

Метод согласованных квадрик для коллимированных пучков

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

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 рассматривается задача расчета преломляющего элемента с двумя поверхностями, формирующего плоский фронт и заданное распределение освещенности. Формулируется метод согласованных квадрик для расчета данного оптического элемента, и показывается, что данный метод совпадает с градиентным методом для некоторого функционала, связанного с задачей перемещения масс Монжа–Канторовича. Это дает возможность адаптивного выбора шага в методе согласованных квадрик. В конце статьи приводится расчетный пример.

Ключевые слова: геометрическая оптика, неизображающая оптика, обратная задача, задача Монжа–Канторовича о перемещении масс.

Цитирование: Мингазов, А.А. Метод согласованных квадрик для коллимированных пучков / А.А. Мингазов, Л. Л. Досколович, Д. А. Быков // Компьютерная оптика. – 2021. – Т. 45, № 1. – С. 29-37. – DOI: 10.18287/2412-6179-CO-783.

Citation: Mingazov AA, Doskolovich LL, Bykov DA. Supporting quadric method for collimated beams. Computer Optics 2021; 45(1): 29-37. DOI: 10.18287/2412-6179-CO-783.

Введение

Задача расчета преломляющей или отражающей оптической поверхности из условия формирования заданного распределения освещенности 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.

Одним из подходов к данной задаче является метод согласованных квадрик [1–3]. Он состоит в следующем. Требуемое распределение освещенности приближается распределением, сосредоточенным в конечном числе точек некоторой сетки. Оптический элемент представляется в виде кусочно-гладкой поверхности, состоящей из фрагментов оптических элементов, каждый из которых фокусирует распределение освещенности в одну точку. В зависимости от задачи эти поверхности могут быть параболоидами вращения, эллипсоидами, гиперболоидами или более сложными поверхностями (например, [3]). Далее параметры фрагментов подбираются итерационно в зависимости от того, насколько отличается количество энергии, приходящее в каждую точку, от требуемого количества.

Мы описываем метод согласованных квадрик для задачи расчета оптического элемента из двух преломляющих поверхностей, формирующих световой пучок с плоским волновым фронтом и заданным распределением освещенности. Далее мы пользуемся

формулировкой данной задачи как задачи перемещения масс Монжа–Канторовича [4], чтобы описать метод согласованных квадрик как градиентный спуск функционала, связанного с функционалом Лагранжа для задачи перемещения масс.

В статьях [1–3] в качестве критерия качества используется среднеквадратичное отклонение, которое в процессе работы алгоритма меняется нетривиальным образом. Результат данной статьи позволяет использовать другой критерий качества, который в процессе работы алгоритма согласованных квадрик изменяется монотонно (до тех пор пока значение шага метода не станет слишком большим). Это позволяет адаптивно выбирать шаг, что может ускорять сходимость алгоритма. Хотя в статье рассматривается задача расчета преломляющего элемента, формирующего изображение с плоским фронтом, результаты могут быть легко распространены на любую задачу неизображающей оптики, которая может быть переформулирована как задача перемещения масс Монжа–Канторовича.

1. Постановка задачи

Рассмотрим трехмерное пространство E^3 с координатами (x_1, x_2, z) . Пусть в области G плоскости $z=0$ задана функция $I(x)$, определяющая распределение интенсивности коллимированного светового пучка, распространяющегося в направлении $k=(0, 0, 1)$. Геометрия задачи приведена на рис. 1.

Зададим оптический элемент с двумя поверхностями свободной формы. Пусть первая преломляющая поверхность R_1 задается функцией $u_1(x)$, $x=(x_1, x_2) \in G$. Тогда

$$R_1 = \{(x, u_1(x)) | x \in G\}.$$

Зафиксируем некоторую плоскость $z=l$, декартовы координаты на этой плоскости будем обозначать (y_1, y_2) . Вторая преломляющая поверхность R_2 будет задаваться с помощью функции $u_2(y)$ следующим образом

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

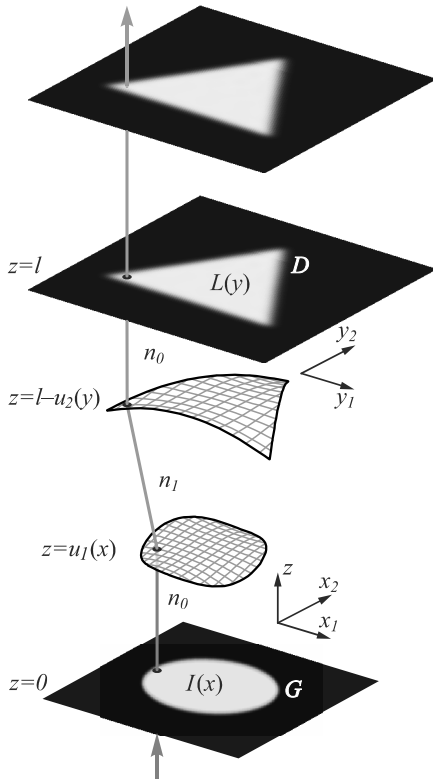


Рис. 1. Геометрия задачи

Мы предполагаем, что показатель преломления среды при $z \leq u_1(x)$ равен n_0 , при $u_1(x) \leq z \leq l - u_2(x)$ показатель преломления равен n_1 , при $l - u_2(x) \leq z$ показатель преломления снова равен n_0 .

Луч, выходящий из точки $x \in G$, преломляясь на двух поверхностях R_1, R_2 , попадает в некоторую точку y области D в плоскости $z=l$. Таким образом, определяется лучевое отображение $\gamma: G \rightarrow D$. Отображение γ формирует в плоскости $z=l$ распределение освещенности, задаваемое функцией $L(x)$, которая определяется условием сохранения светового потока:

$$\int_{\gamma^{-1}(B)} I(x) dx = \int_B L(y) dy,$$

где $B \subset D$ – произвольное борелевское подмножество.

Задача, которую мы рассматриваем, в классической постановке формулируется следующим образом. Пусть задано распределение интенсивности источника $I(x)$, $x \in G$, и требуемое распределение освещенности $L(x)$ в области D плоскости $z=l$. Требуется найти такие дифференцируемые функции $(u_1(x), u_2(y))$, чтобы выполнялись два условия:

1) Сформированное распределение освещенности имеет плоский фронт, параллельный плоскости $z=l$.

2) Отображение $\gamma: G \rightarrow D$ взаимно однозначно и удовлетворяет условию сохранения светового потока:

$$\int_{\gamma^{-1}(B)} I(x) dx = \int_B L(y) dy, \tag{1}$$

где $B \subset D$ – произвольное борелевское подмножество.

Лемма 1. Условие сохранения энергии для отображения $\gamma: G \rightarrow D$ равносильно любому из условий:

1) $I(x) = L(\gamma(x)) \cdot J_\gamma(x)$, а $J_\gamma(x)$ – якобиан отображения γ , вычисленный в точке x ;

2) для любой непрерывной на D функции h выполнено

$$\int_D h(y) L(y) dy = \int_G h(\gamma(x)) I(x) dx.$$

Proof. Доказано в [5, lemma 4.11] или [6, VI.1.1]. □

В силу леммы 1.1 классическое решение данной проблемы – это решение дифференциального уравнения Монжа–Ампера для функции отображения γ специального вида. Мы сформулируем понятие слабого решения для данной задачи, не предполагающее гладкости рефракторов.

2. Функция стоимости

Пусть луч из точки x плоскости $z=0$, преломляясь на двух поверхностях R_1, R_2 , приходит в точку y плоскости $z=l$. Длина оптического пути F между этими точками не зависит от x и y . Число F является параметром оптического элемента, который требуется выбрать. Он неявно связан с толщиной оптического элемента (чем больше F , тем толще оптический элемент). При слишком маленьком параметре F , например, при $F \leq l$ решения рассматриваемой задачи не существует. Итак, вычисляя длину оптического пути, получим соотношение

$$n_0 u_1(x) + n_1 \sqrt{\|y - x\|^2 + (l - u_1(x) - u_2(y))^2} + n_0 u_2(y) = F. \tag{2}$$

Обозначим

$$K(x, y) = u_1(x) + u_2(y).$$

Тогда соотношение (2) является уравнением на $K(x, y)$:

$$n_0 K(x, y) + n_1 \sqrt{\|y - x\|^2 + (l - K(x, y))^2} = F. \tag{3}$$

Решая уравнение (2) относительно K , получим

$$K(x, y) = \frac{n_1^2 l - n_0 F}{n_1^2 - n_0^2} \pm \frac{n_1}{n_1^2 - n_0^2} \times \sqrt{(F - n_0 l)^2 - (n_1^2 - n_0^2) \|y - x\|^2}.$$

В [7] показано, что оптическому элементу соответствует только решение

$$K(x, y) = \frac{n_1^2 l - n_0 F}{n_1^2 - n_0^2} - \frac{n_1}{|n_1^2 - n_0^2|} \times \sqrt{(F - n_0 l)^2 - (n_1^2 - n_0^2) \|y - x\|^2}.$$

Итак, имеет место следующая теорема.

Теорема 1. Если система из преломляющих поверхностей $(u_1(x), u_2(y))$ отображает луч, выходящий из точки x в точку y , то выполнено соотношение

$$u_1(x) + u_2(y) = K(x, y), \text{ где}$$

$$K(x, y) = \frac{n_1^2 l - n_0 F}{n_1^2 - n_0^2} - \frac{n_1}{|n_1^2 - n_0^2|} \times \sqrt{(F - n_0 l)^2 - (n_1^2 - n_0^2) \|y - x\|^2}.$$

□

Нам понадобится представление дифференцируемых рефракторов R_1, R_2 в виде огибающих семейств поверхностей специального вида. Пусть лучевое соответствие $\gamma: G \rightarrow D$ переводит точку $x \in G$ в точку $y \in D$. Для $x \in G$ луч, приходящий в точку $(x, u_1(x))$ преломляющей поверхности R_1 , далее приходит в точку $(y, l - u_2(y))$ второй поверхности. Тогда поверхность R_1 может быть представлена в виде огибающей семейства поверхностей R^y , фокусирующих лучи в точки $(y, l - u_2(y))$ второй поверхности R_2 .

Уравнение рефрактора R^y , фокусирующего в точку $(y, l - u_2(y))$, нетрудно получить, записав условие равенства оптического пути из точки плоскости $z = 0$ до точки $(y, l - u_2(y))$ величине $L - u_2(y)$. Заметим, что в теореме 1 проведено вычисление именно этого оптического пути. В результате функция $u_1^y(x)$, задающая рефрактор R^y , удовлетворяет уравнению:

$$u_1^y(x) + u_2(y) = K(x, y), \text{ то есть}$$

$$u_1^y(x) = K(x, y) - u_2(y).$$

Итак, функция $u_1(x)$ удовлетворяет уравнениям огибающей семейства поверхностей $u_1^y(x)$, то есть выполнено:

$$\begin{cases} u_1(x) = K(x, y) - u_2(y), \\ \frac{\partial}{\partial y_i} [K(x, y) - u_2(y)] = 0, i = \overline{1, 2}. \end{cases} \quad (4)$$

Второй рефрактор R_2 аналогично можно рассматривать как огибающую семейства поверхностей, фокусирующих при обратном распространении все лучи в одной точке $(x, u_1(x))$.

$$\begin{cases} u_2(y) = K(x, y) - u_1(x), \\ \frac{\partial}{\partial x_i} [K(x, y) - u_1(x)] = 0, i = \overline{1, 2}. \end{cases}$$

Мы будем рассматривать частный случай огибающей (4), когда второе условие соответствует не просто некоторой экстремальной точке, а глобальному минимуму. В этом случае соотношения примут вид

$$u_1(x) = \min_{y \in D} (K(x, y) - u_2(y)). \quad (5)$$

Лемма 1. Пусть для системы рефракторов (u_1, u_2) , индуцирующих взаимно однозначное лучевое соответствие, имеет место представление (5). Тогда для функции u_2 выполнено

$$u_2(y) = \min_{x \in G} (K(x, y) - u_1(x)).$$

Proof. Из (5) сразу следует, что

$$u_1(x) \leq K(x, y) - u_2(y) \quad \forall x \in G, y \in D.$$

Значит, выполнено

$$u_2(y) \leq K(x, y) - u_1(x) \quad \forall x \in G, y \in D.$$

Отсюда, переходя к минимуму по x , получаем неравенство

$$u_2(y) \leq \min_{x \in G} (K(x, y) - u_1(x)).$$

Предположим теперь, что существует $y_0 \in D$, для которого

$$u_2(y_0) < \min_{x \in G} (K(x, y_0) - u_1(x)).$$

Значит,

$$u_2(y_0) < K(x, y_0) - u_1(x) \quad \forall x \in G,$$

или

$$u_1(x) + u_2(y_0) < K(x, y_0) \quad \forall x \in G.$$

Последнее равенство означает, что ни для какой точки $x \in G$ оптический путь до y_0 не совпадает с $K(x, y_0)$, то есть никакая точка не отображается в y_0 , что противоречит биективности γ . □

3. Слабое решение и метод согласованных квадрик

Мы сформулируем понятие слабого решения, не требующее дифференцируемости функций $(u_1(x), u_2(y))$. Это понятие, в частности, определяет, что мы будем понимать под решением в случае дискретного распределения освещенности, задаваемого не непрерывной функцией L , а набором точек $\{y \in D\}$ с заданным в каждой из них световым потоком L_y . В этом случае функция u_2 является дискретным набором значений, а u_1 является кусочно-гладкой.

Определение 1. K -выпуклой парой $(u_1(x), u_2(y))$ будем называть пару непрерывных функций $u_1(x)$ на $G, u_2(y)$ на D таких, что

$$u_1(x) = \min_{y \in D} (K(x, y) - u_2(y));$$

$$u_2(y) = \min_{x \in G} (K(x, y) - u_1(x)).$$

Лемма 1. Пусть $(u_1(x), u_2(y))$ – K -выпуклая пара. Тогда функции u_1, u_2 липшицевы.

Proof. Аналогично доказательству [8, lemma 3.2]. □

По теореме Радемахера [6, XI.4.2] липшицева функция является почти всюду дифференцируемой. Поэтому, если $(u_1(x), u_2(y))$ – K -выпуклая пара, то $u_1(x)$ и $u_2(y)$ почти всюду дифференцируемы. Значит, они определяют почти всюду дифференцируемые преломляющие поверхности. Поэтому определено (почти всюду) лучевое соответствие $\gamma: G \rightarrow D$, как отображение преломления на двух поверхностях R_1, R_2 (см. раздел 1).

Лемма 2. Пусть $(u_1(x), u_2(y))$ – K -выпуклая пара, $\gamma: G \rightarrow D$ — лучевое соответствие, определенное почти всюду. Тогда минимум в формуле

$$u_1(x) = \min_{y \in D} (K(x, y) - u_2(y))$$

достигается в точке $y = \gamma(x)$ (там, где отображение γ определено). Аналогично для функции $u_2(y)$.

Proof. Пусть $x_0 \in G$ – точка, в которой лучевое соответствие γ определено, y_0 – точка, в которой достигается минимум. Тогда в точке x_0 функция $u_1(x)$ касается функции $K(x, y_0) - u_2(y_0)$, которая соответствует поверхности, фокусирующей весь световой пучок в точке $(y_0, l - u_2(y_0))$. А значит, $\gamma(x_0) = y_0$. □

Поскольку, как уже было отмечено, функции $u_1(x), u_2(y)$ почти всюду дифференцируемы, многозначное отображение

$$\gamma(x) = \{y \in D | u_1(x) + u_2(y) = K(x, y)\}$$

почти всюду однозначно. Следовательно, можно говорить о том, удовлетворяет ли оно условию сохранения энергии или нет. Многозначное отображение γ для K -выпуклой пары можно записать также следующим образом

$$\gamma(x) = \arg \min_{y \in D} (K(x, y) - u_2(y)). \tag{6}$$

Определение 2. Слабым решением называется K -выпуклая пара функций $(u_1(x), u_2(y)) \in C(G) \times C(D)$, для которой обобщенное отображение преломления

$$\gamma(x) = \arg \min_{y \in D} (K(x, y) - u_2(y))$$

удовлетворяет условию сохранения энергии: для любого борелевского $B \subset D$ выполнено

$$\int_B L(y) dy = \int_{\gamma^{-1}(B)} I(x) dx. \tag{7}$$

Замечание 1. Описанное понятие слабого решения 2 имеет смысл в случае, когда требуемое распределение освещенности $L(y)$ является дискретным, то есть имеет вид

$$\sum_{y \in D} L_y \delta_y,$$

где δ_y – дельта-функция, сосредоточенная в точке y . Но в этом случае функция u_2 задана на дискретном множестве D , а поверхность R_2 представляет собой дискретный набор точек. «Преломление» на системе понимается в данном случае в смысле обобщенного отображения преломления (6).

Мы опишем геометрически метод расчета преломляющих поверхностей R_1, R_2 , являющийся аналогом метода согласованных квадрик, описанного в статьях [1–3] для других задач неизображающей оптики. Он позволяет приблизиться к слабому решению в случае, когда формируемое распределение освещенности дискретно. В дальнейшем мы докажем, что данный метод является градиентным, а потому сходится при достаточно малом шаге.

Пусть формируемое распределение освещенности представлено дискретным распределением

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

$y_i \in D$. Тогда функцию $u_2(y)$, как уже было отмечено в замечании 1, заменяет набор из k чисел u_{2i} , где $i = \overline{1, k}$. Их мы будем варьировать в процессе работы алгоритма. Рефрактор R_1 представляет собой кусочно-гладкую поверхность, состоящую из фрагментов, фокусирующих исходящий световой пучок в точки $(y_i, l - u_{2i})$. Более точно, из определения слабого решения функция $u_1(x)$, задающая поверхность R_1 представляема в виде

$$u_1(x) = \min_i (K(x, y_i) - u_{2i}). \tag{8}$$

Теперь для фиксированных значений $u_{2i}, i = \overline{1, k}$, можно вычислить дискретное распределение, формируемое парой (u_1, u_2) . В точку u_{2i} фокусируется световой поток с области $C(y_i) \subset G$, которая определяется следующим условием

$$C(y_i) = \{x \in G | K(x, y_i) - u_{2i} \leq K(x, y_j) - u_{2j}, \forall j = \overline{1, k}\}.$$

Соответственно, световой поток, фокусирующийся в точку u_{2i} , можно вычислить как интеграл

$$\int_{C(y_i)} I(x) dx.$$

Нетрудно понять, что при увеличении u_{2i} (при условии сохранения остальных u_{2j} для $j \neq i$) область $C(y_i)$ увеличивается, соответственно, увеличивается световой поток.

Нам известно, что требуемый световой поток в точке u_{2i} равен L_i . Соответственно, если

$$L_i > \int_{C(y_i)} I(x) dx,$$

требуется увеличить u_{2i} и уменьшить в противном случае. Метод согласованных квадрик предполагает

вычисление следующего приближения u_2^{n+1} для второй поверхности по предыдущему u_2^n таким образом

$$u_{2i}^{n+1} = u_{2i}^n + \eta \cdot (L_i - \int_{C(y_i)} I(x) dx), \quad i = \overline{1, k}, \quad (9)$$

где η — шаг алгоритма, выбираемый отдельно.

В дальнейшем мы покажем, что метод согласованных квадрик является градиентным спуском некоторого функционала, связанного с задачей перемещения масс.

4. Задача перемещения масс

Задача о нахождении слабого решения оказывается эквивалентна задаче перемещения масс Монжа–Канторовича. Данная теорема аналогична классическим результатам [5, 8, 9], также схожий результат для функции эйконала используется для расчета оптического элемента, формирующего изображение в ближней зоне. Переформулировка рассматриваемой задачи в виде задачи перемещения масс была получена в статье [4].

Точная формулировка приведена в следующей теореме. Мы приводим ее, поскольку доказательство является отправной точкой для дальнейшего рассуждения.

Теорема 1.

Рассмотрим функционал

$$F(T) = \int_G K(x, T(x)) I(x) dx, \quad (10)$$

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

$$I(x) = L(T(x)) \cdot J_T(x). \quad (11)$$

Пусть отображение γ является точкой минимума данного функционала. Тогда данное отображение является лучевым соответствием для пары рефракторов R_1, R_2 .

Proof. Рассмотрим функционал Лагранжа для рассматриваемой вариационной задачи. Возьмем непрерывную на D функцию μ , и пусть

$$G(T, \mu) = \int_G [K(x, T(x)) I(x) + \mu(x) (L(T(x)) \times J_T(x) - I(x))] dx. \quad (12)$$

Так как μ — произвольная функция, в силу биективности T , вместо нее можно записать $\lambda(T(x))$, где λ — произвольная непрерывная функция на G . Функционал G мы переобозначим через

$$H(T, \lambda) = \int_G [K(x, T(x)) I(x) + \lambda(T(x)) (L(T(x)) \times J_T(x) - I(x))] dx.$$

Разбивая на два интеграла и делая замену, получим

$$H(T, \lambda) = \int_G [K(x, T(x)) - \lambda(T(x))] I(x) dx + \int_D L(y) \lambda(y) dy. \quad (13)$$

Минимум функционала (10) при ограничениях (11) соответствует некоторой критической точке функционала (12). Вариация функционала G легко вычисляется. Условие $\delta_\lambda H = 0$, как нетрудно понять из формулы (13), равносильно условию сохранения светового потока. Условие равенства нулю вариации по T $\delta_T H = 0$, равносильно

$$\frac{\partial}{\partial y_i} [K(x, y) - \lambda(y)]|_{y=T(x)} = 0, \quad i = \overline{1, 2}. \quad (14)$$

Как нетрудно увидеть, это в точности условие огибающей (4), то есть отображение T задается системой рефракторов, и функция λ определяет поверхность второго рефрактора. \square

5. Модификация функционала Лагранжа и градиентный метод

В доказательстве теоремы 1 мы видели, что если (T, λ) является критической точкой функционала Лагранжа $H(T, \lambda)$, то отображение $T: G \rightarrow D$ удовлетворяет уравнению Монжа–Ампера и задается системой рефракторов (u_1, u_2) , при этом λ совпадает с u_2 . То есть в критической точке отображение $T: G \rightarrow D$ определяется функцией λ условием (14), а именно

$$\begin{cases} T^\lambda(x_0) = y_0; \\ \frac{\partial}{\partial y_i} [K(x_0, y) - \lambda(y)]|_{y=y_0} = 0, \quad i = \overline{1, 2}. \end{cases} \quad (15)$$

Аналогично тому, как мы ограничивались рассмотрением огибающих специального вида (5), мы будем рассматривать критические точки специального вида, а именно: те, для которых второе соотношение не просто определяет критическую точку, а точку глобального минимума. В этом случае отображение T можно записать аналогично тому, как описывалось лучевое соответствие ранее (6):

$$T^\lambda(x) = \arg \min_{y \in D} (K(x, y) - \lambda(y)).$$

В результате мы можем ограничиться рассмотрением функционала

$$H(\lambda) = \int_G [K(x, T^\lambda(x)) - \lambda(T^\lambda(x))] I(x) dx + \int_D L(y) \lambda(y) dy. \quad (16)$$

Рассмотрим теперь дискретизацию этого функционала. До этого мы рассматривали только биективные отображения из G в D . В описанной постановке задачи это являлось разумным, поскольку для распределений освещенности, задаваемых обычными функциями, искомое лучевое соответствие является именно таким. Но для вычислений мы будем аппроксимировать непрерывное распределение интенсивности $L(y)$ дискретным распределением

$$\sum_{y \in D} L_y \delta_y.$$

В этом случае функционал (16) сохраняет смысл, D – конечное множество точек. Тогда функция λ определяется набором чисел $\lambda_y, y \in D$. Отображение T^λ в этом случае является отображением области G в дискретное множество точек.

Определение 1. Для фиксированного набора чисел $\lambda_y, y \in D$, обобщенными клетками Вороного будем называть области $C^\lambda(y) \subset G, y \in D$, определяемые следующим условием. Точка $x \in G$ лежит в клетке $C^\lambda(y_0)$ тогда и только тогда, когда

$$K(x, y_0) - \lambda_{y_0} \leq K(x, y) - \lambda_y \quad \forall y \in D.$$

Замечание 1. Клетка $C^\lambda(y)$ совпадает с прообразом $(T^\lambda)^{-1}(y)$. Соответственно, разбиение на обобщенные клетки Вороного – это разбиение G на прообразы точек $y \in D$ при отображении T^λ . Поэтому разбиение области G на клетки Вороного для разных весов можно воспринимать как комбинаторно-геометрический способ перебирать отображения вида T^λ .

Теперь функционал $H(\lambda)$ становится функцией на R^k :

$$h(\lambda) = \sum_{y \in D} \int_{C^\lambda(y)} [K(x, y) - \lambda_y] I(x) dx + \sum_{y \in D} \lambda_y L_y, \quad (17)$$

где λ – набор весов обобщенных клеток Вороного.

Аналогичный функционал для случая, когда функция стоимости K является квадратичной, используется для расчета оптического элемента в параксиальном приближении в статье [10], связь этого функционала с задачей перемещения масс с квадратичной функцией стоимости была освещена в статьях [11, 12].

Предложение 1. Функция $h(\lambda)$ является вогнутой.

Proof. Доказательство аналогично доказательству соответствующего утверждения для квадратичной стоимости [12]. Достаточно показать вогнутость первого слагаемого, которое обозначим

$$\begin{aligned} \Psi(\lambda) &= \sum_{y \in D} \int_{C^\lambda(y)} [K(x, y) - \lambda_y] I(x) dx = \\ &= \int_G [K(x, T^\lambda(x)) - \lambda(T^\lambda(x))] I(x) dx. \end{aligned}$$

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

$$K(x, y) - \lambda(y) \leq K(x, T(x)) - \lambda(T(x)),$$

значит, для любого $x \in G$ верно

$$K(x, T^\lambda(x)) - \lambda(T^\lambda(x)) \leq K(x, T(x)) - \lambda(T(x)).$$

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

$$\Psi(\lambda) \leq \Psi_T(\lambda),$$

где

$$\Psi_T(\lambda) = \int_G [K(x, T(x)) - \lambda(T(x))] I(x) dx.$$

Поскольку для $T = T^\lambda$ величины $\Psi(\lambda)$ и $\Psi_T(\lambda)$ совпадают, мы можем записать

$$\Psi(\lambda) = \inf_T \Psi_T(\lambda).$$

Функции $\Psi_T(\lambda)$ являются линейными по λ , а значит, вогнутыми. Нетрудно показать, что инфимум вогнутых функций является вогнутой функцией. Соответственно, функция Ψ является вогнутой, следовательно, h является вогнутой, как сумма двух вогнутых. \square

В силу вогнутости функции h она имеет не более одного экстремума, который может быть только ее глобальным максимумом. Вопрос существования этого максимума остается при этом нетривиальным, мы не будем его касаться.

Далее мы собираемся использовать градиентный метод для поиска максимума функции h . В следующей теореме вычисляется градиент функции $h(\lambda)$.

Теорема 1. Частная производная функции $h(\lambda)$ по переменной λ_y имеет вид

$$\frac{\partial h}{\partial \lambda_y} = L_y - \int_{C^\lambda(y)} I(x) dx.$$

Proof. Для доказательства достаточно вычислить производную функции Ψ , введенной ранее в доказательстве утверждения (1):

$$\Psi(\lambda) = \sum_{y \in D} \int_{C^\lambda(y)} [K(x, y) - \lambda_y] I(x) dx.$$

Для удобства введем обозначение

$$M(x, \lambda) = \min_{y \in D} (K(x, y) - \lambda_y).$$

Тогда, пользуясь определением клеток Вороного, можно записать

$$\begin{aligned} \Psi(\lambda) &= \sum_{y \in D} \int_{C^\lambda(y)} [K(x, y) - \lambda_y] I(x) dx = \\ &= \int_G M(x, \lambda) I(x) dx. \end{aligned}$$

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

$$\begin{aligned} \Psi(\lambda + \delta_y) - \Psi(\lambda) &= \\ &= \int_G (M(x, \lambda + \delta_y) - M(x, \lambda)) I(x) dx. \end{aligned}$$

Предположим для определенности, что $\delta < 0$ (для $\delta > 0$ можно провести аналогичное рассуждение, но области разбиения будут несколько другие). Тогда клетка Вороного $C^\lambda(y)$, вес которой мы варьируем, уменьшается. Значит, имеется цепочка из двух вложений

$$C^{\lambda+\delta_y}(y) \subset C^\lambda(y) \subset G.$$

Разобьем последний интеграл на три, продолжив равенство:

$$\begin{aligned} & \Psi(\lambda + \delta_y) - \Psi(\lambda) = \\ & = \int_{C^{\lambda+\delta_y}(y)} (M(x, \lambda + \delta_y) - M(x, \lambda))I(x)dx + \\ & + \int_{C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)} (M(x, \lambda + \delta_y) - M(x, \lambda))I(x)dx + \\ & + \int_{G \setminus C^\lambda(y)} (M(x, \lambda + \delta_y) - M(x, \lambda))I(x)dx. \end{aligned}$$

Рассмотрим отдельно каждый интеграл, начиная с первого. Так как $C^{\lambda+\delta_y}(y) \subset C^\lambda(y)$, оба минимума достигаются в точке y :

$$\begin{aligned} M(x, \lambda + \delta_y) &= K(x, y) - (\lambda_y + \delta); \\ M(x, \lambda) &= K(x, y) - \lambda_y. \end{aligned}$$

Тогда мы получаем

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

Для второго интеграла мы уже не можем сказать, в какой точке достигаются оба минимума. Тем не менее, для любой точки $x \in C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)$ взвешенное расстояние $K(x, \tilde{y}) - \lambda_{\tilde{y}}$ до произвольной точки $\tilde{y} \in D$ либо не изменилось, либо изменилось на $|\delta|$. Соответственно, минимум этих расстояний изменился не более, чем на $|\delta|$. То есть с учетом знаков можно записать

$$0 \leq M(x, \lambda + \delta_y) - M(x, \lambda) \leq -\delta.$$

В результате для второго интеграла получаем оценку

$$\begin{aligned} 0 \leq & \int_{C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)} (M(x, \lambda + \delta_y) - \\ & - M(x, \lambda))I(x)dx \leq -\delta \cdot \int_{C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)} I(x)dx. \end{aligned}$$

Заметим, что при $\delta \rightarrow 0$ разность $C^\lambda(y) / C^{\lambda+\delta_y}(y)$ стремится к пустому множеству, а значит, и интеграл

$$\int_{C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)} I(x)dx$$

также стремится к нулю. То есть мы установили, что

$$\lim_{\delta \rightarrow 0} \left(\frac{1}{\delta} \times \int_{C^\lambda(y) \setminus C^{\lambda+\delta_y}(y)} (M(x, \lambda + \delta_y) - M(x, \lambda))I(x)dx \right) = 0.$$

Третий интеграл равен нулю, поскольку оба минимума достигаются в других клетках, а там значение функций под знаком минимума одинаково.

Теперь легко вычислим производную:

$$\frac{\partial \Psi}{\partial \lambda_y} = \lim_{\delta \rightarrow 0} \frac{\Psi(\lambda + \delta_y) - \Psi(\lambda)}{\delta} = - \int_{C^\lambda(y)} I(x)dx.$$

□

Зная градиент функции $h(\lambda)$, можем искать максимум, пользуясь градиентным методом. Более точно, следующее приближение λ^{n+1} вычисляется по предыдущему λ^n следующим образом:

$$\lambda_y^{n+1} = \lambda_y^n + \eta \cdot (L_y - \int_{C^{\lambda^n}(y)} I(x)dx), \tag{18}$$

где η — шаг градиентного метода. В качестве критерия качества имеет смысл рассматривать функцию $h(\lambda)$. К сожалению, предсказать значение $h(\lambda)$ в точке максимума невозможно, тем не менее, разумно уменьшать шаг градиентного метода, когда значение $h(\lambda)$ перестает увеличиваться при операции (18).

Поскольку функция λ , как мы видели в доказательстве теоремы 1, является поверхностью второго рефрактора, алгоритм градиентного спуска в точности совпадает с методом согласованных квадратик, описанным ранее (9).

6. Расчетный пример

Рассмотрим расчет оптического элемента, преобразующего пучок с круглым сечением и равномерным распределением

$$I(x) = \frac{1}{\pi R^2}, \quad x \in G = \{(x_1, x_2) | x_1^2 + x_2^2 \leq R^2\} \tag{19}$$

в плоский пучок с распределением в виде квадрата. Более точно, формируемое распределение определяется формулой

$$L(y) = \frac{1}{w^2}, \quad y \in D, \tag{20}$$

где $D = \{(x_1, x_2) | \max\{|x_1|, |x_2|\} \leq w\}$, $i=1,2$ — два квадрат со стороной w . Расчет проводился по следующим параметрам: расстояние $l=10$ мм; оптическая длина пути $F=11$ мм; радиус падающего луча $R=1$ мм; размер формируемого квадрата $w=1$ мм; индексы преломления материала оптического элемента и окружающей среды $n_1=1,5$ и $n_0=1$.

Для вычисления оптического элемента требуемое распределение освещенности $L(y)$ было аппроксимировано дискретным распределением, определенным на квадратной сетке из $N=100^2$ точек. Для вычисления второй поверхности (то есть значений $u_{2,i}$, $i=1, \dots, N$) используется метод согласованных квадратик (формула 9), описанный в параграфе 5. После вычисления значений $u_{2,i}$, $i=1, \dots, N$ вторая поверхность аппроксимируется, используя метод наименьших квадратов, двумерным B -сплайном. После вычисления второй поверхности первая поверхность восстанавливается с помощью соотношения (8). Для расчета приведенного примера проделано 20000

итераций (9). Время работы на персональном компьютере (Intel Core i9-7940X CPU) 4,5 минуты, вос-

становление поверхностей рефрактора заняло около 3 минут.

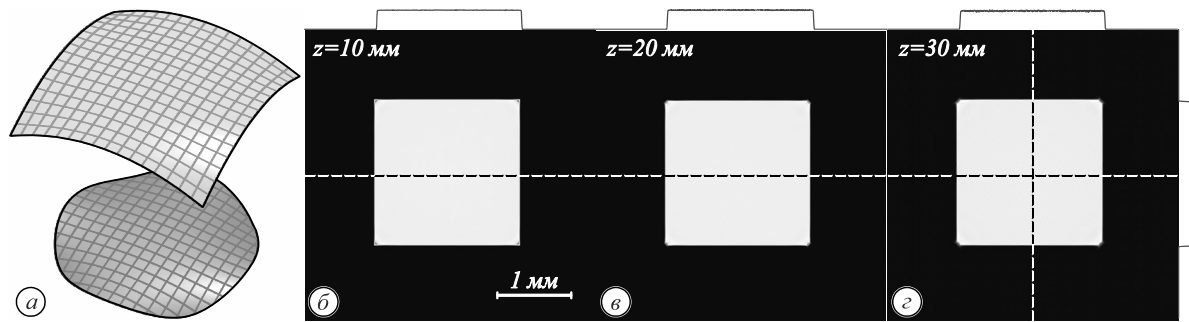


Рис. 2. Формирование плоского пучка с распределением освещенности в виде квадрата

Полученные поверхности показаны на рис. 1а. На рис. 1б–д приведены нормализованные распределения освещенности, сформированные оптическим элементом в плоскостях $z=l=10$ мм, $z=20$ мм и $z=30$ мм соответственно.

Представленные результаты моделирования демонстрируют формирование прямоугольной рамки с необходимыми размерами и практически равномерной освещенностью. Нормализованные среднеквадратичные отклонения полученных распределений освещенности от равномерного составляют 7%, 10% и 15% при $z=10$ мм, $z=20$ мм и $z=30$ мм соответственно. Отметим, что размеры квадрата на разных расстояниях от оптического элемента остаются неизменными, поэтому волновой фронт генерируемого выходного луча является плоским.

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 расчета, а также грантов РФФИ (№ 18-29-03067, 18-07-00982) в части формулировки метода согласованных квадрик и доказательства совпадения с градиентным методом для соответствующего функционала.

Литература

1. **Kochengin, S.A.** Computational algorithms for constructing reflectors / S.A. Kochengin, V.I. Olikier // Computing and Visualization in Science. – 2003. – Vol. 6. – P. 15-21. – DOI: 10.1007/s00791-003-0103-2.
2. **Досколович, Л.Л.** О применении метода согласованных квадрик к расчёту дифракционных оптических элементов / Л.Л. Досколович, М.А. Моисеев, Н.Л. Казанский // Компьютерная оптика. – 2015. – Т. 39, № 3. – С. 339-346. – DOI: 10.18287/0134-2452-2015-39-3-339-346.

3. **Андреева, К.В.** Метод расчёта оптических элементов с поверхностью свободной формы, работающей по принципу полного внутреннего отражения / К.В. Андреева, М.А. Моисеев, С.В. Кравченко, Л.Л. Досколович // Компьютерная оптика. – 2016. – Т. 40, № 4. – С. 467-474. – DOI: 10.18287/2412-6179-2016-40-4-467-474.
4. **Rubinstein, J.** Intensity control with a free-form lens / J. Rubinstein, G. Wolansky // Journal of the Optical Society of America A. – 2007. – Vol. 24, Issue 2. – P. 463-469. – DOI: 10.1364/JOSAA.24.000463.
5. **Glimm, T.** Optical design of two-reflector systems, the Monge-Kantorovich mass transfer problem and Fermat’s principle / T. Glimm, V.I. Olikier // Indiana University Mathematics Journal. – 2004. – Vol. 53, Issue 5. – P. 1255-1277. – DOI: 10.1512/iumj.2004.53.2455.
6. **Макаров, Б.М.** Лекции по вещественному анализу / Б.М. Макаров, А.Н. Подкорытов. – Санкт-Петербург: БХВ-Петербург, 2011. – 688 с.
7. **Yadav, N.K.** Monge-Ampere problems with non-quadratic cost function: application to freeform optics / N.K. Yadav. – Eindhoven: Technische Universiteit Eindhoven, 2018. – 152 p.
8. **Glimm, T.** Optical design of single reflector systems and the Monge-Kantorovich mass transfer problem / T. Glimm, V.I. Olikier // Journal of Mathematical Sciences. – 2003. – Vol. 117. – P. 4096-4108. – DOI: 10.1023/A:1024856201493.
9. **Wang, X.-J.** On design of a reflector antenna II / X.-J. Wang // Calculus of Variations and Partial Differential Equations. – 2004. – Vol. 20. – P. 329-341. – DOI: 10.1007/s00526-003-0239-4.
10. **Schwartzburg, Y.** High-contrast computational caustic design / Y. Schwartzburg, R. Testuz, A. Tagliasacchi, M. Pauly // ACM Transactions on Graphics. – 2014. – Vol. 33, Issue 4. – 74. – DOI: 10.1145/2601097.2601200.
11. **Aurenhammer, F.** Minkowski-type theorems and least-squares clustering / F. Aurenhammer, F. Hoffmann, B. Aronov // Algorithmica. – 1998. – Vol. 20, Issue 1. – P. 61-76. – DOI: 10.1007/PL00009187.
12. **Merigot, Q.** A multiscale approach to optimal transport / Q. Merigot // Computer Graphics Forum. – 2011. – Vol. 30, Issue 5. – P. 1584-1592. – DOI: 10.1111/j.1467-8659.2011.02032.x.

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.

ГРНТИ: 29.31.01

Поступила в редакцию 18 июля 2020 г. Окончательный вариант – 14 октября 2020 г.

Supporting quadric method for collimated beams

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

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

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

Abstract

We consider the problem of calculating a refractive element with two surfaces, forming a flat front and a given distribution of illumination. The supporting quadrics method is formulated for calculating a given optical element and it is shown that this method coincides with the gradient method for some functional related to the problem of the Monge-Kantorovich mass transfer problem. This enables adaptive selection of the step in the supporting quadric method. At the end of the article a design example is given .

Keywords: geometric optics, non-imaging optics, inverse problem, Monge-Kantorovich mass transfer problem.

Citation: Mingazov AA, Doskolovich LL, Bykov DA. Supporting quadric method for collimated beams. *Computer Optics* 2021; 45(1): 29-37. DOI: 10.18287/2412-6179-CO-783.

Acknowledgements: This work was supported by Ministry of Science and Higher Education of the Russian Federation (FSRC “Crystallography and Photonics” RAS) in part of parts of the numerical implementation of the calculation algorithm and Russian Foundation for Basic Research (18-29-03067, 18-07-00982) in part of of formulating the supporting quadric method and proving the coincidence with the gradient method for the corresponding functional.

References

- [1] Kochengin SA, Oliker VI. Computational algorithms for constructing reflectors. *Comput Vis Sci* 2003; 6: 15-21. DOI: 10.1007/s00791-003-0103-2.
- [2] 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.
- [3] 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.
- [4] Rubinstein J, Wolansky G. Intensity control with a free-form lens. *J Opt Soc Am A* 2007; 24(2): 463-469. DOI: 10.1364/JOSAA.24.000463.
- [5] 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.
- [6] Makarov B, Podkorytov A. *Real analysis: Measures, Integrals and applications: Measures, integrals and applications*. London: Springer-Verlag; 2013.
- [7] Yadav NK. Monge-Ampere problems with non-quadratic cost function: application to freeform optics. Eindhoven: Technische Universiteit Eindhoven; 2018.
- [8] 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.
- [9] Wang X-J. On design of a reflector antenna II. *Calc Var Partial Differ Equ* 2004; 20: 329-341. DOI: 10.1007/s00526-003-0239-4.
- [10] Schwartzburg Y, Testuz R, Tagliasacchi A, Pauly M. High-contrast computational caustic design. *ACM Trans Graph* 2014; 33(4): 74. DOI: 10.1145/2601097.2601200.
- [11] Aurenhammer F, Hoffmann F, Aronov B. Minkowski-type theorems and least-squares clustering. *Algorithmica* 1998; 20(1): 61-76. DOI: 10.1007/PL00009187.
- [12] 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.

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. Peterburg 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 depart-

ment 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@smr.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.

Received July 18, 2020. The final version – October 14, 2020.
