

ДИФРАКЦИОННАЯ ОПТИКА, ОПТИЧЕСКИЕ ТЕХНОЛОГИИ

Метод опорных квадрик в задачах неизображающей оптики, допускающих переформулировку в виде задачи перемещения масс

А.А. Мингазов¹, Л.Л. Досколович^{1,2}, Д.А. Быков^{1,2}, Е.В. Бызов¹

¹ ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН,
443001, Россия, г. Самара, ул. Молодогвардейская, д. 151;

² Самарский национальный исследовательский университет имени академика С.П. Королёва,
443086, Россия, г. Самара, Московское шоссе, д. 34

Аннотация

В статье рассматриваются обратные задачи расчёта оптических элементов для формирования заданных распределений освещённости, которые могут быть переформулированы как задачи перемещения масс Монжа–Канторовича. Для всех задач такого типа мы единообразно формулируем метод опорных квадрик и показываем, что он совпадает с градиентным методом для нахождения максимума некоторой вогнутой функции.

Ключевые слова: неизображающая оптика, геометрическая оптика, обратная задача, задача Монжа–Канторовича, метод опорных квадрик.

Цитирование: Мингазов, А.А. Метод опорных квадрик в задачах неизображающей оптики, допускающих переформулировку в виде задачи перемещения масс / А.А. Мингазов, Л.Л. Досколович, Д.А. Быков, Е.В. Бызов // Компьютерная оптика. – 2022. – Т. 46, № 3. – С. 353-365. – DOI: 10.18287/2412-6179-CO-1055.

Citation: Mingazov AA, Doskolovich LL, Bykov DA, Byzov EV. Support quadric method in non-imaging optics problems that can be reformulated as a mass transfer problem. Computer Optics 2022; 46(3): 353-365. DOI: 10.18287/2412-6179-CO-1055.

Введение

Задача расчёта преломляющей или отражающей оптической поверхности из условия формирования заданного распределения освещённости в некоторой области относится к классу обратных задач неизображающей оптики и является крайне сложной. В приближении геометрической оптики данная задача сводится к решению нелинейного дифференциального уравнения (НДУ) эллиптического типа. Решение этого НДУ является сложной теоретической и вычислительной задачей. Несмотря на то, что для решения данного НДУ были предложены различные конечно-разностные методы [1–7], расчёт оптических элементов в рамках такого подхода является сложным и имеет ограничения. В частности, формулировка задачи расчёта оптического элемента в виде НДУ предполагает, что рассчитываемые поверхности оптического элемента являются гладкими. Требование гладкости ограничивает класс распределений освещённости, которые могут быть сформированы оптическим элементом. Например, оптический элемент с гладкими поверхностями не позволяет сформировать распределение освещённости, определённое в несвязной области или в области с негладкими границами.

Одним из наиболее универсальных подходов к данной задаче является метод опорных квадрик [8–10]. Он состоит в следующем. Требуемое распределение освещённости приближается дискретным рас-

пределением, сосредоточенным в конечном числе точек. Оптический элемент представляется в виде кусочно-гладкой поверхности, состоящей из фрагментов оптических элементов, каждый из которых фокусирует падающий на элемент пучок в одну точку. В зависимости от задачи эти поверхности могут быть параболами вращения, эллипсоидами, гиперболоидами или более сложными поверхностями (например, [10]). Далее параметры фрагментов подбираются итерационно в зависимости от того, насколько отличается количество энергии, приходящее в каждую точку, от требуемого значения.

Мы рассматриваем задачи неизображающей оптики, формулируемые в виде задачи перемещения масс Монжа–Канторовича. В первом параграфе данной статьи мы приводим описание всех оптических элементов, для которых известно к настоящему моменту, что задача расчёта элемента из условия формирования заданного распределения освещённости может быть представлена в виде задачи перемещения масс. Для всех этих задач мы единообразно формулируем метод опорных квадрик (МОК), доказывая его совпадение с градиентным методом для нахождения максимума некоторой вогнутой функции. Это переносит результаты статьи [11] на широкий класс задач.

В статьях [8, 9] в качестве критерия качества в МОК используется среднеквадратичное отклонение между формируемым и требуемым распределениями, которое в процессе работы алгоритма меняется не-

тривиальным образом. Результат данной статьи позволяет использовать другой критерий качества для описанного ряда задач, который имеет некоторые преимущества по сравнению со среднеквадратичным отклонением, в частности, в процессе работы алгоритма опорных квадрик изменяется монотонно.

Статья организована следующим образом. В первом параграфе описываются оптические элементы, для которых задача расчёта из условия формирования заданного распределения освещённости может быть переформулирована как задача перемещения масс Монжа–Канторовича, а также их представления в виде огибающих специального вида. Во втором параграфе мы единообразно формулируем задачу расчёта оптического элемента из условия формирования заданного распределения освещённости. В третьем параграфе мы формулируем итерационный метод опорных квадрик для расчёта описанных оптических элементов. Затем в четвёртом параграфе показываем, что МОК является градиентным методом для поиска максимума некоторой выпуклой функции. В последнем параграфе мы приводим два примера расчёта преломляющих элементов, формирующих заданные распределения освещённости в дальней зоне. На первом из них мы иллюстрируем преимущества рассмотрения в качестве критерия сходимости метода опорных квадрик предложенной выпуклой функции перед среднеквадратичным отклонением. В статье использована единая сквозная нумерация определений, теорем, лемм и утверждений.

1. Описание рассматриваемых задач и представление поверхностей оптических элементов в виде огибающих

В данном параграфе мы кратко опишем несколько оптических элементов, для которых задачи формирования заданных распределений освещённости допускают переформулировку в виде задачи перемещения масс (ЗПМ) Монжа–Канторовича с некоторой функцией стоимости \mathcal{K} . Данные оптические элементы рассмотрены в отдельных работах [11–23]. Для каждого элемента мы введём понятие прямой φ и т.н. двойственной функции ψ и лучевого отображения T . В следующем параграфе мы, используя введённые обозначения, приводим общую для всех оптических элементов формулировку проблемы в терминах функций ψ , T и функций I , L , представляющих интенсивность источника и требуемое распределение освещённости (или интенсивности).

1.1. Преломляющий (отражающий) элемент, формирующий заданное распределение интенсивности при точечном источнике

Пусть имеется точечный источник света, расположенный в начале координат в трёхмерном пространстве с координатами (x_1, x_2, x_3) . Рассмотрим еди-

ничную сферу \mathbb{S} с центром в начале координат. Положение точки

$$\bar{x} = (x_1, x_2, \sqrt{1 - x_1^2 - x_2^2}) \in \mathbb{S}$$

на полусфере $x_3 > 0$ будем задавать координатами $x = (x_1, x_2)$. Источник света будем характеризовать функцией интенсивности $I'(x_1, x_2)$, $x \in G$, равной световому потоку, излучаемому в единицу телесного угла. Обозначим $d\sigma$ элемент телесного угла, тогда световой поток через область $\omega \subset \mathbb{S}$ вычисляется как

$$\Phi_s(\omega) = \int_{\omega} I'(x) d\sigma = \int_{\omega} \frac{I'(x)}{\sqrt{1 - x_1^2 - x_2^2}} dx_1 dx_2.$$

Далее нам будем удобно характеризовать излучение источника функцией

$$I(x) = \frac{I'(x)}{\sqrt{1 - x_1^2 - x_2^2}}, x \in G. \tag{1}$$

Пусть излучение от точечного источника падает на преломляющую поверхность W . Будем считать, что показатель преломления среды под поверхностью W (в области расположения источника) равен n_1 , а в области над поверхностью – n_2 .

Рассмотрим некоторый луч, вышедший из начала координат (из источника) с направлением \bar{x} . Данный луч преломляется на поверхности W и приобретает направление единичного вектора

$$\bar{y} = (y_1, y_2, \sqrt{1 - y_1^2 - y_2^2}) \in \mathbb{S}.$$

Это определяет «лучевое отображение» $T: G \rightarrow \mathbb{S}$. Отображение T задаёт на сфере \mathbb{S} функцию распределения интенсивности $L'(y)$ для преломлённого пучка. Функция $L'(y)$ однозначно определяется условием сохранения светового потока, который в интегральной форме имеет вид:

$$\Phi_S(T^{-1}(\omega)) = \Phi_{im}(\omega) \quad \forall \omega \subset \mathbb{S}, \tag{2}$$

где

$$\Phi_{im}(\omega) = \int_{\omega} L'(y) d\sigma = \int_{\omega} \frac{L'(y_1, y_2)}{\sqrt{1 - y_1^2 - y_2^2}} dy -$$

световой поток преломлённого пучка через область $\omega \subset \mathbb{S}$. Далее мы будем характеризовать распределение интенсивности преломлённого пучка функцией

$$L(y) = \frac{L'(y)}{\sqrt{1 - y_1^2 - y_2^2}}. \tag{3}$$

Преломляющую поверхность W удобно задать функцией $\varphi(x) = -\ln \rho(x)$, где $\rho(x)$ – расстояние от начала координат до поверхности в направлении

$$\bar{x} = (x_1, x_2, \sqrt{1 - x_1^2 - x_2^2}).$$

Рассматриваемая в данном пункте задача состоит в расчёте функции $\varphi(x)$, задающей преломляющую поверхность, из условия формирования заданного распределения интенсивности $L(y), y \in D$.

Для решения данной задачи искомую поверхность удобно представить в виде огибающей семейства эллипсоидов $\varphi^{p,y}(x) = \ln(1 - n\langle \bar{x}, \bar{y} \rangle) - \ln p(y)$, где $n = n_2/n_1$, $p(y)$ – фокальный параметр [14], по параметрам $y = (y_1, y_2) \in D$. Эллипсоид $\varphi^{p,y}(x)$ преобразует падающий сферический пучок от источника в коллимированный пучок с направлением $y = (y_1, y_2)$. Уравнение огибающей семейства эллипсоидов имеет вид [14]:

$$\begin{cases} \varphi(x) = \mathcal{K}(x, y) - \psi(y), \\ \frac{\partial}{\partial y_i} [\mathcal{K}(x, y) - \psi(y)] = 0, \quad i = \overline{1, 2}, \end{cases} \quad (4)$$

где $\psi(y) = \ln p(y)$ – так называемый двойственный параметр, а функция $\mathcal{K}(x, y)$ имеет вид

$$\mathcal{K}(x, y) = \ln(1 - n\langle \bar{x}, \bar{y} \rangle). \quad (5)$$

В [12–15] показано, что задача расчёта преломляющей поверхности из условия формирования заданного распределения интенсивности эквивалента ЗПМ Монжа–Канторовича с функцией стоимости $\mathcal{K}(x, y)$, заданной уравнением (5). Формулировка ЗПМ приведена в теореме 4 следующего параграфа.

Формально подставив в (5) $n = -1$, мы получим верные формулы для случая отражающей оптической поверхности. Подробно данная задача рассмотрена в [12–15].

1.2. Преломляющий (отражающий) элемент, формирующий заданное распределение интенсивности при освещающем пучке с плоским волновым фронтом

Пусть в трёхмерном пространстве с координатами (x_1, x_2, x_3) в направлении $k = (0, 0, 1)$ распространяется световой пучок с плоским волновым фронтом, имеющий в плоскости Π , задаваемой уравнением $z = 0$, распределение освещённости $I(x), x = (x_1, x_2) \in G$. Данный пучок падает на преломляющую поверхность W , задаваемую уравнением $z = \varphi(x)$. Будем считать, что в области $z < \varphi(x)$ показатель преломления среды равен n_1 , а в области $z > \varphi(x)$ равен n_2 . Луч, вышедший из точки $x = (x_1, x_2)$ плоскости $z = 0$, преломляется на поверхности W и приобретает направление единичного вектора

$$\bar{y} = (y_1, y_2, \sqrt{1 - y_1^2 - y_2^2}) \in \mathbb{S}.$$

Это определяет лучевое отображение $T: G \rightarrow \mathbb{S}$. Данное отображение задаёт на сфере \mathbb{S} функцию распределения интенсивности $L'(y)$ для преломлённого пучка. Аналогично предыдущему случаю распределение интенсивности преломлённого пучка будем за-

давать функцией $L(y)$, определённой формулой (3). Функция $L(y)$ определяется условием сохранения светового потока

$$\Phi_{\Pi}(T^{-1}(\omega)) = \Phi_{im}(\omega) \quad \forall \omega \subset \mathbb{S}, \quad (6)$$

где

$$\Phi_{im}(\omega) = \int_{\omega} L(y) dy -$$

световой поток преломленного пучка через область на сфере \mathbb{S} , а

$$\Phi_{\Pi}(T^{-1}(\omega)) = \int_{T^{-1}(\omega)} I(x) dx -$$

световой поток падающего пучка через соответствующую область плоскости $z = 0$.

Рассматриваемая в данном пункте задача состоит в расчёте преломляющей поверхности W из условия формирования заданного распределения $L(y), y \in D$. Аналогично предыдущему случаю представим поверхность $z = \varphi(x)$ в виде огибающей семейства поверхностей, преобразующих падающий пучок в пучок с плоским волновым фронтом и заданным направлением $y = (y_1, y_2) \in D$. В данной задаче такими поверхностями являются плоскости [16] вида

$$z = \mathcal{K}(x, y) - \psi(y), \quad (7)$$

где

$$\mathcal{K}(x, y) = \frac{nx_1y_1}{1 - n\sqrt{1 - y_1^2 - y_2^2}} + \frac{nx_2y_2}{1 - n\sqrt{1 - y_1^2 - y_2^2}}, \quad (8)$$

где $n = n_2/n_1$. Двойственный параметр $\psi(y)$ в данной задаче имеет геометрический смысл смещения плоскости (7) относительно начала координат. В результате, так же как и в предыдущем случае, преломляющая поверхность $z = \varphi(x)$ может быть представлена в виде огибающей семейства плоскостей, задаваемой уравнениями (4).

Данная задача также эквивалента ЗПМ Монжа–Канторовича с функцией стоимости (8). Формулировка ЗПМ приведена в теореме 4 следующего параграфа. Заменой координат можно добиться того, что функция стоимости станет квадратичной. Подробнее данная задача рассмотрена в [15, 16].

Формально подставив $n = -1$ в выражение (8) для функции стоимости \mathcal{K} , можно получить формулы для случая отражающего оптического элемента, формирующего заданное распределение интенсивности.

1.3. Элемент из двух отражающих поверхностей, формирующий заданное распределение освещённости и плоский волновой фронт при освещающем пучке с плоским волновым фронтом

Рассмотрим трёхмерное пространство с координатами (x_1, x_2, z) . Пусть в области G плоскости $z = 0$ задана функция $I(x), x = (x_1, x_2) \in G$, определяющая рас-

пределение освещённости коллимированного светового пучка, распространяющегося в направлении $k=(0, 0, 1)$. Пучок падает на оптический элемент с двумя отражающими поверхностями. Пусть первая поверхность элемента R_1 задаётся функцией $z=\varphi(x)$, $x \in G$. Зафиксируем некоторую «выходную» плоскость $z=l$, в которой будет рассматриваться преобразованный элемент пучок, и обозначим $y=(y_1, y_2)$ декартовы координаты на этой плоскости. Вторую отражающую поверхность R_2 зададим относительно выходной плоскости с помощью функции $\psi(y)$ следующим образом $z=l-\psi(y)$.

Мы предполагаем, что луч, выходящий из точки $x \in G$ плоскости $z=0$ с направлением $k=(0, 0, 1)$, последовательно отражается от двух поверхностей R_1, R_2 и затем попадает в некоторую точку y области D в выходной плоскости $z=l$, сохраняя при этом направление $k=(0, 0, 1)$. Это определяет лучевое отображение $T: G \rightarrow D$. В результате на плоскости $z=l$ формируется некоторое распределение освещённости $L(y)$, определяемое условием сохранения светового потока, которое в данной задаче принимает вид:

$$\int_{T^{-1}(\omega)} I(x) dx = \int_{\omega} L(y) dy, \quad \forall \omega \subset D. \quad (9)$$

Задача состоит в расчёте функций $\varphi(x)$ и $\psi(y)$, задающих поверхности элемента, из условия формирования выходного пучка, имеющего в плоскости $z=l$ заданное распределение освещённости $L(y)$ и плоский волновой фронт. В данной задаче мы будем рассматривать первую поверхность R_1 как огибающую семейства поверхностей, фокусирующих в точки второй поверхности R_2 [15, 17]. В этом случае поверхность семейства должна отражать падающий коллимированный пучок в некоторую точку с координатами $(y, l-\psi(y))$ на второй поверхности. Нетрудно понять, что такими поверхностями являются параболоиды вращения с фокусами в точках второй поверхности, задаваемые уравнением

$$\varphi^{y, \psi(y)}(x) = \mathcal{K}(x, y) - \psi(y),$$

где

$$\mathcal{K}(x, y) = \frac{Q^2 - l^2 - \|x - y\|^2}{2(L - l)}, \quad (10)$$

где Q – оптическая длина пути лучей между плоскостями $z=0$ и $z=l$ [17]. Отметим, что оптическая длина пути лучей между плоскостями $z=0$ и $z=l$ является постоянной в силу того, что волновые фронты падающего и выходного являются плоскими. Уравнения огибающей аналогично записываются в виде (4). Данная задача также эквивалента ЗПМ с функцией стоимости \mathcal{K} [17]. Формулировка ЗПМ приведена в следующем параграфе. Подробное описание данной задачи дано в [15, 17].

1.4. Элемент из двух преломляющих поверхностей, формирующий заданное распределение освещённости и плоский волновой фронт при освещающем пучке с плоским волновым фронтом

Рассмотрим трёхмерное пространство с координатами (x_1, x_2, z) . Пусть в области G плоскости $z=0$ задана функция $I(x)$, $x=(x_1, x_2) \in G$, определяющая распределение освещённости коллимированного светового пучка, распространяющегося в направлении $k=(0, 0, 1)$. Пучок падает на оптический элемент с двумя преломляющими поверхностями. Первая поверхность R_1 задаётся функцией $z=\varphi(x)$, $x \in G$. Как и в предыдущей задаче, вторую поверхность R_2 зададим относительно некоторой выходной плоскости $z=l$ через функцию $\psi(y)$ в виде $z=l-\psi(y)$, где $y=(y_1, y_2)$ – декартовы координаты на выходной плоскости. Будем считать, что показатель преломления среды при $z \leq \varphi(x)$ равен n_1 , при $\varphi(x) \leq z \leq l-\psi(x)$ равен n_2 , а при $l-\psi(x) \leq z$ снова равен n_1 .

Мы предполагаем, что луч, выходящий из точки $x \in G$ с направлением $k=(0, 0, 1)$, последовательно преломляется на двух поверхностях R_1, R_2 и затем попадает в некоторую точку y области D в выходной плоскости $z=l$, сохраняя направление $k=(0, 0, 1)$. Это определяет лучевое соответствие $T: G \rightarrow D$. В результате на плоскости $z=l$ формируется некоторое распределение освещённости $L(y)$ выходного пучка, определяемое условием сохранения светового потока (9).

Задача состоит в расчёте функций $\varphi(x)$ и $\psi(y)$, задающих поверхности элемента, из условия формирования выходного пучка, имеющего в плоскости $z=l$ заданное распределение освещённости $L(y)$ и плоский волновой фронт.

Как и для случая двух отражающих поверхностей, мы будем рассматривать первую поверхность R_1 как огибающую семейства поверхностей, фокусирующих в точки второй поверхности R_2 [11, 19]. В этом случае имеют место уравнения огибающей (4) с функцией

$$\mathcal{K}(x, y) = \frac{n_2^2 l - n_1 Q}{n_2^2 - n_1^2} - \frac{n_2}{|n_2^2 - n_1^2|} \times \sqrt{(Q - n_1 l)^2 - (n_2^2 - n_1^2) \|y - x\|^2}, \quad (11)$$

где Q – оптическая длина пути лучей между плоскостями $z=0$ и $z=l$. Данная задача также эквивалента ЗПМ Монжа–Канторовича с функцией стоимости (11) [11, 15, 18]. Формулировка ЗПМ приведена в следующем параграфе. Подробное описание данной задачи приведено в [11, 15, 18, 19].

1.5. Элемент из двух отражающих поверхностей, формирующий заданное распределение освещённости и плоский фронт при точечном источнике

Рассмотрим трёхмерное пространство с координатами (x_1, x_2, z) . Пусть в начале координат расположен точечный источник света. Обозначим через \mathbb{S} сферу единичного радиуса с центром в источнике. Интен-

сивность источника, как и в случае первого элемента, описанного в пункте 1.1, будем характеризовать функцией $I(x_1, x_2)$, $x = (x_1, x_2) \in G$, которая связана с фактической интенсивностью $I'(x)$ соотношением (1).

Сферический пучок от точечного источника падает на оптический элемент с двумя отражающими поверхностями. Пусть первая отражающая поверхность R_1 задаётся функцией $u(x)$, $x \in G$:

$$R_1 = \{u_1(x) \bar{x} \mid x \in G\},$$

где

$$\bar{x} = (x_1, x_2, \sqrt{1 - x_1^2 - x_2^2}).$$

Зафиксируем некоторую «выходную» плоскость $z = l$, в которой будет рассматриваться преобразованный элемент пучок, и обозначим $y = (y_1, y_2)$ декартовы координаты на этой плоскости. Вторую поверхность R_2 зададим функцией $u_2(y)$ следующим образом

$$R_2 = \{(y, z) \in \mathbb{E}^3 \mid z = l - u_2(y)\}.$$

Мы предполагаем, что луч, исходящий из источника в направлении \bar{x} , последовательно отражается от поверхностей R_1, R_2 и затем попадает в точку y некоторой области D на плоскости $z = l$, имея при этом направление $k = (0, 0, 1)$. Это определяет лучевое соответствие $T: G \rightarrow D$. Оно определяет в области D плоскости $z = l$ распределение освещённости $L(y)$ выходного пучка, определяемое условием сохранения светового потока

$$\int_{T^{-1}(\omega)} I(x) dx = \int_{\omega} L(y) dy, \quad \forall \omega \subset D. \quad (12)$$

Задача состоит в расчёте отражающих поверхностей R_1, R_2 , задаваемых функциями u_1, u_2 , из условия формирования выходного пучка, имеющего в плоскости $z = l$ заданное распределение освещённости $L(y)$ и плоский волновой фронт.

Для получения соотношения вида (4) в данной задаче поверхности R_1, R_2 следует задавать другими функциями $\varphi(x)$ и $\psi(y)$, выражающимися через $u_1(x), u_2(y)$, достаточно сложным образом [20, 21]. Пусть Q – оптическая длина пути лучей от источника до выходной плоскости. Отметим, что данная оптическая длина пути является постоянной, поскольку волновой фронт выходного пучка является плоским. В рассматриваемой задаче прямая и двойственная функции вводятся следующим образом [21]:

$$\begin{aligned} \varphi(x) &= \ln \left(\frac{1}{2u_1(x)(1 - \langle \bar{x}, k \rangle)} - \frac{1}{2(Q-l)} \right), \\ \psi(y) &= \ln \left(\frac{u_2(y)}{Q^2 - l^2 - \|y\|^2} - \frac{1}{2(Q-l)} \right). \end{aligned} \quad (13)$$

В терминах функций φ и ψ снова можно записать соотношение (4), рассматривая φ как огибающую се-

мейства функций $\varphi^{\psi(y)}(x) = \mathcal{K}(x, y) - \psi(y)$, соответствующих поверхностям, фокусирующим падающий пучок в точки второй поверхности. В данной задаче функция стоимости имеет вид [20, 21]

$$\begin{aligned} \mathcal{K}(x, y) &= \\ &= \ln \left(\frac{1}{4(Q-l)^2} - \frac{Q - \langle \bar{x}, \eta \rangle}{2(Q-l)(Q^2 - l^2 - \|y\|^2)(1 - \langle \bar{x}, k \rangle)} \right), \end{aligned} \quad (14)$$

где $\eta = (y, l)$. Условие огибающей снова может быть записано в виде (4). Данная задача эквивалента [20, 21] ЗПМ Монжа–Канторовича с функцией стоимости (14). Формулировка ЗПМ приведена в следующем параграфе. Подробнее данная задача рассмотрена в [20, 21].

1.6. Элемент из двух отражающих поверхностей, формирующий заданное распределение освещённости и сферический фронт при точечном источнике

Рассмотрим трёхмерное пространство с координатами (x_1, x_2, z) . Пусть в начале координат расположен точечный источник света. Обозначим через \mathbb{S}_1 сферу единичного радиуса с центром в источнике. Интенсивность источника света, как и для элемента в пункте 1.1, будем характеризовать функцией $I(x_1, x_2)$, $x = (x_1, x_2) \in G$, заданной соотношением (1).

Сферический пучок от точечного источника падает на оптический элемент с двумя отражающими поверхностями. Пусть первая поверхность R_1 задаётся функцией $u_1(x)$, $x \in G$:

$$R_1 = \{u_1(x) \bar{x} \mid x \in G\},$$

где

$$\bar{x} = (x_1, x_2, \sqrt{1 - x_1^2 - x_2^2}).$$

Рассмотрим также точку \mathcal{F} с координатами $(0, 0, F)$ и единичную сферу \mathbb{S}_2 с центром в этой точке. Координаты на \mathbb{S}_2 обозначим $y = (y_1, y_2)$, эти координаты определяют точку $\mathcal{F} + \bar{y}$ сферы \mathbb{S}_2 , где

$$\bar{y} = (y_1, y_2, \sqrt{1 - y_1^2 - y_2^2}).$$

Вторую отражающую поверхность зададим функцией $u_2(y)$ следующим образом

$$R_2 = \{\mathcal{F} + u_2(y) \bar{y} \mid y \in D\}.$$

Мы предполагаем, что луч, исходящий из источника в направлении \bar{x} , последовательно отражается от поверхностей R_1, R_2 и затем приходит в точку \mathcal{F} , пересекая сферу \mathbb{S}_2 в некоторой точке с координатами (y_1, y_2) . Это определяет лучевое отображение $T: G \rightarrow D$. Данное отображение задаёт на сфере \mathbb{S}_2 функцию распределения интенсивности $L'(y)$ для преобразованного элемента пучка. Как и для элемента в пункте 1.1, формируемое на сфере \mathbb{S}_2 распределение интенсивности мы будем задавать функцией

$L(y)$, $y \in D$, которая связана с фактическим распределением интенсивности $L'(y)$ формулой (3).

Задача состоит в расчёте отражающих поверхностей R_1, R_2 , задаваемых функциями u_1, u_2 , из условия формирования выходного пучка, имеющего заданное распределение интенсивности $L(y)$ и сферический волновой фронт с центром в точке \mathcal{F} .

Как и в предыдущем случае, для получения соотношения вида (4) в данной задаче поверхности R_1, R_2 следует задавать другими функциями $\varphi(x)$ и $\psi(y)$, выражающимися через u_1, u_2 достаточно сложным образом [15].

Введём прямую и двойственную параметризацию системы отражающих поверхностей следующим образом [15]:

$$\begin{aligned} \varphi(x) &= \ln \left(\frac{1}{2u_1(x)(Q - F\langle k, \bar{x} \rangle)} - \frac{1}{Q^2 - F^2} \right), \\ \psi(y) &= \ln \left(\frac{1}{2u_2(y)(Q + F\langle k, \bar{y} \rangle)} - \frac{1}{Q^2 - F^2} \right), \end{aligned} \quad (15)$$

где Q – оптическая длина пути лучей от источника до точки \mathcal{F} .

Представляя одну из отражающих поверхностей в виде огибающей семейства поверхностей, фокусирующих в точки другой поверхности, можно получить представление (4) с функцией стоимости [15]

$$\mathcal{K}(x, y) = \ln \left(\frac{1}{(L^2 - F^2)^2} - \frac{1 - \langle \bar{x}, \bar{y} \rangle}{2(Q - F\langle k, \bar{x} \rangle)(Q + F\langle k, \bar{y} \rangle)(Q^2 - F^2)} \right). \quad (16)$$

Данная задача эквивалента ЗПМ Монжа–Канторовича с функцией стоимости (16) [15]. Формулировка ЗПМ приведена в следующем параграфе. Подробно данная задача рассмотрена в [15].

1.7. Функция эйконала на поверхности, обеспечивающая формирование заданного распределения освещённости на другой поверхности

Пусть в трёхмерном пространстве с координатами (x_1, x_2, z) расположены две поверхности: W_1 , заданная уравнением $z = f(x)$, и W_2 , заданная уравнением $z = g(x)$. На поверхности W_1 заданы функция эйконала $\varphi(x)$, $x = (x_1, x_2) \in G$ и распределение освещённости $I'(x)$, $x \in G$. Световой поток в области $\omega \subset W$ вычисляется как

$$\Phi_{W_1}(\omega) = \int_{\omega} I'(x) dS,$$

где dS – элемент площади поверхности. Нам будет удобно характеризовать распределение освещённости функцией

$$I(x) = I'(x) \sqrt{1 + f_{x_1}^2 + f_{x_2}^2},$$

где

$$f_{x_i} = \frac{\partial f}{\partial x_i}, \quad i = 1, 2.$$

В этом случае световой поток Φ_{W_1} вычисляется следующим образом

$$\Phi_{W_1}(\omega) = \int_{\omega} I(x) dx.$$

Функция эйконала $\varphi(x)$, $x = (x_1, x_2) \in G$ определяет направления лучей, исходящих из точек первой поверхности. Будем считать, что поверхности расположены в среде с единичным показателем преломления. В этом случае направления лучей, распространяющихся в положительном направлении оси z , задаются единичным вектором

$$p(x) = (p_1(x), p_2(x), \sqrt{1 - (p_1(x))^2 - (p_2(x))^2}),$$

который определяется из следующей системы уравнений [23]:

$$\begin{cases} \frac{\partial \varphi}{\partial x_1} = p_1 + f_{x_1} \cdot \sqrt{1 - p_1^2 - p_2^2}; \\ \frac{\partial \varphi}{\partial x_2} = p_2 + f_{x_2} \cdot \sqrt{1 - p_1^2 - p_2^2}. \end{cases} \quad (17)$$

Функция эйконала задаёт лучевое отображение T между точками поверхностей W_1, W_2 . Данное отображение определяется следующим образом. Образом точки $x \in G$ является точка $y = (y_1, y_2) \in D$, соответствующая пересечению поверхности W_2 и прямой, выходящей из точки $(x, f(x))$ поверхности W_1 , с единичным направляющим вектором $p = (p_1, p_2, \sqrt{1 - p_1^2 - p_2^2})$, определяемым из системы уравнений (17).

Отображение T задаёт на поверхности W_2 некоторое распределение освещённости. Аналогично поверхности W_1 будем характеризовать распределение освещённости функцией $L(y)$, $y \in D$, которая связана с фактическим распределением освещённости $L'(y)$ следующим соотношением:

$$L(y) = L'(y) \sqrt{1 + g_{y_1}^2 + g_{y_2}^2}.$$

Тогда результирующий световой поток для $\omega \subset D$ вычисляется

$$\Phi_{W_2}(\omega) = \int_{\omega} L(y) dy,$$

и функция $L(y)$ определяется условием сохранения светового потока

$$\int_{T^{-1}(\omega)} I(x) dx = \int_{\omega} L(y) dy, \quad \forall \omega \subset D. \quad (18)$$

Задача состоит в расчёте функции эйконала $\varphi(x)$, $x \in G$, из условия формирования заданного распределения освещённости $L(y)$, $y \in D$, на второй поверхности W_2 .

Аналогично предыдущим задачам, мы представим функцию эйконала $\varphi(x)$ в виде огибающей семейства функций эйконалов сферических пучков с фокусами в точках $(y, g(y))$ второй поверхности W_2 . Данные функции имеют вид [22, 23]

$$\varphi^{y,\psi(y)}(x) = \mathcal{K}(x, y) - \psi(y), \tag{19}$$

где

$$\mathcal{K}(x, y) = \sqrt{\|x - y\|^2 + (f(x) - g(y))^2}. \tag{20}$$

Огибающая семейства функций (19), (20) по параметрам $y = (y_1, y_2) \in D$ определяется соотношением (4). При этом функция $-\psi(y)$ совпадает с формируемым на поверхности W_2 эйконалом.

Данная задача эквивалентна задаче перемещения масс Монжа–Канторовича с функцией стоимости (20) [23]. Формулировка задачи перемещения масс приведена в следующем параграфе. Подробнее данная задача рассмотрена в [22, 23].

2. Формулировка обратной задачи

Сформулируем более явно роль представления поверхности оптического элемента в виде огибающей. Будем говорить, что произвольное лучевое отображение $T: G \rightarrow D$ задаётся соответствующим оптическим элементом, если существует некоторая функция $\psi(y)$, для которой выполнено:

$$\frac{\partial}{\partial y_i} [\mathcal{K}(x, y) - \psi(y)]|_{y=T(x)} = 0, \quad i = \overline{1, 2}. \tag{21}$$

Прямой и двойственный параметры оптического элемента при этом связаны соотношением:

$$\varphi(x) + \psi(T(x)) = \mathcal{K}(x, T(x)). \tag{22}$$

Определение 1. Будем называть отображение $T: G \rightarrow D$ интегрируемым, если оно удовлетворяет условию (21).

Определение 2. Будем называть отображение $T: G \rightarrow D$ удовлетворяющим условию сохранения светового потока, если для любого борелевского подмножества $\omega \subset D$ выполнено

$$\int_{T^{-1}(\omega)} I(x) dx = \int_{\omega} L(y) dy. \tag{23}$$

Условие сохранения светового потока может быть переформулировано следующим образом.

Лемма 3. Условие сохранения энергии для обратимого отображения $T: G \rightarrow D$ равносильно любому из условий:

- 1) для любого $\omega \subset D$ выполнено

$$\int_{T^{-1}(\omega)} I(x) dx = \int_{\omega} L(y) dy;$$

- 2) для любой функции $h \in C(D)$ выполнено

$$\int_D h(y)L(y) dy = \int_G h(T(x))I(x) dx;$$

- 3) $I(x) = L(T(x)) \cdot J_T(x)$, где $J_T(x)$ – якобиан отображения T в точке x .

Доказательство. Доказано в [17, lemma 4.11] или [24, VI.1.1].

Задачи расчёта всех рассмотренных в предыдущем параграфе оптических элементов можно сформулировать в одном общем виде следующим образом. Для заданных распределений $I(x)$ и $L(y)$ требуется найти лучевое отображение $T: G \rightarrow D$, которое удовлетворяет условию сохранения светового потока (23) и условию интегрируемости (21), а также восстановить прямой и двойственный параметры оптического элемента.

Данная задача может быть сформулирована как задача перемещения масс Монжа–Канторовича.

Теорема 4. Рассмотрим функционал

$$\mathcal{F}(T) = \int_G \mathcal{K}(x, T(x))I(x) dx, \tag{24}$$

заданный на пространстве отображений $T: G \rightarrow D$, удовлетворяющих условию сохранения светового потока (23).

Пусть отображение T_0 является точкой минимума данного функционала. Тогда оно является интегрируемым отображением.

Доказательство. Аналогично доказательству [11, Theorem 1].

3. Метод опорных квадрик

Мы будем рассматривать решения специального вида, для которых условие интегрируемости (21) задаёт не просто критическую точку функции $\mathcal{K}(x, y) - \psi(y)$ по переменной y , а точку минимума. Геометрически это соответствует тому, что поверхность оптического элемента, задаваемая функцией φ , расположена по одну сторону от всех касающихся её поверхностей $\varphi^{y,\psi(y)}(x) = \mathcal{K}(x, y) - \psi(y)$. В этом случае условия (21), (22) можно записать в виде

$$\varphi(x) = \min_{y \in D} (\mathcal{K}(x, y) - \psi(y)). \tag{25}$$

При этом экстремум достигается в точке $y = T(x)$, то есть соответствующее отображение $T = T^\psi: G \rightarrow D$ имеет вид

$$T^\psi(x) = \arg \min_{y \in D} (\mathcal{K}(x, y) - \psi(y)). \tag{26}$$

При такой формулировке возможно рассмотрение дискретного варианта задачи, когда формируемое распределение освещённости представлено обобщённой функцией

$$\sum_{i=1}^k L_i \delta_{y_i},$$

$y_i \in D$. В этом случае функция ψ представляет собой набор чисел, а функция φ является непрерывной и кусочно-гладкой. Далее мы будем рассматривать такую постановку задачи, которая соответствует аппроксимации непрерывного распределения $L(y)$ дискретным, заданным в конечном числе точек $\{y_i\}_{i=1}^k$.

Определение 5. Пусть $D = \{y_i\}_{i=1}^k$, $\psi_i = \psi(y_i)$. Обобщёнными взвешенными клетками Вороного с весами $\psi = (\psi_1, \dots, \psi_k)$ будем называть области $C^\psi(y_i)$ равные прообразам $(T^\psi)^{-1}(y_i)$. Из представления (26) следует, что клетки $C^\psi(y_i)$ могут быть определены следующим условием: точка $x \in G$ лежит в клетке $C^\psi(y_i)$ тогда и только тогда, когда

$$\mathcal{K}(x, y_i) - \psi_i \leq \mathcal{K}(x, y_j) - \psi_j \quad \forall j \neq i.$$

Обозначим световой поток, фокусирующийся в точку y_i , через

$$G_i(\psi) = \int_{C^\psi(y_i)} I(x) dx. \quad (27)$$

Теперь условие сохранения светового потока (23) может быть записано так:

$$\begin{cases} G_1(\psi) = L_1, \\ \dots \\ G_n(\psi) = L_n. \end{cases} \quad (28)$$

Система (28) является системой нелинейных уравнений относительно значений $\psi = (\psi_1, \dots, \psi_k)$. Перепишем её следующим образом:

$$\begin{cases} \psi_1 = \psi_1 + \varepsilon \cdot (L_1 - G_1(\psi)), \\ \dots \\ \psi_n = \psi_n + \varepsilon \cdot (L_n - G_n(\psi)), \end{cases} \quad (29)$$

где $\varepsilon > 0$ – некоторая константа. Рассматривая метод простой итерации для решения системы (29), мы получим следующий итерационный алгоритм.

$$\psi_i^{n+1} = \psi_i^n + \varepsilon \cdot (L_i - G_i(\psi)), \quad i = \overline{1, n}. \quad (30)$$

Представленный алгоритм имеет понятный физический смысл. Пусть требуемый световой поток в точке y_i равен L_i , а текущий сформированный поток $G_i(\psi)$ для функции ψ меньше L_i , то есть $L_i > G_i(\psi)$. Тогда нужно увеличить размер клетки $C^\psi(y_i)$, а значит, нужно увеличить вес ψ_i . Если $L_i < G_i(\psi)$, то размер клетки $C^\psi(y_i)$ нужно уменьшить за счёт уменьшения веса ψ_i . Алгоритм (30) увеличивает или уменьшает ψ_i на величину, пропорциональную разнице требуемого и формируемого световых потоков.

В следующем параграфе мы покажем, что метод (30) является градиентным методом для нахождения максимума функции

$$h(\psi) = \sum_{i=1}^n \int_{C^\psi(y_i)} [\mathcal{K}(x, y_i) - \psi_i] I(x) dx + \sum_{i=1}^n \psi_i L_i. \quad (31)$$

В силу вогнутости функции $h(\psi)$ она имеет не более одного экстремума, который может быть только её глобальным максимумом. Совпадение метода (30) с градиентным гарантирует его сходимость при достаточно малом значении шага ε , а также даёт возможность адаптивного выбора шага метода.

Функция $h(\psi)$ тесно связана с функционалом Лагранжа для задачи перемещения масс Монжа–Канторовича. В данной статье мы ограничимся лишь вычислением её градиента и доказательством вогнутости. Происхождение данной функции описано в статье [11].

4. Вогнутость и вычисление градиента

В этом параграфе мы покажем, что функция (31) является вогнутой, а также вычислим её частные производные.

Утверждение 6. Функция $h(\psi)$ является вогнутой.

Доказательство. Доказательство полностью аналогично доказательству соответствующего утверждения для квадратичной стоимости [25]. Мы приводим доказательство для полноты изложения.

Для этого сначала покажем вогнутость первого слагаемого, которое обозначим

$$\begin{aligned} R(\psi) &= \sum_{i=1}^n \int_{C^\psi(y_i)} [\mathcal{K}(x, y_i) - \psi_i] I(x) dx = \\ &= \int_G [\mathcal{K}(x, T^\psi(x)) - \psi(T^\psi(x))] I(x) dx. \end{aligned}$$

Для произвольного отображения $T: G \rightarrow D$ и любого $x \in C^\psi(y_i)$ выполнено

$$\mathcal{K}(x, y_i) - \psi_i \leq \mathcal{K}(x, T(x)) - \psi(T(x)),$$

значит, для любого $x \in G$ верно

$$\mathcal{K}(x, T^\psi(x)) - \psi(T^\psi(x)) \leq \mathcal{K}(x, T(x)) - \psi(T(x)).$$

Интегрируя по $I(x) dx$, получим соотношение

$$R(\psi) \leq R_T(\psi),$$

где

$$R_T(\psi) = \int_G [\mathcal{K}(x, T(x)) - \psi(T(x))] I(x) dx.$$

Поскольку для $T = T^\psi$ величины $R(\psi)$ и $R_T(\psi)$ совпадают, мы можем записать

$$R(\psi) = \inf_T R_T(\psi).$$

Функции $R_T(\psi)$ являются линейными по ψ , а значит, вогнутыми. Нетрудно показать, что инфимум вогнутых функций является вогнутой функцией. Соответственно, функция R является вогнутой, следовательно, $h(\psi)$ также является вогнутой, как сумма двух вогнутых функций.

Теорема 7. Частная производная функции $h(\psi)$ по переменной ψ_i имеет вид

$$\frac{\partial h}{\partial \psi_i} = L_i - \int_{C^\psi(y_i)} I(x) dx.$$

Доказательство. Для доказательства достаточно вычислить производную функции $R(\psi)$, введённой ранее в доказательстве утверждения 6. Для удобства введём обозначение:

$$M(x, \psi) = \min_i (\mathcal{K}(x, y_i) - \psi_i).$$

Тогда, пользуясь определением клеток Вороного, можно записать

$$R(\psi) = \sum_{i=1}^n \int_{C^\psi(y_i)} [\mathcal{K}(x, y_i) - \psi_i] I(x) dx = \int_G M(x, \psi) I(x) dx.$$

Пусть δ_i – вектор $(0, \dots, 0, \delta, 0, \dots, 0)$, имеющий ненулевую компоненту только на i -м месте. Рассмотрим разность

$$R(\psi + \delta_i) - R(\psi) = \int_G (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx.$$

Предположим для определённости, что $\delta > 0$ (случай $\delta < 0$ рассматривается аналогично). Тогда клетка Вороного $C^\psi(y_i)$, вес которой мы варьируем, увеличивается, то есть $C^\psi(y_i) \subset C^{\psi+\delta_i}(y_i) \subset G$. Разобьём последний интеграл на три:

$$\begin{aligned} R(\psi + \delta_i) - R(\psi) &= \\ &= \int_{C^\psi(y_i)} (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx + \\ &+ \int_{C^{\psi+\delta_i}(y_i) \setminus C^\psi(y_i)} (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx + \\ &+ \int_{G \setminus C^{\psi+\delta_i}(y_i)} (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx. \end{aligned}$$

Теперь рассмотрим отдельно каждый интеграл, начиная с первого. Так как $C^\psi(y_i) \subset C^{\psi+\delta_i}(y_i)$, то оба минимума достигаются в точке y_i , и поэтому

$$\begin{aligned} \int_{C^\psi(y_i)} (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx &= \\ &= -\delta \cdot \int_{C^\psi(y_i)} I(x) dx. \end{aligned}$$

Рассмотрим второй интеграл. Для любой точки $x \in C^{\psi+\delta_i}(y_i) \setminus C^\psi(y_i)$ взвешенное расстояние $\mathcal{K}(x, y_i) - \psi_i$ до произвольной точки $y_i \in D$ либо не изменилось, либо изменилось на δ . Соответственно, минимум этих расстояний изменился не более, чем на δ . То есть

$$|M(x, \psi + \delta_i) - M(x, \psi)| \leq \delta.$$

В результате для второго интеграла получаем оценку

$$\begin{aligned} \left| \int_{C^{\psi+\delta_i}(y_i) \setminus C^\psi(y_i)} (M(x, \psi + \delta_i) - M(x, \psi)) I(x) dx \right| &\leq \\ &\leq \delta \cdot \int_{C^{\psi+\delta_i}(y_i) \setminus C^\psi(y_i)} I(x) dx. \end{aligned}$$

Заметим, что при $\delta \rightarrow 0$ разность $C^{\psi+\delta_i}(y_i) \setminus C^\psi(y_i)$ стремится к пустому множеству, а значит, второй интеграл имеет порядок малости $o(\delta)$.

Третий интеграл равен нулю, поскольку оба минимума достигаются в других клетках, а там значение функций под знаком минимума одинаково. Переходя к пределу при $\delta \rightarrow 0$, мы получаем утверждение теоремы.

5. Расчётные примеры

Теоретические положения данной статьи мы проиллюстрируем на примере расчёта преломляющего оптического элемента, формирующего заданное распределение интенсивности при точечном источнике. Данный оптический элемент описан в параграфе 1.1. Будут рассмотрены два расчётных примера. Первый из них имеет небольшую расчётную сложность, на нём хорошо видно преимущество критерия сходимости $h(\psi)$ (31) по сравнению со среднеквадратичным отклонением. Во втором примере демонстрируется возможность рассчитывать с помощью данного метода оптические элементы, формирующие сложные распределения освещённости.

5.1. Рефракционная поверхность, формирующая постоянное распределение освещённости на контуре квадрата

В качестве первого модельного примера рассмотрим расчёт преломляющей поверхности W , формирующей при точечном источнике излучения постоянное распределение освещённости на кривой в виде контура квадрата с размером стороны d в удалённой плоскости $z=f$. Преломляющая поверхность разделяет среды с показателями преломления $n_1=1,49$ (полиметилметакрилат) и $n_2=1$. Понятие «удалённая плоскость» предполагает, что размерами преломляющей поверхности по сравнению с расстоянием до плоскости $z=f$ можно пренебречь. Данную задачу можно переформулировать [16] как задачу расчёта преломляющей поверхности из условия формирования заданного распределения интенсивности (см. подпараграф 1.1). Пусть (m_1, m_2) – декартовы координаты на плоскости $z=f$, а требуемое распределение задаётся функцией $E(m)$. Точка на этой плоскости однозначно соответствует единичному вектору

$$s(m) = \frac{(m_1, m_2, f)}{\sqrt{m_1^2 + m_2^2 + f^2}},$$

который мы отождествляем с точкой на сфере. Первые две координаты вектора $s(m)$ соответствуют введённым в подпараграфе 1.1 координатам

$$(y_1, y_2) = \frac{(m_1, m_2)}{\sqrt{m_1^2 + m_2^2 + f^2}}.$$

Функция $L(y(m))$, с помощью которой мы характеризуем формируемое распределение интенсивности, связана с $E(m)$ следующим образом [16]:

$$L(y(m)) = E(m) \frac{f^2 + m_1^2 + m_2^2}{f}$$

В качестве источника излучения будем использовать точечный ламбертовский источник, излучающий в телесный угол, соответствующий конусу с вершиной в начале координат и углом при вершине $2\theta_0$. В этом случае интенсивность ламбертовского источника имеет вид

$$I(x_1, x_2) = \sqrt{1 - x_1^2 - x_2^2}, \quad x = (x_1, x_2) \in G,$$

где

$$G = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq \sin^2 \theta_0\}.$$

Расчёт поверхности производился при следующих параметрах: угол при вершине «конуса излучения» $2\theta_0 = 90^\circ$, расстояние от источника до удалённой плоскости $f = 1000$ мм, размер стороны квадрата $d = f \tan(60^\circ) = 1376$ мм.

В соответствии с описанием метода опорных квадратиков (МОК) в параграфе 3, требуемое распределение освещённости в виде контура квадрата было аппроксимировано дискретным распределением, содержащим $N = 240$ точек $\{y_i\}_{i=1}^{240}$. Точки были расположены равномерно на контуре с угловым шагом в 1° , то есть все требуемые распределения освещённости L_i одинаковы. В этом случае преломляющая поверхность W представляет собой кусочно-гладкую поверхность из фрагментов $N = 240$ квадратиков (эллипсоидов вращения), фокусирующих весь световой поток в точки y_i , задающихся уравнением в полярных координатах

$$\rho(x) = \frac{p_i}{1 - n \langle \vec{e}(x), \vec{y}_i \rangle}.$$

Поверхность W тогда задается в полярных координатах функцией

$$\rho(x) = \max_i \frac{p_i}{1 - n \langle \vec{e}(x), \vec{y}_i \rangle}. \quad (32)$$

После применения \ln с учётом обозначений из параграфа 1 данное соотношение эквивалентно формуле

(25) параграфа 3. Поверхность зависит только от набора фокальных параметров p_i . Расчёт фокальных параметров эллипсоидов проводился итерационно градиентным методом (30)

$$\psi_i^{n+1} = \psi_i^n + \varepsilon \cdot (L_i - G_i(\psi)),$$

где $\psi_i = \ln p_i$, а $G_i(\psi)$ – световой поток, фокусирующийся оптическим элементом в точку y_i . Это соответствует такому преобразованию значений p_i :

$$p_i^{n+1} = p_i^n e^{\varepsilon \cdot (L_i - G_i(\psi))}.$$

При этом на каждой итерации методом трассировки лучей [11] осуществлялся расчёт значений светового потока $G_i(\psi)$ в точках дискретного распределения, и затем выполнялась коррекция значений ψ_i , $i = 1, N$. При расчёте было выполнено 1600 итераций (этого оказалось достаточно для сходимости метода). Общее время расчёта преломляющей поверхности на персональном компьютере (Core i9-7940X CPU) составило около 3 минут.

Расчитанная преломляющая поверхность W приведена на рис. 1а, где чёрной точкой отмечено положение источника излучения. В данном примере расстояние от источника до вершины поверхности составляет 3,1 мм. На рис. 1б приведено формируемое распределение освещённости в плоскости $z = f = 1000$ мм. Данное распределение было рассчитано методом трассировки лучей [11] и с хорошей точностью совпадает с требуемым распределением: нормированное среднеквадратичное отклонение (СКО) сформированного распределения от заданного постоянного распределения на контуре квадрата составляет всего 2,5%. Следует отметить, что полученная рефракционная поверхность является непрерывной и кусочно-гладкой. Изломы (разрывы в производных) хорошо видны на рис. 1а и обусловлены тем, что формируемый контур квадрата не является гладкой кривой. Для проверки правильности выполненных расчётов было проведено моделирование работы расчитанной поверхности в программе для светотехнических расчётов TracePro [26]. Результаты моделирования приведены на рис. 1в и также показывают формирование равномерно освещённого контура.

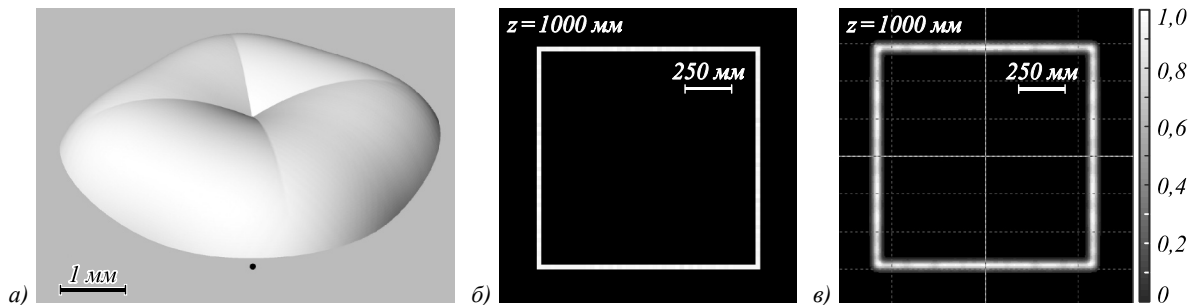


Рис. 1. Рефракционная поверхность, рассчитанная из условия формирования постоянного распределения освещённости на контуре квадрата (а); нормированное распределение освещённости, формируемое в плоскости $z = f = 1000$ мм (б); результат моделирования работы преломляющей поверхности в программе для светотехнических расчетов TracePro (в)

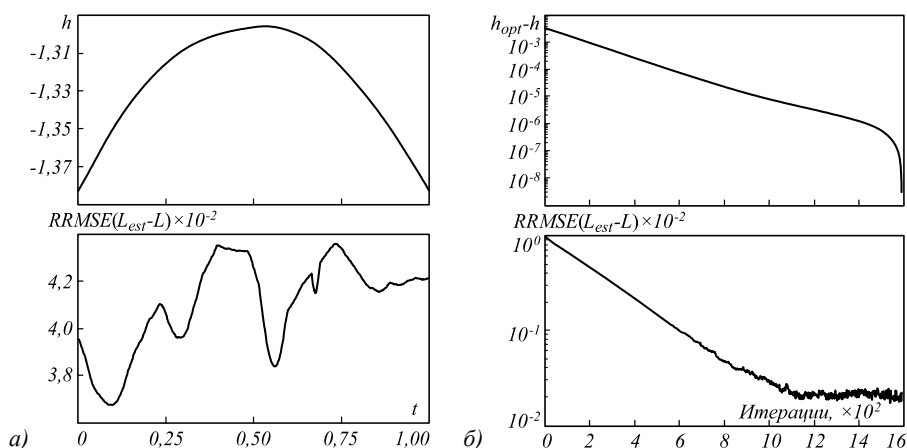


Рис. 2. График функции $h(\psi)$ (сверху) и график нормированного СКО между расчетным и требуемым распределением в виде контура квадрата (снизу) на отрезке $(1-t)\psi_1 + t\psi_2$, $t \in [0, 1]$ (а); графики сходимости МОК при расчете поверхности на рис. 1а: зависимость ошибки $|h(\psi) - h_{opt}|$ (сверху) и СКО-критерий (снизу) от номера итерации (б)

Как показано в параграфе 4, использованный для расчёта преломляющей поверхности МОК (30) соответствует градиентному методу максимизации вогнутой функции $h(\psi)$ (31). Для иллюстрации вогнутости функции $h(\psi)$ мы выбрали два случайных вектора ψ_1 и ψ_2 , представляющих параметры эллипсоидов в уравнении для преломляющей поверхности (35), и построили функцию $h(\psi)$ на отрезке $(1-t)\psi_1 + t\psi_2$, $t \in [0, 1]$ соединяющем эти точки. Полученная функция $h(1-t)\psi_1 + t\psi_2$ показана на верхнем графике на рис. 2а и является вогнутой. Для сравнения на нижнем графике на рис. 2а мы показали график нормированной среднеквадратической ошибки (СКО) между требуемыми и расчётными значениями освещённости на контуре квадрата при параметрах эллипсоидов, лежащих на том же самом отрезке $(1-t)\psi_1 + t\psi_2$. На графике СКО хорошо видно наличие нескольких экстремумов. Это показывает, что СКО-критерий, широко используемый для контроля сходимости в МОК [1–3], является менее подходящим по сравнению с функцией $h(\psi)$. Это также иллюстрируется графиками сходимости МОК на рис. 2б, полученными при расчёте преломляющей поверхности на рис. 1а. На верхнем графике рис. 2б приведена величина $h_{opt} - h(\psi)$, где h_{opt} – значение функции, полученное в конце расчёта (на последней итерации). Из рис. 2б видно монотонное уменьшение величины $h_{opt} - h(\psi)$. Для сравнения в нижней части рис. 2б показан график СКО-критерия. Из сравнения графиков на рис. 2б видно, что СКО, по сравнению с $h_{opt} - h(\psi)$, является более плохим критерием для контроля сходимости итерационного процесса. В частности, можно видеть, что СКО-критерий после примерно тысячи итераций имеет зашумлённый и осциллирующий вид.

5.2. Рефракционная поверхность, формирующая распределение освещённости в виде символа «Триглав»

В качестве более сложного примера, иллюстрирующего возможности предложенного МОК в зада-

чах формирования сложных распределений освещённости, была рассчитана преломляющая поверхность, формирующая постоянное распределение освещённости в области в виде символа «Триглав» (см. врезку к рис. 3в). Как и в примере 1, в качестве источника излучения был использован точечный ламбертовский источник, излучающий в телесный угол, соответствующий конусу с углом при вершине $2\theta_0 = 90^\circ$. При этом угловой размер формируемого распределения (угловой размер квадрата, описанного вокруг символа «Триглав») составляет 100° .

При расчёте требуемое распределение освещённости в виде символа «Триглав» было аппроксимировано дискретным распределением, содержащим $N = 20890$ точек. Точки были расположены в узлах квадратной угловой сетки (шаг сетки $0,26^\circ$), покрывающей область символа. Расчёт преломляющей поверхности (параметров квадрик) проводился градиентным методом (30). Как и в примере 1, на каждой итерации методом трассировки лучей [11] осуществлялся расчёт значений светового потока в точках дискретного распределения и затем выполнялась коррекция значений $\psi_i, i = 1, N$ по формуле (30). При расчёте было выполнено 25000 итераций. Общее время расчёта преломляющей поверхности составило около 3,5 часов.

Рассчитанная преломляющая поверхность приведена на рис. 3а, где чёрной точкой отмечено положение источника излучения (расстояние от источника до вершины поверхности составляет 3,0 мм). На рис. 3б приведено формируемое распределение освещённости в удалённой плоскости $z = f = 1000$ мм. Данное распределение было рассчитано методом трассировки лучей [11] при числе лучей $N_r = 106$. СКО сформированного распределения от заданного постоянного распределения в области символа «Триглав» составляет 3,6%. Можно видеть, что рассчитанная рефракционная поверхность является непрерывной и кусочно-гладкой. Изломы у рассчитанной поверхно-

сти обусловлены «разрывным» видом формируемого распределения, заданного в многосвязной области. Важно отметить, что при распределениях такого вида широко используемые методы расчёта оптических поверхностей [1–7], основанные на численном решении нелинейного дифференциального уравнения эллиптического типа, оказываются неприменимыми

или являются численно неустойчивыми. Для проверки правильности выполненных расчётов, было проведено моделирование работы рассчитанной поверхности в программе TracePro. Результаты моделирования приведены на рис. 3в и показывают формирование примерно постоянного распределения освещённости в области символа.

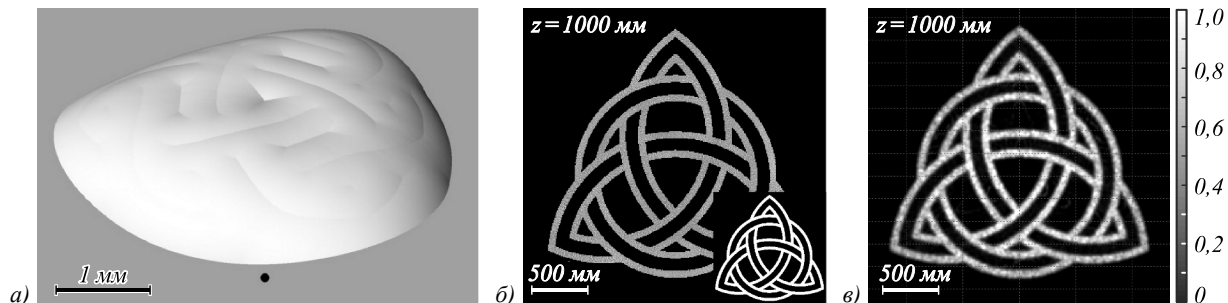


Рис. 3. Рефракционная поверхность, рассчитанная из условия формирования постоянного распределения освещённости в области в виде символа «Триглав» (а); нормированное распределение освещённости, формируемое в плоскости $z = f = 1000$ мм (б); результат моделирования работы преломляющей поверхности в программе для светотехнических расчетов TracePro (в). На врезке показано требуемое распределение

Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию ФНИЦ «Кристаллография и фотоника» РАН в части численной реализации алгоритма расчёта, а также гранта РНФ18-19-00326 в части доказательства совпадения метода опорных квадрик с градиентным методом для соответствующего функционала.

References

- [1] Chang S, Wu R, Li A, Zheng Z. Design beam shapers with double freeform surfaces to form a desired wavefront with prescribed illumination pattern by solving a Monge–Ampère type equation. *J Opt* 2016; 18(12): 125602. DOI: 10.1088/2040-8978/18/12/125602.
- [2] Feng Z, Froese BD, Huang C-Y, Ma D, Liang R. Creating unconventional geometric beams with large depth of field using double freeform-surface optics. *Appl Opt* 2015; 54(20): 6277-6281. DOI: 10.1364/AO.54.006277.
- [3] Feng Z, Froese BD, Liang R, Cheng D, Wang Y. Simplified freeform optics design for complicated laser beam shaping. *Appl Opt* 2017; 56(33): 9308-9314. DOI: 10.1364/AO.56.009308.
- [4] Bösel C, Worku NG, Gross H. Ray-mapping approach in double freeform surface design for collimated beam shaping beyond the paraxial approximation. *Appl Opt* 2017; 56(13): 3679-3688. DOI: 10.1364/AO.56.003679.
- [5] Bösel C, Gross, H. Double freeform illumination design for prescribed wavefronts and irradiances. *J Opt Soc Am A* 2018; 35(2): 236-243. DOI: 10.1364/JOSAA.35.000236.
- [6] Mao X, Li J, Wang F, Gao R, Li X, Xie Y. Fast design method of smooth freeform lens with an arbitrary aperture for collimated beam shaping. *OSA Technical Digest* 2019; JT5A.2. Washington: DC United States; 2019. ISBN: 978-1-943580-60-6.
- [7] Wei S, Zhu Z, Fan Z, Yan Y, Ma D. Double freeform surfaces design for beam shaping with non-planar wavefront using an integrable ray mapping method. *Opt Express* 2019; 27(19): 26757-26771. DOI: 10.1364/OE.27.026757.
- [8] Kochengin SA, Oliker VI. Computational algorithms for constructing reflectors. *Comput Vis Sci* 2003; 6: 15-21. DOI: 10.1007/s00791-003-0103-2.
- [9] Doskolovich LL, Moiseev MA, Kazanskiy NL. On using a supporting quadric method to design diffractive optical elements. *Computer Optics* 2015; 39(3): 339-346. DOI: 10.18287/0134-2452-2015-39-3-339-346.
- [10] Andreeva KV, Moiseev MA, Kravchenko SV, Doskolovich LL. Design of optical elements with TIR freeform surface. *Computer Optics* 2016; 40(4): 467-474. DOI: 10.18287/2412-6179-2016-40-4-467-474.
- [11] Mingazov AA, Bykov DA, Bezus EA, Doskolovich LL. On the use of the supporting quadric method in the problem of designing double freeform surfaces for collimated beam shaping. *Opt Express* 2020; 28(15): 22642-22657. DOI: 10.1364/OE.398990.
- [12] Glimm T, Oliker VI. Optical design of single reflector systems and the Monge-Kantorovich mass transfer problem. *J Math Sci* 2003; 117: 4096-4108. DOI: 10.1023/A:1024856201493.
- [13] Wang X-J. On design of a reflector antenna II. *Calc Variations Partial Differ Equ* 2004; 20: 329-341. DOI: 10.1007/s00526-003-0239-4.
- [14] Doskolovich LL, Bykov DA, Mingazov AA, Bezus EA. Optimal mass transportation and linear assignment problems in the design of freeform refractive optical elements generating far-field irradiance distributions. *Opt Express* 2019; 27(9): 13083-13097. DOI: 10.1364/OE.27.013083.
- [15] Yadav NK. Monge-Ampère problems with non-quadratic cost function: application to freeform optics. Eindhoven: Technische Universiteit Eindhoven; 2018.
- [16] Bykov DA, Doskolovich LL, Mingazov AA, Bezus EA. Optimal mass transportation problem in the design of freeform optical elements generating far-field irradiance distributions for plane incident beam. *Appl Opt* 2019; 58(33): 9131-9140. DOI: 10.1364/AO.58.009131.
- [17] Glimm T, Oliker VI. Optical design of two-reflector systems, the Monge-Kantorovich mass transfer problem and Fermat’s principle. *Indiana Univ Math J* 2004; 53(5): 1255-1277. DOI: 10.1512/iumj.2004.53.2455.
- [18] Rubinstein J, Wolansky G. Intensity control with a freeform lens. *J Opt Soc Am A* 2007; 24(2): 463-469. DOI: 10.1364/JOSAA.24.000463.

- [19] Olikier V, Rubinstein J, Wolansky G. Supporting quadric method in optical design of freeform lenses for illumination control of a collimated light. *Adv Appl Math* 2015; 62: 160-183. DOI: 10.1016/j.aam.2014.09.009.
- [20] Glimm T. A rigorous analysis using optimal transport theory for a two-reflector design problem with a point source. *Inverse Probl* 2010; 26(4): 045001. DOI: 10.1088/0266-5611/26/4/045001.
- [21] Mingazov AA, Doskolovich LL, Bykov DA, Kazanskiy NL. The two reflector design problem for forming a flat wavefront from a point source as an optimal mass transfer problem. *Computer Optics* 2019; 43(6): 968-975. DOI: 10.18287/2412-6179-2019-43-6-968-975.
- [22] Doskolovich LL, Mingazov AA, Bykov DA, Andreev ES, Bezus EA. Variational approach to calculation of light field eikonal function for illuminating a prescribed region. *Opt Express* 2017; 25(22): 26378-26392. DOI: 10.1364/OE.25.026378.
- [23] Bykov DA, Doskolovich LL, Mingazov AA, Bezus EA, Kazanskiy NL. Linear assignment problem in the design of freeform refractive optical elements generating prescribed irradiance distributions. *Opt Express* 2018; 26(21): 27812-27825. DOI: 10.1364/OE.26.027812.
- [24] Makarov B, Podkorytov A. *Real analysis: Measures, integrals and applications*. London: Springer-Verlag; 2013.
- [25] Merigot Q. A multiscale approach to optimal transport. *Comput Graph Forum* 2011; 30(5): 1584-1592. DOI: 10.1111/j.1467-8659.2011.02032.x.
- [26] TracePro – software for design and analysis of illumination and optical systems. Source: <https://www.lambdare.com/tracepro/>.

Сведения об авторах

Мингазов Альберт Айдарович, в 2010 году окончил Самарский государственный университет (СамГУ, ныне – Самарский национальный исследовательский университет имени академика С.П. Королёва) по специальности «Математика». В 2012 году окончил магистратуру НИУ ВШЭ по направлению «Математика», в 2015 окончил аспирантуру Санкт-Петербургского отделения математического института имени В.А. Стеклова. Кандидат физико-математических наук (2015), научный сотрудник лаборатории дифракционной оптики Института систем обработки изображений РАН – филиала ФНИЦ «Кристаллография и фотоника» РАН. E-mail: mingazov88@gmail.com.

Досколович Леонид Леонидович, в 1989 году с отличием окончил Куйбышевский авиационный институт (КуАИ, ныне – Самарский национальный исследовательский университет имени академика С.П. Королёва) по специальности «Прикладная математика». Доктор физико-математических наук (2001 год), профессор, главный научный сотрудник лаборатории дифракционной оптики Института систем обработки изображений РАН (ИС-ОИ РАН) — филиала ФНИЦ «Кристаллография и фотоника» РАН, профессор кафедры технической кибернетики Самарского университета и ведущий научный сотрудник научно-исследовательской лаборатории прорывных технологий дистанционного зондирования Земли Самарского университета. Специалист в области дифракционной оптики, лазерных информационных технологий, нанофотоники. E-mail: leonid@ipsiras.ru.

Быков Дмитрий Александрович, в 2009 году с отличием окончил Самарский государственный аэрокосмический университет имени академика С.П. Королёва (СГАУ, ныне — Самарский национальный исследовательский университет имени академика С.П. Королёва) по специальности «Прикладная математика и информатика». Доктор физико-математических наук (2017 г.), старший научный сотрудник лаборатории дифракционной оптики Института систем обработки изображений РАН (ИСОИ РАН) – филиала ФНИЦ «Кристаллография и фотоника» РАН и доцент Самарского университета. Области научных интересов: оптика резонансных дифракционных структур, магнитооптика, электромагнитная теория дифракции. E-mail: bykovd@gmail.com.

Бызов Егор Владимирович, 1988 года рождения. В 2014 году с отличием окончил обучение в магистратуре Самарского государственного аэрокосмического университета имени академика С.П. Королёва (СГАУ, ныне Самарский национальный исследовательский университет имени академика С.П. Королёва) по направлению «Прикладная математика и физика». Область научных интересов: методы расчетов формирующей неизображающей оптики для светодиодов. E-mail: egor.byзов@gmail.com.

ГРНТИ: 29.31.29

Поступила в редакцию 29 сентября 2021 г. Окончательный вариант – 26 ноября 2021 г.

Support quadric method in non-imaging optics problems that can be reformulated as a mass transfer problem

A.A. Mingazov¹, L.L. Doskolovich^{1,2}, D.A. Bykov^{1,2}, E.V. Byzov¹

¹ IPISI RAS – Branch of the FSRC “Crystallography and Photonics” RAS,
443001, Samara, Russia, Molodogvardeyskaya 151;

² Samara National Research University, , 443086, Samara, Russia, Moskovskoye Shosse 34

Abstract

The article deals with problems of generating desired illumination patterns, formulated in a special way. More precisely, we consider problems that can be reformulated as a Monge–Kantorovich mass transfer problem with some cost function. For all problems of this type, we uniformly formulate the support quadric method and show that it coincides with the gradient method for finding the maximum of a certain concave function.

Keywords: non-imaging optics, geometric optics, inverse problem, Monge–Kantorovich problem, support quadric method.

Citation: Mingazov AA, Doskolovich LL, Bykov DA, Byzov EV. Support quadric method in non-imaging optics problems that can be reformulated as a mass transfer problem. *Computer Optics* 2022; 46(3): 353-365. DOI: 10.18287/2412-6179-CO-1055.

Acknowledgements: The work was funded by the Ministry of Science and Higher Education of the Russian Federation under a government project of the FSRC “Crystallography and Photonics” RAS (numerical implementation of the calculation algorithm) and the Russian Science Foundation under grant #18-19-00326 (proof of the coincidence with the gradient method for the corresponding functional).

Authors' information

Albert Aidarovich Mingazov graduated from Samara State University (presently, Samara National Research University) in 2010, majoring in Mathematics. In 2012 he got the master degree in Mathematics at National Research University Higher School of Economics, in 2015 he graduated from the post-graduate course of the St. Petersburg department of Steklov Mathematical Institute of Russian Academy of Sciences. Candidate of Physics and Mathematics (2015). Currently he is a researcher at Diffractive Optics laboratory of the Image Processing Systems Institute RAS – Branch of the FSRC “Crystallography and Photonics” RAS. E-mail: mingazov88@gmail.com.

Leonid Leonidovich Doskolovich graduated with honours (1989) from S.P. Korolyov Kuibyshev Aviation Institute (presently, Samara National Research University), majoring in Applied Mathematics. He received his Doctor in Physics & Maths (2001) degree from Samara State Aerospace University. Main researcher of the Image Processing Systems Institute RAS – Branch of the FSRC “Crystallography and Photonics” RAS, professor at Technical Cybernetics department of National Research University, senior researcher at the Breakthrough Technologies for Earth’s Remote Sensing laboratory at SSAU. His leading research interests include diffractive optics, laser information technologies, nanophotonics. E-mail: leonid@ipsiras.ru.

Dmitry Alexandrovich Bykov graduated with honors (2009) from Samara State Aerospace University (SSAU), majoring in Applied Mathematics and Computer Science. Doctor of Physics and Mathematics (2017). Senior researcher at the Samara University and at the Diffractive Optics Laboratory of the Image Processing Systems Institute RAS – Branch of the FSRC “Crystallography and Photonics” RAS. His current research interests include optics of resonant diffractive structures, magneto-optics of nanostructured materials, and electromagnetic diffraction theory. E-mail: bykovd@gmail.com.

Egor Vladimirovich Byzov (b. 1988) graduated with honors (2014) from Samara State Aerospace University named after S.P. Korolyov (now – Samara National Research University named after academician S.P. Korolyov), majoring in Applied Mathematics and Physics. Research interests: design methods of nonimaging optics for LEDs. E-mail: egor.byzov@gmail.com.

Received September 29, 2021. The final version – November 26, 2021.
