МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОПТИЧЕСКИХ ПОВЕРХНОСТЕЙ ПРИ АВТОМАТИЗИРОВАННОМ ПРОЕКТИРОВАНИИ

С.А. РОДИОНОВ, А.А.ШЕХОНИН

Одним из основных функциональных элементов оптической системы является отражающая или преломляющая поверхности, а основным инструментом при компьютерном проектировании служит расчет хода лучей через систему таких поверхностей. Важную роль при этом играет математическая модель, используемая для их описания.

Как известно, оптические поверхности принято описывать в локальной системе декартовых координат, начало которой помещено в вершину поверхности, а ось Z направлена по нормали к ней (так называемая система Федера [1]). Существовавшие ранее и изредка встречающиеся до сих пор попытки использовать другие системы координат, например, сферические, декартовы с началом в центре кривизны поверхности, обладают столь явными недостатками, что здесь не рассматриваются.

Наиболее часто в оптических системах встречаются поверхности вращения, для описания которых в отечественной литературе [2,3] используются два типа уравнений, представляющие одну из переменных z или $u = x^2+y^2$ в виде разложения по степеням другой:

$$u = c_1 z + c_2 z^2 + c_3 z^3 + \ldots, \tag{1}$$

$$z = d_1 u + d_2 u^2 + d_3 u^3 + \ldots. \tag{2}$$

Наиболее широкое употребление получило уравнение (1), так как позволяет описывать поверхности второго порядка с любым эксцентриситетом, в том числе сферы, эллипсоиды, гиперболоиды и параболоиды вращения с добавлением деформаций высшею порядка, определяемых членами, содержащими z в степени выше второй. Уравнение (2) применимо лишь для параболоидов вращения второго и высшего порядков. С другой стороны, уравнение (1) не пригодно для описания планоидных поверхностей типа пластинок Шмидта, у которых u неоднозначная функция от z.

Более универсальным является уравнение, используемое в зарубежной литературе [4], в котором z есть функция от u, как и в выражении (2). Однако в отличие от него целесообразно выделить член, позволяющий описывать любую поверхность второю порядка:

$$z = \frac{\rho_0 u}{1 + \sqrt{1 - \rho_0^2 (1 - e^2) u}} + P, \tag{3}$$

где ρ_0 – кривизна при вершине, r_0 – радиус кривизны, e^2 – квадрат эксцентриситета образующей поверхности второго порядка, $P = \sum_{k=2} P_k u^k$.

Уравнение (3), как и выражение (2), не позволяет описывать глубокие поверхности, у которых u неоднозначная функция от z. Кроме того, в процессе расчета хода лучей возникают сложности, когда выражение под радикалом становится отрицательным. Легко увидеть, что члены уравнения (1) со степенями z выше третьей описывают деформацию поверхности высшею порядка относительно поверхности второго порядка в радиальном направлении, а соответствующие члены высшею порядка уравнений (2) и (3) – в осевом направлении.

В качестве универсальной математической модели оптической поверхности может быть взято предложенное автором работы [5] следующее уравнение, описывающее поверхность:

$$f(r) = f_0(r) - Q(r) = 0, \tag{4}$$

где r – радиус-вектор точки, $r^T = (x,y,z)$ $f_0(r)$ – функция, описывающая базовую поверхность в виде

$$f_0(r) = 0; \tag{5}$$

$Q(r)$ – член добавочной деформации.

Уравнение базовой поверхности (5) должно быть приведено к такому виду, чтобы нормаль к поверхности в начале координат совпадала с ортом k оси z. Так как вектор нормали есть вектор градиента функции $f(r)$, то это условие имеет вид

$$g = \nabla f_0(r) = \begin{vmatrix} \partial f_0/\partial x \\ \partial f_0/\partial y \\ \partial f_0/\partial z \end{vmatrix} = k, \quad \text{при} \ \ r = 0. \tag{6}$$

В частности, уравнение сферы

$$f_0(r) = z - \rho_0/2(x^2 + y^2 + z^2) = 0. \tag{7}$$

Нормаль к сфере

$$g = k - \rho_0 r, \tag{8}$$

причем длина этой нормали всегда равна единице: $\|g\|^2 = 1$.

Уравнение поверхности вращения второю порядка

$$f_0(r) = z - \rho_0/2\left[u + (1-e^2)z^2\right] = 0, \tag{9}$$

где e^2 – квадрат эксцентриситета. Длина нормали

$$\|g\|^2 = 1 + \rho^2 e^2 u. \tag{10}$$

Для цилиндрических поверхностей второю порядка достаточно положить в формуле (9) $u = y^2$ для цилиндров с осью, параллельной оси X и $u = x^2$ для цилиндров с осью, параллельной оси Y.

Рассмотрим связь члена Q в выражении (4) с деформацией поверхности относительно базовой, например, относительно поверхности второго порядка.

Пусть точка r_0 лежит на бязевой поверхности, т.е. r_0 удовлетворяет уравнению поверхности (5):

$$f_0(r_0) = 0. \tag{11}$$

Пусть точка $r = r_0 + \Delta r$, где Δr – вектор деформации, лежащий на деформированной поверхности, описываемой уравнением (4):

$$f(r) = f_0(r) - Q(r) = f_0(r_0 + \Delta r) - Q(r) = 0. \tag{12}$$

При малых величинах Δr $f_0(r_0 + \Delta r)$ в выражении (12) можно разложить и ряд Тейлора по степеням Δr:

$$f(r) = f_0(r) + g^T \Delta r + 1/2 \Delta r^T H \Delta r - Q(r) = 0, \tag{13}$$

где $g = \nabla f_0(r)$ – нормаль к базовой поверхности, H – матрица Гессе уравнения (5), т.е.

$$H = \frac{\partial^2 f_0(r)}{\partial r^2}. \tag{14}$$

Подставим Δr в виде

$$\Delta r = g \frac{l}{\|g\|}, \tag{15}$$

т. е. будем измерять деформацию вдоль нормали к базовой поверхности. Подставляя формулу (15) в выражение (13) с учетом уравнения (12), получим

$$Q(r) = l\|g\| \left(1 + \tfrac{1}{2} g^T H g \frac{l}{\|g\|^2} \right) \tag{16}$$

Оценим второй член в скобках уравнения (16). Возьмем наиболее распространенный случай базовой поверхности, близкой к сферической. В других случаях будут более сложные выражения, но порядок величины этого члена останется таким же. Для сферы матрица Гессе (14) $H = \rho_0 I$, где ρ_0 – кривизна. I – единичная матрица. Таким образом, с учетом выражений (7) и (8) имеем

$$Q(r) = l\|g\| \left(1 + \frac{l}{2r_0} \right). \tag{17}$$

при деформациях l, не превышающих несколько процентов радиуса криви ты. Членом $\frac{l}{2r_0}$ можно пренебречь, и тогда из формулы (17) следует, что член $Q(r)$ в уравнении (4) описывает деформацию относительно базовой поверхности вдоль нормали к этой поверхности, умноженную на длину нормали, которая определяется выражением (10). Именно это свойство делает уравнение (4) более универсальным и соответствующим практике измерения деформацией при контроле несферических поверхностей по сравнению с уравнениями (1) – (3). Для описания поверхностей вращения высшего порядка член $Q(r)$ можно представить в двух формах:

$$Q(r) = a_2\left(\frac{u}{h_y^2}\right)^2 + a_3\left(\frac{u}{h_y^2}\right)^3 + \ldots \qquad (18)$$

и

$$Q(r) = b_3\left(\frac{z}{h_z}\right)^3 + b_4\left(\frac{u}{h_z}\right)^4 + \ldots, \qquad (19)$$

где h_y и h_z. – нормировочные факторы, которые не равны нулю и выбраны вследствие их целесообразности. Легко установить связь коэффициентов a_k и b_k универсального уравнения (4) с коэффициентами выражений (1), (2), а именно [5]:

$$r_0 = c_1/2; \quad e^2 = 1 + c_2; \quad b_k = -c_1^{-1} c_k h_z^k, \quad k = 3,4\ldots \qquad (20)$$

$$r_0 = 1/(2d_2); \quad e^2 = 1; \quad b_k = d_{2k} h_y^{2k}, \quad k = 2,3,\ldots \qquad (21)$$

Однако для уравнения (3) эта связь не столь проста. Путем несложных алгебраических преобразовании можно получить

$$(z - P) = \rho_0/2\left[u + (1 - e^2)(z - P)^2\right] = 0, \qquad (22)$$

или

$$z - \rho_0/2\left[u + (1 - e^2)z^2\right] - \left\{-\rho(1 - e^2)Pz + \frac{\rho_0}{2}(1 - e^2)P^2 + P\right\} = 0. \qquad (23)$$

Сравнивая выражения (23) и (4) с формулой (9), приходим к выводу, что

$$Q = P\left\{1 - \rho(1 - e^2)\left[z - \tfrac{1}{2}P\right]\right\}, \qquad (24)$$

но из формулы (3) можно выразить z в виде

$$\rho_0(1 - e^2)z = 1 - \sqrt{1 - \rho_0^2(1 - e^2)u} + \rho_0(1 - e^2)P. \qquad (25)$$

Подставляя в уравнение (24), получим

$$Q = P\sqrt{1 - \rho_0^2(1 - e^2)u} - P^2\frac{\rho_0}{2}(1 - e^2)$$

или, обозначая $c = (1 - e^2)$,

$$Q = P\sqrt{1 - \rho_0^2 cu} - P^2\frac{\rho_0}{2}c. \qquad (26)$$

Пусть $P = \sum\limits_{k=2} p_k u^k$; $Q = \sum\limits_{k=2} q_k u^k$ в соответствии с выражениями (3) и (18). Представим также радикал в формуле (26) в виде ряда Тейлора $\sqrt{1-\alpha} = 1 - \sum\limits_{k=1} r_k \alpha^k$, где r_k - коэффициенты ряда Тейлора для радикала $\sqrt{1-\alpha}$. Тогда из уравнения (26)

$$\sum q_k u^k = \sum p_k u^k (1 - \sum r_i \rho^{2i} c^i u^i) - \sum\sum p_e p_j u^{i+j}\frac{\rho_0}{2}c. \qquad (27)$$

Раскрывая скобки и приравнивая затем почленно, получаем соотношения для связи коэффициентов:

$$\left.\begin{aligned}
&q_2 = p_2;\\
&q_3 = p_3 - p_2 r_1 \rho^2 c;\\
&q_4 = p_4 - p_3 r_1 \rho^2 c - p_2 r_2 \rho^4 c^2 - p_2^2 \frac{\rho}{2} c;\\
&q_5 = p_5 - p_4 r_1 \rho^2 c - p_3 r_2 \rho^4 c^2 - p_2 r_3 \rho^6 c^3 - p_3 p_2 \rho c;\\
&q_6 = p_6 - p_5 r_1 \rho^2 c - p_4 r_2 \rho^4 c^2 - p_3 r_3 \rho^6 c^3 - p_2 r_4 \rho^8 c^4 - p_4 p_2 \rho c - p_3^2 \frac{\rho}{2} c.
\end{aligned}\right\} \quad (28)$$

И обратно:

$$\begin{aligned}
&p_2 = q_2;\\
&p_3 = q_3 + p_2 r_1 \rho^2 c; \qquad (29)\\
&p_4 = q_4 + p_3 r_1 \rho^2 c + p_2 r_2 \rho^4 c^2 + p_2^2 \frac{\rho}{2} c
\end{aligned}$$

и т.д.

Рекурсивные соотношения (28) и (29) вместе с формулами (20) и (21) позволяют осуществить точный переход от любого из применяемых в настоящее время уравнений к универсальному уравнению (4) и обратно.

Такие возможности реализованы в пакете OPAL-PC, где расчет лучей производится по стандартным формулам на основе универсального уравнения (4), а конструктор-оптик имеет возможность задать поверхность в любом виде, используя уравнения (1) – (3) или (4), и в любой момент перейти к другому виду уравнения, с автоматическим пересчетом коэффициентов.

ЛИТЕРАТУРА

1. Feder D.P. Optical calculation with automatic computing mashinery // JOSA, 1951. Vol.41. P.630-641.
2. Леонова И.Б. Автоматизирование расчетов оптических систем. Л.: Машиностроение. 1987. 200с.
3. Слюсарев Г.Г. Методы расчета оптических систем. Л.: Машиностроение. 1967. 670с.
4. Feder D.P. Calculation of an optical merit function and its derivatives with respect to the system parameters //JOSA. 1957. Vol.47. P.913-925.
5. Родионов С.А. Автоматизация проектирования оптических систем. Л.: Машиностроение. 1982. 270с.