

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)**
Институт прикладной математики и информатики био- и нанотехнологий
Кафедра физики и прикладной математики

Седов Евгений Сергеевич
Давыдов Николай Николаевич

Методические указания к практическим занятиям
по дисциплине **«Системы адаптивной оптики и их приложения»**
для студентов ВлГУ, обучающихся
по направлению
12.04.05 (200500.68) **«Лазерная техника и лазерные технологии»**

Владимир-2015 г.

Оглавление

Список используемых сокращений	3
ТЕМА 1. Принципиальная схема адаптивной оптической системы. Датчики волнового фронта	4
ТЕМА 2. Мембранные и биморфные зеркала	15
ТЕМА 3. Алгоритмы фазового сопряжения	23
ТЕМА 4. Метод покоординатного спуска	29
ТЕМА 5. Генетический и гибридный алгоритмы определения управляющих напряжений	36
ТЕМА 6. Применение адаптивной оптики в оптических системах мощных лазеров	47
ТЕМА 7. Применение адаптивной оптики в оптике атмосферы и лазерных технологиях	53
Список литературы	59

Список используемых сокращений

АОС	адаптивная оптическая система
ДВФ	датчик волнового фронта
ДЗ	деформируемое зеркало
КВФ	корректор волнового фронта
метод ПВР	метод последовательной верхней релаксации
СКУ	система контроля и управления

ТЕМА 1. Принципиальная схема адаптивной оптической системы. Датчики волнового фронта

Основные вопросы темы: назначение и принципиальная схема систем адаптивной оптики. Описание атмосферной турбулентности. Описание искажений волнового фронта излучения, распространяющегося в турбулентной атмосфере.

Цель и задачи практического занятия состоят в ознакомлении студента с основными компонентами систем адаптивной оптики, их назначением и взаимным расположением, с алгоритмом описания турбулентной атмосферы на основе многослойной модели, с физическими ограничениями построения изображений оптическими системами, с методом декомпозиции волнового фронта в базисе ортонормированных функций для описания вклада турбулентности атмосферы в распространение излучения.

Основной понятийно-терминологический аппарат: адаптивная оптическая система, волновой фронт, датчик волнового фронта, деформируемое зеркало, длина когерентности, корректор волнового фронта, многослойная модель атмосферы, полиномы Цернике, разрешающая способность, система контроля и управления, структурная постоянная флуктуаций показателя преломления, структурная функция.

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, представлять принципиальную схему системы адаптивной оптики, назначение её компонентов, понимать необходимость использования систем адаптивной оптики, быть знакомым с многослойной моделью атмосферы, уметь проводить процедуру дискретизации непрерывного профиля турбулентной атмосферы, на её основе находить зависимости структурной постоянной показателя преломления и скорости ветра от высоты, понимать роль атмосферной турбулентности в построении изображения оптической системой, уметь проводить разложение волнового фронта по полиномам Цернике и понимать физический смысл этого разложения.

Теоретический материал и примеры заданий по теме 1

1.1. Схема адаптивной оптической системы

Методы адаптивной оптики применяются для улучшения качества изображения, получаемого при помощи оптических систем. Роль адаптивных оптических систем (АОС)

в построении изображения заключается в исправлении искажений волнового фронта излучения, возникающих в виду различных причин, как то: турбулентность среды распространения излучения (в частности, земной атмосферы), особенности структуры и испускательных свойств источников излучения, изменение оптических свойств и геометрических характеристик компонентов оптической системы в результате нагрева, под воздействием вибраций и др.

Принципиальная схема АОС представлена на рис. 1. Излучение от источника подвергается внешним воздействиям, и его волновой фронт искажается. Излучение попадает на корректор волнового фронта (КВФ), в роли которого чаще всего выступает деформируемое зеркало (ДЗ), которое компенсирует внесённые искажения и выправляет волновой фронт. От ДЗ излучение поступает на светоделитель, с которого часть излучения направляется на приёмник, где формируется изображение источника. Оставшееся излучение отводится на датчик волнового фронта (ДВФ); он измеряет не исправленные зеркалом искажения и отправляет сигнал с этой информацией с систему контроля и управления (СКУ), которая преобразует полученную информацию в сигнал для ДЗ. В соответствии с сигналом меняется форма ДЗ с целью исправить оставшиеся искажения волнового фронта.

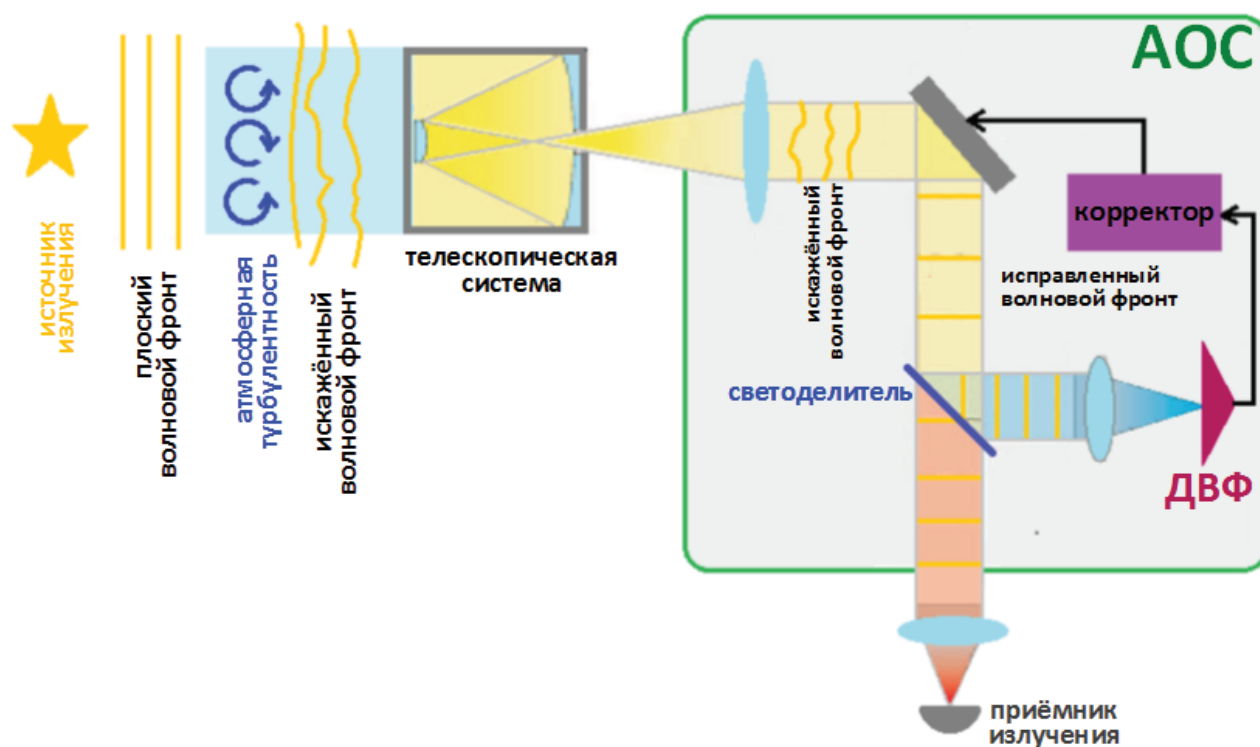


Рис. 1. Принципиальная схема системы адаптивной оптики.

Деформируемое зеркало — это один из ключевых компонентов системы коррекции волнового фронта. Его форма может изменяться в зависимости от сигнала СКУ. В идеальном случае форма зеркала должна соответствовать инвертированной форме искажённого волнового фронта: волновой фронт излучения, отражённого от такого зеркала, соответствует исходному волновому фронту излучения источника (плоскому или сферическому). Помимо ДЗ система коррекции волнового фронта может содержать также поворотные зеркала, которые позволяют исправить простейшие аберрации, связанные с наклоном и поворотом волнового фронта излучения.

При помощи датчиков волнового фронта оцениваются искажения волнового фронта, внесённые в ходе распространения излучения от источника. В общем случае данные с ДВФ позволяют создать трёхмерную модель поверхности с одинаково фазой; на основе этой модели задаётся форма ДЗ.

1.2. Многослойная модель атмосферы

Одной из областей научных исследований, где АОС нашли наиболее широкое применение, является астрономия. Свет от удалённых астрономических объектов по мере распространения подвергается целому ряду воздействий, приводящих к искажениям волнового фронта. Основной причиной искажения является атмосферная турбулентность. Для её описания построен целый ряд математических моделей, одна из которых — многослойная модель. В рамках этой модели полагается, что турбулентные эффекты сосредоточены в бесконечно тонких статистически независимых фазовых экранах, разделяющих между собой набор слоёв с однородной (нетурбулентной) атмосферой. Данное приближение подразумевает, что волна при прохождении фазового экрана приобретает набег фазы в направлении распространения излучения и затем распространяется до следующего экрана в однородном слое атмосферы. Рассмотрение набора слоёв позволяет провести процедуру дискретизации непрерывного профиля турбулентной атмосферы. Процедура состоит из следующих этапов.

- 1) Толща атмосферы разделяется на N_{ph} слоёв одинаковой толщины.
- 2) Эквивалентная высота j -го слоя вычисляется в соответствии с выражением

$$h_j^{\text{eq}} = \frac{\int_{h_{j-1}}^{h_j} z C_n^2(z) dz}{\int_{h_{j-1}}^{h_j} C_n^2(z) dz}, \quad j = 1, \dots, N_{\text{el}},$$

где h_j – высота границы слоя, z – координата вдоль оси наблюдения, C_n^2 – структурная постоянная флуктуаций показателя преломления, которая количественно определяет вклад турбулентности. Величина C_n^2 оценивается в соответствии со статистическими моделями: в частности, это модели Модель Хафнагеля-Волли, SLC, Гринвуда. Согласно модели Хафнагеля-Волли величина C_n^2 определяется выражением

$$C_n^2(z) = 5.94 \left(\frac{v_m}{27} \right)^2 (10^{-5.3} z)^{10} e^{-z/1000} + 2.7 \times 10^{-16} e^{-z/1500} + C_{n,0}^2 e^{6H_0/7 - z/100},$$

где z – высота над уровнем моря, v_m – среднее квадратическое значение скоростей воздушного потока в диапазоне высот от 5 до 20 км, H_0 – высота точки наблюдения над уровнем моря, $C_{n,0}^2 = C_n^2(z = H_0)$.

3) Эквивалентное (усреднённое) значение коэффициента C_n^2 находится по формуле

$$(C_{\text{eq},n}^j)^2 = \frac{1}{\Delta h_j} \int_{h_{j-1}}^{h_j} C_n^2(z) dz,$$

где Δh_j – толщина j -го слоя. Вклад каждого слоя в искажения волнового фронта оценивается выражением

$$W_j = \frac{\int_{h_{j-1}}^{h_j} C_n^2(z) dz}{\int_{h_0}^{h_{\text{max}}} C_n^2(z) dz},$$

при этом $\sum_{j=1}^{N_{\text{eq}}} W_j = 1$.

4) Эквивалентная скорость ветра в каждом слое находится как

$$|\mathbf{v}_j^{\text{eq}}| = \left[\frac{\int_{h_{j-1}}^{h_j} C_n^2(z) |\mathbf{v}(z)|^{5/3} dz}{\int_{h_{j-1}}^{h_j} C_n^2(z) dz} \right]^{3/5}.$$

На рис. 2 представлены зависимости коэффициента C_n^2 и скорости ветра $|\mathbf{v}(z)|$ от высоты, оцененные на основе данных обсерватории “Gemini” (Чили) с использованием многослойной модели атмосферы [1].

Задание для самостоятельного выполнения: при помощи системы компьютерной математики (MATLAB, Mathcad, Mathematica, др.) построить в общей системе координат зависимость структурной постоянной показателя преломления C_n^2 от высоты z на основе а) модели Хафнагеля-Волли, б) многослойной модели атмосферы, используя 4 слоя, в) 10 слоёв, г) 100 слоёв. Сделать вывод об адекватности

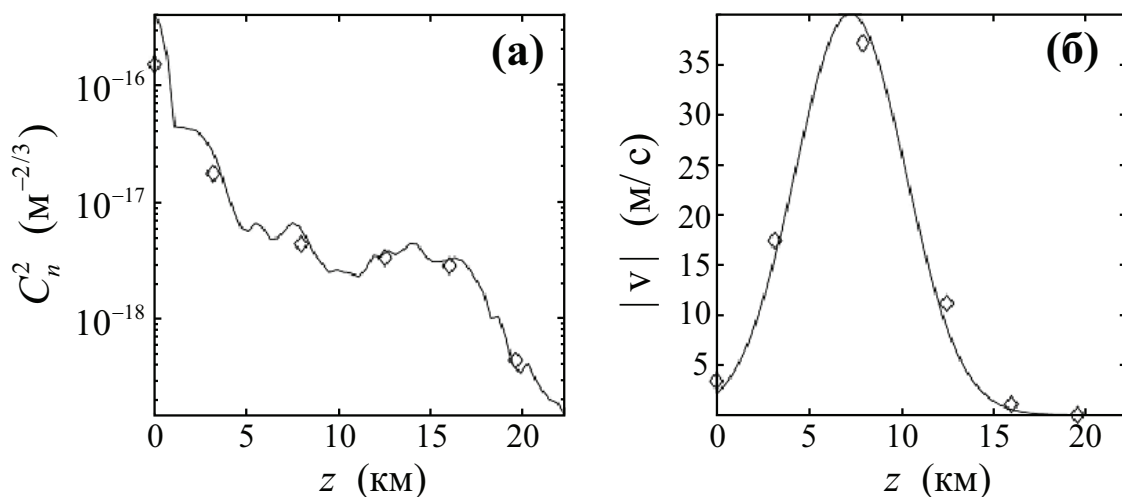


Рис. 2. Зависимости (а) — коэффициента C_n^2 и (б) — скорости ветра $|\mathbf{v}(z)|$, оцененные на основе данных обсерватории “Gemini” (Чили) от высоты с разбиением на пятьдесят четыре слоя (сплошные линии) и шесть слоёв (маркеры в форме ромбов) [1] .

четырёхслойной модели для описания турбулентной атмосферы в сравнении с моделью, использующей 100 слоёв. Для модели Хафнагеля-Волли среднее квадратическое значение скоростей воздушного потока v_m принять равным 21 м/с, построение проводить от высоты $H_0 = 600$ м/с до $h = 30$ км. Значение постоянной $C_{n,0}^2$ принять равным 1.7×10^{-13} м^{-2/3}.

1.3. Описание искажённого волнового фронта

Получение изображения в дифракционном пределе

При получении изображения в оптических системах имеется ряд ограничений. Одно из таких ограничений накладывают дифракционные эффекты, который возникают, например, при прохождении излучения через оптические элементы конечных размеров, которые по сути представляют собой диафрагмы.

В общем виде плоская электромагнитная волна $E(\mathbf{r})$ может быть задана как

$$E(\mathbf{r}) = A(\mathbf{r})e^{-i\varphi(\mathbf{r})}, \quad (1.1)$$

где $A(\mathbf{r})$ описывает амплитуду волны, а $\varphi(\mathbf{r})$ — фазу.

Введём понятие волнового фронта. Волновой фронт — это гипотетическая поверхность, все точки которой характеризуются одинаковой фазой. На рис. 3 схематично представлен волновой фронт одиночного точечного источника. В идеальном случае

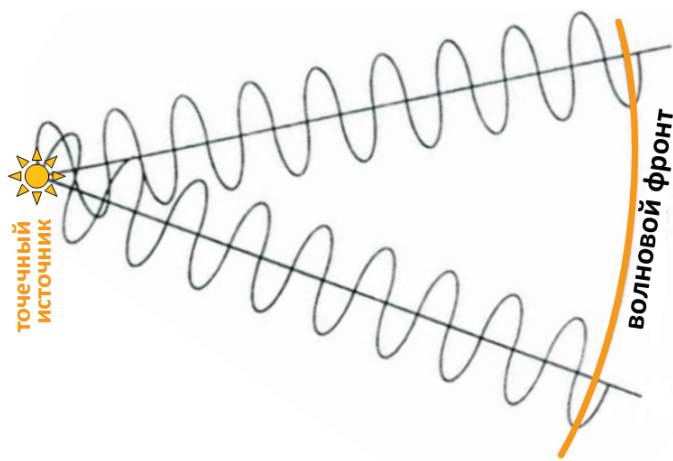


Рис. 3. Формирование сферического волнового фронта излучения одиночного точечного источника.

(в отсутствии aberrаций) волновой фронт точечного источника, расположенного на конечном расстоянии, является сферическим. Если источник находится в бесконечности, его волновой фронт образует плоскость. Сферический и плоский волновые фронты, являющиеся идеальными, в оптике выступают в роли шаблонов, с которыми осуществляется сравнение реальных, подвергнутых искажениям за счёт aberrаций волновых фронтов.

На рис. 4а схематично представлено получение в плоскости экрана изображения точечного источника через круглую апертурную диафрагму в отсутствии aberrаций. В рассматриваемом случае свет от источника приходит на диафрагму в виде плоской волны. В плоскости диафрагмы, характеризующейся радиус-вектором \mathbf{r}_\perp , такой волне присущи постоянные (не зависящие от \mathbf{r}_\perp) амплитуда и фаза. При этом интенсивность излучения составляет $|E|^2 = A^2$ в пределах диафрагмы и $|E|^2 = 0$ за её пределами. В результате дифракции Фраунгофера на круглой диафрагме полученное на экране изображение представляет собой дифракционную картину Эйри — см. схематичный рисунок 4б. Центральная область дифракционной картины называется диском Эйри.

Тот факт, что изображение точечного источника в оптической системе с диафрагмой представляет собой световое пятно конечного размера, говорит о том, что существует ограничение в получении различимых изображения двух таких источников, находящихся на некотором расстоянии друг от друга. Характеристикой этого ограничения является разрешающая способность — это минимальное угловое расстояние между точечными объектами, которые можно различить в оптическую систему. Очевидно,

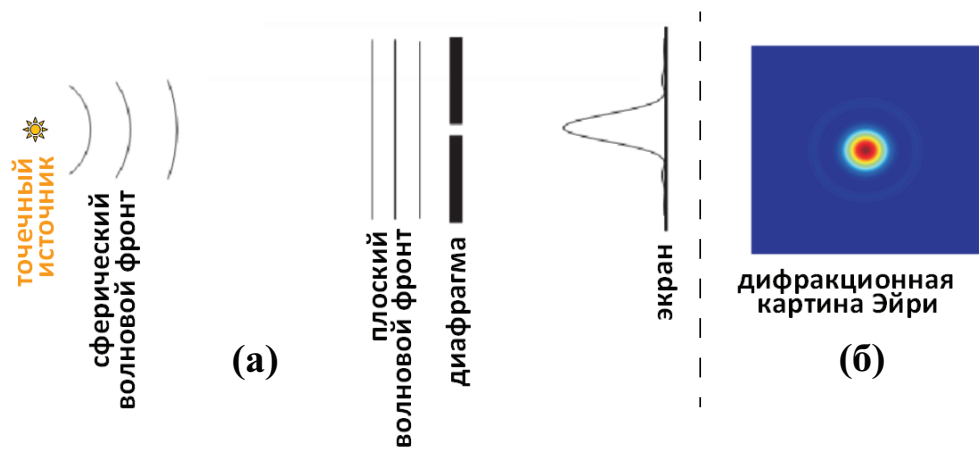


Рис. 4. (а) – Формирование изображения точечного источника через круглую диафрагму и (б) — дифракционная картина Эйри.

что угловое разрешение ограничено размерами диска Эйри. В соответствии с критерием Рэлея, два источника являются различимыми, если центральный максимум дифракционной картины одного источника совпадает с первым минимумом другого. Когда угловое разрешение оптической системы для заданной длины волны определяется радиусом диска Эйри, такая система называется дифракционно ограниченной.

На рис. 5 схематично представлено получение в плоскости экрана изображения точечного источника через круглую апертурную диафрагму при прохождении излучения сквозь турбулентную атмосферу. В этом случае в изначально плоский волновой фронт вносятся случайные возмущения, и он перестаёт быть плоским; изображение источника на экране расплывается, и, как следствие, ухудшается разрешающая способность оптической системы.

Длина когерентности Фрайда

Вклад атмосферной турбулентности в распространение излучения может быть описан при помощи пространственного распределения задержки фазы $\varphi(\mathbf{r})$, которую вносит турбулентность в изначально плоскую волну.

Одним из основных параметров описания волнового фронта при распространении волны через турбулентную атмосферу является длина когерентности или параметр Фрайда (англ.: Fried) r_0 [2]. Воздух, как и любая среда, обладает дисперсионной способностью. Для большинства вычислений поправки на дисперсионные эффекты не вносятся, поскольку являются пренебрежимо малыми. Однако это приближение неприменимо при расчёте фазы волны. Фаза $\varphi(\mathbf{r})$ электромагнитной волны (1.1)

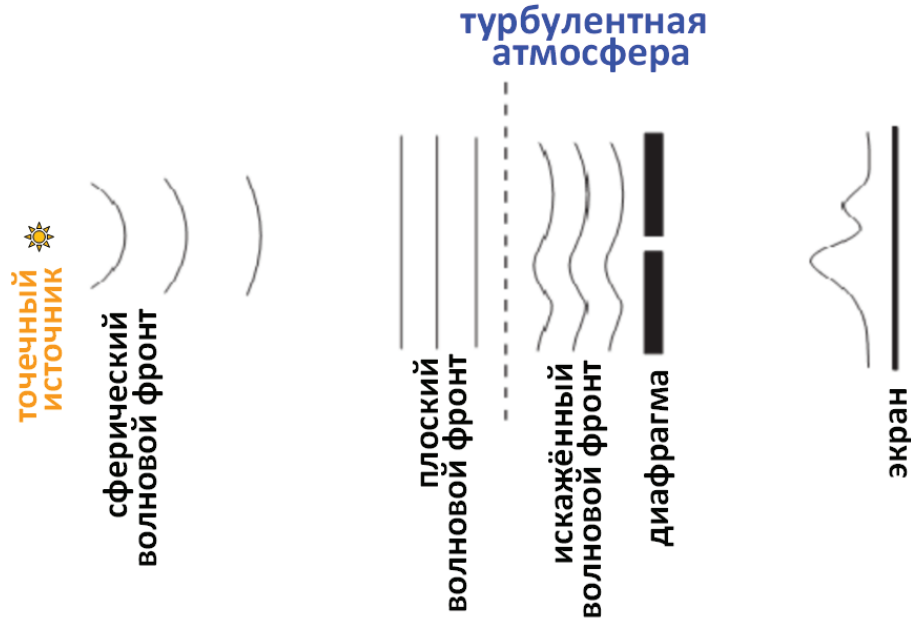


Рис. 5. Формирование изображения оптической системой в присутствии аберраций.

существенно зависит от длины волны света λ . Эта зависимость задаётся выражением

$$\varphi(\mathbf{r}_\perp) = \frac{2\pi}{\lambda} \int_0^\infty n(\mathbf{r}_\perp, z) dz, \quad (1.2)$$

где координата z задаёт высоту, \mathbf{r}_\perp – радиус-вектор в плоскости, перпендикулярной z , $\mathbf{r} = (\mathbf{r}_\perp, z)$; $n(\mathbf{r}_\perp, z)$ – показатель преломления, в общем случае статистически зависящий от высоты. Длина оптического пути вдоль оси z определяется интегралом $l = \int n(z) dz$. По выражению (1.2) легко проследить, что наибольшее влияние искажения оказывают на излучение более коротких длин волн. Так, например, при оптической длине пути l в несколько миллиметров искажение фазы радиоволн оказывается несущественным, поскольку они обладают большой, превышающей величину l длиной волны λ . В этом случае волновой фронт волны от источника, находящегося на удалении, по мере распространения остаётся плоским. Кардинально противоположная ситуация наблюдается для излучения оптического диапазона при той же длине оптического пути l . По мере распространения такая волна перестаёт быть плоской, волновой фронт искривляется, поскольку величина l на порядки превосходит оптическую длину волны λ .

Для описания фазы применяется структурная функция, которая задаётся как

$$D_\varphi(\boldsymbol{\rho}) = \langle [\varphi(\mathbf{r}_\perp) - \varphi(\mathbf{r}_\perp + \boldsymbol{\rho})]^2 \rangle, \quad (1.3)$$

где угловые скобки $\langle \dots \rangle$ обозначают усреднение по ансамблю реализаций. С учётом (1.2) выражение (1.3) преобразуется к виду

$$D_\varphi(\rho) = \frac{2.91(2\pi)^2}{\lambda^2 \cos \theta} \int C_n^2(z) dz \rho^{5/3}, \quad (1.4)$$

где θ задаёт угол между линией наблюдения и зенитом. Зачастую выражение (1.4) записывают следующим образом:

$$D_\varphi(\rho) = 6.88 \left(\frac{\rho}{r_0} \right)^{5/3},$$

где r_0 и определяет длину когерентности (параметр Фрайда):

$$r_0 = \left(\frac{0.423(2\pi)^2}{\lambda^2 \cos \theta} \int C_n^2(z) dz \right)^{-3/5}. \quad (1.5)$$

С физической точки зрения, параметр Фрайда r_0 определяет размер диафрагмы $\rho = r_0$, при котором среднеквадратичное значение aberrации фазы составляет один радиан. Длина когерентности часто используется для описания атмосферной турбулентности. В частности, при помощи параметра r_0 может быть задана спектральная плотность флуктуаций фазы:

$$\Phi_\varphi(K) = \frac{0.023}{r_0^{5/3}} |K|^{-11/3}.$$

Разложение волнового фронта

Более полного представления о воздействии турбулентности атмосферы на волновой фронт позволяет достичь приём, заключающийся в разложении распределения фазы в базисе ортонормированных функций. Случайным образом меняющаяся во времени фаза $\varphi(\mathbf{r}, t)$ раскладывается по детерминированным функциям $Z(\mathbf{r})$ со случайными коэффициентами $C(t)$:

$$\varphi(\mathbf{r}, t) = \sum_{j=1}^{\infty} C_j(t) \phi_j(\mathbf{r}).$$

Базисные функции $Z(\mathbf{r})$ называются полиномами Цернике (англ.: Zernike), которые представляют собой бесконечные ряды ортогональных на единичном круге многочленов. Полиномы Цернике удовлетворяют условию ортонормированности: $\int Z_i(\mathbf{r}) Z_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}$, где δ_{ij} – символ Кронекера; δ_{ij} принимает значение 1 при $i = j$ и значение 0 в противном случае.

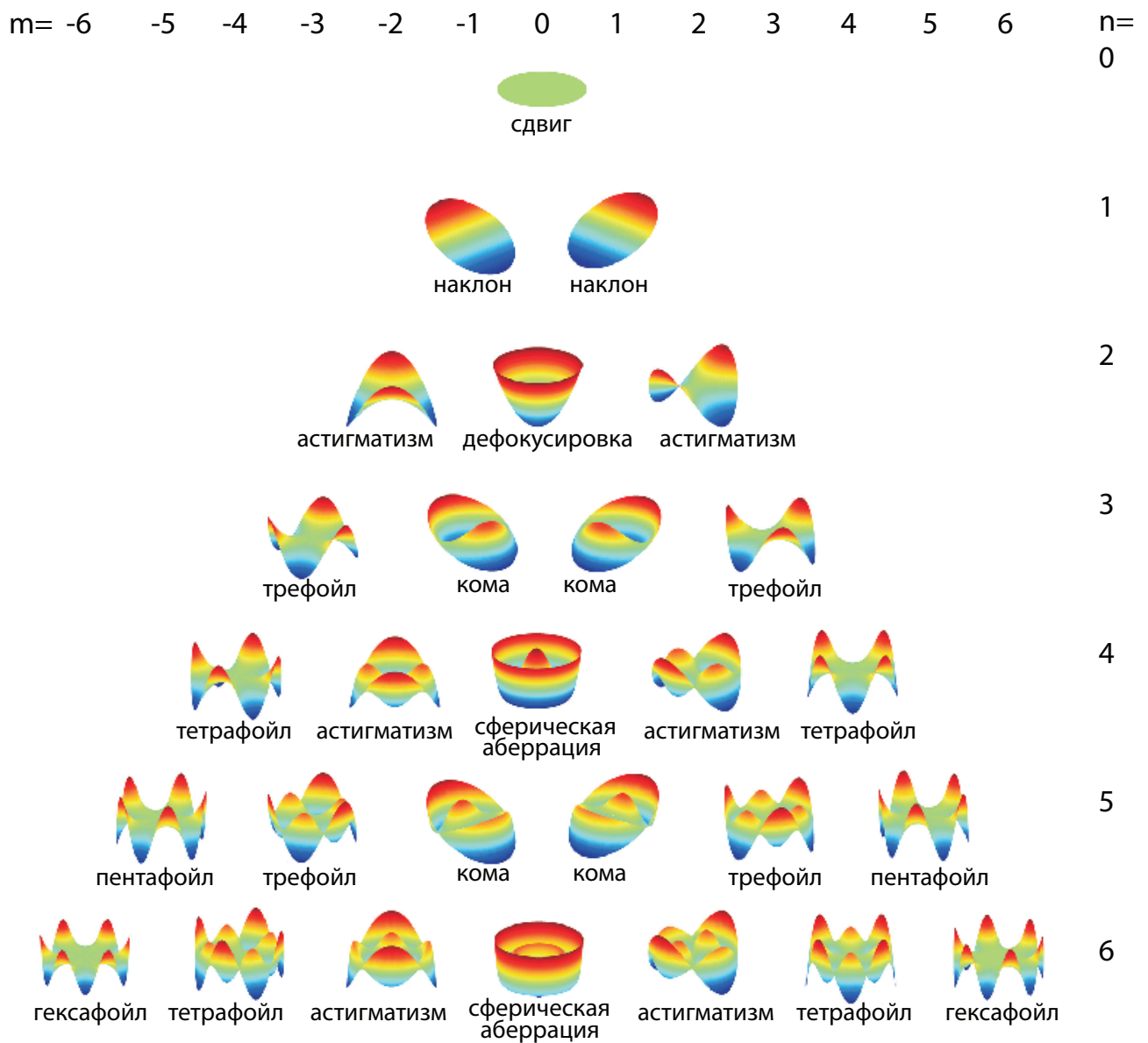


Рис. 6. Визуальное представление полиномов Цернике.

Различают чётные и нечётные полиномы Цернике, которые задаются следующим образом:

$$Z_n^m(r_\perp, \gamma) = R_l^m(r_\perp) \cos(m\gamma), \quad m \neq 0, \quad (1.6a)$$

$$Z_n^{-m}(r_\perp, \gamma) = R_l^m(r_\perp) \sin(m\gamma), \quad m \neq 0, \quad (1.6б)$$

$$Z_0^0(r_\perp, \gamma) = 1, \quad m = 0, \quad (1.6в)$$

где l и m – неотрицательные целые числа, при этом $l \geq m$; $r_\perp = |\mathbf{r}_\perp|$ – радиальное расстояние, $0 \leq r_\perp \leq 1$; γ – азимутальный угол. Коэффициенты $R_l^m(r_\perp)$ называются радиальными полиномами. Когда число $(l - m)$ является нечётным, полином $R_l^m(r_\perp)$ тождественно равен нулю. Для чётных значений $(l - m)$ коэффициенты $R_l^m(r_\perp)$ находятся

в соответствии с выражением

$$R_l^m(r_\perp) = \sum_{s=0}^{(l-m)/2} \frac{(-1)^s (l-s)!}{s! \left[\frac{l+m}{2} - s\right]! \left[\frac{l-m}{2} - s\right]!} r_\perp^{l-2s}. \quad (1.7)$$

Задание для самостоятельного выполнения: записать радиальные полиномы $R_l^m(r_\perp)$ для m и l , принимающих значения от 1 до 6. При помощи системы компьютерной математики воспроизвести представленные на рис. 6 зависимости $Z_n^m(r_\perp, \gamma)$ для заданных значений m и l .

Физический смысл полиномов Цернике заключается в том, что они описывают aberrации волнового фронта — см. рис. 6. Нулевая мода, Z_0^0 , описывает смещение волнового фронта вдоль оси распространения излучения. Моды $Z_1^{\pm 1}$ описывают наклон волнового фронта по осям x и y . Остальные моды, представленные на рис. 6, описывают дефокусировку (в том числе трёх-, четырёх-, пяти- и шестилепестковую), а также астигматизм, сферическую aberrацию, кому третьего и пятого порядков.

ТЕМА 2. Мембранные и биморфные зеркала

Основные вопросы темы: ключевые характеристики деформируемых зеркал. Мембранное зеркало — корректор волнового фронта на основе тонкой отражающей пластины. Биморфное зеркало — корректор волнового фронта на основе двухслойной пьезоэлектрической пластины с отражающим покрытием.

Цель и задачи практического занятия состоят в ознакомлении студента с основными характеристиками деформируемых зеркал вне зависимости от их типа, со структурой и принципом действия мембранных и биморфных зеркал, с математическим описанием механических свойств мембранных зеркал, а также пьезоэлектрических и механических свойств биморфных зеркал.

Основной понятийно-терминологический аппарат: биморфное зеркало, изгибающий момент, мембранное зеркало, пьезоэлектрический эффект, функция Бесселя.

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, быть готовым назвать основные характеристики деформируемых зеркал и раскрыть и суть, быть способным описать общую структуру и принцип функционирования мембранных и биморфных зеркал, уметь построить математическую модель мембранного зеркала простейшей формы и оценивать на её основе прогиб зеркала по действию внешней силы, формулировать задачу оценки прогиба зеркала в удобной для компьютерных вычислений форме, оценить величину деформации биморфного зеркала в зависимости от приложенного электрического напряжения

Теоретический материал и примеры заданий по теме 2

2.1. Характеристики деформируемых зеркал

Деформируемое зеркало — один из основных компонентов АОС, роль которого заключается в исправлении искажений волнового фронта излучения, поступившего в оптическую систему. В зависимости от назначения оптической системы ДЗ могут представлять собой как миниатюрные электромагнитные устройства, размерами не превышающие нескольких миллиметров, так и большие установки

с отражающей поверхностью более чем метрового диаметра. Основные параметры, характеризующие ДЗ, следующие.

1) Размер зеркала.

Часто зеркала оптических систем являются осесимметричными, и их размер характеризуется диаметром D .

2) Количество приводов.

Изменение формы ДЗ осуществляется при помощи механических приводов, количество которых определяет число степеней свободы зеркала, что существенно влияет на его пространственное разрешение. Требуемое количество приводов N_{DOF} возрастает с увеличением диаметра зеркала D и уменьшается с увеличением длины волны излучения λ в соответствии с зависимостью

$$N_{\text{DOF}} \sim \frac{D^2}{\lambda^{12/5}}. \quad (2.1)$$

Данная оценка получена из следующих соображений. Деформируемое зеркало условно делится вдоль диаметра на домены, размер которых определяется параметром Фрайда r_0 . Для исправления аберрации в каждом домене требуется по крайней мере один привод. Таким образом, число приводов вдоль диаметра можно оценить отношением D/r_0 . В соответствии с (1.5) имеем $r_0 \sim \lambda^{6/5}$. Таким образом, для зеркала площадью $S \sim D^2$ для оценки числа приводов получаем выражение (2.1).

3) Запас хода.

Именно этот параметр в наибольшей степени определяет величину возможных деформаций зеркала. Запас хода не зависит от длины волны излучения: он определяется размерами зеркала, механизмом воздействия приводов на зеркало, а также граничными условиями.

4) Временная пропускная способность.

Искажение волнового фронта — это не стационарный, а динамический процесс. В простейшем случае можно рассматривать изменение во времени фазы излучения, распространяющегося через турбулентную атмосферу, как результат воздействия ветра. Для оценки временного поведения искажений используется время когерентности τ_0 , которое связано с r_0 выражением $\tau_0 = r_0/v$, где v — скорость ветра. Временная задержка между изменением величины флуктуаций и реакцией компонентов АОС на это изменение и определяет временную пропускную способность этих компонентов.

5) Гистерезис и линейность.

Для АОС важна предсказуемость отклика её компонентов на внешние воздействия. В идеальном случае сдвиг, который обеспечивает привод ДЗ, линейно зависит от входного сигнала. Нелинейные эффекты могут быть устранены путём калибровки ДЗ. Гистерезис имеет место для пьезоэлектрических приводов. Этот эффект подразумевает, что в ответ на одинаковые по модулю, но разные по знаку управляющие сигналы привод обеспечивает неодинаковое по модулю перемещение. Компенсация этого эффекта осуществляется введением в АОС дополнительной обратной связи.

Деформируемые зеркала подразделяются на сегментированные, представляющие собой решётку элементарных зеркал, которые образуют единую отражающую поверхность, и непрерывные зеркала, отражающая поверхность которых является непрерывной, и их форма изменяется за счёт локального искривления поверхности под внешним воздействием. К непрерывным относятся мембранные и биморфные зеркала.

2.2. Мембранное зеркало

Зеркало такого типа представляет собой тонкую пластину, передняя поверхность которой является отражающей. К задней поверхности пластины подведены механические приводы, а также датчики смещения: они позволяют изменять форму пластины, а также контролировать вносимые изменения. На рис. 7а схематично представлено размещение приводов и датчиков на мембранном зеркале.

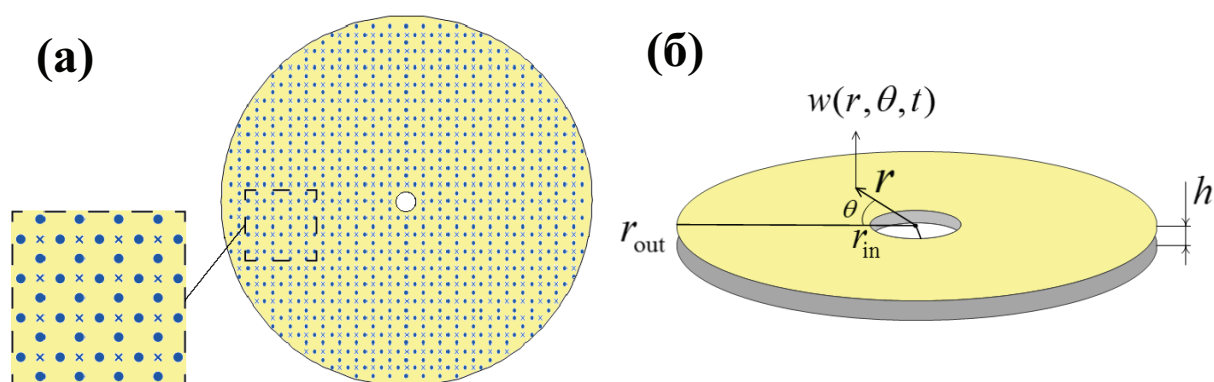


Рис. 7. (а) — Пример размещения приводов (×) и датчиков (•) на мембранном зеркале. (б) — Схематичное изображение мембранного зеркала.

Математическое моделирование динамики пластины мембранного зеркала может быть выполнено на основе часто используемого в теории упругости принципа

виртуальной работы, либо на основе законов Ньютона с использованием метода конечных элементов [3].

Аналитическая модель мембранного зеркала

Собственные моды зеркала

Рассмотрим мембранное зеркало как тонкую (толщины h), изотропную, однородную пластину в форме диска с круглым отверстием в центре — см. рис. 7б. Радиус диска составляет r_{out} , радиус отверстия — r_{in} . Изменение во времени изгиба пластины в направлении, перпендикулярном поверхности зеркала, в полярных координатах описывается следующим дифференциальным уравнением [4]:

$$D\nabla_r^2\nabla_r^2w + \rho h\frac{\partial^2w}{\partial t^2} = l, \quad (2.2)$$

где D — жёсткость при изгибе, $w \equiv w(r, \theta, t)$ — отклонение (прогиб) по отношению к недеформированной поверхности, (r, θ) — полярные координаты, t — время, ρ — плотность материала, $l \equiv l(r, \theta, t)$ — нагрузка на единицу площади; $\nabla_r^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}$. Деформации в плоскости пластины пренебрежимо малы и не учитываются в (2.2) (см. [5]).

Ввиду того что края пластины мембранного зеркала не свободны, а закреплены в оправе, введём граничные условия:

$$D \left(\frac{\partial^2 w}{\partial r^2} + \frac{\nu}{r} \frac{\partial w}{\partial r} + \frac{\nu}{r^2} \frac{\partial^2 w}{\partial \theta^2} \right) \Big|_{r=r_{\text{out}}} = 0, \quad (2.3a)$$

$$D \left(\frac{\partial}{\partial r} \nabla_r^2 w - (1 - \nu) \frac{\partial^2}{\partial \theta^2} \left(\frac{1}{r^2} \frac{\partial w}{\partial r} - \frac{w}{r^2} \right) \right) \Big|_{r=r_{\text{out}}} = 0, \quad (2.3б)$$

$$w|_{r=r_{\text{in}}} = 0, \quad (2.3в)$$

$$\frac{\partial w}{\partial r} \Big|_{r=r_{\text{in}}} = 0, \quad (2.3г)$$

где ν — коэффициент Пуассона.

Осуществим разделение переменных, представив w как $w(r, \theta, t) = W(r, \theta)e^{-i\omega t}$. В этом случае для свободной пластины (без нагрузки) можем переписать уравнение (2.2) для переменной W в виде

$$(\nabla_r^2 - \beta^2) W = 0, \quad (2.4)$$

где $\beta^4 = \frac{\rho h \omega^2}{D}$. Решение уравнения (2.4) в соответствии с [6] может быть представлено в виде $W_k = W_k^{(1)} + W_k^{(2)}$, где

$$W_k^{(1)} = \left[A_k^{(1)} J_k(\beta_k r) + A_k^{(3)} Y_k(\beta_k r) \right] \sin(k\theta) + \left[A_k^{(2)} J_k(\beta_k r) + A_k^{(4)} Y_k(\beta_k r) \right] \cos(k\theta), \quad (2.5a)$$

$$W_k^{(2)} = \left[B_k^{(1)} I_k(\beta_k r) + B_k^{(3)} K_k(\beta_k r) \right] \sin(k\theta) + \left[B_k^{(2)} I_k(\beta_k r) + B_k^{(4)} K_k(\beta_k r) \right] \cos(k\theta). \quad (2.5б)$$

В (2.5) J_k и Y_k – функции Бесселя первого и второго рода порядка k ; I_k и K_k – модифицированные функции Бесселя первого и второго рода порядка k . Модальный анализ, позволяющий определить W_k и ω_k , выполняется на основе граничных условий (2.3). С учётом (2.5) можем переписать граничные условия в виде

$$\mathfrak{M}_{\sin}(\beta_k) \begin{Bmatrix} A_k^{(1)} \\ A_k^{(3)} \\ B_k^{(1)} \\ B_k^{(3)} \end{Bmatrix} = 0, \quad \mathfrak{M}_{\cos}(\beta_k) \begin{Bmatrix} A_k^{(2)} \\ A_k^{(4)} \\ B_k^{(2)} \\ B_k^{(4)} \end{Bmatrix} = 0,$$

где $\mathfrak{M}_{\sin}(\beta_k)$ и $\mathfrak{M}_{\cos}(\beta_k)$ – матрицы размерности 4×4 , элементы которых определяются граничными условиями (2.3). Собственные частоты ω_k находятся из уравнения $\det(\mathfrak{M}(\beta_k)) = 0$.

Задание для самостоятельного выполнения: получить матрицы $\mathfrak{M}_{\sin}(\beta_k)$ и $\mathfrak{M}_{\cos}(\beta_k)$ на основе граничных условий (2.3) и решения (2.5).

Модовый состав в присутствии внешней нагрузки

Пусть к пластине зеркала в точке с координатами (R, Θ) приложена внешняя сила u . В этом случае правая часть уравнения (2.2) принимает вид

$$l = u \frac{\delta(r - R)\delta(\theta - \Theta)}{r} e^{-i\omega t}, \quad (2.6)$$

где $\delta(\eta)$ – дельта-функция Дирака, $\int_0^{2\pi} d\theta \int_0^\infty \frac{\delta(r-R)\delta(\theta-\Theta)}{r} r dr = 1$.

Прогиб зеркала w можно записать как суперпозицию собственных мод

$$w(r, \theta, t) = \sum_{k=1}^{\infty} c_k W_k(r, \theta) e^{-i\omega t}, \quad (2.7)$$

где c_k – весовые коэффициенты. Для $W_k(r, \theta)$ выполняется условие ортонормированности $\int_0^{2\pi} d\theta \int_{r_{\text{in}}}^{r_{\text{out}}} W_k(r, \theta) W_l(r, \theta) r dr = \delta_{kl} S$, где S – площадь пластины зеркала, δ_{kl} – символ Кронекера. Подставив (2.6) и (2.7) в (2.2), с учётом (2.4) получим уравнение

$$\rho h \sum_{k=1}^{\infty} (\omega_k^2 - \omega^2) c_k W_k = u \frac{\delta(r - R)\delta(\theta - \Theta)}{r}. \quad (2.8)$$

Домножим обе части уравнения (2.8) на W_l и проинтегрируем по всей поверхности пластины. Воспользовавшись условием ортонормированности W_k и свойством дельта-функции, мы получим выражение для весовых коэффициентов:

$$c_l = \frac{uW_l(R, \Theta)}{m(\omega_l^2 - \omega^2)}, \quad (2.9)$$

где $m = \rho Sh$ – масса пластины зеркала.

Поскольку в модели не рассматривались эффекты затухания, выражение (2.9) описывает физически нереализуемый предел $c_l \rightarrow \infty$ при $\omega = \pm\omega_l$. С учётом затухания выражение (2.9) переписывается в виде

$$c_l = \frac{uW_l(R, \Theta)}{m(\omega_l^2 - \omega^2 + 2i\zeta_l\omega\omega_l)},$$

где ζ_l – коэффициент модального демпфирования.

Задание для самостоятельного выполнения: на основе описанной процедуры из уравнения (2.2) получить промежуточное уравнение (2.8) и конечное выражение (2.9).

2.3. Биморфное зеркало

Схематично сечение биморфного зеркала показано на рис. 8а. Зеркало состоит из двух активных слоёв, по обе стороны которых прикреплены электроды. При помощи электродов в направлении, перпендикулярном плоскости зеркала, создаётся электрическое поле. Прикладывая к выбранным электродам определённое напряжение, можно вызывать удлинение или укорачивание активных слоёв в заданной области, что влечёт изгиб поверхности зеркала.

Два наиболее распространённых способа расположения электродов представлены на рис. 8б. Трапецеидальное расположение (слева) применяется для зеркал с круглой апертурой. Расположение в виде сот обеспечивает однородный изгиб по всей поверхности зеркала и может применяться для зеркал иной формы.

Основные характеристики биморфных зеркал следующие: толщина активного слоя h_{al} , диаметр зеркала D , пьезоэлектрический коэффициент d_{31} , максимальная напряжённость электрического поля E_{max} .

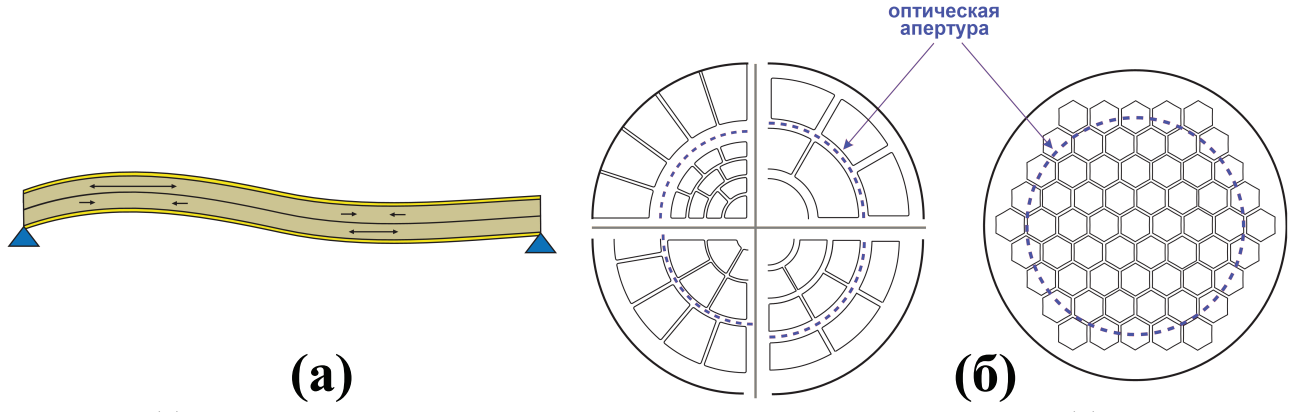


Рис. 8. (а) — Схематичное изображение сечения биморфного зеркала. (б) — Примеры расположения электродов на активном слое биморфного зеркала.

При прикладывании напряжения V к электродам в пьезоэлектрическом слое вдоль границ электрода возникает изгибающий момент (см. рис. 9а), который задаётся выражением [7]

$$M = \frac{Y_{al}}{1 - \nu_{al}} d_{31} V z_m, \quad (2.10)$$

где Y_{al} – модуль Юнга, ν_{al} – коэффициент Пуассона материала слоя, $z_m = h_{al}/2$ – расстояние между центральной плоскостью активного слоя и плоскостью раздела активных слоёв (см. рис. 9б).

Поскольку биморфное зеркало содержит два одинаковых активных слоя, которые подвергаются деформациям в противоположных направлениях (при растяжении одного слоя в плоскости зеркала второй слой сжимается), изгибающий момент для биморфного зеркала рассчитывается как удвоенное значение момента M , заданного выражением (2.10). Длину z_m можем переписать как $z_m = h/4$, где h – толщина зеркала. Таким образом, полный изгибающий момент для зеркала задаётся выражением

$$M_{mirr} = \frac{h Y_{mirr}}{2(1 - \nu_{mirr})} d_{31} V.$$

Величина деформации зависит от приложенного напряжения V и определяется выражением [8]

$$W(r) = \frac{3M_{mirr} D^2 (1 - \nu_{mirr})}{2h^3 Y_{mirr}} \left[1 - \left(\frac{2r}{D} \right)^2 \right] = \frac{3d_{31} V D^2}{4h^2} \left[1 - \left(\frac{2r}{D} \right)^2 \right].$$

Деформация имеет форму параболоида, радиус кривизны которого находится как

$$R = \left(\frac{\partial^2 W(r)}{\partial r^2} \right)^{-1} = \frac{h^2}{6d_{31} V}.$$

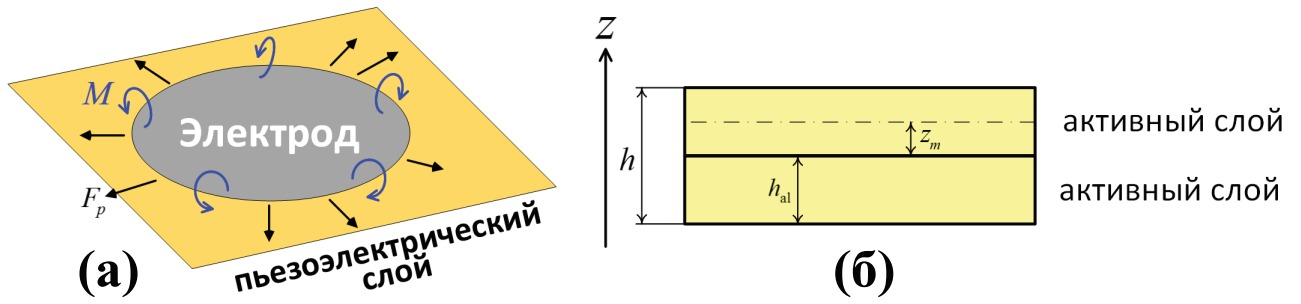


Рис. 9. Схематичное изображение (а) — распределения продольных сил F_p и изгибающих моментов M в пьезоэлектрическом слое и (б) — поперечного сечения биморфного зеркала.

Наибольшее отклонение от исходной формы наблюдается в центре параболоида, и его величина определяется выражением

$$W_{\max} = \frac{3d_{31}VD^2}{4h^2}.$$

ТЕМА 3. Алгоритмы фазового сопряжения

Основные вопросы темы: принцип фазового сопряжения. Измерение и реконструкция волнового фронта излучения.

Цель и задачи практического занятия состоят в ознакомлении студента с математическими и физическими основами принципа фазового сопряжения, а также методами восстановления формы волнового фронта на основе результатов измерения наклона датчиком Шака-Гартмана.

Основной понятийно-терминологический аппарат: датчик Шака-Гартмана, метод наименьших квадратов, метод минимизации отклонений, разностная схема, реконструкция волнового фронта.

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, понимать принцип фазового сопряжения и быть готовым объяснить его на примере прохождения излучения через возмущающую среду с последующим отражением от зеркала, иметь представление о физических ограничениях, препятствующих идеальному сопряжению фазы, быть знакомым с методологией реконструирования формы волнового фронта на основе данных о наклоне с датчика Шака-Гартмана, уметь формулировать задачу реконструкции формы в матричном виде.

Теоретический материал и примеры заданий по теме 3

Принцип фазового сопряжения лежит в основе большинства методов исправления aberrаций волнового фронта в современных АОС. Когда излучение с плоским волновым фронтом проходит через среду с aberrациями, его волновой фронт искажается. В общем случае напряжённость такого поля может быть представлена в виде $|E|e^{-i\varphi}$. Для того чтобы скорректировать волновой фронт, т. е. обеспечить для прошедшего излучения распределение фазы, аналогичное таковому для падающего излучения, необходимо выполнить операцию фазового сопряжения: с математической точки зрения это означает умножение поля с искажённым волновым фронтом $|E|e^{-i\varphi}$ на комплексно сопряжённую величину $|E|e^{+i\varphi}$.

Для понимания принципа фазового сопряжения аналогично [9] рассмотрим двухпроходную отражающую систему с aberrациями — см. рис. 10. Плоская

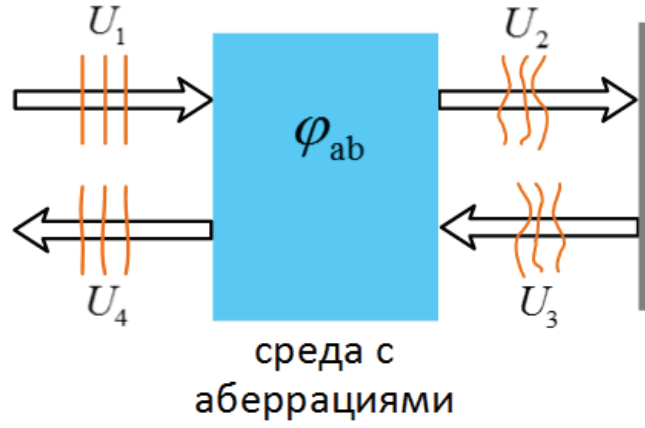


Рис. 10. К описанию принципа фазового сопряжения.

волна с постоянной амплитудой $|U_1|$ проходит через среду с абберациями, которая представляет собой плоскопараллельную пластину конечных размеров и в числе прочего выполняет роль диафрагмы площадью S_{ab} . На выходе из среды поле характеризуется напряжённостью $U_{ab} = |U_1|e^{ik\varphi_{ab}}$ с дополнительной фазовой добавкой. После выхода из среды излучение попадает на зеркало: напряжённость поля на зеркале задаётся как

$$U_2 = \frac{ike^{ikd}}{2\pi d} \iint_{S_{ab}} U_{ab} \exp\left(\frac{ikr_{a2}^2}{2d}\right) dx_a dy_a,$$

где $k = 2\pi/\lambda$ – волновое число, d – расстояние от пластины до зеркала; $r_{a2} = \sqrt{(x_a - x_m)^2 + (y_a - y_m)^2}$, (x_a, y_a) – поперечные координаты в плоскости пластины, (x_m, y_m) – поперечные координаты в плоскости зеркала.

Зеркало вносит коррекцию φ_c в волновой фронт отражённого излучения, которое описывается напряжённостью $U_2e^{ik\varphi_c}$. После отражения от зеркала излучение возвращается к абберационной пластине; на входе в пластину напряжённость поля задаётся как

$$U_3 = \frac{ike^{ikd}}{2\pi d} \iint_{S_{mirr}} U_2 e^{ik\varphi_c} \exp\left(\frac{ikr_{a2}^2}{2d}\right) dx_m dy_m.$$

После прохождения абберационной среды поле на выходе описывается сложным выражением

$$U_4 = e^{ik\varphi_{ab}} \frac{-k^2}{(2\pi d)^2} |U_1| \iint_{S_{mirr}} e^{ik\varphi_c} \exp\left(\frac{ikr_{a2}^2}{2d}\right) dx_m dy_m \iint_{S_{ab}} e^{ik\varphi_{ab}} \exp\left(\frac{ikr_{a2}^2}{2d}\right) dx_a dy_a. \quad (3.1)$$

В общем случае поле на входе в систему не равняется полю на выходе, т. е. $U_1 \neq U_4$. Рассмотрим теоретический предел, когда абберационная пластинка и зеркало обладают

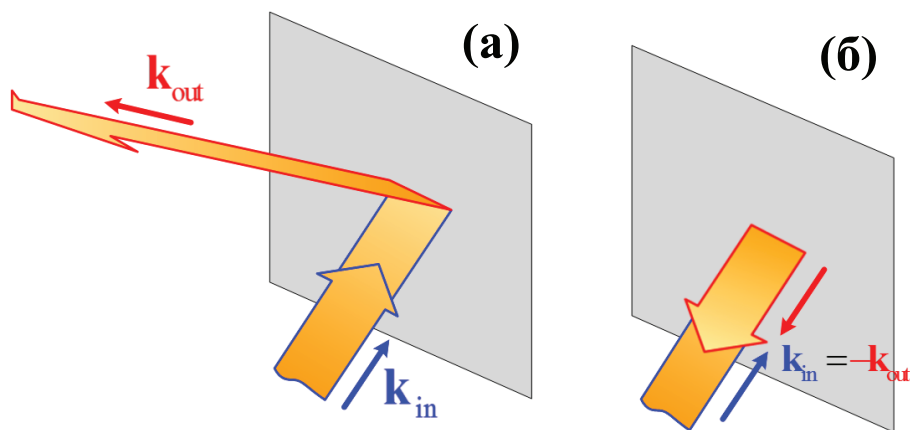


Рис. 11. Иллюстрация отражения излучения (а) — от обычного зеркала и (б) — в результате фазового сопряжения; \mathbf{k}_{in} и \mathbf{k}_{out} — волновые векторы падающего и отражённого излучения соответственно.

бесконечными размерами в плоскости, перпендикулярной оси распространения излучения. В этом случае дифракционные эффекты не вносят вклад в построение изображения. В данном пределе выражение (3.1) существенно упрощается: $U_4 = |U_1|e^{ik(2\varphi_{ab}+\varphi_c)}$. Очевидно, что при выполнении равенства $\varphi_c = -2\varphi_{ab}$ результирующее поле U_4 равняется исходному U_1 .

Важно отметить, что на практике идеальное сопряжение фазы нереализуемо. Причины тому следующие. Во-первых, все оптические элементы обладают конечными размерами, что подразумевает присутствие в системе дифракционных эффектов. Во-вторых, эффекты затухания, связанные с потерями энергии, приводят к уменьшению амплитуды отражённой волны.

В результате сопряжения фазы в отличие от обычного отражения изменяется как направление распространения излучения, так и фазовый множитель напряжённости электромагнитной волны. При простом отражении изменяется только знак компоненты поля, перпендикулярной к поверхности зеркала: как следствие, отражённое излучение распространяется под углом, равным углу падения. При фазовом сопряжении изменяются на противоположные знаки всех компонент поля, в результате отражённое излучение распространяется вдоль направления распространения падающего излучения — ср. схематичные изображения на рис. 11.

Также важно отметить, что метод фазового сопряжения позволяет в динамике отслеживать искажения волнового фронта и вносить требуемые коррекции.

Реконструкция волнового фронта

Для того, чтобы исправить искажения волнового фронта, необходимо сначала их измерить. Измерение искажений волнового фронта осуществляется при помощи ДВФ. Важно отметить тот факт, что фаза волны оптического диапазона не может быть измерена напрямую ввиду высокой частоты изменения напряжённости светового излучения. Поэтому в основе функционирования ДВФ лежат косвенные измерения. Так, например, в наиболее широко распространённом ДВФ — датчике Шака-Гартмана — измеряется пространственный градиент фазы светового пучка. Форма волнового фронта затем реконструируется численными методами на основе результатов измерений. Датчик Шака-Гартмана включает в себя два компонента. Первый компонент — это линзовый растр, представляющий собой двумерную матрицу микролинз. Второй компонент — это приёмник излучения. Линзы фокусируют излучение в плоскости приёмника, который фиксирует положение точек фокусов каждой микролинзы. По положению точек фокусов линз раstra оценивается градиент фазы падающего излучения. Схематично принцип действия датчика Шака-Гартмана представлен на рис. 12 а и б.

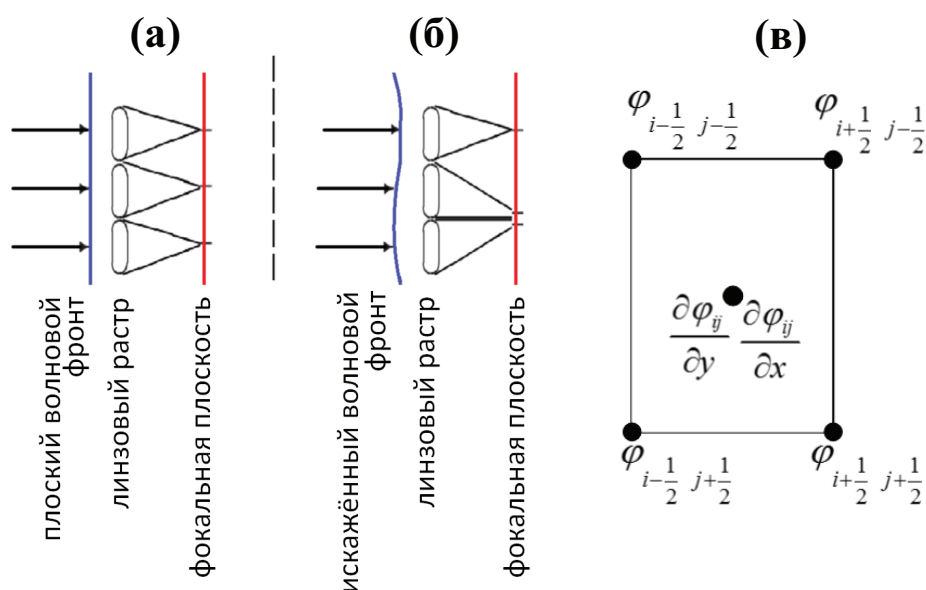


Рис. 12. (а) и (б) — Иллюстрация к описанию принципа действия ДВФ типа Шака-Гартмана: прохождение излучения с (а) — плоским волновым фронтом и (б) — искажённым волновым фронтом через линзовый растр. (в) — К описанию принципа реконструкции волнового фронта.

Методология реконструирования формы волнового фронта на основе данных с датчика Шака-Гартмана была предложена в работе [10]. Плоскость изображения разбивается на области одинаковой геометрии (чаще всего, квадратной формы), которые

соответствуют изображениям, полученным от микролинз растра — см. рис. 12в. На основе результатов измерения градиента фазы в центре каждой области вычисляется величина фазы на краю области. В дискретной системе производные по пространственным координатам задаются следующим образом:

$$\frac{\partial \varphi_{ij}}{\partial x} \approx \frac{1}{2l_x} \left[\left(\varphi_{i+\frac{1}{2} j-\frac{1}{2}} - \varphi_{i-\frac{1}{2} j-\frac{1}{2}} \right) + \left(\varphi_{i+\frac{1}{2} j+\frac{1}{2}} - \varphi_{i-\frac{1}{2} j+\frac{1}{2}} \right) \right], \quad (3.2a)$$

$$\frac{\partial \varphi_{ij}}{\partial y} \approx \frac{1}{2l_y} \left[\left(\varphi_{i-\frac{1}{2} j+\frac{1}{2}} - \varphi_{i-\frac{1}{2} j-\frac{1}{2}} \right) + \left(\varphi_{i+\frac{1}{2} j+\frac{1}{2}} - \varphi_{i+\frac{1}{2} j-\frac{1}{2}} \right) \right], \quad (3.2б)$$

где l_x и l_y — шаг сетки в соответствующих направлениях.

При размерах линзового растра $n \times n$ будем иметь вектор значений фазы \mathbf{w} длины n^2 . Правые части уравнений (3.2) можем переписать в матричной форме как $\mathbf{S}_x \cdot \mathbf{w}$ и $\mathbf{S}_y \cdot \mathbf{w}$. Зададим вектор \mathbf{d} , содержащий данные с результатами измерений градиента фазы; его размерность — $n^2 \times 2$. Он связан с вектором дискретных значений фазы \mathbf{w} линейным уравнением

$$\mathbf{d} = \mathbf{S} \cdot \mathbf{w} + \mathbf{n}, \quad (3.3)$$

где $\mathbf{S} = [\mathbf{S}_x, \mathbf{S}_y]^T$ — матрица отклика ДВФ, матричный оператор, описывающий линейные преобразования между значениями фазы и данным с ДВФ. Вектор \mathbf{n} описывает шумы.

Метод наименьших квадратов

Искомый вектор значений \mathbf{w} формально может быть задан выражением [11, 12]

$$\mathbf{w} = \arg \left[\min_w |\mathbf{S} \cdot \mathbf{w} - \mathbf{d}|^2 \right]. \quad (3.4)$$

Задача минимизации (3.4) сводится к определению значений фазы w , являющихся компонентами вектора \mathbf{w} , при которых производные $\partial |S w - d|^2 / \partial w$ обращаются в ноль.

В данном случае решения уравнение (3.3) будут эквивалентны решениям уравнения

$$\mathbf{S}^T \cdot \mathbf{S} \cdot \mathbf{w} = \mathbf{S}^T \cdot \mathbf{d}.$$

Метод минимизации отклонений

Оценка формы волнового фронта на основе данных, которые содержит вектор \mathbf{d} , представляет собой обратную задачу. В простейшем случае решение линейно зависит от входных данных \mathbf{d} , и мы можем записать вектор $\tilde{\mathbf{w}}$, описывающий реконструированный (восстановленный) волновой фронт, как [12, 13]

$$\tilde{\mathbf{w}} = \mathbf{R} \cdot \mathbf{d},$$

где \mathbf{R} – матрица восстановления. Очевидно, что различия между векторами \mathbf{w} и $\tilde{\mathbf{w}}$, описывающим исходный (истинный) и восстановленный волновые фронты соответственно, должны быть минимальны. Величина $\langle |\tilde{\mathbf{w}} - \mathbf{w}|^2 \rangle$ выступает в роли критерия качества восстановления формы волнового фронта. В соответствии с данным критерием матрица восстановления определяется выражением

$$\tilde{\mathbf{R}} = \arg \left[\min_R \langle |\mathbf{R} \cdot \mathbf{d} - \mathbf{w}|^2 \rangle \right].$$

Учитывая тот факт, что фаза \mathbf{w} и шумы \mathbf{n} являются независимыми, и их средние значения равны нулю, в соответствии с [13–15] можем записать $\tilde{\mathbf{R}}$ в виде

$$\tilde{\mathbf{R}} = (\mathbf{S}^T \cdot \mathbf{C}_n^{-1} \cdot \mathbf{S} + \mathbf{C}_w^{-1})^{-1} \cdot \mathbf{S}^T \cdot \mathbf{C}_n^{-1},$$

где $\mathbf{C}_n = \langle \mathbf{n} \cdot \mathbf{n}^T \rangle$ – ковариационная матрица шумов, $\mathbf{C}_w = \langle \mathbf{w} \cdot \mathbf{w}^T \rangle$ – ковариационная матрица фазы. Подействовав матричным оператором $\tilde{\mathbf{R}}$ на вектор данных \mathbf{d} , мы получим систему линейных уравнений, число которых соответствует числу измерений.

Учёт формы зрачка оптической системы

Как указано ранее, плоскость изображения для упрощения расчётов делится на квадратные ячейки. Однако зачастую оптические элементы обладают круглым сечением, что соответствующим образом сказывается на формируемом изображении. Учёт формы зрачка в модели восстановления волнового фронта осуществляется введением в уравнение (3.3) матрицы \mathbf{P} , элементы которой принимают значение $P_{i,j} = 1$, если i, j находятся в пределах зрачка, и $P_{i,j} = 0$, если i, j – за пределами зрачка. Уравнение (3.3) переписывается в виде

$$\mathbf{P} \cdot \mathbf{d} = \mathbf{P} \cdot \mathbf{S} \cdot \mathbf{w} + \mathbf{n}. \quad (3.5)$$

Поиск решения уравнений (3.5) осуществляется одним из уже описанных методов.

ТЕМА 4. Метод покоординатного спуска

Основные вопросы темы: общие сведения о методе покоординатного спуска и его суть. Отличи метода Гаусса-Зейделя от метода Якоби. Метод Гаусса-Зейделя для восстановления формы волнового фронта излучения

Цель и задачи практического занятия состоят в ознакомлении студента с подходами к восстановлению формы волнового фронта излучения, с методом покоординатного спуска как одним из методов модальной реконструкции волнового фронта.

Основной понятийно-терминологический аппарат: зональная реконструкция, метод Гаусса-Зейделя, метод последовательной верхней релаксации, метод Якоби, модальная реконструкция, .

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, осознавать различия между модальной и зональной реконструкцией волнового фронта, понимать принцип метода покоординатного спуска, его преимущество перед методом Якоби, быть готовым продемонстрировать суть метода покоординатного спуска на примере решения системы линейных уравнений, уметь применить методы Якоби, покоординатного спуска и последовательной верхней релаксации для решения задачи восстановления формы волнового фронта и быть готовым выполнить сравнение результатов решения задачи.

Теоретический материал и примеры заданий по теме 4

Можно выделить два подхода к восстановлению (реконструкции) формы волнового фронта на основе результатов измерений.

– Модальная реконструкция. Этот подход основан на разложении волнового фронта в базисе ортонормированных функций. Весовые коэффициенты при членах разложения оцениваются по результатам измерений градиента волнового фронта при помощи ДВФ. Примером разложения может служить разложение по полиномам Цернике.

– Зональная реконструкция (прямое численное моделирование). Значение фазы оценивается непосредственно на основе результатов измерений градиента. Данный подход расценивается как более точный, поскольку не требует приближённого

разложения по базисным функциям. Для реализации данного подхода имеется набор итерационных алгоритмов.

Рассмотрим метод покоординатного спуска (Гаусса-Зейделя) для восстановления формы волнового фронта на основе дискретных данных измерения градиента фазы.

4.1. Суть метода покоординатного спуска

Для начала проиллюстрируем суть метода на примере решение системы трёх линейных уравнений:

$$A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = B_1, \quad (4.1a)$$

$$A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = B_2, \quad (4.1б)$$

$$A_{31}x_1 + A_{32}x_2 + A_{33}x_3 = B_3. \quad (4.1в)$$

Положим коэффициенты A_{ij} отличными от нуля. Выразим переменные $x_{1,2,3}$ из первого, второго и третьего уравнений соответственно:

$$x_1 = \frac{1}{A_{11}}(B_1 - A_{12}x_2 - A_{13}x_3), \quad (4.2a)$$

$$x_2 = \frac{1}{A_{22}}(B_2 - A_{21}x_1 - A_{23}x_3), \quad (4.2б)$$

$$x_3 = \frac{1}{A_{33}}(B_3 - A_{31}x_1 - A_{32}x_2). \quad (4.2в)$$

Поиск решения уравнений (4.1) осуществляется в несколько итераций. В начале задаётся нулевое приближение переменных $x_{1,2,3}$: $x_1 = x_1^{(0)}$, $x_2 = x_2^{(0)}$, $x_3 = x_3^{(0)}$. На первом шаге (итерации) в соответствии с выражением (4.2a) находим первое приближение переменной x_1 :

$$x_1^{(1)} = \frac{1}{A_{11}}(B_1 - A_{12}x_2^{(0)} - A_{13}x_3^{(0)}). \quad (4.3a)$$

Далее в соответствии с (4.2б), пользуясь результатом вычисления (4.3a), находим первое приближение переменной x_2 :

$$x_2^{(1)} = \frac{1}{A_{22}}(B_2 - A_{21}x_1^{(1)} - A_{23}x_3^{(0)}). \quad (4.3б)$$

Очевидно, что первое приближение переменной x_3 вычисляется в соответствии с выражением (4.2в) и с учётом результатов вычисления (4.3a,б):

$$x_3^{(1)} = \frac{1}{A_{33}}(B_3 - A_{31}x_1^{(1)} - A_{32}x_2^{(1)}). \quad (4.3в)$$

На данном этапе можем показать отличие метода Гаусса-Зейделя от метода Якоби. Оно заключается в том, что в методе Якоби результаты вычисления значений переменных используются только в следующей итерации. Применительно к рассматриваемому примеру это означает, что первое приближение переменных $x_{2,3}$ будет задаваться выражениями $x_2^{(1)} = \frac{1}{A_{22}}(B_2 - A_{21}x_1^{(0)} - A_{23}x_3^{(0)})$ и $x_3^{(1)} = \frac{1}{A_{33}}(B_3 - A_{31}x_1^{(0)} - A_{32}x_2^{(0)})$ — ср. с (4.3в,г).

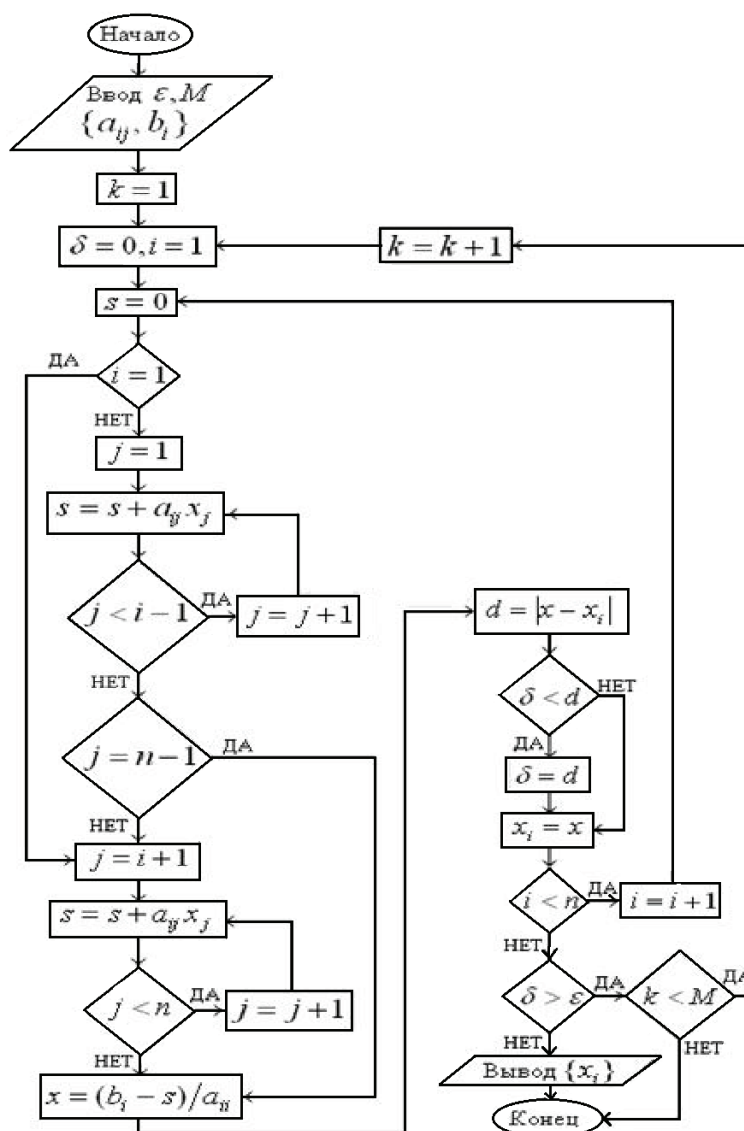


Рис. 13. Блок-схема метода Гаусса-Зейделя.

Аналогичным образом проводятся вторая и последующие итерации, в результате которых находятся второе, $x_{1,2,3}^{(2)}$, и последующие приближения к решению. Для k -го приближения можем записать выражения

$$\begin{aligned}
x_1^{(k)} &= \frac{1}{A_{11}}(B_1 - A_{12}x_2^{(k-1)} - A_{13}x_3^{(k-1)}), \\
x_2^{(k)} &= \frac{1}{A_{22}}(B_2 - A_{21}x_1^{(k)} - A_{23}x_3^{(k-1)}), \\
x_3^{(k)} &= \frac{1}{A_{33}}(B_3 - A_{31}x_1^{(k)} - A_{32}x_2^{(k)}).
\end{aligned}$$

Повторение итераций осуществляется до тех пор, пока приближения двух соседних итераций, $x_{1,2,3}^{(k)}$ и $x_{1,2,3}^{(k-1)}$, не станут близки друг к другу с точностью до требуемой погрешности ϵ : $\Delta = \max_j |x_j^{(k)} - x_j^{(k-1)}| < \epsilon$. На рис. 13 приведена блок-схема алгоритма решения системы линейных уравнений методом Гаусса-Зейделя. Исходные данные: коэффициенты A_{ij} при искомым переменных, величина погрешность ϵ , максимально допустимое число итераций M , нулевые приближения переменных $x_i^{(0)}$. Используемые обозначения: k – номер итерации; i и j нумеруют искомые переменные и коэффициенты при них. Окончание цикла наступает при выполнении условия $\Delta < \epsilon$, либо при $k = M$.

Задание для самостоятельного выполнения: при помощи метода Гаусса-Зейделя найти с точностью $\epsilon = 10^{-3}$ решение системы уравнений

$$\begin{aligned}
20.1w + 1.3x + 3.3y + 0.5z &= 22.4, \\
2.1w + 18.4x + 1.6y + 3.0z &= 26.6, \\
2.2w + 1.4x + 22.7y + 1.3z &= 30.1, \\
0.5w + 2.8x + 1.8y + 29.5z &= 29.9.
\end{aligned}$$

Найти решение указанной системы методом Якобы с такой же точностью. Указать, для какого метода требуется большее число итераций.

4.2. Реконструкция волнового фронта на основе метода покоординатного спуска

Рассмотрим задачу восстановления формы волнового фронта на основе результатов измерений градиента фазы при помощи ДВФ Шака-Гартмана [16]. Рассмотрим конфигурацию, отличную от приведённой в указаниях к предыдущей теме. Плоскость

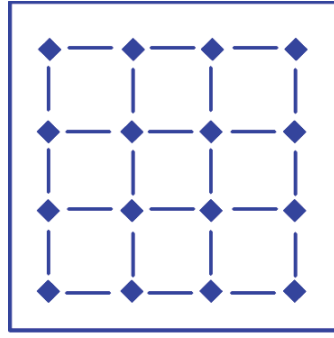


Рис. 14. Сетка измерения наклонов волнового фронта и расчёта фазы. Горизонтальные и вертикальные линии обозначают позиции, в которых осуществляется измерение наклона волнового фронта в направлениях x и y соответственно. Маркеры \blacklozenge обозначают точки, в которых осуществляется расчёт фазы.

изображения также разбивается на квадратные ячейки — см. рис. 14. Однако значение фазы вычисляется в тех же точках, в которых измеряется градиент.

Наклоны волнового фронта в направлениях x и y определяются разностными производными

$$s_x^{(ij)} = \frac{\partial \phi_{ij}}{\partial x} \simeq \frac{\varphi_{i+1j} - \varphi_{ij}}{\ell}, \quad i = 1, \dots, N-1, j = 1, \dots, N, \quad (4.4a)$$

$$s_y^{(ij)} = \frac{\partial \phi_{ij}}{\partial y} \simeq \frac{\varphi_{ij+1} - \varphi_{ij}}{\ell}, \quad i = 1, \dots, N, j = 1, \dots, N-1, \quad (4.4b)$$

где ℓ – шаг сетки, N – число точек, в которых рассчитывается фаза. Значение фазы в точке (i, j) связано со значениями фазы и наклоном волнового фронта в четырёх соседних точках с номерами $(i \pm 1, j)$ и $(i, j \pm 1)$. Используя соотношения (4.4), можем записать для фазы в точке (i, j) следующее выражение:

$$\varphi_{ij} = \frac{1}{g_{ij}} [\ell (s_x^{(i-1j)} - s_x^{(ij)} + s_y^{(ij-1)} - s_y^{(ij)}) + (\varphi_{i+1j} + \varphi_{i-1j} + \varphi_{ij+1} + \varphi_{ij-1})], \quad (4.5)$$

где $g_{ij} = 2$ при $i = 1$ или N и $j = 1$ или N ; $g_{ij} = 3$ при $i = 1$ или N , $2 \leq j \leq N-1$ либо при $j = 1$ или N , $2 \leq i \leq N-1$; в остальных случаях $g_{ij} = 4$.

Можем переписать (4.5) в упрощённом виде

$$\varphi_{ij} = \tilde{\varphi}_{ij} + \frac{b_{ij}}{g_{ij}}, \quad (4.6)$$

где введены обозначения $\tilde{\varphi}_{ij} = \frac{1}{g_{ij}} (\varphi_{i+1j} + \varphi_{i-1j} + \varphi_{ij+1} + \varphi_{ij-1})$, $b_{ij} = \ell (s_x^{(i-1j)} - s_x^{(ij)} + s_y^{(ij-1)} - s_y^{(ij)})$. Решение уравнения (4.6) осуществляется итерационным методом. Для k -й итерации решение принимает вид

$$\varphi_{ij}^{(k)} = \tilde{\varphi}_{ij}^{(k-1)} + \frac{b_{ij}}{g_{ij}}. \quad (4.7)$$

Аналогично сказанному ранее, если результаты вычислений, содержащиеся в левой части уравнения (4.7), не используются вплоть до следующей итерации, такой метод решения — метод Якоби. Если результаты в левой части используются в ходе текущей итерации для нахождения значений других неизвестных — это метод Гаусса-Зейделя.

Выбор нулевого приближения значений φ_{ij} может осуществляться случайным образом либо на основе априорных данных. Следует отметить, что $\varphi_{ij}^{(0)}$ могут принимать и нулевые значения.

Развитием метода Гаусса-Зейделя стал метод последовательной верхней релаксации (ПВР; англ.: successive over-relaxation). Для этого метода решение на k -й итерации записывается в виде

$$\varphi_{ij}^{(k)} = \varphi_{ij}^{(k-1)} + \omega \left[\tilde{\varphi}_{ij}^{(k-1)} + \frac{b_{ij}}{g_{ij}} - \varphi_{ij}^{(k-1)} \right],$$

где $\omega = 2 / (1 + \sin[\pi / (N + 1)])$. Повторение итераций осуществляется до тех пор, пока выражение в скобках не станет меньше заданной погрешности ϵ .

Задание для самостоятельного выполнения: на основе приведённых данных о наклоне волнового фронта восстановить форму волнового фронта а) методом Якоби, б) методом покоординатного спуска, в) методом ПВР. Наклоны в матричном виде задаются следующим образом:

$$s^x = \begin{pmatrix} -7.842 & -3.921 & 3.921 & 7.842 \\ -7.842 & -3.921 & 3.921 & 7.842 \\ -7.842 & -3.921 & 3.921 & 7.842 \\ -7.842 & -3.921 & 3.921 & 7.842 \end{pmatrix}, \quad s^y = \begin{pmatrix} -7.842 & -7.842 & -7.842 & -7.842 \\ -3.921 & -3.921 & -3.921 & -3.921 \\ 3.921 & 3.921 & 3.921 & 3.921 \\ 7.842 & 7.842 & 7.842 & 7.842 \end{pmatrix}.$$

Поиск решения осуществлять с точностью $\epsilon = 10^{-4}$. К выбору начальных значений фазы (в нулевом приближении) требований не предъявляется.

Задание для самостоятельного выполнения: записать решение на k -й итерации для конфигурации, представленной на рис. 12в.

Решение:

$$\varphi_{ij}^{(k)} = \frac{1}{g_{ij}} \left[\frac{\ell}{2} (s_x^{(i+1j)} - s_x^{(i-1j)} + s_y^{(ij+1)} - s_y^{(ij-1)}) + (\varphi_{i+1j}^{(k-1)} + \varphi_{i-1j}^{(k-1)} + \varphi_{ij+1}^{(k-1)} + \varphi_{ij-1}^{(k-1)}) \right],$$
$$g_{ij} = \begin{cases} 2 & \text{при } i = 1, N \text{ и } j = 1, N; \\ 3 & \text{при } i = 1, N \text{ и } 2 \leq j \leq N - 1 \text{ либо при } j = 1, N \text{ и } 2 \leq i \leq N - 1; \\ 4 & \text{в остальных случаях.} \end{cases}$$

ТЕМА 5. Генетический и гибридный алгоритмы определения управляющих напряжений

Основные вопросы темы: генетический алгоритм — стохастический метод решения задач оптимизации. Применение генетического алгоритма в задачах адаптивной оптики.

Цель и задачи практического занятия состоят в ознакомлении студента с сутью генетического алгоритма, его основными этапами и операциями, с гибридным алгоритмом как модификацией генетического алгоритма, с возможностью использования генетического алгоритма для целей оптимизации расположения механических приводов деформируемого зеркала и величины прикладываемых к ним электрических напряжений.

Основной понятийно-терминологический аппарат: генетический алгоритм, гибридный алгоритм, функция приспособленности.

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, понимать суть генетического алгоритма уметь оперировать специфической терминологией, используемой при описании алгоритма, иметь представление об основных эволюционных операторах (отбор, скрещивание, мутация), основных этапах алгоритма, условиях завершения выполнения алгоритма, быть готовым продемонстрировать умение пользования алгоритмом на примере простейших задач оптимизации, осознавать возможности использования алгоритма в задачах адаптивной оптики, в частности в задаче проектирования деформируемых зеркал.

Теоретический материал и примеры заданий по теме 5

5.1. Описание генетического алгоритма

Генетический алгоритм — это один из стохастических методов решения проблемы оптимизации; он входит в более крупное семейство эволюционных алгоритмов. Генетический алгоритм работает с набором индивидуальных объектов, представляющих возможные решения конкретной задачи. В ходе выполнения алгоритма производится отбор объектов (решений), наилучшим образом удовлетворяющих заданным

критериям. На основе отобранных объектов генерируется новое поколение объектов, наследующих их свойства.

Независимо от области применения алгоритма используется терминология, принятая в генетике [17]. Так генетический алгоритм оперирует хромосомами и генами. Хромосомой называют одно из возможных решений поставленной проблемы; в математическом смысле хромосома представляет собой вектор переменных, которые удовлетворяют заданным критериям. Каждая переменная кодируется отдельным геном. В зависимости от решаемой задачи, от параметров, которые требуется оптимизировать, гены могут представлять собой числа, булевы значения, строки переменных, а также любые комбинации описанных вариантов. Однако чаще всего гены задаются в виде битов, а хромосомы — в виде битовых строк. Набор различных хромосом (возможных решений) образует популяцию. Путём воздействия на хромосомы эволюционными операторами, такими как отбор, скрещивание и мутация, генерируется новое поколение.

Отбор.

На данном этапе отбираются наиболее пригодные хромосомы, которые участвуют в дальнейших вычислениях. Пригодность оценивается по соответствию хромосом некоторому критерию. Для объективной оценки пригодности хромосом используется функция приспособленности (англ.: fitness function). Чем ближе значение функции приспособленности для данной хромосомы к целевому значению, тем более пригодна данная хромосома для дальнейших вычислений. Примером функции приспособленности может служить сумма среднеквадратичных отклонений результатов измерений от требуемого значения. В этом случае отбор заключается в минимизации функции приспособленности для набора хромосом.

Скрещивание.

На этом этапе оставшиеся после этапа отбора хромосомы разбиваются на пары; затем из фрагментов родительских хромосом создаются хромосомы следующего поколения. Каждая хромосома может неоднократно участвовать в скрещивании с другими хромосомами. Математически этот процесс можно записать следующим образом:

$$C_1 = M_1[P_1] + M_2[P_2],$$

$$C_2 = M_2[P_1] + M_1[P_2],$$

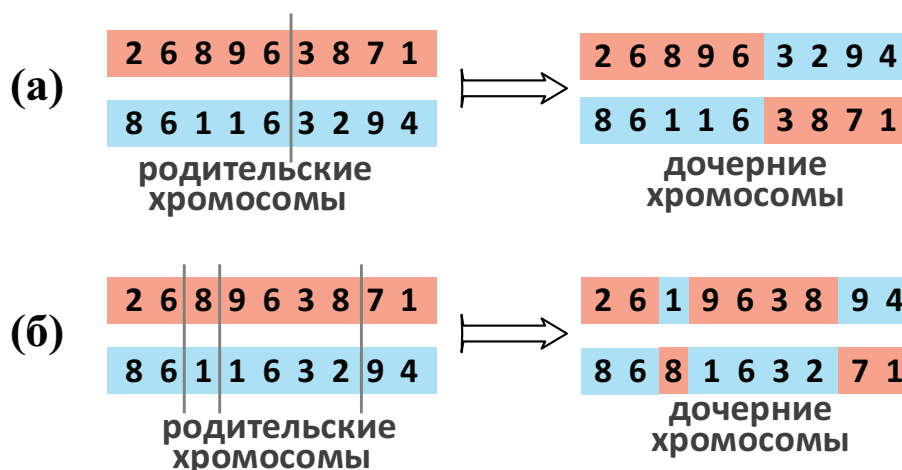


Рис. 15. Примеры скрещивания хромосом (а) — с одной и (б) — тремя точками скрещивания.

где $P_{1,2}$ – родительские хромосомы, $C_{1,2}$ – хромосомы следующего поколения; операторы $M_{1,2}$ задают шаблон, по которому разделяются гены родительских хромосом. На рис. 15 приведены два примера скрещивания хромосом — с одной (а) и несколькими (б) точками скрещивания.

Мутации.

Оператор мутаций случайным образом заменяет один или несколько генов в хромосоме. Вероятность мутаций выбирается достаточно малой, обычно это величина порядка 0.001. Пример мутации приведён на рис. 16.

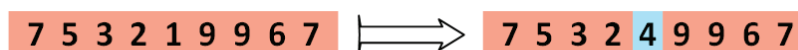


Рис. 16. Пример мутации хромосомы.

Блок-схема простейшего генетического алгоритма представлена на рис. 17. Алгоритм работает следующим образом.

1. Генерация начальной популяции из N объектов. Генерация может осуществляться как случайным образом, так и с учётом априорных данных. Хромосомы строятся путём выбора генов из разрешённого набора.

2. Вычисление функции приспособленности f_j для каждой хромосомы j в популяции.

3. Проверка на необходимость завершения алгоритма.

Алгоритм завершается, если выполняется одно из следующих условий.

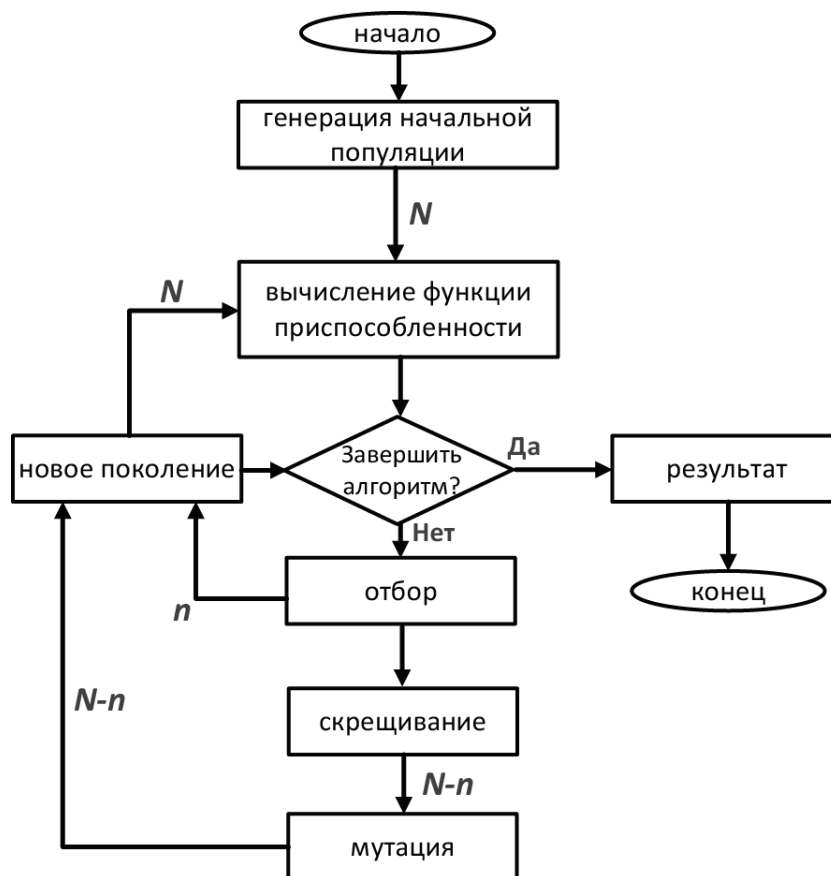


Рис. 17. Блок-схема генетического алгоритма.

– Значение функции приспособленности одной из хромосом находится в заданном диапазоне. Данный исход выполнения алгоритма означает, что решение поставленной задачи найдено.

– Достигнуто максимально разрешённое число итераций. Независимо от того, достигла ли функция приспособленности требуемого значения, после выполнения заданного числа итераций выполнение алгоритма прерывается.

– Не произошло улучшения значения функции приспособленности на протяжении заданного числа итераций.

4. Отбор. Из числа хромосом текущей популяции отбираются хромосомы, которые будут участвовать в дальнейших вычислениях. Отбор осуществляется на основе значения функции приспособленности для каждой хромосомы. На данном этапе также могут быть отобраны n наиболее приспособленных хромосом, которые, минуя последующие этапы, сразу попадут в новое поколение.

5. Скрещивание. На основе отобранных представителей поколения родительских хромосом генерируются новые дочерние хромосомы. Поскольку цель скрещивания состоит

в получении дочерних объектов, наследующих наилучшее сочетание генов родительских объектов, из дочерних хромосом выбирается $N - n$ наиболее приспособленных.

6. Мутация. Часть генов нового поколения объектов случайным образом изменяется. Даже если предпочтительный для решаемой задачи ген не присутствовал в родительских хромосомах, операция мутации позволяет с некоторой вероятностью внедрить этот ген в объекты дочернего поколения.

7. Замена предыдущего поколения новым. Новое поколение представляет собой n объектов, напрямую отобранных на этапе 4, и $N - n$ объектов, являющихся дочерними предыдущему поколению.

8. Переход к этапу 2.

Рассмотрим выполнение генетического алгоритма на следующем примере.

Задание: на основе генетического алгоритма найти набор целых положительных значений коэффициентов A, B, C , при которых удовлетворяется равенство $A+B+2C = 70$.

Выполнение осуществляется на основе предложенного алгоритма. В соответствии с алгоритмом в начале сгенерируем первое поколение объектов. Число объектов выберем $N = 9$. Выбор генов (которые описывают значения коэффициентов A, B, C) будем осуществлять случайным образом, но с учётом того факта, что они принимают значения, не превышающие 70.

1) Итак, первое поколение объектов, представленное в виде $\{A, B, C\}$, следующее:

- | | | |
|-----------------|-----------------|-----------------|
| 1) {53, 36, 36} | 2) {24, 4, 14} | 3) {36, 19, 8} |
| 4) {36, 50, 2} | 5) {63, 67, 61} | 6) {25, 68, 45} |
| 7) {42, 64, 22} | 8) {69, 21, 14} | 9) {38, 49, 16} |

2) Функцию приспособленности зададим в виде

$$f(A, B, C) = |A + B + C - 70|;$$

чем меньше значение $f(A, B, C)$ для заданной хромосомы $\{A, B, C\}$, тем более приспособленной она является. Значения функций приспособленности для хромосом первого поколения следующие:

1) 91	2) 14	3) 1
4) 20	5) 182	6) 113
7) 80	8) 48	9) 49

3) Будем считать алгоритм завершённым, когда для одной из хромосом функция приспособленности примет значение 0. Для всех объектов первого поколения f отлична от нуля.

4) Проведём отбор наиболее приспособленных объектов. Хромосома №3 обладает наилучшей приспособленностью, поскольку значение $f(A_3, B_3, C_3) = 1$ наименьшее из всех представленных; эта хромосома напрямую попадает в новое поколение объектов.

Исключим из дальнейших вычислений хромосомы, функции f которых особенно велики. В качестве критерия исключения зададим условие $f(A, B, C) > 70$. Таким образом, в следующем этапе скрещивания будут принимать участие хромосомы с номерами 2, 4, 8 и 9.

5) Составим восемь пар родительских хромосом, каждая из которых будет давать по одной дочерней хромосоме. Эти пары следующие:

1) 2 – 4	2) 2 – 8	3) 4 – 9	4) 8 – 9
5) 4 – 2	6) 9 – 2	7) 9 – 8	8) 8 – 4

В данном случае все отобранные хромосомы выступают в роли родительских одинаково число раз (по четыре раза). Однако возможно увеличивать или уменьшать шанс хромосомы выступить в роли родителя в зависимости от степени приспособленности, которая определяется функцией f .

В результате скрещивания дочерняя хромосома берёт часть генов от каждого из родителей, при этом наследуемые фрагменты при каждом скрещивании могут быть разными. Будем отделять наследуемые фрагменты символом «|»: от стоящего первым родителя наследуются гены слева от разделителя «|», а от второго родителя — стоящие справа от «|». Итак в результате скрещивания получены следующие дочерние хромосомы:

1) $\{24 4, 14\} + \{36 50, 2\} \rightarrow \{24, 50, 2\}$	2) $\{24, 4 14\} + \{69, 21 14\} \rightarrow \{24, 4, 14\}$
3) $\{36 50, 2\} + \{38 49, 16\} \rightarrow \{36, 49, 16\}$	4) $\{69 21, 14\} + \{38 49, 16\} \rightarrow \{69, 49, 16\}$
5) $\{36, 50 2\} + \{24, 4 14\} \rightarrow \{36, 50, 14\}$	6) $\{38 49, 16\} + \{24 4, 14\} \rightarrow \{38, 4, 14\}$
7) $\{38 49, 16\} + \{69 21, 14\} \rightarrow \{38, 21, 14\}$	8) $\{69, 21 14\} + \{36, 50 2\} \rightarrow \{69, 21, 2\}$

6) После применения оператора мутации одна из дочерних хромосом подверглась следующему изменению:

$$\{36, 49, 16\} \rightarrow \{36, 54, 16\},$$

т. е. ген B случайным образом изменил значение с 49 на 54.

7) С учётом пунктов 4–6 сего перечня составим новое поколение хромосом:

- 1) $\{36, 19, 8\}$ 2) $\{24, 50, 2\}$ 3) $\{24, 4, 14\}$
 4) $\{36, 54, 16\}$ 5) $\{69, 49, 16\}$ 6) $\{36, 50, 14\}$
 7) $\{38, 4, 14\}$ 8) $\{38, 21, 14\}$ 9) $\{69, 21, 2\}$

8) Найдём значения функции приспособленности для второго поколения хромосом:

- 1) 1 2) 8 3) 14
 4) 52 5) 80 6) 44
 7) 0 8) 17 9) 24

9) Как мы видим, значение функции f для хромосомы №7 второго поколения равно 0. Это означает, что набор значений $\{A, B, C\} = \{38, 4, 14\}$ удовлетворяет заданному равенству и является решением поставленной задачи. На данном этапе выполнение алгоритма прекращается.

Задание для самостоятельного выполнения:

Задан треугольник $A_3B_3C_3$ (см. рис. 18а), длины его сторон следующие: $A_3B_3 = 2$ см, $B_3C_3 = 2$ см, $A_3C_3 = 3$ см.

У трапеции $A_4B_4C_4D_4$ заданы внутренние углы: $\angle B_4A_4D_4 = 30^\circ$, $\angle A_4B_4C_4 = 150^\circ$, $\angle B_4C_4D_4 = 120^\circ$, $\angle A_4D_4C_4 = 60^\circ$ (см. рис. 18б). Используя генетический алгоритм, подобрать длины сторон трапеции так, чтобы площадь трапеции отличалась от площади заданного треугольника не более чем на 3%.

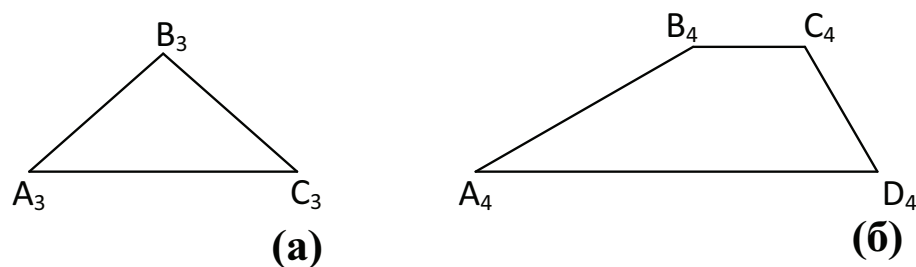


Рис. 18. Для задания.

При выполнении учесть следующее. Площадь треугольника определяется выражением $S_{\Delta} = \sqrt{p(p-a)(p-b)(p-c)}$, где $p = \frac{1}{2}(a+b+c)$ – полупериметр треугольника, a, b, c – длины сторон треугольника. Площадь трапеции определяется выражением $S_{tr} = \frac{a+b}{4|a-b|} \sqrt{(a+c+d-b)(a+d-b-c)(a+c-b-d)(b+c+d-a)}$, где a и b – основания, c и d – боковые стороны трапеции. При расчётах необходимо предусмотреть тот факт, что при заданных углах длины боковых сторон, а также меньшего основания ограничены длиной большего основания.

5.2. Проблема оптимизации для приводов деформируемых зеркал

Изменение формы деформируемых зеркал осуществляется при помощи механических приводов. Наилучшую точность позиционирования и скорость работы обеспечивают приводы, работающие на основе пьезоэлектрического эффекта [18]. Важной и необходимой при разработке и конструировании КВФ является задача оптимизации таких параметров, как число приводов, положение, размеры, разность потенциалов. Широкое применение для решения задач оптимизации, в том числе применительно к АОС, нашёл генетический алгоритм [19], а также его модификации, повышающие скорость вычислений, корректность и обоснованность результатов вычислений.

Конкретная проблема в рамках общей задачи оптимизации, которая ставится перед разработчиком – при заданном количестве пьезоэлектрических приводов с известными геометрическими и электрическими свойствами определить наилучшее их взаиморасположение, а также прикладываемое электрическое напряжение, при которых обеспечивается наилучшая для коррекции искажений волнового фронта форма КВФ.

Объективно оценить сложность проблемы позволяет параметр ${}^{n_p}C_{n_a}$, который задаётся выражением

$${}^{n_p}C_{n_a} = \frac{n_p!}{n_a!(n_p - n_a)!}$$

и показывает возможное число комбинаций взаимного расположения приводов. В представленном выражении n_a – количество пьезоприводов, n_p – число возможных местоположений приводов.

В работе [20] рассматривалась проблема размещения $n_a = 30$ пьезоприводов на задней поверхности зеркала гексагональной формы. Схематично возможные

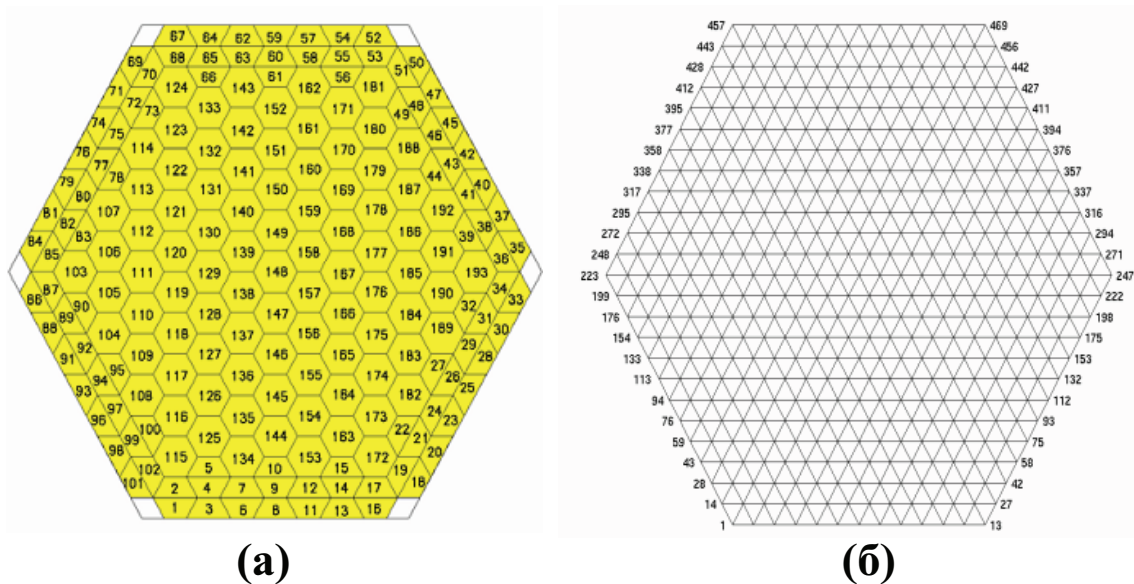


Рис. 19. Иллюстрация к описанию проблемы оптимизации для приводов деформируемых зеркал [20]: (а) — возможные местоположения приводов зеркала, (б) — разбиение поверхности зеркала на фрагменты треугольной формы.

местоположения приводов обозначены на рис. 19а; число местоположений достигало $n_p = 193$. В результате показано, что количество возможных комбинаций расположения приводов для описанной задачи составляет ${}^{193}C_{30} \approx 1.29 \times 10^{35}$. Очевидно, что при таком числе комбинаций выбор наилучшей комбинации невозможен прямым перебором, и возникает необходимость использования методов оптимизации.

Если для решения задачи оптимизации генетический алгоритм работает в паре с иным алгоритмом, такой алгоритм называется гибридным. Для решения поставленной задачи вначале осуществляется моделирование зеркала методом конечных элементов. Поверхность зеркала условно разбивается на фрагменты треугольной формы — см. рис. 19. Смещение каждого фрагмента, u_i , в результате воздействия пьезопроводов полагается достаточно малым, таким, что его можно представить в виде

$$u_i = \sum_{j=1}^n \alpha_{ij} V_j,$$

где V_j — напряжение, прикладываемое к j -му пьезопроводу, α_{ij} — коэффициент воздействия, который определяет величину смещения i -го фрагмента в результате приложения к j -му проводу единичного напряжения (1 В). Матрица коэффициентов воздействия α имеет размерность $m \times n$, где m — число фрагментов, n — число

пьезоприводов. Элементы α_{ij} матрицы воздействия определяются эмпирически путём и задают результат измерения величины смещения фрагмента при поочерёдном приложении единичного напряжения к пьезоприводам.

На основе данных о требуемой форме зеркала рассчитывается требуемое смещение каждого фрагмента w_i . Степень соответствия смещения u_i требуемому значению w_i оценивается на основе расчёта среднеквадратичного отклонения:

$$\sigma_{uw} = \sqrt{\frac{1}{m} \sum_{i=1}^m (w_i - u_i)^2} = \sqrt{\frac{1}{m} \sum_{i=1}^m \left(w_i - \sum_{j=1}^n \alpha_{ij} V_j \right)^2}.$$

Наилучшее соответствие требуемой форме зеркала достигается при $\partial \sigma_{uw} / \partial V_k = 0$:

$$\sum_{i=1}^m \left(w_i - \sum_{j=1}^n \alpha_{ij} V_j \right) \alpha_{ik} = 0.$$

Последнее выражение удобно переписать в матричной форме:

$$\mathbf{A} \cdot \mathbf{V} = \mathbf{b}, \quad (5.1)$$

где $A_{kj} = \sum_{i=1}^m \alpha_{ij} \alpha_{ik}$, $b_k = \sum_{i=1}^m w_i \alpha_{ik}$. Система линейных уравнений (5.1) представляет собой задачу оптимизации.

Каждый набор пьезоприводов можно охарактеризовать оптимальными значениями приложенных к ним напряжений, при которых среднеквадратичное отклонение σ_{uw} минимально. Разные наборы пьезоприводов характеризуются разными значениями оптимальных напряжений и минимальных σ_{uw} .

Задача оптимизации заключается том, чтобы выбрать набор пьезоприводов с таким взаимным расположением \mathbf{J} и оптимальными напряжениями \mathbf{V} , при которых среднеквадратичное отклонение смещений от заданных значений является минимальным:

$$\sigma = \min_{\mathbf{J}} \min_{\mathbf{V}} [\sigma_{uw}(\mathbf{J}, \mathbf{V})]. \quad (5.2)$$

Форма зеркала, характеризуемая требуемыми отклонениями w_i , определяется aberrациями, которые испытывает попадающее на зеркало излучение, и которые необходимо исправить. Если в искажение волнового определяющий вклад вносит одна aberrация, то для подбора расположения приводов и определения оптимальных управляющих напряжений решения задачи оптимизации (5.2) достаточно. Однако при большем числе aberrаций требуется решение более сложной задачи:

$$\sigma = \min_{\mathbf{J}} \max_{\mathbf{w}} \min_{\mathbf{V}} [\sigma_{uw}(\mathbf{J}, \mathbf{V}, \mathbf{w})], \quad (5.3)$$

где вектор \mathbf{w} описывает требуемую форму зеркала для исправления рассматриваемой аберрации. Задача оптимизации (5.3) подразумевает подбор расположения приводов и определение оптимальных управляющих напряжений, при которых самое большое отклонение σ_{uw} для одной из рассматриваемых аберраций минимизируется.

ТЕМА 6. Применение адаптивной оптики в оптических системах мощных лазеров

Основные вопросы темы: области применения систем мощных лазеров с адаптивной оптикой. Математическое описание лазерного излучения. Оценка вклада aberrаций в распространение лазерного излучения

Цель и задачи практического занятия состоят в ознакомлении студента со способом математического описания лазерного излучения, в том числе с учётом возникающих в ходе распространения искажений.

Основной понятийно-терминологический аппарат: гауссов пучок, полиномы Эрмита, уравнение Гельмгольца, число Штреля, .

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, осознавать необходимость использования адаптивной оптики в системах мощного лазерного излучения, уметь на основе решения уравнения Гельмгольца показать правомерность описания напряжённости лазерного излучения в отсутствие искажений гауссовым распределением, используя отношение Штреля, уметь описать искажения волнового фронта лазерного пучка.

Теоретический материал и примеры заданий по теме 6

Сфера применения оптических систем, построенных на основе мощных лазеров, чрезвычайно широка. Лазерные системы с адаптивной оптикой нашли широкое применение в промышленности, особенно в области метрологии и контроля качества. В частности, методы измерения искажений волнового фронта излучения использовались для контроля плоскостности кремниевых пластин, которые лежат в основе большинства современных микроэлектронных приборов. В области оптики атмосферы лазерные АОС служат для целей экологического мониторинга, а именно: для измерения скоростей потока твёрдых частиц в воздухе, газов, а также тепла. Принцип действия таких систем основан на измерении флуктуаций величины показателя преломления воздуха. С помощью лазерных АОС в настоящее время обеспечивается передача информации на средние расстояния, например, для объединения двух отстоящих друг от друга волоконно-оптических систем связи.

Во всех случаях качество выполнения поставленных перед лазерными системами задач существенно зависит от качества лазерного излучения, степени искажений, внесённых в излучение внешними факторами. В связи с этим, для улучшения работы лазерных систем важным является оценка качества лазерного излучения и разработка методов компенсации его искажений.

6.1. Лазерный гауссов пучок

В свободном пространстве в отсутствии аберраций напряжённость поля лазерного пучка $E(\mathbf{r})$ подчиняется уравнению Гельмгольца

$$\nabla^2 E + k^2 E = 0, \quad (6.1)$$

где $k = \frac{2\pi}{\lambda}$ – волновое число; $\mathbf{r} = (x, y, z)$. В свободном пространстве данное уравнение имеет решение в виде плоских и сферических волн. Будем искать решение данного уравнения в виде $E(\mathbf{r}) = \Psi(\mathbf{r})e^{-ikz}$, где z – координата распространения лазерного излучения, $\mathbf{r}_\perp = (x, y)$ – координаты в плоскости, перпендикулярной оси пучка. Полагая, что изменение Ψ в направлении z мало в сравнении с изменением в направлении \mathbf{r}_\perp , можем переписать уравнение Гельмгольца для Ψ в виде

$$\nabla_\perp^2 \Psi - 2ik \frac{\partial \Psi}{\partial z} = 0, \quad (6.2)$$

где $\nabla_\perp^2 \Psi = \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2}$; слагаемое $\frac{\partial^2 \Psi}{\partial z^2}$ исключено из уравнения в связи с его малостью. Одно из решений уравнения (6.2) имеет вид

$$\Psi_{00}(r_\perp, z) = A(z)e^{-\frac{ikr_\perp^2}{2q(z)}}, \quad (6.3)$$

где $r_\perp^2 = x^2 + y^2$. Функция (6.3), описывающая пространственное распределение лазерного поля, имеет гауссову форму в направлении \mathbf{r}_\perp . Если мы примем за начало координат на оси z точку, в которой поперечный размер пучка w_0 имеет наименьшее значение, то можем записать для параметра $q(z)$ следующие выражения:

$$q(z) = q_0 + z = i\frac{\pi}{\lambda}w_0^2, \quad (6.4a)$$

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i\frac{\lambda}{\pi w^2(z)}, \quad (6.4b)$$

в которых заданы такие важные параметры как радиус кривизны волнового фронта $R(z)$ и радиус пучка $w(z)$, который определяет удалённость от оси пучка, на котором

напряжённость поля падает в e раз относительно максимального значения. Они задаются выражениями

$$w(z) = w_0 \sqrt{1 + (\lambda z / \pi w_0^2)^2}, \quad (6.5a)$$

$$R(z) = z \left[1 + (\pi w_0^2 / \lambda z)^2 \right], \quad (6.5b)$$

Используя выражения (6.3)–(6.5), с учётом условия нормировки $2\pi \int_0^\infty |E|^2 r dr = |A|^2$, где $A \equiv A(z=0)$, получим выражение для $E(\mathbf{r})$ в следующем виде:

$$E(\mathbf{r}) = A \sqrt{\frac{2}{\pi w^2(z)}} \exp \left[-\frac{r_\perp^2}{w^2(z)} - i \frac{kr_\perp^2}{2R(z)} - ikz + i\varphi_0 \right]. \quad (6.6)$$

Интенсивность гауссова пучка задаётся выражением

$$I(\mathbf{r}) = |E(\mathbf{r})|^2 = I_0 e^{-\frac{2r_\perp^2}{w^2(z)}}.$$

Гауссова мода (6.3) является не единственным решением уравнения (6.2). В книге [21] показано, что решениями (6.2) являются также функции вида

$$\Psi_{nm}(\mathbf{r}) = \Psi_{00}(\mathbf{r}) H_n \left[\frac{\sqrt{2}x}{w(z)} \right] H_m \left[\frac{\sqrt{2}y}{w(z)} \right] e^{i(n+m)\varphi_0}, \quad (6.7)$$

где H_n и H_m – полиномы Эрмита порядка n и m соответственно. Также в выражение (6.7) введена фазовая добавка $i(n+m)\varphi_0$. Таким образом, напряжённость лазерного поля в общем виде может быть описана суперпозицией решений (6.7):

$$\Psi(\mathbf{r}) = \sum_{n,m=0}^{\infty} c_{nm} \Psi_{nm}(\mathbf{r}),$$

причём весовой коэффициент c_{00} существенно превышает коэффициенты $c_{nm}|_{n,m \neq 0}$.

Задание для самостоятельного выполнения: на основе описанной процедуры получить выражение (6.6) из уравнения (6.1).

6.2. Лазерное излучение в присутствии аберраций

Внешние воздействия, связанные, например, с турбулентностью среды, с искажением формы и изменением оптических свойств элементов оптической системы,

оказывают существенное влияние на распространение лазерного излучения. Для оценки этого влияния вводится число Штреля или отношение Штреля (англ.: Strehl ratio). Для оптической системы с круглым зрачком оно задаётся следующим образом [22, 23]:

$$S_R = \frac{\left| \int_0^{2\pi} d\theta \int_0^{\bar{r}} g(r) \exp [ik\Phi(r, \theta)] r dr \right|^2}{\left| 2\pi \int_0^{\bar{r}} g(r) r dr \right|^2}, \quad (6.8)$$

где r и θ – радиальная и угловая координаты в плоскости изображения, для простоты обозначения мы опустили нижний индекс « \perp » при координате r ; $g(r)$ – радиально симметричное пространственное распределение интенсивности лазерного излучения, функция $\Phi(r, \theta)$ описывает искажение волнового фронта; верхний предел интегрирования \bar{r} определяется индивидуально для каждой оптической системы.

При однородном распределении интенсивности в плоскости изображения, $g(r) = \text{const}$, число Штреля (6.8) принимает вид

$$S_{\text{UF}} = \frac{1}{\pi^2 a^4} \left| \int_0^{2\pi} d\theta \int_0^a \exp [ik\Phi(r, \theta)] r dr \right|^2,$$

где a – радиус выходного зрачка оптической системы.

Если через оптическую систему распространяется лазерный пучок, то в отсутствии aberrаций распределение интенсивности в плоскости изображения будет задаваться выражением $g(r) \propto \exp(r^2/w^2)$, где w – ширина пучка в плоскости изображения. Если гауссов пучок обладает достаточно малой расходимостью, такой что $w < a/2$, то верхний предел \bar{r} в определении (6.8) можно заменить бесконечностью, $\bar{r} \rightarrow \infty$. Число Штреля для гауссова лазерного пучка:

$$S_{\text{GF}} = \frac{1}{\pi^2 w^4} \left| \int_0^{2\pi} d\theta \int_0^\infty \exp \left[-\frac{r^2}{w^2} \right] \exp [ik\Phi(r, \theta)] r dr \right|^2.$$

В работах [24, 25] показано, что, выражение (6.8) для числа Штреля может быть преобразовано к виду

$$S_R \approx 1 - k^2 \sigma_\Phi^2 \approx e^{-k^2 \sigma_\Phi^2},$$

где σ_Φ – среднеквадратичное отклонение, σ_Φ^2 – взвешенная дисперсия отклонений фазы, которая задаётся выражением

$$\sigma_\Phi^2 = \langle \Phi^2 \rangle - \langle \Phi \rangle^2; \quad (6.9)$$

$$\langle \Phi^n \rangle = \frac{\int_0^{2\pi} d\theta \int_0^{\bar{r}} g(r) \Phi^n(r, \theta) r dr}{\int_0^{2\pi} d\theta \int_0^{\bar{r}} g(r) r dr}, \quad (6.10)$$

где верхний предел интегрирования \bar{r} принимает значение a для однородного излучения и значение ∞ для гауссова пучка. Описывающая искажения функция $\Phi(r, \theta)$ может быть задана в виде ряда [22]:

$$\Phi(r, \theta) = \sum_{n,m} c_{nm} \Phi_{nm}(r, \theta) = \sum_{n,m} c_{nm} A_{nm} \left(\frac{r}{r_0} \right)^n \cos^m(\theta), \quad (6.11)$$

где коэффициент A_{nm} характеризует величину искажений вдоль базовой окружности радиуса r_0 ; для однородного излучения $r_0 = a$, для лазерного гауссова пучка $r_0 = \kappa w$, где κ – случайный постоянный коэффициент; c_{nm} – весовой коэффициент, описывающий вклад аберрации порядка nm в искажение волнового фронта.

Задание: получить выражения для среднеквадратичного отклонения $\sigma_{\Phi, nm}$ искажений волнового фронта в случае а) однородного и б) гауссова излучения.

Выполнение:

а) Подставляем (6.11) в (6.10) и, полагая для однородного излучения $g(r) = const$, $\bar{r} = a$, для $\langle \Phi_{nm}^{UF} \rangle$ и $\langle (\Phi_{nm}^{UF})^2 \rangle$ получаем выражения

$$\langle \Phi_{nm}^{UF} \rangle = \frac{2A_{nm}\zeta_m}{n+2}, \quad \langle (\Phi_{nm}^{UF})^2 \rangle = \frac{A_{nm}^2 (2m!)}{2^{2m}(n+1)(m!)^2},$$

где $\zeta_m = 0$ для нечётных m и $\zeta_m = \frac{m!}{2^m[(0.5m)!]^2}$ для чётных m . Подставив полученные выражения в (6.9) для однородного пучка в результате получаем (ср. с [22])

$$\sigma_{UF, nm} = A_{nm} \sqrt{\frac{(2m!)}{2^{2m}(n+1)(m!)^2} - \frac{4\zeta_m^2}{(n+2)^2}}.$$

б) Для гауссова пучка аналогично подставляем (6.11) в (6.10) с учётом $g(r) \propto \exp(r^2/w^2)$ и $\bar{r} \rightarrow \infty$. В результате получаем

$$\langle \Phi_{nm}^{GF} \rangle = A_{nm} \kappa^{-n} \zeta_m \Gamma \left[\frac{n+2}{2} \right], \quad \langle (\Phi_{nm}^{GF})^2 \rangle = A_{nm}^2 \kappa^{-2n} \frac{n!(2m!)}{2^{2m}(m!)^2},$$

где $\Gamma(x)$ – гамма-функция, которая определяется следующим образом: $\Gamma(x) = (x-1)!$. Среднеквадратичное отклонение для гауссова пучка задаётся выражением (ср. с [22])

$$\sigma_{GF, nm} = A_{nm} \kappa^{-n} \sqrt{\frac{n!(2m!)}{2^{2m}(m!)^2} - \Gamma^2 \left[\frac{n+2}{2} \right] \zeta_m^2}.$$

Задание для самостоятельного выполнения: для заданных aberrаций (первая колонка таблицы 1) записать аналитическое выражение для искажения фазы Φ_{nm} (вторая колонка), среднеквадратичного отклонения $\sigma_{\Phi, nm}$ (третья колонка) и числа Штреля S_R (четвёртая колонка), характеризующих однородное излучение. Функции $S(x)$ и $C(x)$ в четвёртой колонке – интегралы Френеля, которые задаются в виде $S(x) = \int_0^x \sin(t^2)dt$, $C(x) = \int_0^x \cos(t^2)dt$.

Задание для самостоятельного выполнения: выполнить предыдущее задание для гауссова лазерного излучения — см. табл. 2. Коэффициент χ в таблице определяется как $\chi = \frac{\kappa^2}{\sqrt{2\pi k|A_{40}|}}$.

Таблица 1. Характеристики искажений волнового фронта однородного излучения [22, 23]

Тип aberrации, (n, m)	Φ_{nm}^{UF}	$\sigma_{UF, nm}$	S_{UF}
1	2	3	4
Сферическая, (4,0)	$A_{40} \left(\frac{r}{a}\right)^4$	$\frac{2}{3\sqrt{5}} A_{40}$	$\frac{\pi}{2kA_{40}} \left[C^2 \left(\sqrt{\frac{2kA_{40}}{\pi}} \right) + S^2 \left(\sqrt{\frac{2kA_{40}}{\pi}} \right) \right]$
Кома, (3,1)	$A_{31} \left(\frac{r}{a}\right)^3 \cos(\theta)$	$\frac{\sqrt{2}}{4} A_{31}$	$\left \int_0^1 J_0(kA_{31}x^{3/2}) dx \right ^2$
Астигматизм, (2,2)	$A_{22} \left(\frac{r}{a}\right)^2 \cos^2(\theta)$	$\frac{1}{4} A_{22}$	$J_0^2\left(\frac{1}{2}kA_{22}\right) + J_1^2\left(\frac{1}{2}kA_{22}\right)$
Кривизна поля, (2,0)	$A_{20} \left(\frac{r}{a}\right)^2$	$\frac{1}{2\sqrt{3}} A_{20}$	$\left(\frac{\sin(kA_{20}/2)}{kA_{20}/2} \right)^2$
Дисторсия, (1,1)	$A_{11} \left(\frac{r}{a}\right) \cos(\theta)$	$\frac{1}{2} A_{11}$	$\left(\frac{2J_1(kA_{11})}{kA_{11}} \right)^2$

Таблица 2. Характеристики искажений волнового фронта гауссова пучка [22]

Тип aberrации, (n, m)	Φ_{nm}^{GF}	$\sigma_{GF, nm}$	S_{GF}
1	2	3	4
Сферическая, (4,0)	$A_{40} \left(\frac{r}{\kappa w}\right)^4$	$\frac{2\sqrt{5}}{\kappa^4} A_{40}$	$(\pi\chi)^2 \left[\left(\frac{1}{2} - S(\chi)\right)^2 + \left(\frac{1}{2} - C(\chi)\right)^2 \right]$
Кома, (3,1)	$A_{31} \left(\frac{r}{\kappa w}\right)^3 \cos(\theta)$	$\frac{\sqrt{3}}{\kappa^3} A_{31}$	$\left \int_0^\infty \exp(-x) J_0 \left(kA_{31} \left(\frac{\sqrt{x}}{\kappa}\right)^3 \right) dx \right ^2$
Астигматизм, (2,2)	$A_{22} \left(\frac{r}{\kappa w}\right)^2 \cos^2(\theta)$	$\frac{1}{\sqrt{2}\kappa^2} A_{22}$	$\frac{1}{\sqrt{1+(kA_{22}/\kappa^2)^2}}$
Кривизна поля, (2,0)	$A_{20} \left(\frac{r}{\kappa w}\right)^2$	$\frac{1}{\kappa^2} A_{20}$	$\frac{1}{1+(kA_{20}/\kappa^2)^2}$
Дисторсия, (1,1)	$A_{11} \left(\frac{r}{\kappa w}\right) \cos(\theta)$	$\frac{1}{\sqrt{2}\kappa} A_{11}$	$\exp \left[-2 \left(\frac{kA_{11}}{2\kappa} \right)^2 \right]$

ТЕМА 7. Применение адаптивной оптики в оптике атмосферы и лазерных технологиях

Основные вопросы темы: эффекты самовоздействия лазерного излучения как источник искажений формы волнового фронта. Математическое описание эффектов самовоздействия.

Цель и задачи практического занятия состоят в ознакомлении студента с причинами самовоздействия лазерного излучения высокой мощности при распространении в атмосфере, с математическим описанием изменения показателя преломления атмосферного воздуха во времени на основе базовых принципов термодинамики, с методом описания искажений, основанным на нахождении высших порядков решения уравнения Гельмгольца для лазерного поля.

Основной понятийно-терминологический аппарат: адвекция, дефокусировка, моды Гаусса-Эрмита, полиномы Цернике, самовоздействие лазерного излучения.

Требования к уровню подготовленности студента: студент должен ориентироваться в основных вопросах/понятиях темы, понимать логические связи между ними, понимать механизм самовоздействия распространяющегося в среде лазерного излучения, причины искажений волнового фронта в результате самовоздействия, используя принципы термодинамики, уметь построить математическую модель временной динамики величины показателя преломления атмосферного воздуха, уметь показать, что искажения волнового фронта лазерного излучения математически могут быть описаны модами Гаусса-Эрмита, быть готовым провести аналогию между модами Гаусса-Эрмита и полиномами Цернике в описании искажений.

Теоретический материал и примеры заданий по теме 7

7.1. Тепловое самовоздействие лазерного излучения при распространении в атмосфере

В оптике атмосферы эффективность использования лазерных систем сопряжена с обязательностью коррекции неизбежно возникающих искажений излучения, что напрямую диктует необходимость применения АОС. В указаниях к теме 1 настоящей работы показано, что одним из основных источников искажений волнового фронта

излучения, распространяющегося через атмосферу, является атмосферная турбулентность. Если речь идёт о мощном лазерном излучении, существенное влияние на распространение излучения начинают оказывать нелинейные процессы. Один из таких процессов — тепловое самовоздействие — также приводит к искажению волнового фронта. Как схематично показано на рис. 20, температура атмосферного воздуха вдоль оси лазерного пучка повышается в результате поглощения излучения, что приводит к локальному уменьшению показателя преломления. Вследствие этого процесса происходит рассеяние (дефокусировка) лазерного пучка аналогично прохождению излучения через отрицательную линзу.

В дополнение к дефокусировке адвекция (перемещение воздуха в горизонтальном направлении и перенос вместе с ним его свойств, в том числе температуры) приводит к смещению луча в направлении, перпендикулярном оси распространения излучения.

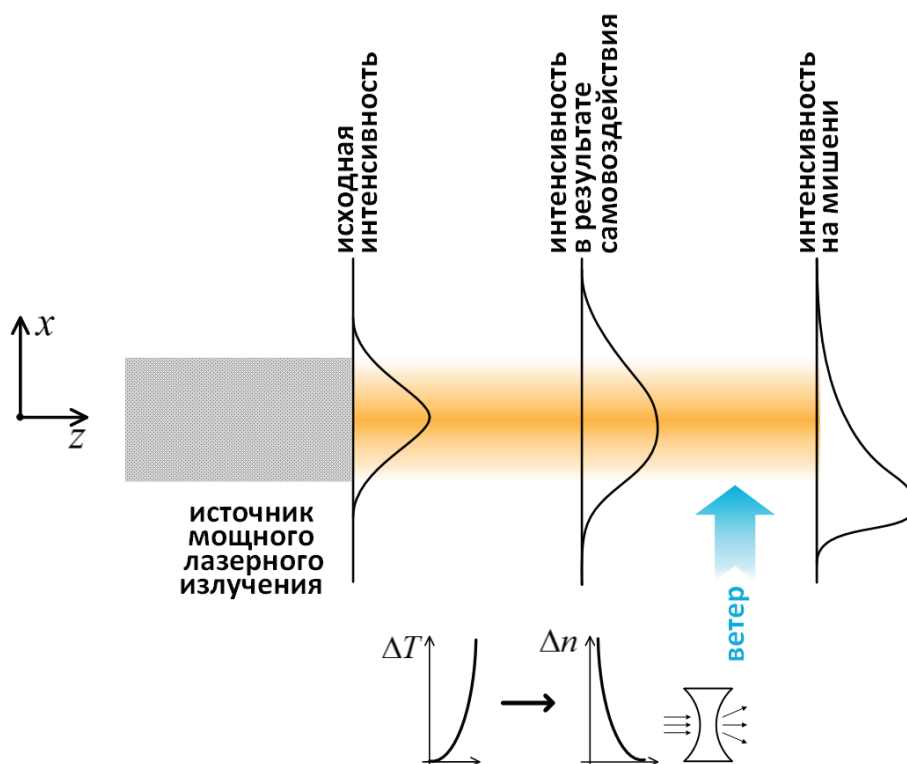


Рис. 20. Схематичное изображение распространения мощного лазерного излучения с учётом самовоздействия. Повышение температуры воздуха приводит к уменьшению его показателя преломления. Действие нагретого воздуха на излучение эквивалентно действию отрицательной линзы.

Изменение величины показателя преломления воздуха в результате теплового самовоздействия описывается на основе принципов термодинамики. В соответствии с

первым законом термодинамики изменение внутренней энергии атмосферного воздуха ΔE связано с его нагревом Q и работой W выражением $\Delta E = Q - W$. Внутренняя энергия E включает в себя кинетическую и потенциальную энергию молекул воздуха. С физической точки зрения, результатом изменения внутренней энергии ΔE является изменение особенностей движения молекул, а именно — их скорости. Так нагрев Q приводит к увеличению скорости частиц воздуха. Что касается работы W , то для объёма воздуха ΔV , находящегося при давлении P , она задаётся выражением $W = P\Delta V$. Таким образом, мы можем переписать первый закон термодинамики в виде $dE = dQ - PdV$.

Для дальнейшего анализа необходимо ввести две константы: это удельная теплоёмкость при постоянном объёме $C_V = \left(\frac{dQ}{dT}\right)_{dV=0}$ и удельная теплоёмкость при постоянном давлении $C_P = \left(\frac{dQ}{dT}\right)_{dP=0}$. Для сухого атмосферного воздуха их величины равны $C_V \approx 718 \text{ Дж}^{-1}\text{К}^{-1}$ и $C_P \approx 1005 \text{ Дж}^{-1}\text{К}^{-1}$.

Происходящие в атмосфере процессы справедливо рассматривать как изохорные, т. е. полагать $dV = 0$. Нагрев на единицу молекулярной массы dQ/m задаётся выражением $\frac{dQ}{m} = C_V dT - \frac{1}{\rho} dP + R_S dT$ (см. [26]), где $\rho = V/m$ — плотность атмосферного воздуха; $R_S = C_P - C_V$, для сухого атмосферного воздуха $R_S \approx 287 \text{ Дж}^{-1}\text{К}^{-1}$. Вновь перепишем первый закон термодинамики с учётом приведённых выражений:

$$\frac{dE}{m} = C_P dT - \frac{1}{\rho} dP. \quad (7.1)$$

Зададим скорость изменения внутренней энергии (скорость нагрева) как $J = dE/dt$ и на основе (7.1) запишем уравнение, описывающее изменение температуры атмосферного воздуха во времени в результате внешнего нагрева:

$$\frac{dT}{dt} = \frac{1}{\rho C_P} \frac{dP}{dt} + \frac{J}{C_P}. \quad (7.2)$$

Распространение лазерного излучения в атмосфере можно рассматривать как изобарный процесс, т. е. полагать $dP = 0$. С учётом того факта, что изменение температуры происходит за счёт самовоздействия лазерного излучения в результате нелинейных процессов, можем записать уравнение (7.2) в виде

$$\frac{dT(\mathbf{r}, t)}{dt} = \mu_T(z) I(\mathbf{r}, t), \quad (7.3)$$

где $\mu_T(z) = \alpha(z)/C_P \rho z$, $\alpha(z)$ — коэффициент поглощения, $I(\mathbf{r}, t)$ — интенсивность лазерного излучения; $J/C_P = \mu_T(z) I(\mathbf{r}, t)$.

Как было ранее указано, изменение температуры воздуха приводит к изменению показателя преломления. Показатель преломления воздуха задаётся следующей эмпирической зависимостью от температуры и давления [26, 27]:

$$n(\mathbf{r}, t) = 1 + 77.6 \times 10^{-6} (1 + 7.52 \times 10^{-3} \lambda^{-2}) R_S \frac{\rho(z) T_0(z)}{T(\mathbf{r}, t)},$$

где $T_0(z)$ – температура окружающего воздуха. Скорость изменения показателя преломления во времени можно выразить через скорость изменения температуры как $\frac{dn}{dt} = \frac{dn}{dT} \frac{dT}{dt}$. Величину dn/dT можно оценить в приближении $T = T_0$. Для этого будем полагать показатель преломления n вещественным, и представим его в виде $n(\mathbf{r}, t) = n_0 + \Delta n(\mathbf{r}, t)$, где n_0 – показатель преломления окружающего воздуха, $\Delta n(\mathbf{r}, t)$ – изменение показателя преломления. Для $\frac{dn}{dT}$ получим выражение

$$\left. \frac{dn(\mathbf{r}, t)}{dT(\mathbf{r}, t)} \right|_{T(\mathbf{r}, t) = T_0(z)} = \left. \frac{d\Delta n(\mathbf{r}, t)}{dT(\mathbf{r}, t)} \right|_{T(\mathbf{r}, t) = T_0(z)} = -\frac{n_0 - 1}{T_0(z)}.$$

С учётом выполненных приближений на основе уравнения (7.2) запишем уравнение, описывающее изменение во времени показателя преломления атмосферного воздуха в результате самовоздействия мощного лазерного излучения:

$$\frac{d\Delta n(\mathbf{r}, t)}{dt} = -\mu(z) I(\mathbf{r}, t), \quad (7.4)$$

где параметр поглощения задаётся выражением $\mu(z) = (n_0 - 1)\alpha(z)/\rho_0 C_P T_0(z)$. Учёт вклада поперечного смещения пучка за счёт адвекции осуществляется путём добавления в уравнение (7.4) дополнительного слагаемого, учитывающего пространственные эффекты. Модифицированное уравнение в этом случае принимает вид

$$\frac{\partial \Delta n(\mathbf{r}, t)}{\partial t} + \mathbf{v}_\perp(z) \cdot \vec{\nabla}_\perp \Delta n(\mathbf{r}, t) = -\mu(z) I(\mathbf{r}, t).$$

Задание для самостоятельного выполнения: на основе описанной процедуры и с учётом указанных приближений из (7.3) получить уравнение (7.4).

7.2. Анализ волнового фронта лазерного пучка

Рассмотрим гауссов лазерный пучок, подвергшийся влиянию аберраций двух типов: астигматизма и дефокусировки [28]. Введём новую систему координат (\bar{x}, \bar{y}) в

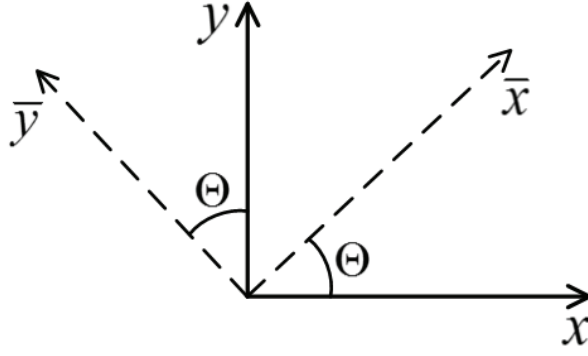


Рис. 21. К анализу волнового фронта лазерного излучения.

плоскости, перпендикулярной оси пучка, связанную с главными осями астигматического изображения и повёрнутую относительно исходной системы координат (x, y) на угол Θ — см. рис. 21. Связь координат задаётся следующим образом:

$$\bar{x} = x \cos(\Theta) + y \sin(\Theta), \quad \bar{y} = x \sin(\Theta) - y \cos(\Theta).$$

На основе (6.6) запишем амплитуду поля в присутствии искажений $\Psi(\bar{x}, \bar{y})$:

$$\Psi(\bar{x}, \bar{y}) = \sqrt{\frac{2}{\pi}} \frac{A}{w_0} \exp \left(-\frac{\bar{x}^2 + \bar{y}^2}{w_0^2} - i \frac{\pi}{\lambda} \left(\frac{\bar{x}^2}{R_{\bar{x}}} + \frac{\bar{y}^2}{R_{\bar{y}}} \right) \right), \quad (7.5)$$

где $R_{\bar{x}, \bar{y}}$ — радиусы кривизны волнового фронта в плоскостях, содержащих ось пучка и главные оси астигматического изображения; $R_{\bar{x}}$ и $R_{\bar{y}}$ представляют собой наибольший и наименьший радиусы кривизны. Средний радиус R определим следующим выражением: $R^{-1} = 0.5 (R_{\bar{x}}^{-1} + R_{\bar{y}}^{-1})$. Введём также параметр $\delta = 0.5 (R_{\bar{x}}^{-1} - R_{\bar{y}}^{-1})$, характеризующий степень астигматизма. Выразим радиусы кривизны следующим образом: $R_{\bar{x}}^{-1} = R^{-1} + \delta^{-1}$, $R_{\bar{y}}^{-1} = R^{-1} - \delta^{-1}$. С учётом заданных параметров на основе выражения (7.5) запишем амплитуду лазерного поля $\Psi(x, y)$ в виде

$$\Psi(x, y) = \sqrt{\frac{2}{\pi}} \frac{A}{w_0} \exp \left(-\frac{x^2 + y^2}{w_0^2} - i \frac{\pi w_0^2}{2\lambda R} \left(2 \frac{x^2 + y^2}{w_0^2} - 1 \right) - i \frac{\pi w_0^2}{2\lambda R} - i \frac{\pi \delta w_0^2}{\lambda} \frac{x^2 - y^2}{w_0^2} \cos(2\Theta) - i \frac{\pi \delta w_0^2}{\lambda} \frac{2xy}{w_0^2} \sin(2\Theta) \right). \quad (7.6)$$

Разложив в выражении (7.6) экспоненту в ряд, мы получим

$$\Psi(x, y) \approx \sqrt{\frac{2}{\pi}} \frac{A}{w_0} \exp \left(-\frac{x^2 + y^2}{w_0^2} \right) \left(1 - i \frac{\pi w_0^2}{2\lambda R} - i \frac{\pi w_0^2}{2\lambda R} \left(2 \frac{x^2 + y^2}{w_0^2} - 1 \right) - i \frac{\pi \delta w_0^2}{\lambda} \frac{x^2 - y^2}{w_0^2} \cos(2\Theta) - i \frac{\pi \delta w_0^2}{\lambda} \frac{2xy}{w_0^2} \sin(2\Theta) \right). \quad (7.7)$$

Воспользовавшись выражением (6.7), задающим моды Гаусса-Эрмита, можем преобразовать амплитуду поля (7.7) к виду

$$\Psi(x, y) \approx A \left[\left(1 - i \frac{\pi w_0^2}{2\lambda R} \right) \Psi_{00} - i \frac{\pi w_0^2}{2\sqrt{2}\lambda R} (\Psi_{20} + \Psi_{02}) - \right. \\ \left. - i \frac{\pi \delta w_0^2}{\sqrt{2}\lambda} \cos(2\Theta) (\Psi_{20} - \Psi_{02}) - i \frac{\pi \delta w_0^2}{\lambda} \sin(2\Theta) \Psi_{11} \right]. \quad (7.8)$$

Выразим из (7.6) мнимую (фазовую) часть, которая описывает волновой фронт:

$$\Phi(\rho, \theta) \approx \frac{\pi w_0^2}{2\lambda R} + \frac{\pi w_0^2}{2\lambda R} (2\rho^2 - 1) + \\ + \frac{\pi \delta w_0^2}{\lambda} \rho^2 \cos(2\Theta) \cos(2\theta) + \frac{\pi \delta w_0^2}{\lambda} \rho^2 \sin(2\Theta) \sin(2\theta).$$

Воспользовавшись выражениями (1.6) и (1.7), можем выразить форму волнового фронта $\Phi(\rho, \theta)$ через полиномы Цернике следующим образом:

$$\Phi(\rho, \theta) \approx \frac{\pi w_0^2}{2\lambda R} Z_0^0 + \frac{\pi w_0^2}{2\lambda R} Z_2^0 + \frac{\pi \delta w_0^2}{\lambda} Z_2^2 \cos(2\Theta) + \frac{\pi \delta w_0^2}{\lambda} Z_2^{-2} \sin(2\Theta). \quad (7.9)$$

Сравнивая выражения (7.8) и (7.9), видим, что имеется прямое соответствие между модами Гаусса-Эрмита, описывающими излучение в присутствии возмущений, и полиномами Цернике, описывающими искажённый волновой фронт. Так, например, мода $\frac{1}{\sqrt{2}} (\Psi_{20} + \Psi_{02})$ соответствует полиному Z_2^0 , описывающему дефокусировку; моды $\frac{1}{\sqrt{2}} (\Psi_{20} - \Psi_{02})$ и Ψ_{11} соответствуют полиномам Z_2^2 и Z_2^{-2} , и описывают астигматизм; графически указанные полиномы приведены на рис. 6.

Задание для самостоятельного выполнения: На основе приведённой процедуры из (7.5) получить выражение (7.7). При выполнении использовать следующие равенства: $\bar{x}^2 + \bar{y}^2 = x^2 + y^2$, $\bar{x}^2 - \bar{y}^2 = (x^2 - y^2)(\cos^2(\Theta) - \sin^2(\Theta)) + 4xy \cos(\Theta) \sin(\Theta)$, — а также известные тригонометрические преобразования $\cos^2(\Theta) - \sin^2(\Theta) = \cos(2\Theta)$, $2 \sin(\Theta) \cos(\Theta) = \sin(2\Theta)$.

Список литературы

1. Vernin, J. Gemini CP Site Characterization Report / J. Vernin, A. Agabi, R. Avila, Azouit, R. Conan, F. Martin, E. Masciadri, L Sanchez, A. Ziad. – Internal Report RPT-AO-G0094, Gemini Observatory, January 2000.
2. Fried, D. L. Statistics of a geometric representation of wavefront distortion / D. L. Fried // Journal of the Optical Society of America. – 1965. – Vol. 55. – P. 1427–1435.
3. Da Costa Rodrigues, G. M. Adaptive optics with segmented deformable bimorph mirrors [Text]: Thesis submitted in candidature for the degree of Doctor in Engineering Sciences: February, 2010 / Gonçalo Mendes da Costa Rodrigues. – Bruxelles, 2010. – 149 pp.
4. Timoshenko, S. P. Theory of Plates and Shells / S. P. Timoshenko and S. Woinowsky-Krieger. – 2nd edition. – USA: McGraw-Hill Book Company, 1987. – 580 pp.
5. Reddy, J. N. Theory and Analysis of Elastic Plates and Shells / J. N. Reddy. – 2nd edition. – Boca Raton, USA: CRC Press. – 547 pp.
6. Ruppel, T. Model-based feedforward control of large deformable mirrors / T. Ruppel, W. Osten, O. Sawodny // European Journal of Control. – 2011. – Vol. 17. – P. 261–272.
7. Preumont, A. Vibration Control of Active Structures / A. Preumont. – Verlages, Berlin Heidelberg: Springer Science & Business Media, 2011. – 456 pp.
8. Pilkey, W. D. Formulas of Stress, Strain and Structural Matrices / W. D. Pilkey. – 2nd edition. – NY, USA: John Wiley, 2004. – 1536 pp.
9. Tyson, R. Principles of Adaptive Optics / R. Tyson. – 3rd edition. – NY, USA: CRC Press Taylor & Francis Group, 2011. – 315 pp.
10. Fried, D. L. Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements / D. L. Fried // Journal of the Optical Society of America. – 1977. – Vol. 67. – P. 370–375.
11. Ellerbroek, B. L. Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques / B. L.. Ellerbroek // Journal of the Optical Society of America. – 2002. – Vol. 19. – P. 1803–1816.

12. Bardsley, J. M. Wavefront Reconstruction Methods for Adaptive Optics Systems on Ground-Based Telescopes / J. M. Bardsley // *SIAM Journal on Matrix Analysis and Applications*. – 2008. – Vol. 30. – P. 67–83.
13. Thiébaud, E. Fast minimum variance wavefront reconstruction for extremely large telescopes / E. Thiébaud and M. Tallon // *arXiv:math-ph/1003.0274v1*.
14. Tarantola, A. Inverse Problem Theory and Methods for Model Parameter Estimation // A. Tarantola. – Philadelphia, USA: SIAM, 2005. – xii + 342 pp.
15. Tarantola, A. Inverse problems=quest for information / A. Tarantola, B. Valette // *Journal of Geophysics*. – 1982. – Vol. 50. – P. 159–170.
16. Southwell, W. H. Wave-front estimation from wave-front slope measurements / W. H Southwell // *Journal of the Optical Society of America*. – 1980. – Vol. 70. – P. 998–1006.
17. Popov, A. Genetic algorithms for optimization / A. Popov. – Hamburg, 2005. – 21 pp.
18. Irschik, H. A review on static and dynamic shape control of structures by piezoelectric actuation / H. Irschik // *Engineering Structures*. – 2002. – Vol. 24. – P. 5–11.
19. Goldberg, D. Genetic Algorithms in Search, Optimization, and Machine Learning / D. Goldberg. – NY, USA: Addison-Wesley Professional, 1989. – 372 pp.
20. Sheng, L. Genetic Algorithms for the Optimization of Piezoelectric Actuator Locations / L. Sheng, R. K. Kapania // *Proceedings of AIAA/ASME/ASCE/AHS/AHC 41st Structures, Structural Dynamics and Materials Conference, Atlanta, GA, April 3-6, 2000*. – Presented as Paper AIAA-2000-1581.
21. Svelto, O. Principles of lasers / O. Svelto. – 5th edition. – New York, Dordrecht, Heidelberg, London: Springer, 2010. – 620 pp.
22. Herloski, R. Strehl ratio for untruncated aberrated Gaussian beams / R. Herloski // *Journal of the Optical Society of America*. – 1985. – Vol. 2. – P. 1027–1030.
23. Mahajan, V. N. Strehl ratio for primary aberrations: some analytical results for circular and annular pupils / V. N. Mahajan // *Journal of the Optical Society of America*. – 1982. – Vol. 72. – P. 1258–1266.
24. Mahajan, V. N. Zernike annular polynomials for imaging systems with annular pupils / V. N. Mahajan // *Journal of the Optical Society of America*. – 1981. – Vol. 71. – P. 75–85.

25. Mahajan, V. N. Strehl ratio for primary aberrations in terms of their aberration variance / V. N. Mahajan // Journal of the Optical Society of America. – 1983. – Vol. 73. – P. 860–861.
26. Spencer, M. F. Branch point mitigation of thermal blooming phase compensation instability, Master's Thesis, Department of the air force air university, Air force institute of technology, March, 2011, 266 pp.
27. Clifford, S. F. The Classical Theory of Wave Propagation in the Atmosphere / S. F. Clifford // in Laser Beam Propagation in the Atmosphere, J. W. Strohbehn, Ed. New York, USA: Springer Verlag, 1978, ch. 2, pp. 9–41.
28. Born, M. Principles of Optics / M. Born, E. Wolf. – 7th(expanded) edition. – Cambridge, UK: Cambridge University Press, 1999. – 952 pp.