду под углом $\alpha \geq \arcsin(1/n_n)$ наблюдается оптический эффект *полного внутреннего отражения*, при котором преломленный луч отсутствует.

Таким образом, алгоритм преломления луча на **ограниченной** поверхности раздела сред имеет следующий вид.

☞ **Шаг 1.** Решается задача пересечения луча $q + V$ с ограниченной поверхностью. Если точка q' отсутствует, то луч проходит мимо поверхности.

Шаг 2. По (1.61), (1.62) вычисляется нормаль N к поверхности в точке q' .

Шаг 3. Если не выполняется условие (1.143), то преломленный луч отсутствует.

Шаг 4. По (1.142) находится направляющий вектор R отраженного луча $q' + Rt$. ☐

☞ *Пример 1.9.* Из точки $q = [6 \ 0 \ 0]$ воздушной среды ($n_1 = 1$) выпущен луч в направлении $V = [-4 \ 0 \ 1]$. Найти точку его пересечения со стеклянным ($n_2 \approx 1.52$) эллипсоидом, имеющим полуоси $r_x = 1$, $r_y = 2$, $r_z = 3$ и центр в точке $c = [0 \ -1 \ -1]$, а также направления отраженного и преломленного лучей (рис. 1.58).

Решение. Составим неявное уравнение $f(x, y, z) = 0$ и матрицу F эллипсоида по его каноническому уравнению из табл. 1.3:

$$\left(\frac{x-c_x}{r_x}\right)^2 + \left(\frac{y-c_y}{r_y}\right)^2 + \left(\frac{z-c_z}{r_z}\right)^2 - 1 = 0 \Rightarrow 36x^2 + 9y^2 + 4z^2 + 18y + 8z - 23 = 0,$$

$$F = \left[\begin{array}{ccc|c} 36 & 0 & 0 & 0 \\ 0 & 9 & 0 & 9 \\ 0 & 0 & 4 & 4 \\ \hline 0 & 9 & 4 & -23 \end{array} \right], F_{11} = \begin{bmatrix} 36 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{bmatrix}, F_{21} = [0 \ 9 \ 4], D = 23.$$

Вычислим коэффициенты и дискриминант квадратного уравнения (1.55):

$$a = 580, b = -860, c = 1273, d = 1260.$$

Так как $\{a \neq 0\} \cap \{d > 0\}$, то прямая $p(t) = q + Vi$ пересекается с поверхностью эллипсоида в двух различных точках с параметрами

$$t_1 = \frac{-b - \sqrt{d}}{a} = 1.422, t_2 = \frac{-b + \sqrt{d}}{a} = 1.544.$$

Ближайшая к источнику точка пересечения имеет координаты

$$q' = p(t_1) = [0.314 \ 0 \ 1.422].$$

Для расчета отраженного и преломленного лучей вычислим по (1.61, я) вектор нормали в точке падения:

$$N = [72x' \ 18y' + 18 \ 8z' + 8] = [22.59 \ 18 \ 19.37].$$

На рис. 1.58 этот вектор, имеющий длину $N = 34.78$, изображен для удобства в четыре раза короче.

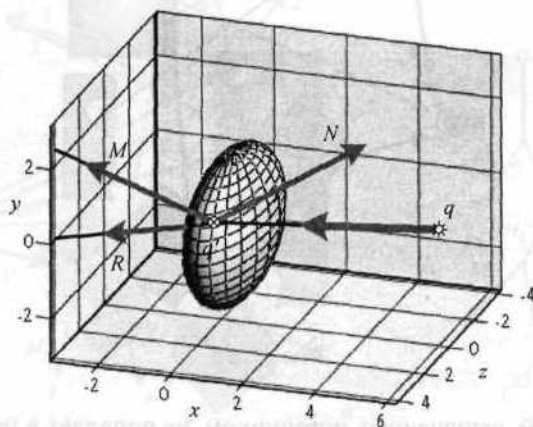


Рис. 1.58

По (1.140) и (1.142) определяем векторы направлений отраженного и преломленного лучей, предварительно вычислив значения

$$n_n = 1/1.52 \approx 0.658, \quad \bar{N} = [0.65 \quad 0.518 \quad 0.557], \quad V \circ \bar{N} = -2.041 < 0,$$

$$V \times \bar{N} = [-0.518 \quad 2.878 \quad -2.07], \quad |V \times \bar{N}| = 3.582, \quad |V|^2 - n_n^2 |V \times \bar{N}|^2 = 11.445 > 0:$$

$$M = V - 2(V \circ \bar{N})\bar{N} = [-1.348 \quad 2.113 \quad 3.274],$$

$$R = n_n^2 V - n_n \left(\sqrt{|V|^2 - n_n^2 |V \times \bar{N}|^2} + n_n V \circ \bar{N} \right) \bar{N} = [-2.603 \quad -0.695 \quad -0.315]. \quad \square$$

1.3.4. Прямая и обратная трассировка лучей

На основе *закона обратимости* можно построить путь светового луча как в *прямом* направлении от источника L к объекту и от него в приемник S , так и в *обратном* — от приемника к объекту и источнику. Каждый из методов трассировки лучей моделирует соответствующий процесс распространения света и имеет свои отличительные черты.

В *методе прямой трассировки* генерируется пучок лучей $L + V_{ij}$, выходящих из источника во всевозможных направлениях V_1, V_2, \dots . На рис. 1.59 показаны варианты распространения лучей в сцене, включающей отражающие, преломляющие и поглощающие свет объекты.

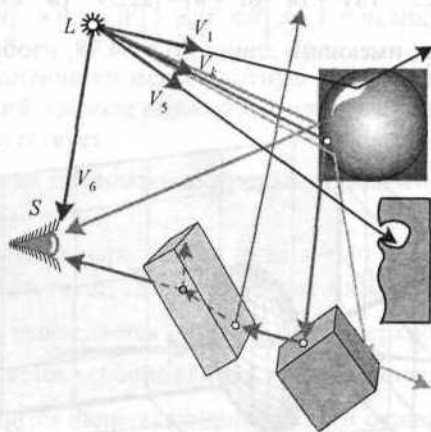


Рис. 1.59

Большинство лучей, испущенных источником, не попадает в приемник, а значит, и не влияет на формируемое в нем изображение. Лишь очень малая часть лучей

после всех отражений и преломлений в конце концов попадает в приемник, создавая изображение сцены в его рецепторах. На шероховатых поверхностях возникает множество диффузно отраженных лучей. Все их нужно программно генерировать и отслеживать, что лавинообразно увеличивает сложность задачи трассировки.

Прохождение луча в неидеальной среде сопровождается рассеянием и поглощением световой энергии на ее микрочастицах. Эти физические процессы чрезвычайно сложно адекватно моделировать на ЭВМ с ее конечными вычислительными ресурсами. Практически ограничиваются применением коэффициента затухания энергии луча на единицу пройденного им расстояния. Аналогично вводятся коэффициенты уменьшения энергии луча при его отражении и преломлении на поверхности раздела сред. С учетом этих коэффициентов отслеживается уменьшение энергии всех первичных и вторичных лучей в процессе их блуждания в пространстве сцены. Как только энергия некоторого луча становится меньше заданного абсолютного уровня или уменьшается в заданное число раз, трассировка данного луча прекращается.

Таким образом, главным недостатком метода прямой трассировки является его большая трудоемкость и малая эффективность. При реализации данного метода большая часть работы по расчету пересечений лучей с объектами оказывается проделанной впустую. В следующем разделе предпринимаются некоторые попытки сокращения числа рассматриваемых прямых лучей путем решения уравнений, полученных на основе законов оптики и моделей поверхностей.

После недолгих размышлений становится ясно, что для отсекаания лучей, не попавших в приемник, достаточно рассматривать наблюдателя S в качестве источника *обратных лучей*. В соответствии с **методом обратной трассировки** проложим трассу луча $S + Vt$ в направлении V от точки S к какой-нибудь точке на поверхности некоторого объекта (рис. 1.60, а). По вышерассмотренным методикам рассчитываются вторичные, третичные и т. д. лучи.

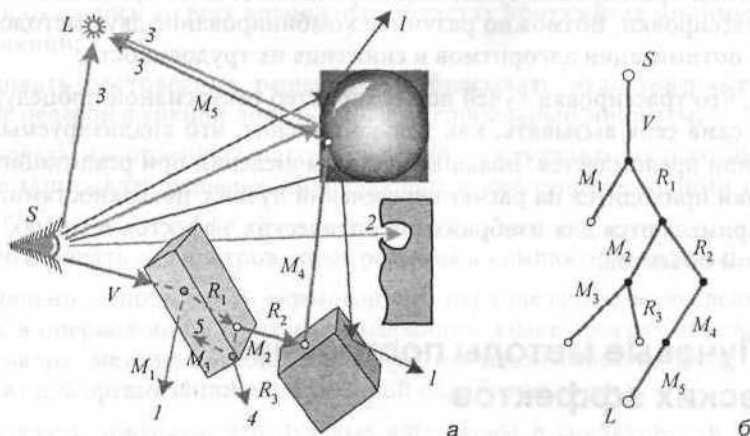


Рис. 1.60

В результате для каждого первичного луча строится *дерево трассировки* (рис. 1.60,5), ветви которого составляют вторичные лучи. Ветвление трассы заканчивается, когда луч (У) выходит за пределы сцены или (2) встречается с непрозрачным телом, поглощающим свет, или (3) попадает в источник света, или (4) когда его интенсивность падает ниже порога чувствительности, или, наконец, (5) когда число расщеплений первичного луча становится слишком большим для имеющихся машинных ресурсов. В результате прямая световая энергия (цвет и интенсивность), попавшая в приемник из направления V , складывается из энергий терминальных вершин дерева с учетом их потерь при распространении в оптических средах.

Метод обратной трассировки фактически аккумулирует все лучи, в действительности *приходящие* в приемник из определенного направления независимо от их начала. Это позволяет видеть и изображать на экране:

- непрозрачные объекты, поглощающие обратные лучи;
- прозрачные объекты, через которые благодаря преломлению наблюдателю видны другие объекты;
- отражения объектов на зеркальных поверхностях, в том числе и блики, соответствующие попаданию обратных лучей в источник света;
- тени, образующиеся в точках поверхности, заслоненных от источника другими объектами;
- другие разнообразные оптические эффекты.

Количество "зондирующих" обратных лучей, подвергаемых трассировке, ограничено числом точек на поверхностях объектов сцены, видимых из точки S и перебираемых с конечным шагом пространственного квантования, обусловленным разрешением экрана. Благодаря этому объем вычислительных затрат в методе обратной трассировки существенно уменьшается по сравнению с методом прямой трассировки. Возможно разумное комбинирование двух методов трассировки для оптимизации алгоритмов и снижения их трудоемкости.

Очевидно, что трассировка лучей носит характер рекурсивной процедуры, которая будет сама себя вызывать, как только выяснит, что анализируемый луч отражается или преломляется. Большая часть вычислений при реализации методов трассировки приходится на расчет пересечений лучей с поверхностями, в связи с чем они применяются для изображения оптических эффектов в сценах с небольшим числом объектов.

1.3.5. Лучевые методы построения оптических эффектов

Применим методы прямой и обратной трассировки лучей для решения задач построения оптических эффектов: тени, отражения и преломления. Сразу отметим

некоторые *геометрические и вычислительные особенности*, порождаемые нелинейностью поверхности, на которой строится оптический эффект:

- эффект может быть *фрагментарным*, т. е. созданным не всеми точками объекта, а лишь теми, в которых выполнены все условия существования эффекта на некотором участке поверхности;
- эффект непрерывного объекта может быть *разрывным*, когда тень от двух близких точек объекта падает на разные участки поверхности, удаленные друг от друга в силу ее криволинейности и самоэкранирования;
- эффект одной точки может быть *множественным*, когда его видно не в одной, а в нескольких точках поверхности, где одновременно выполнены все условия существования эффекта;
- большинство численных методов решения нелинейных уравнений имеют *локальный* характер сходимости, а результат их работы зависит от выбора начальных приближений неизвестных. Следовательно, для получения достоверного результата — всех и именно тех точек поверхности, в которых наблюдается эффект, — необходимо выбирать начальные приближения в окрестности сходимости к *глобальному* (нулевому) минимуму целевой функции. С учетом первых трех геометрических особенностей становится ясно, что глобальных минимумов целевой функции может быть несколько.

Таким образом, качественный алгоритм построения оптического эффекта на нелинейной поверхности должен быть способен:

- выполнять декомпозицию объекта на множество не связанных между собой точек, каждая из которых обрабатывается независимо от других;
- генерировать вычислительный процесс с множества различных начальных приближений неизвестных параметров. Период пространственной дискретизации (расстояние между элементами этого множества) выбирается на основе 3d-аналога теоремы Шеннона—Котельникова, т. е. так, чтобы узлы дискретизации оказались во всех возможных областях притяжения минимумов целевой функции;
- накапливать достоверные решения и отбрасывать недостоверные, доставляющие целевой функции локальный, а не глобальный минимум;
- фильтровать накопленный массив решений, т. е. удалять из него либо сразу туда не записывать решения, совпадающие с уже существующими или близкие к ним;
- фрагментировать отфильтрованные решения в компактные группы;
- максимально использовать возможности параллельных вычислений, заложенные в операционной системе и выбранном языке программирования. Вообще говоря, методы трассировки лучей по самой своей природе идеально подходят для организации параллельной обработки данных.

В целом следует признать, что лучевые алгоритмы в совокупности с другими алгоритмами 3d-графики требуют для своей работы значительных машинных ресурсов и на современных персональных компьютерах работают пока еще мед-

ленно, особенно при моделировании динамических сцен. Собственно говоря, основной прогресс в развитии компьютерной техники есть постоянное совершенствование аппаратных и программных ЭС1-ускорителей. Нет сомнений, что в ближайшее время задачи построения самых сложных оптических эффектов будут быстро решаться в реальном времени.

Множество сложных геометрических объектов может быть представлено совокупностью простейших графических элементов — точек и отрезков, а поверхности, на которых строятся оптические эффекты, аппроксимированы системой плоских граней. Отсюда ясна актуальность решения задач построения оптических эффектов точек и отрезков на плоскости. В связи с этим подробнее рассмотрим построение оптических эффектов на *плоскости* для точки и отрезка, заданного двумя концевыми точками.

Желательно иметь описание плоскости элементами нормальной формы $\{p_0, N\}$. Если плоскость задана элементами параметрической формы $\{p_0, V, W\}$, то по (1.69) ее нормаль равна $N = V \times W$. Возможно, на плоскости задан полигон $P = \{p_1 \dots p_n p_1\}$ с числом сторон $n > 3$, по которому должно выполняться отсечение оптического эффекта. Тогда принимаем точку $p_0 = p_1$, а по (1.117) вычисляем нормаль $N = \text{norm}(P)$. Непосредственное использование тестов ориентации и алгоритмов отсечения, описанных в разд. 1.1.2, возможно при совмещении плоскости эффекта с координатной плоскостью xy методом аффинного преобразования системы координат с помощью функций *МАР1*, *МАР2* или *МАРА*, определенных в гл. 2.

При построении тени должно быть задано положение *источника света*. В зависимости от его удаленности от объекта возможны два случая, отличающиеся свойством параллельности лучей (рис. 1.61):

- источник, например, светильник, расположен в *конечно* удаленной точке L . Различные точки пространства q связаны с ней *непараллельными* расходящимися из L прямыми лучами $q + V_q t$ либо сходящимися в L обратными лучами $q - V_q t$ с векторами направления

$$V_q = q - L.$$

Непараллельность лучей приводит к искажению формы тени объекта (I), что делает ее перспективное изображение более естественным;

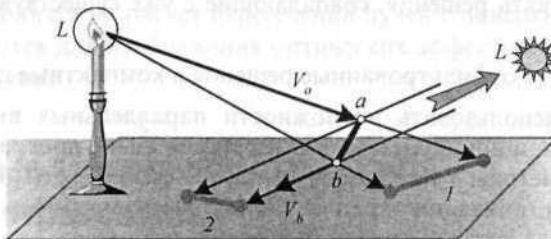


Рис. 1.61

- источник, например, Солнце, *бесконечно* удален от объекта в направлении вектора L . Источник можно считать бесконечно удаленным, если минимальное расстояние между ним и объектом хотя бы на порядок, т. е. в 10 раз, превышает максимальный габарит объекта. При этом условии максимальный угол расхождения лучей не превышает значения $\arctg(0.1) = 5.7^\circ$, что достаточно близко к 0° . Такая *идеализация* позволяет считать прямые $q + V_q t$ и обратные $q - V_q t$ лучи *параллельными* с векторами направления

$$V_q = -L$$

для всех точек q , упрощает построение тени (2), делает возможным применение *групповых* матричных методов обработки точек, описанных в гл. 2, но не дает перспективного изображения.

При построении оптических эффектов необходимо знать положение их *наблюдателя*. В зависимости от его удаленности от объекта также возможны два случая, отличающиеся свойством параллельности лучей (рис. 1.62):

- наблюдатель расположен в *конечно* удаленной точке S . Различные точки пространства q соединены с ней *непараллельными* сходящимися в S прямыми лучами $q + S_q t$ либо расходящимися из S обратными лучами $q - S_q t$ с векторами направления

$$S_q = S - q.$$

Непараллельность лучей делает перспективное изображение оптического эффекта (3) реалистичным, но усложняет его расчет;

- наблюдатель *бесконечно* удален от объекта в направлении вектора S . Конечно удаленного наблюдателя можно считать бесконечно удаленным, если минимальное расстояние между ним и объектом хотя бы на порядок больше максимального габарита объекта. Такая *идеализация* позволяет считать прямые $q + S_q t$ и обратные $q - S_q t$ лучи *параллельными* с равными векторами направления

$$S_q = S$$

для всех точек q , упрощает построение оптических эффектов, делает возможным применение групповых матричных методов обработки точек, но не дает перспективного изображения (4).

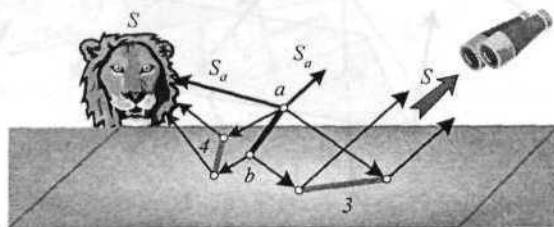


Рис. 1.62

Используя устоявшуюся терминологию [36], будем впредь называть конечно удаленного наблюдателя и источники света "ближними", а их же, удаленных в бесконечность, — "дальними". Каждый из этих режимов, как отмечалось чуть выше, имеет свои преимущества и недостатки, но применять их следует, руководствуясь конкретными геометрическими условиями.

Переходим к изучению лучевых методов построения оптических эффектов точек и отрезков на произвольной поверхности и более детально — на плоскости.

1.3.5.1. Тень

Тень точки q на произвольной поверхности образуется в точке q' пересечения с ней прямого луча $q + V_q t$. При дальнем источнике света рекомендуется задавать вектор его направления в виде $V_q = -\gamma L$ с коэффициентом удаленности γ , во много раз большим отношения максимального габарита сцены к длине вектора L .

Задача построения тени точки на произвольной поверхности решена в разд. 1.3.1. Более сложный, чем точка, графический объект представляется системой точек, для которых тени строятся однотипно.

Возможность пересечения прямого луча $q + V_q t$ с плоскостью $\{p_0, N\}$ в точке q' (рис. 1.63) определяется числом

$$v_q = V_q \circ N.$$

При $v_q = 0$ луч (1) проходит параллельно плоскости, нигде ее не пересекая. Если $v_q \neq 0$, то по (1.102,6) находим параметр и точку пересечения прямой $\{q, V_q\}$ с плоскостью $\{p_0, N\}$:

$$t_q = \frac{(p_0 - q) \circ N}{v_q}, \quad q' = q + \frac{(p_0 - q) \circ N}{V_q \circ N} V_q. \quad (1.144)$$

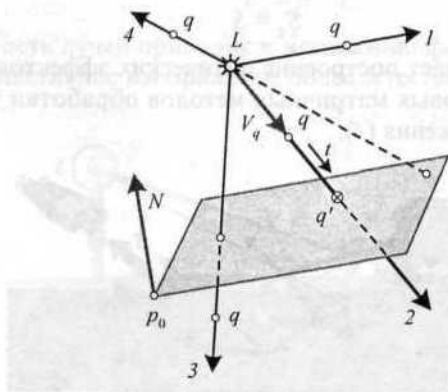


Рис. 1.63

При $t_q = 0$ точка q лежит в плоскости $\{p_0, N\}$ и согласно (1.144) совпадает со своей тенью: $q' = q$. При $t_q > 0$ луч (2) пересекает плоскость впереди точки q и дает действительную тень q' . Условие существования тени $t_q > 0$ геометрически означает, что векторы $p_0 - q$ и V_q одинаково ориентированы относительно плоскости, т. е. образуют с нормалью N односторонние углы: либо оба острые, либо оба тупые. При условии $-1 \leq t_q < 0$ луч (5) пересекает плоскость в точке q' раньше, чем доходит до точки q , а при $t_q < -1$ луч (4) вообще удаляется от плоскости. Таким образом, при $t_q < 0$ тень является мнимой.

Важным аспектом построения тени является ее видимость наблюдателем. Даже если тень точки существует, она будет невидима, когда наблюдатель и падающий луч находятся по разные стороны от плоскости. Тогда отпадает необходимость самого расчета тени. Таким образом, условия видимости и существования тени точки имеют следующий вид:

$$\{v_{p_0} \sigma < 0\} \cap \{t_q \geq 0\}, \text{ где } v_{p_0} = V_{p_0} \circ N, \sigma = S_{p_0} \circ N. \quad (1.145)$$

Первое в (1.145) условие видимости тени геометрически означает, что векторы V_{p_0} и S_{p_0} противоположно ориентированы относительно плоскости, т. е. из двух углов $\angle V_{p_0} N$ и $\angle S_{p_0} N$ один острый, а другой — тупой.

Построение тени отрезка ab на плоскости $\{p_0, N\}$ (рис. 1.64) требует анализа взаимного расположения концевых точек отрезка a и b относительно плоскости и источника. Каждая концевая точка $q = a$ либо $q = b$ может располагаться в одном из четырех положений:

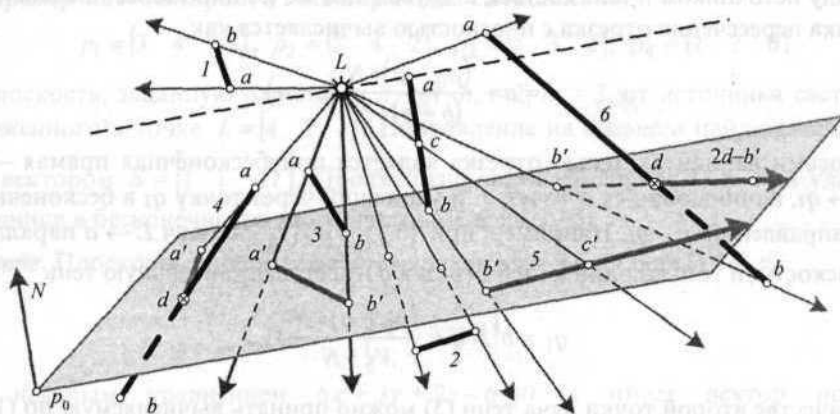


Рис. 1.64

- при $v_q = 0$ на равном с источником удалении и по одну сторону от плоскости;
- при $\{v_q \neq 0\} \cap \{t_q < -1\}$ дальше от плоскости, чем источник;

- при $\{v_q \neq 0\} \cap \{-1 \leq t_q < 0\}$ по разные с источником стороны от плоскости;
- при $\{v_q \neq 0\} \cap \{t_q \geq 0\}$ между источником и плоскостью.

В табл. 1.5 приведено 16 возможных вариантов образования тени отрезка ab :

Таблица 1.5

$a \backslash b$		$v_b = 0$	$v_b \neq 0$		
			$t_b < -1$	$-1 \leq t_b < 0$	$t_b \geq 1$
$v_a = 0$		—	—	$d \rightarrow 2d - b'$	$b' \rightarrow c'$
$v_a \neq 0$	$t_a < -1$	—	—	$d \rightarrow 2d - b'$	$b' \rightarrow 2b' - a'$
	$-1 \leq t_a < 0$	$d \rightarrow 2d - a'$	$d \rightarrow 2d - a'$	—	$[d, b']$
	$t_a \geq 1$	$a' \rightarrow c'$	$a' \rightarrow 2a' - b'$	$[a', d]$	$[a', b']$

- в пяти вариантах тень *отсутствует* (прочерки в таблице): если оба конца ab расположены не ближе к плоскости, чем источник (7), либо оба они лежат по другую сторону от плоскости, чем источник (2);
- в трех случаях тень имеет вид *отрезка* $q \setminus qr$, когда один из концов ab находится между источником и плоскостью, а другой там же (3) либо ниже плоскости (4). Точка пересечения отрезка с плоскостью вычисляется как

$$d = a + \frac{(p_0 - a) \circ N}{(b - a) \circ N} (b - a);$$

- в восьми вариантах тенью отрезка является полубесконечная прямая — *луч* $q_1 \rightarrow q_2$, начинающийся в точке $q \setminus$ и уходящий через точку q_2 в бесконечность в направлении $q_2 - q_1$. Например, при $\{v_a = 0\} \cap \{t_b \geq 0\}$ луч $L \rightarrow a$ параллелен плоскости, и только один конец отрезка b имеет действительную тень

$$q_1 = b' = q + \frac{(p_0 - q) \circ N}{V_b \circ N} V_b.$$

В качестве второй точки луча тени (5) можно принять вычисляемую по (1.144) проекцию $q_2 = c'$ средней точки отрезка $c = 0.5(a + b)$. При $\{t_a < -1\} \cap \{-1 \leq t_b < 0\}$ концы отрезка (6) имеют мнимые тени a' и b' , но, тем не менее, отрезок отбрасывает тень в виде луча, выходящего из точки d в направлении точки $d + (d - b) = 2d - b'$.

В связи с возможностью существования тени отрезка в виде луча необходимо модифицировать описанные в разд. 1.1.2 алгоритмы отсечения *clip2_cross* и *clip2_test* так, чтобы они могли работать не только с отрезком ab , но и с лучом $a \rightarrow b$. Рассмотрим два варианта модификации.

1. Луч $a \rightarrow b$ заменяется отрезком $[a, a + (b - a)\gamma]$ с коэффициентом $\gamma \rightarrow \infty$.
2. В модифицированный вариант функции *clip2_cross*(ab, P, mod) введем дополнительный аргумент *mod*, задаваемый равным 0, если ab — отрезок $[a, b]$, и 1, если ab — луч $a \rightarrow b$. При *mod* = 1 изменения алгоритма на рис. 1.18,б минимальны и заключаются в том, что расчет концов неотсеченной части луча $a \rightarrow b$ вместо табл. 1.1 выполняется по более простой в программировании табл. 1.6.

Таблица 1.6

σ_2 σ_1	$\sigma_2 < 0$	$0 \leq \sigma_2 \leq 1$	$\sigma_2 > 1$
$\sigma_1 < 0$	$[a, a]$	$[a, a + V\sigma_2]$	$[a, a + V\sigma_2]$
$0 \leq \sigma_1 \leq 1$	$[a, a + V\sigma_1]$	$[a + V\sigma_1, a + V\sigma_2]$	$[a + V\sigma_1, a + V\sigma_2]$
$\sigma_1 > 1$	$[a, a + V\sigma_1]$	$[a + V\sigma_1, a + V\sigma_2]$	$[a + V\sigma_1, a + V\sigma_2]$

Пример 1.10. Построить тень пирамиды с вершинами

$$p_1 = [3 \ 4 \ -1], p_2 = [2 \ 4 \ 2], p_3 = [5 \ 3 \ 2], p_4 = [2 \ 2 \ 0]$$

на плоскость, заданную отрезками $h_x = 1, h_y = 2, h_z = 3$, от источника света, расположенного в точке $L = [2 \ 1]$. Направление на *дальнего* наблюдателя задается вектором $S = [1 \ 1 \ \sqrt{7}]$. Проанализировать изменение тени при удалении источника в бесконечность вдоль вектора L (рис. 1.65).

Решение. Плоскость тени описывается уравнением в отрезках (1.74, а)

$$\frac{x}{1} + \frac{y}{2} + \frac{z}{3} = 1$$

или неявным уравнением $6x + 3y + 2z - 6 = 0$ и имеет вектор нормали $N = [6 \ 3 \ 2]$. Выбрав на плоскости точку $p_0 = h_z z^0 = [0 \ 0 \ 3]$, оценим по (1.145) видимость точек тени:

$$V_{p_0} = p_0 - L = [-4 \ -2 \ 2], v_{p_0} = -26, S_{p_0} = S, \sigma = 14.29 \Rightarrow v_{p_0} \sigma < 0.$$

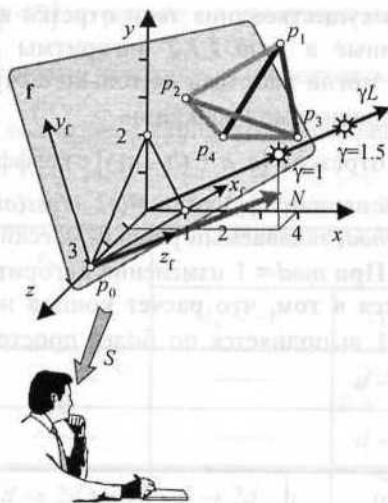


Рис. 1.65

Условие видимости выполнено. Приступаем к расчету тени точек по формулам

$$V_i = p_i - L, v_i = V_i \circ N, t_i = \frac{(p_0 - p_i) \circ N}{v_i}, p'_i = p_i + V_i t_i;$$

$$V_1 = [-1 \ 2 \ -2], v_1 = -4 \neq 0, t_1 = 5.5 > 0, p'_1 = [-2.5 \ 15 \ -12];$$

$$V_2 = [-2 \ 2 \ 1], v_2 = -4 \neq 0, t_2 = 5.5 > 0, p'_2 = [-9 \ 15 \ 7.5];$$

$$V_3 = [1 \ 1 \ 1], v_3 = 11 \neq 0, t_3 = -3.364 < 0, p'_3 = [1.636 \ -0.364 \ -1.364];$$

$$V_4 = [-2 \ 0 \ -1], v_4 = -14 \neq 0, t_4 = 0.857 > 0, p'_4 = [0.286 \ 2 \ -0.857].$$

Удовлетворяющие условию существования тени $t_i > 0$ значения t_1 , t_2 и t_4 означают, что ребра пирамиды $p_0 p_2$, $p_0 p_4$ и $p_1 p_4$ имеют на плоскости действительные тени. Отрицательное значение t_3 сигнализирует, что точка p_3 отстоит от плоскости дальше источника и у нее есть только мнимая тень p'_3 . Следовательно, тени ребер $p_1 p_3$, $p_2 p_3$, и $p_3 p_4$ представляют собой расходящиеся на плоскости лучи $p'_i + W_i \tau$ ($i \in \{1, 2, 4\}$, $\tau \geq 0$) с направляющими векторами $W_i = p'_i - p'_3$ и общей задней точкой схода p'_3 .

На рис. 1.66, а изображена рассчитанная MathCAD-программой сцена, включающая оси координат, объект, источник света, теньевую плоскость и тени ребер объекта на ней. Ракурс, соответствующий направлению на дальнего наблюдателя $S = [1 \ 1 \ \sqrt{7}]$ в системе координат сцены, является *диетрической проекцией*

на ортогональную к наблюдателю плоскость и очень наглядно отображает трехмерный мир на плоскости.

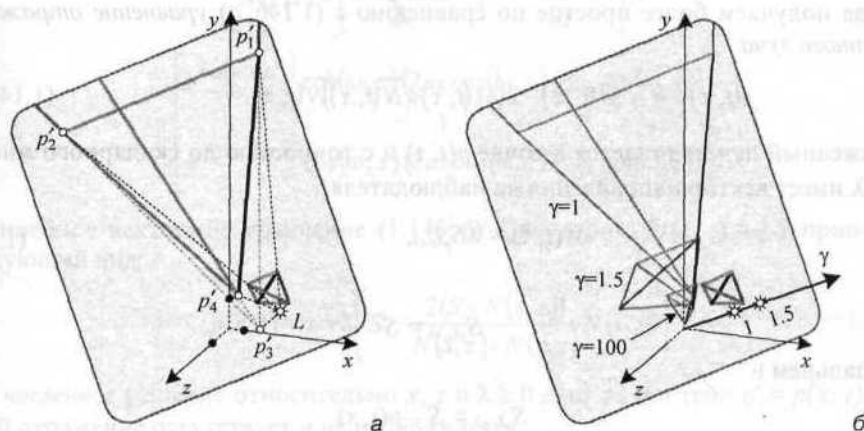


Рис. 1.66

Заменяв в исходных данных положение источника света L на γL , рассчитаем тень объекта при $\gamma \geq 1$ — удалении источника в бесконечность вдоль вектора L . На рис. 1.66,б приведены положения тени пирамиды при $\gamma = 1$, $\gamma = 1.5$ и $\gamma = 100$. Последнее значение γ соответствует практически дальнему источнику света. Анализ изображений тени показывает, что с удалением источника она уменьшается в размерах, достигая в пределе $\gamma \rightarrow \infty$ размеров объекта. При этом уменьшаются и ее перспективные искажения. \square

1.3.5.2. Отражение

Отражение точки q от произвольной поверхности $p(t, \tau)$ наблюдается в такой точке $q' \in p(t, \tau)$ (см. рис. 1.56), где выполняются все перечисленные в разд. 1.3.2 законы отражения. Векторы падающего

$$V_q(t, \tau) = p(t, \tau) - q,$$

отраженного $M(t, \tau)$ лучей и нормали, вычисляемой по достаточно сложной формуле (1.61, б) как

$$N(t, \tau) = \frac{\partial p(t, \tau)}{\partial t} \times \frac{\partial p(t, \tau)}{\partial \tau},$$

связаны уравнением отражения прямого луча (1.140):

$$M(t, \tau) = p(t, \tau) - q - 2((p(t, \tau) - q) \circ \bar{N}(t, \tau)) \bar{N}(t, \tau). \quad (1.146, a)$$

Из принципа обратимости лучей следует соотношение

$$-V_q = -M - 2(-M \circ \bar{N})\bar{N},$$

откуда получаем более простое по сравнению с (1.146, а) уравнение отражения обратного луча

$$p(t, \tau) - q = M(t, \tau) - 2(M(t, \tau) \circ \bar{N}(t, \tau))\bar{N}(t, \tau). \quad (1.146, б)$$

Отраженный луч начинается в точке $p(t, \tau)$ и с точностью до скалярного множителя X имеет вектор направления на наблюдателя

$$M(t, \tau) = \lambda S_{p(t, \tau)}, \quad (1.147)$$

где

$$S_{p(t, \tau)} = S$$

при дальнем и

$$S_{p(t, \tau)} = S - p(t, \tau)$$

при ближнем наблюдателе. Знак числа λ служит индикатором направления отраженного луча относительно наблюдателя. При $\lambda \geq 0$ отражение действительное, а при $\lambda < 0$ — мнимое, т. к. отраженный луч удаляется от наблюдателя в противоположном направлении.

Решение уравнений (1.146, а) либо (1.146, б) совместно с (1.147) относительно всех допустимых переменных $\{t, \tau, \lambda \geq 0\}$ имеет перечисленные в начале разд. 1.3.5 свойства, изучаемые здесь более подробно, и следствия из них:

- отражение может быть *множественным*: решение представляет собой набор разных точек $q'_k = p(t_k, \tau_k)$, $k \geq 1$, в которых наблюдатель видит отражения одной и той же точки q на произвольной нелинейной поверхности $p(t, \tau)$;
- отражение может быть *разрывным*: нельзя соединять точки отражения друг с другом, даже если соответствующие точки объекта были соединены, например, ребром;
- отражение может быть *фрагментарным*: возможно, что решение $\{t, \tau, \lambda \geq 0\}$ пусто, а отражение отсутствует.

□ *Пример 1.11.* Построить отражение пирамиды из примера 1.10 (см. рис. 1.65) от криволинейной поверхности из примера 1.7 (см. рис. 1.53, а), описываемой явным уравнением

$$y = f(x, z) = e^{-a(x^2 + z^2)} \cos(\omega_x x) \cos(\omega_z z), \quad a = 0.02, \omega_x = 1, \omega_z = 0.5.$$

Направление на дальнего наблюдателя задается вектором $S = [1 \ 1 \ 1]$, соответствующим *изометрической проекции*.

Решение. Составим неявное уравнение поверхности $y=f(x, z)=0$ и по (1.61,а) вычислим нормаль к ней в точке $p(x, z) = [x \ f(x,z) \ z]^T$.

$$N(x, z) = \begin{bmatrix} -\frac{\partial f(x, z)}{\partial x} & 1 & -\frac{\partial f(x, z)}{\partial z} \end{bmatrix} = \\ = \begin{bmatrix} e^{-a(x^2+z^2)} \cos(\omega_z z) (2ax \cos(\omega_x x) + \omega_x \sin(\omega_x x)) \\ 1 \\ e^{-a(x^2+z^2)} \cos(\omega_x x) (2az \cos(\omega_z z) + \omega_z \sin(\omega_z z)) \end{bmatrix}^T.$$

Нелинейное векторное уравнение (1.146,6) с вектором $M(x, z) = \lambda S$ принимает следующий вид:

$$p(x, z) - q = \lambda \left(S - \frac{2(S \circ N(x, z))}{N(x, z) \circ N(x, z)} N(x, z) \right).$$

Его численное решение относительно x, z и $\lambda \geq 0$ дает точки тени $q' = p(x, z)$. При $\lambda < 0$ отражение отсутствует и не изображается.

На рис. 1.67 изображена рассчитанная MathCAD-программой изометрическая проекция сцены, включающая объект, зеркальную криволинейную поверхность и отражение в ней объекта. Каждое ребро пирамиды заменялось цепочкой из 21 точки. Начальные приближения $\{x, z\}$ выбирались из всевозможных сочетаний (всего их 24) элементов списков $x \in \{1, 3, 4, 5, 7, 10\}$ и $z \in \{1, 3, 5, 7\}$. Анализ отражения убеждает, что оно, действительно, имеет множественные и фрагментарные свойства. \square

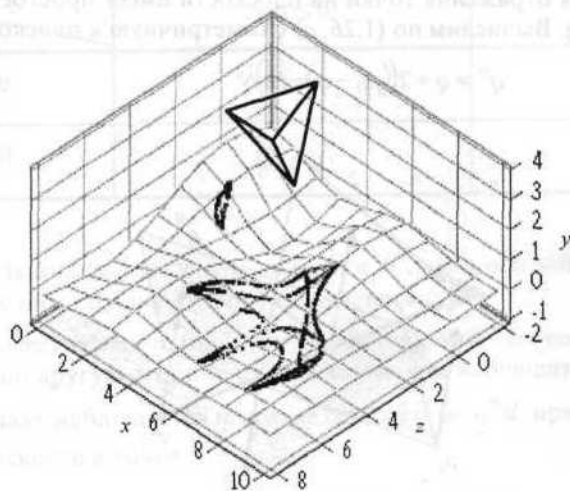


Рис. 1.67

В частном случае *плоской* параметрической поверхности $\{p_0, V, W\}$ с нормалью $N = V \times W = \text{const}$ и ближнего наблюдателя подстановка в (1.146, а) выражения

$$M(t, \tau) = \lambda(S - p(t, \tau))$$

и свойство ортогональности $(Vt + W\tau) \circ \bar{N} = 0$ дают *нелинейное* уравнение

$$\lambda(S - p_0 - Vt - W\tau) = p_0 + Vt + W\tau - q - 2((p_0 - q) \circ \bar{N}) \bar{N}.$$

Его *точное* решение

$$[\alpha \ \beta \ \lambda] = (q - p_0 - 2((q - p_0) \circ \bar{N}) \bar{N}) \begin{bmatrix} V \\ W \\ p_0 - S \end{bmatrix}^{-1} \Rightarrow t = \frac{\alpha}{\lambda + 1}, \quad \tau = \frac{\beta}{\lambda + 1} \quad (1.148, a)$$

определено при векторе $p_0 - S$, не лежащем в отражающей плоскости.

При дальнем наблюдателе подстановка

$$M(t, \tau) = \lambda S$$

в (1.146, б) дает следующее *линейное* уравнение и его точное решение при $S \circ N$:

$$p_0 + Vt + W\tau - q = \lambda S - 2\lambda(S \circ \bar{N}) \bar{N} \Rightarrow [t \ \tau \ \lambda] = (q - p_0) \begin{bmatrix} V \\ W \\ 2(S \circ \bar{N}) \bar{N} - S \end{bmatrix}^{-1}. \quad (1.148, б)$$

В обоих случаях отражение точки q действительно при $\lambda \geq 0$ и наблюдается на поверхности неограниченной плоскости в точке $q' = p_0 + Vt + W\tau$.

Задача построения отражения точки на плоскости имеет простое *геометрическое решение* (рис. 1.68). Вычислим по (1.26, а) симметричную к плоскости точку

$$q^o = q + 2((p_0 - q) \circ \bar{N}) \bar{N} \quad (1.149)$$

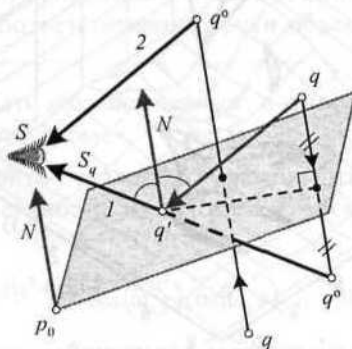


Рис. 1.68

и выпустим из нее луч $q^0 + S_{q^0} t$. Наблюдатель видит отражение точки q в точке

$$q' = q^0 + \frac{(p_0 - q^0) \circ N}{S_{q^0} \circ N} S_{q^0} \quad (1.150)$$

пересечения луча с плоскостью. Тип отражения определяется соотношением знаков чисел $v_q = (q - p_0) \circ N$ и $\sigma = S_{p_0} \circ N$. При $v_q \sigma > 0$ точка q и наблюдатель расположены по одну сторону от плоскости, а вычисленное по (1.150) отражение (7) действительно. При $v_q \sigma < 0$ по одну сторону от плоскости находятся наблюдатель и точка q^0 (2), а отражение является мнимым.

Точки q^0 и q' лежат на одном прямом луче $q^0 \rightarrow S$ и визуально сливаются в одну точку. Зрительная система человека размещает отражение объекта за зеркалом, а не на его поверхности. Контур зеркала является как бы окном, через которое наблюдатель рассматривает виртуальный (кажущийся) объект. Таким образом, для расчета отражения точки q от зеркальной плоскости достаточно найти по (1.149) симметричную точку q^0 и при выполнении условия $v_q \sigma > 0$ считать ее действительным отражением.

Построение отражения отрезка ab на плоскости $\{p_0, N\}$ (рис. 1.69) начинается с анализа взаимного расположения концевых точек отрезка a и b относительно плоскости и наблюдателя. Четыре варианта этого расположения приведены в табл. 1.7:

Таблица 1.7

$a \backslash b$	$v_b \sigma \leq 0$	$v_b \sigma > 0$
$v_a \sigma \leq 0$	—	$[d, b^0]$
$v_a \sigma > 0$	$[a^0, d]$	$[a^0, b^0]$

- наблюдатель видит полное отражение $a^0 b^0$, если оба конца отрезка (1) расположены с ним по одну сторону от плоскости;
- отражение отсутствует (прочерк в таблице), если оба конца отрезка (2) расположены по другую сторону от плоскости, чем наблюдатель;
- в двух случаях наблюдается неполное отражение $q^0 d$ при пересечении отрезком (3) плоскости в точке

$$d = a + \frac{(p_0 - a) \circ N}{(b - a) \circ N} (b - a) \in ab.$$

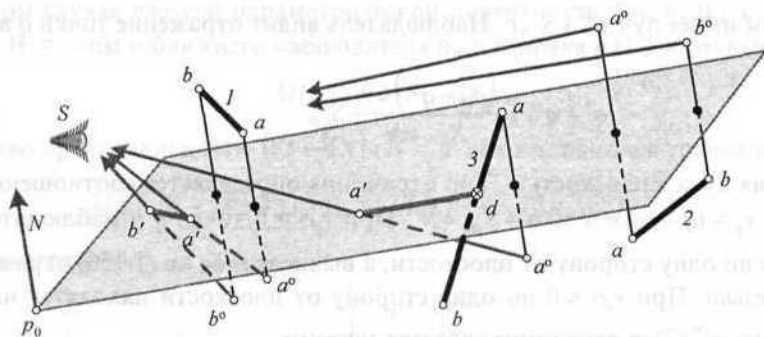


Рис. 1.69

Пример 1.12. Построить отражение пирамиды с вершинами

$$p_1 = [2 \ 2 \ -2], \quad p_2 = [1 \ 3 \ 1], \quad p_3 = [4 \ 2 \ -1], \quad p_4 = [1 \ 0 \ -1]$$

от плоского зеркала в форме параллелограмма, три вершины которого отсекают на осях координат отрезки $h_x = 1$, $h_y = 3$ и $h_z = 2$. Наблюдатель находится в точке $S = [1 \ 1 \ \sqrt{8}]$. Сравнить отражения при удалении наблюдателя в бесконечность вдоль вектора γS при $\gamma \geq 1$ (рис. 1.70).

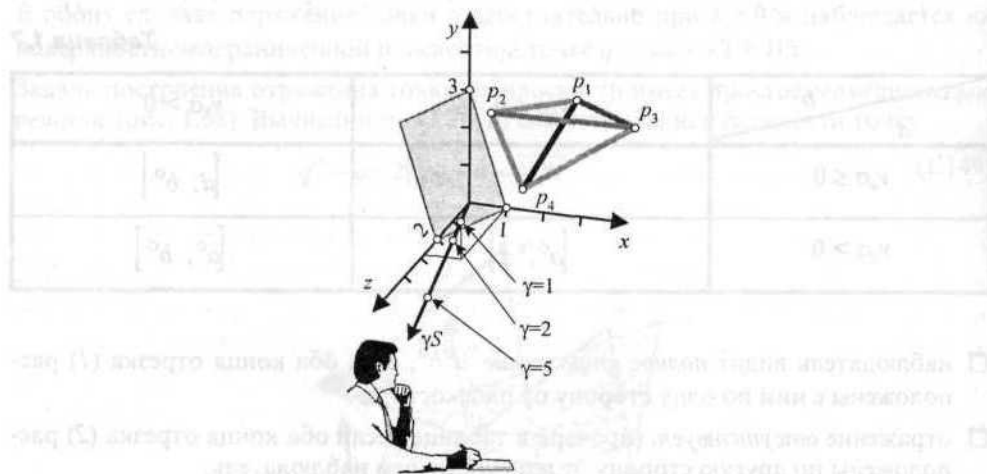


Рис. 1.70

Решение. Плоскость зеркала описывается неявным уравнением $6x + 2y + 3z - 6 = 0$ и имеет нормаль $N = [6 \ 2 \ 3]$. Выбрав на ней точку $p_0 = h_z z^0 = [0 \ 0 \ 2]$,

вычислим индикаторы $v_i = (p_i - p_0) \circ N$ ориентации вершин p_i , относительно плоскости, а по (1.149) — вершины симметричной пирамиды

$$p_i^0 = p_i + \frac{2(p_0 - p_i) \circ N}{N \circ N} N;$$

$$v_1 = 4 > 0, \quad p_1^0 = [1.02 \quad 1.673 \quad -2.49];$$

$$v_2 = 9 > 0, \quad p_2^0 = [-1.204 \quad 2.265 \quad -0.102];$$

$$v_3 = 19 > 0, \quad p_3^0 = [-0.653 \quad 0.449 \quad -3.327];$$

$$v_4 = -3 < 0, \quad p_4^0 = [1.735 \quad 0.245 \quad -0.633].$$

Поскольку $\sigma = (S - p_0) \circ N = 10.485 > 0$, то из всех вершин пирамиды лишь p_4 находится с наблюдателем по разные стороны от плоскости зеркала и не имеет действительного отражения. Вычислим точки

$$d_{i4} = p_i + \frac{(p_0 - p_i) \circ N}{(p_4 - p_i) \circ N} (p_4 - p_i) \in p_i p_4$$

пересечения бесконечной плоскости $\{p_0, N\}$ с ребрами пирамиды, смежными с вершиной p_4 (рис. 1.71, а):

$$d_{14} = [1.429 \quad 0.857 \quad -1.429], \quad d_{24} = [1 \quad 0.75 \quad -0.5], \quad d_{34} = [1.409 \quad 0.273 \quad -1].$$

Таким образом, отражение пирамиды от бесконечной плоскости зеркала составляют отрезки $p_1^0 p_2^0$, $p_1^0 p_3^0$, $p_2^0 p_3^0$, $p_1^0 d_{14}$, $p_2^0 d_{24}$ и $p_3^0 d_{34}$.

Второй этап расчета отражения состоит в проецировании всех объектов сцены на некоторую картинную плоскость, допустим, координатную плоскость $x_0 y_0 z_0$. Проецирование является *центральной*, т. к. как наблюдатель как бы испускает проецирующие лучи из центра — конечно удаленной от сцены точки $S = [s_x \quad s_y \quad s_z]$. Не вдаваясь пока (до подробного изучения в гл. 2) в методику

центрального проецирования, отметим ее сходство с алгоритмом построения тени, в котором глаз наблюдателя является источником обратных лучей.

Заключительный этап состоит в отсеении *проекций* отрезков отражения полигоном *проекции* зеркала. На рис. 1.71, б изображена рассчитанная MathCAD-программой центральная проекция сцены. Пунктиром обозначены отсеженные фрагменты отрезков отражения.

Заменив в исходных данных положение наблюдателя S на γS , рассчитаем отражения пирамиды при удалении наблюдателя в бесконечность вдоль вектора S . Для ряда значений $\gamma = 1$, $\gamma = 2$ (рис. 1.71, в) и $\gamma = 5$ (рис. 1.71, г) углы горизонтального обзора сцены составляют соответственно 106° , 56° и 23° . Последний случай практически эквивалентен дальнему наблюдателю с коэффициентом

$\gamma \rightarrow \infty$, Соответствующий ракурс с вектором направления $S = [1 \ 1 \ \sqrt{8}]$ является кабинетной косоугольной проекцией на плоскость xy и широко применяется для отображения трехмерного мира на плоскости. Анализ изменения отражения показывает, что с удалением наблюдателя уменьшаются как его размеры, так и перспективные искажения. \square

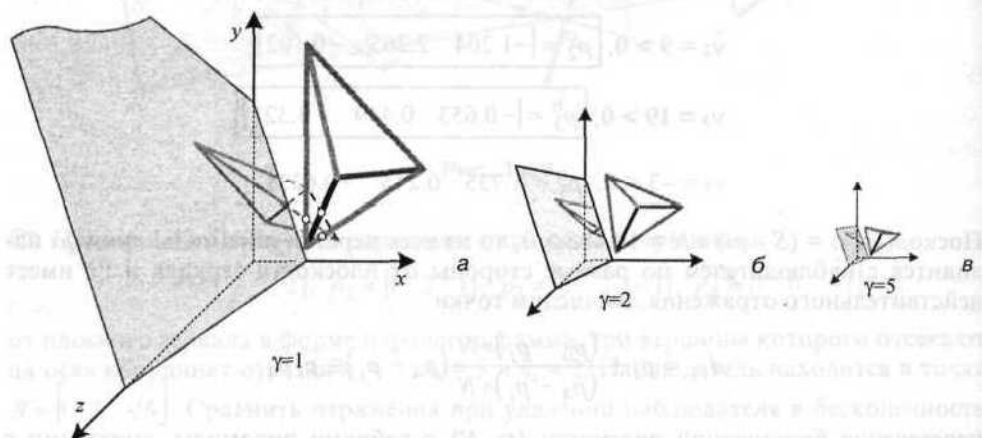


Рис. 1.71

1.3.5.3. Преломление

Преломление точки q на произвольной поверхности $p(t, \tau)$ раздела двух сред с показателями преломления n_1 и n_2 наблюдается в такой точке $q' \in p(t, \tau)$ (см. рис. 1.57), где выполняются все перечисленные в разд. 1.3.3 законы преломления. Векторы падающего

$$V_q(t, \tau) = p(t, \tau) - q,$$

преломленного $R(t, \tau)$ лучей и нормали

$$N(t, \tau) = \frac{\partial p(t, \tau)}{\partial t} \times \frac{\partial p(t, \tau)}{\partial \tau}$$

связаны уравнением преломления прямого луча (1.142):

$$R = n_n^2(p - q) + n_n \left(\operatorname{sgn}((p - q) \circ \bar{N}) \sqrt{|(p - q)|^2 - n_n^2 |(p - q) \times \bar{N}|^2} - n_n(p - q) \circ \bar{N} \right) \bar{N}. \quad (1.151, a)$$

Из принципа обратимости лучей следует уравнение преломления обратного луча:

$$n_n^2(p - q) = R + \left(\operatorname{sgn}(R \circ \bar{N}) \sqrt{n_n^2 |R|^2 - |R \times \bar{N}|^2} - R \circ \bar{N} \right) \bar{N}. \quad (1.151, б)$$

Преломленный луч начинается в точке $p(t, \tau)$ и с точностью до скалярного множителя X имеет вектор направления на наблюдателя

$$R(t, \tau) = \lambda S_{p(t, \tau)}, \quad (1.152)$$

где

$$S_{p(t, \tau)} = S - p(t, \tau)$$

при ближнем и

$$S_{p(t, \tau)} = S$$

при дальнем наблюдателе. Знак числа λ служит индикатором направления преломленного луча относительно наблюдателя. При $\lambda \geq 0$ преломление действительное, а при $\lambda < 0$ — мнимое, т. к. преломленный луч удаляется от наблюдателя в противоположном направлении.

Дополнительные условия существования действительного преломления, означающие отсутствие эффекта полного внутреннего отражения, имеют вид, аналогичный (1.143):

$$|p - q| > n_n |(p - q) \times \bar{N}|. \quad (1.153)$$

Решение уравнений (1.151, а) либо (1.151, б) совместно с (1.34, в) и (1.34, г) относительно всех допустимых переменных $\{t, \tau, \lambda \geq 0\}$ имеет перечисленные в начале разд. 1.3.5 свойства, изучаемые здесь более подробно, и следствия из них:

- преломление может быть множественным — решение представляет собой набор разных точек $q'_k = p(t_k, \tau_k)$, $k \geq 1$, в которых наблюдатель видит преломления одной и той же точки q на произвольной нелинейной поверхности $p(t, \tau)$;
- преломление может быть разрывным — нельзя соединять точки преломления друг с другом, даже если соответствующие точки объекта были соединены, например, ребром;
- преломление может быть фрагментарным — возможно, что решение $\{t, \tau, \lambda \geq 0\}$ пусто, а преломление отсутствует.

▣ **Пример 1.13.** Внутри янтарного эллипсоида с полуосями $a=2$, $b=3$, $c=2$, центром в начале координат и коэффициентом преломления $n_1 = 1.6$ находится пирамида (рис. 1.72, а) с вершинами

$$p_1 = [-1 \ 2 \ -1], \quad p_2 = [1 \ 2 \ 1], \quad p_3 = [-1 \ 0 \ 1], \quad p_4 = [1 \ 0 \ -1].$$

Построить преломленное изображение пирамиды, видимое наблюдателем из точки $S = [4 \ 4 \ 16]$ воздушной среды с коэффициентом преломления $n_2=1$.

Решение. Поверхность действительного эллипсоида описывается следующими неявным (см. табл. 1.3) и параметрическими (1.129, я) уравнениями:

$$f(p) = \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 - 1 = 0;$$

$$p(t, \tau) = [a \sin(t) \sin(\tau) \quad b \cos(t) \quad c \sin(t) \cos(\tau)], \quad t \in [0, \pi], \tau \in [0, 2\pi].$$

Нормаль к поверхности эллипсоида в точке $p(t, \tau)$ получим по (1.61, б):

$$N(t, \tau) = [bc \sin^2(t) \sin(\tau) \quad ac \sin(t) \cos(t) \quad absin^2(t) \cos(\tau)].$$

Все вершины пирамиды лежат внутри эллипсоида, о чем свидетельствуют отрицательные значения неявной функции поверхности:

$$f(p_1) = f(p_2) = -0.0556 < 0, \quad f(p_3) = f(p_4) = -0.5 < 0.$$

Благодаря выпуклости эллипсоида остальные точки ребер пирамиды также являются внутренними и могут иметь преломленные лучи.

На рис. 1.72, б представлен результат расчета преломления MathCAD-программой. Каждое ребро пирамиды заменялось цепочкой из 50 точек. Для каждой такой точки q решением уравнений (1.151, б), (1.152) при $S_{p(t, \tau)} = S - p(t, \tau)$ рассчитывалась точка $p(t, \tau)$ на поверхности эллипсоида, в которой преломляется луч, идущий от q к S . Отсутствие решений для точек вблизи вершин p_1 и p_4 объясняется, по-видимому, невыполнением условий существования преломленных лучей $\lambda > 0$ и $n_n |R(t, \tau)| > |R(t, \tau) \times \overline{N}(t, \tau)|$.

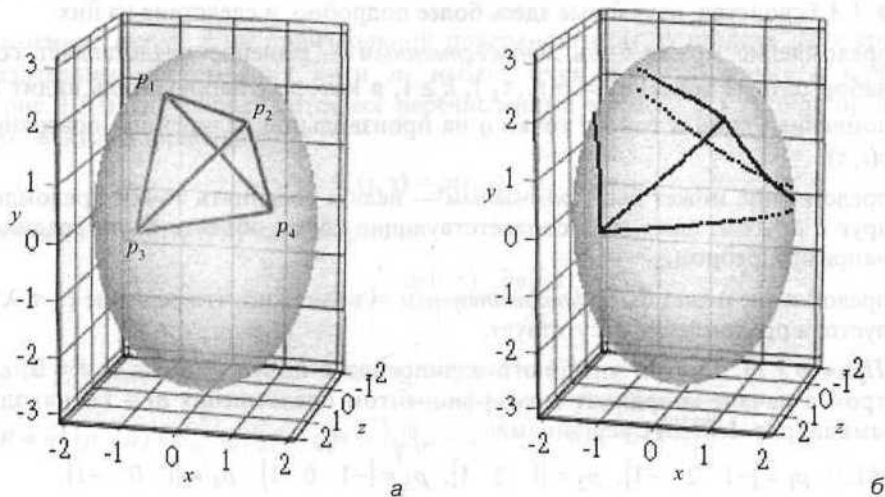


Рис. 1.72