А.А.Гончарский, С.Ю.Романов, Д.В.Туницкий

О НЕКОТОРЫХ ОБРАТНЫХ ЗАДАЧАХ СИНТЕЗА ПЛОСКОЙ КОМПЬЮТЕРНОЙ ОПТИКИ

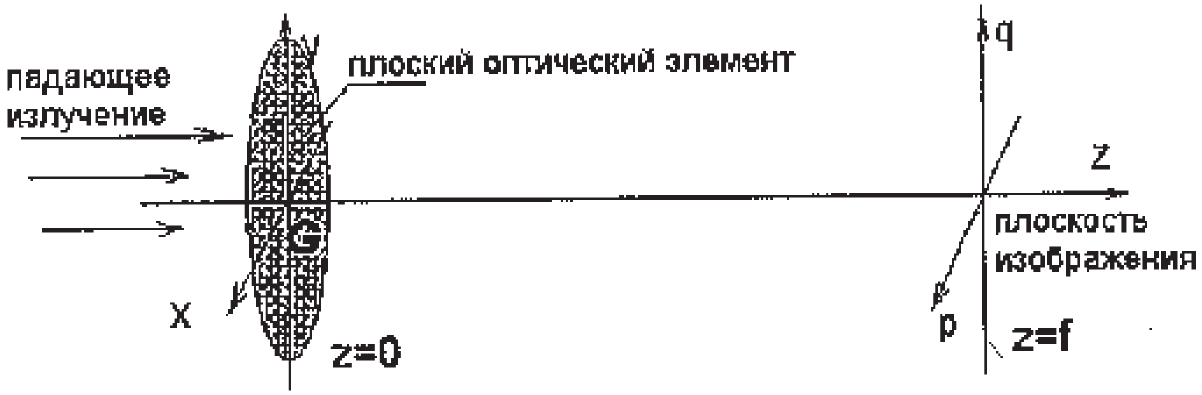
1. Введение

При практическом применении лазеров и других источников света часто возникает задача преобразования их излучения. Возможности традиционных оптических элементов - линз, призм, зеркал и т.п. весьма ограничены. Плоские компьютерные элементы обладают более широкими возможностями. Технология синтеза плоских компьютерных элементов включает в себя выбор математической модели адекватно описывающей распространение излучения в рассматриваемом случае; математическую постановку и решение обратной задачи синтеза; и практическую реализацию полученного решения. Настоящая статья посвящена решению одной обратной задачи синтеза оптических элементов для преобразования когерентного оптического излучения. Среди решенных обратных задач синтеза оптических элементов преобразующих лазерное излучение необходимо отметить решение проблемы фокусировки в произвольную кривую с заданным распределением интенсивности [1]. В настоящее время плоские фазовые оптические элементы находят широкое применение как средство защиты от подделки в компьютерно-синтезированных голограммах [2]. Цель настоящей работы: разработка алгоритмов решения обратной задачи преобразования излучения из области в область; исследование применимости приближения геометрической оптики для решения обратной задачи синтеза оптических элементов для преобразования излучения из области в область при формировании изображений.

2. Модель распространения излучения для решения обратной задачи

Наиболее точной моделью описывающей распространение света являются уравнения Максвелла. Но использование этой модели для решения задач формирования оптического излучения плоскими оптическими элементами представляется весьма проблематичным и возможно только в очень простых случаях. В настоящей работе мы будем рассматривать обратную задачу в приближении геометрической оптики. В этом приближении волновое поле в каждой точке пространства описывается амплитудой и фазой. Пусть излучение распространяется вдоль оси Z, а оптический элемент находится в точке $Z = 0$, на него падает плоская волна с равномерным распределением амплитуды (Рисунок 1);

Рис. 1:



пусть также оптический элемент занимает область G_1. Требуется создать такой элемент который на расстоянии f создаст такое волновое поле что его амплитуда будет равна нулю всюду за пределами области G_2, а внутри области G_2 амплитуда будет равномерной. В дальнейшем будем рассматривать только фазовые оптические элементы, которые преобразуют излучение без потерь энергии. Тогда оптический элемент характеризуется фазовой функцией $\varphi(x, y)$. Если оптический элемент задан, то задано отображение из области G_1 в G_2 [1]:

$$\begin{cases} p = p(x, y) = x + f\varphi_x(x, y), \\ q = q(x, y) = y + f\varphi_y(x, y). \end{cases} \tag{1}$$

Если задано отображение, то образом области G_1 будет некоторая область G_2, при этом якобиан отображения определяет отношение амплитуд. Если отображение найдено, то легко восстановить функцию $\varphi(x, y)$ которая и является искомым решением обратной задачи синтеза. Таким образом требуется найти фазовую функцию $\varphi(x, y)$ оптического элемента занимающего область G_1, который формирует из падающей на него плоской волны на расстоянии f область G_2, в которой сформированная волна имеет постоянную амплитуду. В когерентной оптике явлется важной взаимнооднозначность отображения. С точки зрения практической реализации элементов предпочтительны элементы с гладкой фазовой функцией.

Таким образом, для нахождения фазовой функции такого элемента требуется найти непрерывно дифференцируемое взаимнооднозначное отображение G_1 в G_2. Несмотря на кажущуюся простоту задача относится к нелинейным обратным задачам, поэтому разработка математических алгоритмов решения представляется достаточно сложной математической проблемой. На базе модели геометрической оптики сформулируем математическую постановку обратной задачи.

3. Математическая модель и постановка обратной задачи

Распределение амплитуд в поперечном направлении падающего излучения в плоскости xy и сфокусированного в плоскости pq будем считать постоянными. Формально это приводит к условию постоянства якобиана отображения (1):

$$J(p,q) = p_x q_y - p_y q_x = const. \tag{2}$$

Причем ищется взаимнооднозначное отображение G_1 в G_2. Из системы (1) нетрудно также получить уравнение

$$p_y = q_x. \tag{3}$$

Рассмотрим следующую математическую постановку обратной задачи. Пусть на плоскости параметров xy задана область G_1 , а на плоскости параметров pq - область G_2, причем G_1 и G_2 - компактные выпуклые области. Будем считать, что G_1 ограничена выпуклой кривой

$$g_1(x,y) = 0, \tag{4а}$$

а область G_2 - выпуклой кривой

$$g_2(p,q) = 0, \tag{4б}$$

где функции g_1 и g_2 принадлежат классу $\mathrm{C}^{1+\alpha}$, $0 < \alpha < 1$, и

$$dg_1(x,y) \neq 0, \quad dg_2(p,q) \neq 0, \tag{5}$$

в некоторой окрестности границ ∂G_1 и ∂G_2 соответственно. Требуется найти такой диффеоморфизм

$$p = p(\mathrm{x},\mathrm{y}), \quad q = q(\mathrm{x},\mathrm{y}), \tag{5а}$$

класса $C^{1+\alpha}(G_1)$ области G_1 на область G_2, который удовлетворяет двум условиям:

$$\begin{cases} J(p,q) = p_x q_y - p_y q_x = const > 0, \\ p_y = q_x. \end{cases} \qquad (6)$$

Отметим, что для создания оптических элементов, фокусирующих область G_1 в G_2 необходимо вычисление фазовой функции $\varphi(x,y)$ из уравнений (1). Как можно показать, по найденным функциям $p(x,y)$, $q(x,y)$ из системы (1) с точностью до константы находится фазовая функция $\varphi(x,y)$. Действительно, по условию исходной задачи область G_1 односвязна, и потому, как известно, выполнение условия (2) и (3) равносильно существованию функции $z = z(\mathrm{x},\mathrm{y})$, первые производные которой равны $p(x,y)$ и $q(x,y)$, т.е. для $(\mathrm{x},\mathrm{y}) \in G_1$

$$p(x,y) = z_x(x,y), \quad q(x,y) = z_y(x,y). \qquad (7)$$

Тогда фазовая функция имеет вид

$$\varphi(x,y) = 1/f(z(x,y) - (x^2+y^2)/2). \qquad (8)$$

Для решения задачи (5а)-(6) ее удобно свести к эллиптическому уравнению Монжа-Ампера. Подставляя $p(x,y)$, $q(x,y)$ из (7) в систему (6), приходим к уравнению Монжа-Ампера

$$z_{xx} z_{yy} - {z^2}_{xy} - 1 = 0. \qquad (9)$$

Покажем, что исходная задача (5а)-(6) равносильна решению уравнения (9) с краевыми условиями (4б). Действительно, если отображение (5а) является решением уравнений (6), то с точностью до постоянного слагаемого в G_1 однозначно определена функция $z = z(\mathrm{x},\mathrm{y})$, для которой выполняются (7) и $z(\mathrm{x},\mathrm{y})$ - решение уравнения (9). Обратно, если функция $z = z(\mathrm{x},\mathrm{y})$ - такое решение уравнения (9), что определенное равенствами (7) отображение является диффеоморфизмом области G_1 на G_2 , то этот диффеоморфизм - решение исходной задачи (5а)-(6). Как известно, всякая дважды непрерывно дифференцируемая на компактном выпуклом множестве функция тогда и только тогда сильно выпукла, когда ее гессиан положителен [3]. Кроме того, градиент дважды дифференцируемой сильно выпуклой функции является диффеоморфизмом. Тем самым, задача (5а)-(6) равносильна нахождению в G_1 решения уравнения (9) класса $C^{2+\alpha}(G_1)$, для которого на границе ∂G_1 выполнено нелинейное краевое условие (4б).

Для нелинейной краевой задачи (9) (4б) в рассмотренных выше условиях имеет место теорема существования и единственности [4], а именно, с точностью до постоянного слагаемого существует ровно два решения $z_1 = z_1(x,y)$ и $z_2 = z_2(x,y)$, первое из которых направлено выпуклостью вниз, а второе - выпуклостью вверх. В терминах задачи (5а)-(6) это означает, что существует ровно два диффеоморфизма $p_1(x,y)$, $q_1(x,y)$ и $p_2(x,y)$ $q_2(x,y)$ области G_1 в G_2, первое из которых имеет положительную производную p_{1x} , а второе отрицательную производную p_{2x} . Аналогично существует ровно две фазовые функции из соотношения (1) фокусирующие излучение из области G_1 в G_2. Для случая, когда области G_1 и G_2 центральносимметричны оба решения системы (6) связаны простыми соотношениями:

$$\begin{gathered} p_2(-x,-y) = p_1(x,y), q_2(-x,-y) = q_1(x,y), \\ p_2(x,y) = -p_1(x,y), q_2(x,y) = -q_1(x,y). \end{gathered} \tag{10}$$

Это означает, что оба отображения областей G_1 в G_2 переводят центральносимметричные точки области G_1 в центральносимметричные точки области G_2 . Отличаются отображения между собой лишь дополнительным поворотом на 180^0 в области G_2.

4. Методы решения обратной задачи

Остановимся на методах решения исходной задачи (5а)-(6). Рассмотрим единичный круг

$$\mathrm{K} = \{(u,v)| \quad u^2 + v^2 \le 1\}, \tag{11}$$

и дробно-линейное преобразование

$$\chi = \mathrm{e}^{i\theta} \frac{(w-a)}{(1-w\overline{a})}, \tag{12}$$

где θ - произвольное вещественное число, a - комплексное число, лежащее внутри единичного круга, т.е. $|a| < 1$, и $w = u + iv$.

Найдем решение следующей краевой задачи в единичном круге K

$$\begin{cases} y_u - p_v = 0, \\ q_u - x_v = 0, \\ y_v - p_u = 0, \\ q_v - x_u = 0, \end{cases} \tag{13}$$

с краевыми условиями

$$g_1(x(u,v), \quad y(u,v)) = 0, \quad g_2(p(u,v)), \quad q(u,v)) = 0, \tag{14}$$

на единичной окружности ∂K, где $x(u,v)$, $y(u,v)$, $p(u,v)$, $q(u,v)$ – функции заданные в круге $(u,v) \in K$.

Можно показать, что если дополнительно выполнено условие

$$x_u y_v - x_v y_u \neq 0, \tag{15}$$

то функции $p(u(x,y),v(x,y))$, $q(u(x,y),v(x,y))$, где $(u(x,y),v(x,y)) = (x(u,v),y(u,v))^{-1}$ осуществляют диффеоморфизм выпуклой области G_1 на выпуклую область G_2 и являются решением задачи (5a)-(6).

Для задачи (13)-(14) выполняются теоремы существования и единственности решения с точностью до дробно-линейного преобразования χ в круге. Следующие очевидные свойства будут использованы при реализации численных алгоритмов решения краевой задачи (13)-(14).

Из (13) следует, что функции $y + i$p и $q + ix$ голоморфны внутри единичного круга K.

Далее заметим, что в силу голоморфности $y + i$p и $q + ix$ в замкнутом круге K имеет место равномерная сходимость

$$\lim_{N \to +\infty} \overline{f_N} = y + ip, \lim_{N \to +\infty} \overline{t_N} = q + ix, \tag{16}$$

где $\overline{f_N}, \overline{t_N}$ комплексные многочлены вида

$$\overline{f_N} = \sum_{n=0}^{N} a_n (u+iv)^n, \overline{t_N} = \sum_{n=0}^{N} b_n (u+iv)^n. \tag{17}$$

Кроме того, при фиксированном N функции $y+ip-f_N$ и $q+ix-t_N$ (где f_N, t_N - многочлены вида (17)) голоморфны внутри единичного круга K, а потому

$$x - Im(t_N), y - Re(f_N), p - Im(f_N), q - Re(t_N) \tag{19}$$

аналитичны и достигают наибольшего и наименьшего значения на единичной окружности ∂K. Иначе говоря, модули погрешностей приближений f_N и t_N принимают максимальные значения на границе.

Исходя из вышеизложенного, можно предложить следующий численный алгоритм решения задачи (13)-(14), а следовательно задачи (5a)-(6), причем приведенные выше соображения позволяют надеяться на минимизирующие свойства алгоритма, а при фокусировке из круга в квадрат (G_1 -круг, G_2 -квадрат) сходимость равномерная.

Рассматривается функционал при фиксированной степени N многочленов

$$\Theta(f_N, t_N) = \sum_{k=1}^{M} (g_1(Im(t_{Nk}), Re(f_{Nk}))^2 + (g_2(Im(f_{Nk}), Re(t_{Nk}))^2), \qquad (20)$$

где t_{Nk} и f_{Nk} – значения многочленов t_N и f_N на равномерной сетке из M точек единичной окружности ∂К. Алгоритм состоит в минимизации этого функционала по $2N$ комплексным параметрам – коэффициентам многочленов (17). Полученные при минимизации функционала (20) многочлены f_N, t_N рассматриваются как приближенное значение для $y + ip$ и $q + ix$ соответственно. Заметим, что приближенное решение удовлетворяет уравнениям (2) и (3). Численную реализацию алгоритма удобно осуществлять, например, градиентным методом [5], постепенно повышая степень N многочленов, причем для многочленов $N+1$-ой степени за начальное приближение берутся многочлены степени N, минимизирующие функционал $\Theta(f_N, t_N)$.

Таким образом, задача нахождения фазовой функции оптического элемента $\varphi(x, y)$ решается в три этапа: минимизируется функционал (20), находятся $p(u, v)$, $q(u, v)$, $x(u, v)$, $y(u, v)$ в виде полиномов; обращением находится отображение (5а) $p(x, y)$, $q(x, y)$; из (7) вычисляется функция $z(x, y)$, и получается $\varphi(x, y)$ из (8).

5. Расчет модельных задач

В качестве модельной задачи рассматривалась задача формирования излучения из круга с равномерным распределением амплитуды в квадрат с равномерным распределением аплитуды. Согласно описанному выше алгоритму были рассчитаны две фазовые функции $\varphi_1(x, y)$ и $\varphi_2(x, y)$ оптических элементов преобразующих излучение из круга в квадрат, соответствующие первому и второму решениям (9). Расчеты фазовых функций проводились при следующих параметрах: площади круга и квадрата равны; диаметр круга 1мм; соответственно сторона квадрата $\sqrt{(1/\pi)}$мм; расстояние f между оптическим элементом и изображением 20мм. При минимизации функционала (20) использовались полиномы с числом членов $N = 150$. На рисунке 2 справа показана фазовая функция $\varphi_1(x, y)$ как поверхность, слева показаны ее линии уровня. На рисунке 3 показана фазовая функция $\varphi_2(x, y)$.

Приближение геометрической оптики не учитывает волновых эффектов таких как дифракция и интерференция. Приближение Френеля более

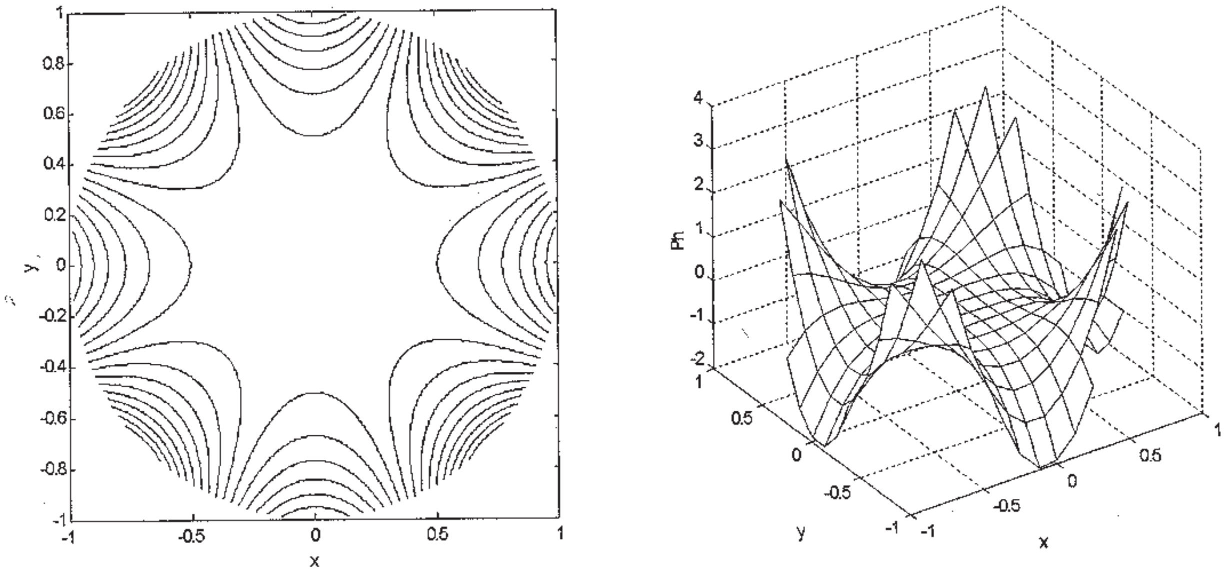


Рис. 2: Фазовая функция 1.

точно описывает процесс распространения излучения. В приближении Френеля [1] существует связь между значениями волнового поля $u_1(x, y)$ после элемента с фазовой функцией $\varphi(x, y)$ в плоскости $Z = 0$ и $u_2(p, q)$ в плоскости изображения $Z = f$ (см Рис.1).

$$u_2(p, q) = A \int\limits_{G_1} e^{-\frac{ik}{2f}((p-x)^2+(q-y)^2)} e^{-ik\varphi(x,y)} dxdy, \tag{21}$$

где $k = \frac{2\pi}{\lambda}$, λ – длина волны излучения, A – некоторая константа. Как показывает практика, точность рассчитанной в этом приближении амплитуды волнового поля в плоскости изображения достаточна при параметрах настоящей задачи. Были проведены расчеты амплитуды волнового поля в приближении Френеля. При расчетах использовалось сведение интеграла в (21) к двумерному преобразованию Фурье, которое вычислялось посредством БДПФ, при этом использовалась сетка 512 на 512 точек. На рисунке 3 показана амплитуда волнового поля созданная оптическим элементом с фазой $\varphi_1(x, y)$ слева и с фазой $\varphi_2(x, y)$ справа. Видно что не смотря на весьма разный вид фазовых функций (см. рис. 1 и 2) рассчитанные амплитуды отличаются мало и они отличаются незначительно от требуемых. На рисунке 4 показана амплитуда в плоскости изображения в разрезе при $q = 0$.

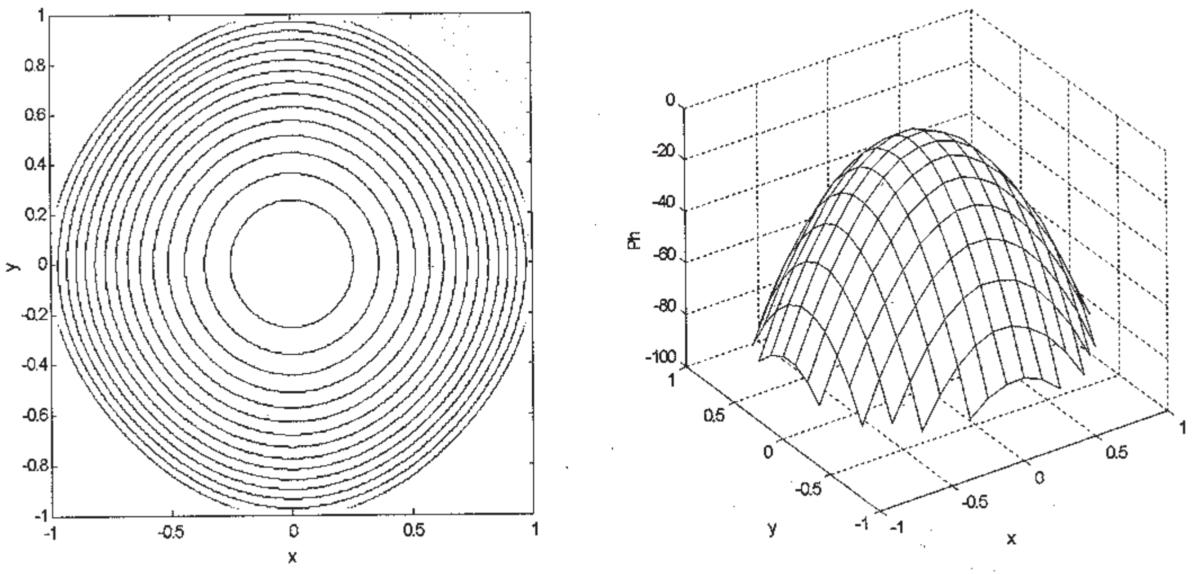


Рис. 3: Фазовая функция 2.

Литература

[1] **Гончарский А.В., Попов В.В., Степанов В.В.** Введение в компьютерную оптику. М.: Изд-во Моск. ун-та, 1991.

[2] **Van Renesse R.L.** Optical Document Security. Boston, London: Artech House, 1998.

[3] **Погорелов А.В.** Внешняя геометрия выпуклых поверхностей. М.: Наука, 1969.

[4] **Васильев Ф.П.** Численные методы решения экстремальных задач. М.: Наука, 1988.

[5] **Корн Г., Корн Т.** Справочник по математике. М.: Наука, 1977.

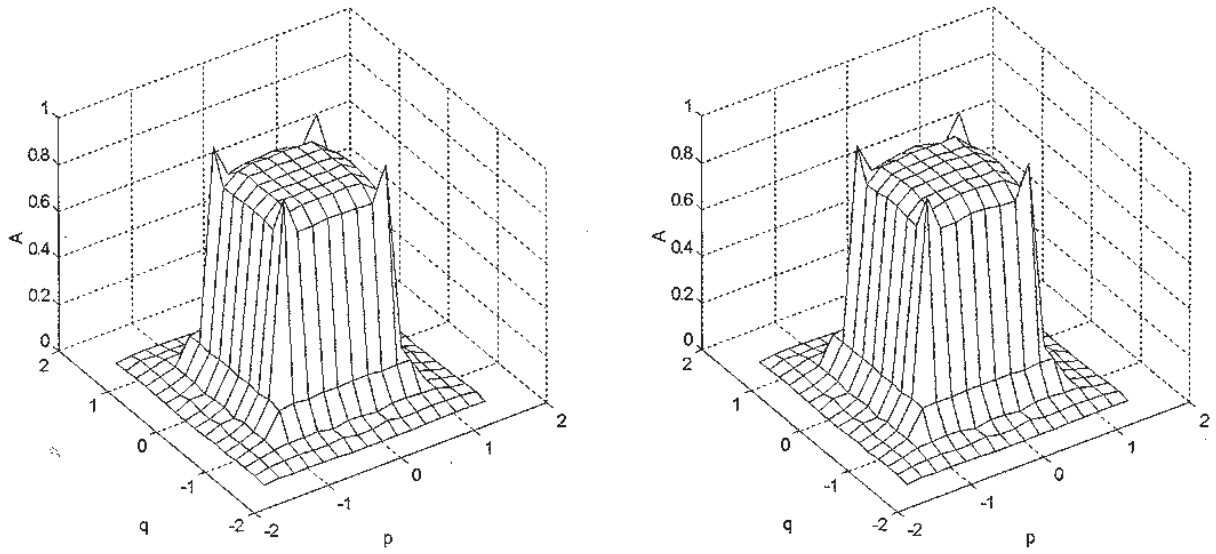


Рис. 4: Амплитуды изображения при $\varphi_1(x, y)$ слева; при $\varphi_2(x, y)$ справа.

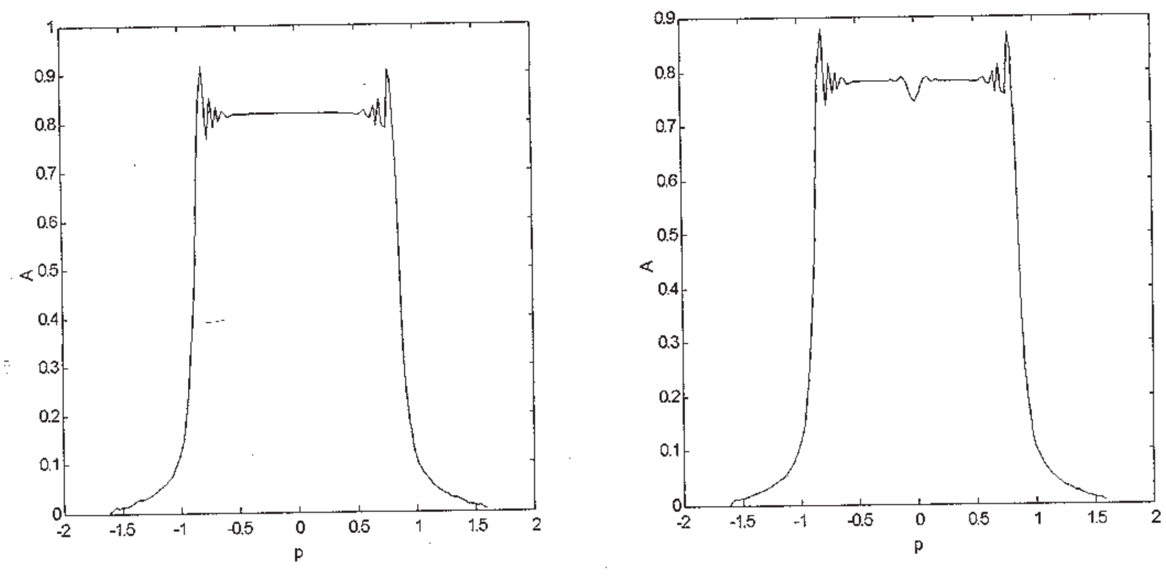


Рис. 5: Амплитуда изображения в разрезе $(q = 0)$ при $\varphi_1(x, y)$ слева; при $\varphi_2(x, y)$ справа.