

Содержание

Байкова А. Т. Метод устранения фона внегалактических источников из наблюдаемых карт реликтового излучения.....	909
Кузнецов А. А., Власов В. Г. Влияние неоднородностей магнитного поля и концентрации плазмы на распространение и усиление радиоизлучения в солнечной короне.....	928
Лесовой С. В. Реконструкция изображений Солнца, полученных на Сибирском солнечном радиотелескопе.....	942
Каневский М. Б. Новая спектральная оценка для РСА-изображения поверхности океана.....	950
Гавриленко В. Г., Сорокин А. В. О трансформации углового спектра мощности рассеянного излучения в турбулентной поглощающей магнитоактивной плазме.....	963
Якубович Е. И., Малыкин Г. Б. О влиянии некомпланарной намотки световода в волоконном кольцевом интерферометре на разность фаз встречных волн.....	975
Флакман А. Г. Пространственное разделение пользователей в ММО-системах, использующих параллельную передачу данных.....	986

УДК 523.1.11+523.1.164+621.391:53.08

МЕТОД УСТРАНЕНИЯ ФОНА ВНЕГАЛАКТИЧЕСКИХ ИСТОЧНИКОВ ИЗ НАБЛЮДАЕМЫХ КАРТ РЕЛИКТОВОГО ИЗЛУЧЕНИЯ

А. Т. Байкова

Предлагается новый метод высокоточного восстановления анизотропии реликтового излучения из наблюдаемых карт, содержащих фон неразрешённых точечных источников и пиксельный шум. Принципиальным отличием предлагаемого метода от известных является использование априорной информации об ограниченности и круговой симметрии пространственного спектра мощности флуктуаций температуры реликтового фона, подчиняющихся гауссовой статистике, при их восстановлении в областях искажений точечными источниками. Для решения задач деконволюции исходных карт, локализации точечных источников и интерполяции значений реликтового фона в «дырах» использованы эффективные модификации хорошо известного метода максимальной энтропии. В ходе моделирования продемонстрирована возможность достижения предельных точностей восстановления.

ВВЕДЕНИЕ

Поскольку будущие эксперименты по измерению анизотропии реликтового фона (см., например, [1]) предназначены для высокоточного определения основных космологических параметров Вселенной, необходимо обеспечить адекватную точность устранения из наблюдаемых карт реликтового излучения фоновых искажений, вызванных галактической пылью, синхротронным излучением, эффектом Сюняева—Зельдовича и т. д., а также внегалактическими точечными источниками и пиксельным шумом. Все перечисленные компоненты можно разделить на две основные группы [2]. К первой относятся так называемые низкомультипольные составляющие, источниками которых являются Галактика и скопления галактик, искажающие спектр мощности реликтового излучения на низких пространственных частотах, или мультиполях, в случае разложения функции по сферическим гармоникам. Ко второй группе относятся высокомультипольные составляющие от неразрешённых внегалактических источников и пиксельного шума, обладающих широким пространственным спектром. Действительно, т. к. пространственный спектр реликтового излучения заметно затухает с увеличением частоты и в конце концов сходит на нет, то эффект влияния точечных источников и пиксельного шума усиливается с ростом номера пространственной гармоники и наиболее сильно проявляется в высокочастотной области спектра.

Как известно, реликтовое излучение обладает чернотельным спектром. Кроме того, в соответствии с большинством космологических моделей будем считать, что флуктуации реликтового фона подчиняются гауссовой статистике [3]. Т. к. фоновые искажения не обладают двумя этими свойствами [2], они, в принципе, могут быть выделены из суммарных измерений.

Так, негауссова природа фоновых компонент может быть использована для удаления участков карты, искажённых яркими источниками и зоной Галактики, например, по критерию 10σ . Частотная зависимость искажений может быть использована для их устранения путём формирования линейных комбинаций карт, измеренных на разных частотах. Такая технология успешно применяется к перечисленным выше низкомультипольным фоновым компонентам [4, 5]. Однако она не может быть применена к внегалактическим точечным источникам, т. к. 1) спектральные индексы внегалактических источников изменяются гораздо сильнее, чем у других фоновых составляющих; 2) многие источники обнаруживают сильную временную переменность; 3) распределение точечных источников в большей части частотного

диапазона реликтового излучения известно довольно плохо. Поэтому для выделения точечных источников используют их негауссовость и факт преобладания сигнала от точечных источников над реликтовым фоном в высокочастотной области пространственного спектра. Действительно, если начиная с некоторого значения пространственной частоты (порядок мультиполя), которую обозначим l_d , вклад точечных источников существенно превосходит вклад реликтового фона, то спектральные компоненты, лежащие между l_d и l_f , где l_f — верхняя частота пропускания передаточной функции антенны, можно использовать в качестве исходных данных для восстановления точечных источников с целью вычитания их вклада из измеренной карты либо локализации участков искажения.

Проблема устранения вклада точечных источников рассматривалась многими авторами, но мы отметим здесь только одну работу, наиболее близкую нам по идеологии. Так, в работе Тегмарка и де Оливейра-Косты [6] предлагается использовать высокочастотную пространственную фильтрацию наблюдаемой карты для усиления вклада точечных источников и подавления флуктуаций реликтового фона с целью последующего удаления из карты областей с центрами в местах обнаруженных источников и диаметром, равным ширине диаграммы направленности антенны, а угловой спектр мощности флуктуаций реликтового фона оценивать по остаточной карте с удалёнными участками.

Хотя метод Тегмарка и де Оливейра-Косты обеспечивает существенно более высокую точность, чем непосредственное ограничение яркости измеренной карты по критерию $\nu\sigma$, он не лишён недостатков. Во-первых, амплитуда детектируемых точечных источников всё ещё сильно ограничена снизу; во-вторых, удаление искажённых областей, достаточно больших при большом числе источников и широкой диаграмме направленности антенны, приводит к значительному уменьшению общей площади, по которой проводится оценка спектра мощности флуктуаций реликтового фона; в-третьих, простое устранение искажённых участков не восстанавливает в полной мере гауссову статистику флуктуаций. Всё вместе это может привести к ощутимым ошибкам в оценке искомого углового спектра мощности, а значит, и в оценке космологических параметров. Данная работа посвящена разработке метода, свободного от указанных недостатков.

В последующих разделах статьи даётся описание модели наблюдаемой карты реликтового излучения, метода разделения компонент, метода максимальной энтропии, являющегося основой выделения точечных источников и восстановления реликтового фона и, наконец, анализируются результаты численного моделирования.

1. МОДЕЛЬ НАБЛЮДАЕМОЙ КАРТЫ

1.1. Статистическое описание анизотропии реликтового фона

Распределение температуры реликтового излучения по небесной сфере может быть представлено в виде следующего разложения по сферическим гармоникам $Y_l^m(\theta, \phi)$ [1, 3]:

$$\frac{\Delta T}{\langle T \rangle} = \sum_{l,m} a_l^m Y_l^m(\theta, \phi),$$

где $\langle T \rangle$ и ΔT — средняя температура и температура флуктуаций реликтового фона соответственно, a_l^m — коэффициенты разложения. Угловой спектр мощности флуктуаций определяется как средний квадрат коэффициентов a_l^m :

$$C_l = \langle |a_l^m|^2 \rangle,$$

где l — порядок разложения (или мультиполя).

Если флуктуации излучения ранней Вселенной удовлетворяют гауссовой статистике в соответствии с большинством космологических теорий, то каждый коэффициент a_l^m является статистически независимым и, таким образом, спектр мощности C_l обеспечивает полное статистическое описание анизотропии реликтового фона, представляя собой фундаментальную характеристику Вселенной. Эта характеристика может быть получена непосредственно из наблюдений с помощью гармонического анализа.

Для гауссовых полей коэффициенты разложения a_l^m , где $l \neq 0$, также представляют собой гауссовы случайные величины со случайными фазами, нулевыми средними и дисперсиями

$$\langle a_l^m a_{l'}^{m'} \rangle = \delta_{ll'} \delta_{mm'} C_l,$$

где δ_{nk} — символ Кронекера.

В данной работе рассматриваются небольшие участки неба. В таких случаях целесообразно оперировать гауссовыми полями на плоскости и разложение по сферическим гармоникам заменить преобразованием Фурье. Тогда температура флуктуаций реликтового фона $\Delta T = T - \langle T \rangle$ может быть сгенерирована путём простого вычисления рядов Фурье [3]:

$$\frac{\Delta T(\theta_x, \theta_y)}{\langle T \rangle} = \sum_{n_u=0}^{N_u-1} \sum_{n_v=0}^{N_v-1} D(n_u, n_v) \exp \left[i \frac{2\pi}{L} (n_u \theta_x + n_v \theta_y) \right], \quad (1)$$

где L — линейный размер выделенного участка в радианах, (θ_x, θ_y) — декартовы координаты на небе (также в радианах), (n_u, n_v) — координаты фурье-компонент $D(n_u, n_v)$ в области пространственных частот, которую в дальнейшем будем называть (u, v) -покрытием. Амплитуды фурье-компонент $D(n_u, n_v)$ удовлетворяют гауссовому распределению с нулевым средним и дисперсией

$$\langle |D(n_u, n_v)|^2 \rangle = C_l, \quad l = \frac{2\pi}{L} \sqrt{n_u^2 + n_v^2}, \quad (2)$$

а фазы равномерно распределены в интервале $(0, 2\pi)$. Здесь C_l — угловой спектр мощности температуры реликтового излучения в случае разложения его по сферическим гармоникам.

В принятой модели реликтового излучения характеристикой инструмента, измеряющего яркость неба на прямоугольной сетке координат, является (u, v) -покрытие, представляющее собой круг с радиусом, определяемым частотой дискретизации измеряемой карты.

Соотношение (2) обеспечивает свойство круговой симметрии спектра мощности, т. е. независимости от угла, задаваемого координатами u и v , которое в дальнейшем используется в качестве априорной информации при восстановлении флуктуаций реликтового фона. Другой априорной характеристикой спектра мощности является ограниченность его ширины. Действительно, как уже говорилось во введении, флуктуации реликтового фона характеризуются так называемой частотой затухания [1], выше которой вклад реликтового излучения в суммарный наблюдаемый сигнал пренебрежимо мал.

1.2. Распределение фоновых компонент и модель карты

При генерировании реализаций фона точечных источников было принято, что их распределение подчиняется закону Пуассона. Пиксельный шум удовлетворяет гауссовому закону распределения с нулевым средним.

Для упрощения дальнейших записей примем следующие обозначения: CMB — карта флуктуаций реликтового излучения (от английского Cosmic Microwave Background), PS — карта фоновых точечных источников (от английского Point Sources), N — карта пиксельного шума (от английского Noise),

W — диаграмма направленности антенны, CMB_{det} — измеренная (детектированная) карта реликтового излучения. Тогда принятая в данной работе модель наблюдаемой карты реликтового излучения имеет вид

$$CMB_{det} = (CMB + PS) * W + N, \quad (3)$$

где знак $*$ обозначает операцию свёртки.

Ниже отдельно исследуются два частных случая: 1) входная карта реликтового излучения свободна от пиксельного шума; 2) входная карта содержит значительный пиксельный шум. Такое разделение сделано для того, чтобы при моделировании лучше понять источники ошибок восстановления.

2. МЕТОД РАЗДЕЛЕНИЯ КОМПОНЕНТ

Предлагаемый метод разделения флуктуаций температуры реликтового излучения и фона точечных источников состоит из следующих четырёх этапов.

1. Выполняется операция обратной свёртки, или деконволюции, измеренной карты соответствующей диаграммой направленности антенны с целью получения оценки, представляющей собой простую сумму флуктуаций реликтового излучения и фона точечных источников. Здесь решается известная обратная задача типа свёртки, которая становится некорректно поставленной при наличии в данных аддитивного шума. Решение такой задачи требует применения алгоритмов восстановления с хорошими стабилизирующими свойствами.

2. Восстанавливается фон точечных источников с использованием высокочастотной части пространственного спектра Фурье измеренной карты, где вклад точечных источников существенно превосходит реликтовое излучение. В качестве исходных данных в этой задаче восстановления используются отсчёты пространственного спектра измеренной карты с координатами (n_u, n_v) , удовлетворяющими соотношению $l_d \leq l \leq l_f$, где $l = (2\pi/L) \sqrt{n_u^2 + n_v^2}$ в случае карт небольшого размера, а пространственный спектр вычисляется как преобразование Фурье исходной карты (см. также (2)). При этом, очевидно, соответствующее (u, v) -покрытие данных представляют собой кольцо с центром в точке $(0, 0)$ и характеризует не инструмент, а свойства спектров разделяемых сигналов. Для решения задачи восстановления в данном случае могут быть использованы как линейные методы (например, хорошо известная процедура чистки CLEAN [7]), так и нелинейные (например, метод максимальной энтропии [8, 9]).

3. Из карты, полученной на первом этапе, устраняются те искажённые участки, в которых яркость точечных источников, восстановленных на втором этапе, превосходит некоторое пороговое значение. В результате мы получаем карту реликтового фона с «дырами».

4. Восстанавливается яркость реликтового фона в образовавшихся «дырах» с использованием дополнительной априорной информации о пространственном спектре флуктуаций, а именно о его конечной ширине (финитности) и круговой симметрии (2), вытекающей из гауссовой природы флуктуаций [2]. Из теории аналитических функций известно, что финитной функции соответствует фурье-образ, представляющий собой аналитическую функцию. А аналитическая функция может быть восстановлена по её известной части с использованием любой дополнительной информации. Для решения этой задачи следует использовать нелинейные методы восстановления изображений, обладающие хорошими интерполирующими и экстраполирующими свойствами.

Оригинальность предложенного подхода определяется в основном наличием четвёртого этапа, осуществление которого и позволяет достичь качественно нового результата в восстановлении реликтового фона. Выполнение предыдущих этапов лишь улучшает метод Тегмарка и де Оливейра-Косты [6], обеспечивая более точную локализацию мест искажения реликтового фона, позволяя понижать уровень амплитудной отсечки. Таким образом, можно ожидать, что предлагаемый метод позволит: 1) локализовать большее количество слабых источников; 2) восстановить флуктуации реликтового фона

в местах искажений, что приводит к более корректному восстановлению их гауссовой природы, а также к увеличению пространственной области, по которой производится оценка углового спектра мощности.

Очевидно, разница между известным и предложенным методами будет тем существеннее, чем большее число слабых источников попадает на измеряемую карту.

Для реализации требуемых операций обратной свёртки, восстановления точечных источников и интерполяции значений реликтового фона был выбран метод максимальной энтропии, который в силу присущей ему нелинейности обладает хорошими стабилизирующими и сверхразрешающими свойствами и, кроме того, может быть легко модифицирован под конкретные задачи [10–12].

3. МЕТОД МАКСИМАЛЬНОЙ ЭНТРОПИИ

Метод максимальной энтропии (ММЭ) известен как один из самых мощных стабилизаторов решения некорректных обратных задач, в частности, связанных с интерполяцией и экстраполяцией в фурье-области. В нашем подходе он используется для реализации следующих трёх основных операций: 1) обратной свёртки измеренной карты диаграммой направленности антенны, что представляет собой, по-существу, операцию инверсной фильтрации входных данных аппаратной функцией измерительной системы; 2) восстановления фона точечных источников с использованием высокочастотной части фурье-спектра измеренной карты, где вклад точечных источников доминирует над реликтовым фоном; 3) восстановления яркости реликтового фона в «дырах», образовавшихся в результате устранения участков, искажённых фоновыми источниками.

Для решения второй задачи достаточно применения стандартного ММЭ, изначально предназначенного для восстановления положительно определённых сигналов. Для решения первой и третьей задач требуется применение более универсальной формы ММЭ, так называемого обобщённого метода максимальной энтропии (ОММЭ) [10–12], предназначенного для восстановления сигналов, принимающих как положительные, так и отрицательные значения. Рассмотрим сначала стандартную форму ММЭ.

3.1. Стандартная форма ММЭ

Пусть x_{ml} , где $m = 0, 1, \dots, M-1$, $l = 0, 1, \dots, M-1$, — неизвестный двумерный сигнал, определённый в дискретном пространстве размера $M \times M$. Тогда решение для x_{ml} по методу максимальной энтропии должно удовлетворять критерию

$$-\sum_m \sum_l x_{ml} \ln x_{ml} \equiv H(x) = \max, \quad (4)$$

где $x_{ml} \geq 0$, при условии выполнения ограничений на x_{ml} , вытекающих из данных измерений. Во всех наших задачах мы имеем дело с линейными ограничениями на данные в фурье-области сигнала, которые могут быть представлены следующим образом:

$$\sum_m \sum_l x_{ml} \cos[2\pi (nm + kl)/M] + \eta_{nk}^r = D_{nk}^r, \quad (5)$$

$$\sum_m \sum_l x_{ml} \sin[2\pi (nm + kl)/M] + \eta_{nk}^i = D_{nk}^i. \quad (6)$$

Здесь D_{nk}^r и D_{nk}^i — данные; верхние индексы r и i обозначают вещественную и мнимую части соответственно; η_{nk}^r и η_{nk}^i — вещественная и мнимая части аддитивного шума соответственно, $0 \leq n \leq M-1$,

$0 \leq k \leq M - 1$. Таким образом, мы ищем решение для x_{ml} , которое удовлетворяет принципу (4) при условии выполнения ограничений (5), (6) на данные наблюдений.

Существуют два подхода к учёту ограничений. Первый предполагает их включение в функцию, определяющую меру качества по критерию χ^2 и представляющую собой среднеквадратическое отклонение между наблюдаемыми данными и решением. Эта функция вычитается из критерия (4), а полученный результат максимизируется [9]. Во втором подходе каждое ограничение включается в функционал (4) с соответствующим множителем Лагранжа, а затем производится максимизация полученной суммы [8]. Мы используем второй подход.

Кроме того, в алгоритм восстановления необходимо встроить механизм учёта шума. Это приводит к дальнейшей регуляризации решения сверх того, чего можно достичь только путём максимизации нелинейного функционала. Предположим наличие в данных кругового гауссового шума. Пусть σ_{nk}^2 — дисперсия шума для вещественной и мнимой частей отсчётов фурье-спектра с координатами (n, k) . Тогда критерий оценки можно записать в виде

$$H(x) - (1/2) \sum_n \sum_k [(\eta_{nk}^r)^2 + (\eta_{nk}^i)^2] / \sigma_{nk}^2 = \max. \quad (7)$$

Это эквивалентно поиску вектора (x, η^r, η^i) , который максимизирует совместную вероятность неизвестного сигнала x_{ml} и составляющих шума η_{nk}^r и η_{nk}^i . Согласно теореме Байеса при априорно известном законе вероятности для x , заданном как $\exp[H(x)]$, выражение (7) также представляет собой максимум апостериорной вероятности оценки неизвестных.

3.2. Обобщённая форма ММЭ

Формулировка (4) принципа максимума энтропии может быть использована исключительно для оценки положительно определённых функций, таких, как плотность вероятности, спектр мощности или интенсивность изображения. Однако нам необходимо работать с шенноновской энтропией флуктуаций реликтового фона, принимающих как положительные, так и отрицательные значения. Подробное описание обобщённого метода максимальной энтропии (ОММЭ), применённого для восстановления функций любого вида, включая знакопеременные и комплексные, даётся в работах [10–12].

Итак, рассмотрим сигнал x_{ml} , принимающий как положительные, так и отрицательные значения. Такой сигнал может быть представлен как разность двух положительно определённых функций:

$$x_{ml} = x_{ml}^+ - x_{ml}^-, \quad (8)$$

где $x_{ml}^+ \geq 0, x_{ml}^- \geq 0$.

Если положительные участки сигнала представляются функцией x_{ml}^+ , а отрицательные — функцией x_{ml}^- и эти функции не перекрываются, то задачу оценки функции x_{ml} можно заменить задачей оценки функций x_{ml}^+ и x_{ml}^- следующим образом. Запишем выражение

$$- \sum_m \sum_l (x_{ml}^+ \ln x_{ml}^+ + x_{ml}^- \ln x_{ml}^-), \quad (9)$$

которое эквивалентно энтропии абсолютных значений функции x_{ml} :

$$- \sum_m \sum_l |x_{ml}| \ln |x_{ml}| \equiv H(x).$$

Таким образом, максимизация энтропии $H(x)$ эквивалентна максимизации суммы мер энтропии функций x_{ml}^+ и x_{ml}^- . Это даёт возможность оценки каждой составляющей x_{ml}^+ и x_{ml}^- искомого решения.

Требование неперекрываемости функций x_{ml}^+ и x_{ml}^- не может быть удовлетворено простой максимизацией функционала (9). Как показано в [10, 11], необходимо ввести дополнительный параметр α , который формирует ассоциативную энтропию

$$H(x, \alpha) \equiv - \sum_m \sum_l [x_{ml}^+ \ln(\alpha x_{ml}^+) + x_{ml}^- \ln(\alpha x_{ml}^-)]. \quad (10)$$

Параметр α выбирается по усмотрению пользователя. Чем выше значение α , тем выше степень неперекрываемости функций x_{ml}^+ и x_{ml}^- , что, в свою очередь, обеспечивает лучшее согласование решения с данными и, соответственно, более высокое разрешение.

3.3. ММЭ-алгоритм разделения компонент

3.3.1. Данные без шумов. В этом случае мы имеем следующее представление наблюдаемой карты реликтового излучения:

$$\text{CMB}_{\text{det}} = (\text{CMB} + \text{PS}) * W.$$

Требуется разделить компоненты CMB и PS с максимальной возможной точностью.

Алгоритм ММЭ, реализующий схему метода, приведённую в разделе 2, заключается в выполнении следующих операций.

1. Производится обратная свёртка исходной карты диаграммой направленности антенны с тем, чтобы получить оценку карты, представляющую собой простую сумму флуктуаций температуры реликтового фона и излучения фоновых точечных источников, которую обозначим как $(\text{CMB} + \text{PS})_{\text{rec}}$:

$$(\text{CMB} + \text{PS})_{\text{rec}} = \text{CMB}_{\text{det}} * W^{-1}.$$

Эта задача решается путём оптимизации обобщённого функционала (10), поскольку искомое решение может принимать как положительные, так и отрицательные значения. В качестве исходных данных используются все фурье-компоненты измеренной карты. Линейные ограничения (5) и (6) в этом случае следует переписать в виде

$$\sum_m \sum_l x_{ml} \cos[2\pi (nm + kl)/M] w^{nk} + \eta_{nk}^r = D_{nk}^r, \quad (11)$$

$$\sum_m \sum_l x_{ml} \sin[2\pi (nm + kl)/M] w^{nk} + \eta_{nk}^i = D_{nk}^i, \quad (12)$$

где w^{nk} — значения передаточной функции антенны в фурье-области, $n = 0, 1, \dots, M-1$, $k = 0, 1, \dots, M-1$.

Отметим, что для реализации обратной свёртки может быть использован альтернативный метод максимальной энтропии с исходными данными в пространственной области измерений.

2. Производится восстановление карты фоновых точечных источников стандартным ММЭ (4) с использованием фурье-данных, расположенных в высокочастотной области (l_d, l_f) спектра измеренной карты.

При этом для достижения как можно более высокой точности восстановления точечных источников предпочтительно использовать в качестве исходных данных фурье-компоненты карты CMB_{det} , а не оценки $(\text{CMB} + \text{PS})_{\text{rec}}$, полученной в результате обратной свёртки, поскольку решение, полученное на первом этапе описываемого алгоритма, уже отягощено ошибками восстановления. Тогда уравнения, описывающие ограничения на данные, будут выглядеть так же, как и на предыдущем этапе

(см. (11), (12)), изменится только их число. Однако, несмотря на предпринимаемые предосторожности, следует ожидать, что точечные источники будут восстановлены с ошибками, поскольку в данных, используемых для их восстановления, содержится шум от реликтового излучения, которым не следует пренебрегать вплоть до $l = 2000$. Поэтому непосредственное вычитание восстановленного фона точечных источников из карты $(\text{CMB} + \text{PS})_{\text{rec}}$ может привести к более ощутимым ошибкам в оценке углового спектра мощности, чем простое удаление искажённых пикселей, как предложено Тегмарком и де Оливейра-Костой [6].

Существенное повышение точности разделения компонент достигается, как уже говорилось во втором разделе, реализацией следующих этапов.

3. Производится удаление из карты $(\text{CMB} + \text{PS})_{\text{rec}}$, полученной на первом этапе, тех пикселей, уровень искажения которых точечными источниками выше некоторого выбранного порога. При этом операция порогового ограничения с целью локализации участков искажения применяется к карте точечных источников, восстановленной на втором этапе алгоритма. Устранение искажённых участков из карты $(\text{CMB} + \text{PS})_{\text{rec}}$ приводит к выходным картам CMB с пустыми участками («дырами»).

Отметим, что порог должен быть выбран сравнительно малым, чтобы обеспечить удаление из карты как можно большего числа искажённых пикселей. Но при этом надо следить, чтобы число удаляемых пикселей не превысило критической величины, при которой последующая интерполяция реликтового фона в «дырах» ещё возможна.

4. Производится восстановление реликтового фона в «дырах» с использованием ОММЭ, модифицированного в соответствии с необходимостью учёта дополнительной информации о конечной ширине и круговой симметрии спектра мощности флуктуаций (2). Учёт конечной ширины фурье-спектра обеспечивается селекцией исходных данных с практически нулевыми величинами в области пространственных частот $l \geq l_d$.

Пусть x_{st} — искомые значения сигнала в «дырах», а a_{pq} — известные значения сигнала вне «дыр». Очевидно, сумма этих двух последовательностей составляет всю карту реликтового излучения.

В этом случае отсчёты фурье-спектра с координатами (n, k) запишутся следующим образом:

$$\sum_s \sum_t x_{st} \cos[2\pi (ns + kt)/M] + \sum_p \sum_q a_{pq} \cos[2\pi (np + kq)/M] + \eta_{nk}^r = D_{nk}^r,$$

$$\sum_s \sum_t x_{st} \sin[2\pi (ns + kt)/M] + \sum_p \sum_q a_{pq} \sin[2\pi (np + kq)/M] + \eta_{nk}^i = D_{nk}^i.$$

Приравнявая вещественные D_{nk}^r и мнимые D_{nk}^i части спектра к нулю, получим следующие ограничения, используемые в задаче интерполяции с использованием ОММЭ:

$$\sum_s \sum_t x_{st} \cos[2\pi (ns + kt)/M] + \eta_{nk}^r = A_{nk}^r, \quad (13)$$

$$\sum_m \sum_l x_{ml} \sin[2\pi (nm + kl)/M] + \eta_{nk}^i = A_{nk}^i, \quad (14)$$

где

$$A_{nk}^r = - \sum_p \sum_q a_{pq} \cos[2\pi (np + kq)/M],$$

$$A_{nk}^i = - \sum_p \sum_q a_{pq} \sin[2\pi (np + kq)/M].$$

При этом надо иметь в виду, что искомая знакопеременная последовательность x_{st} удовлетворяет соотношению (8).

Для реализации условия круговой симметрии спектра мощности требование выполнения соотношения (2) налагается после каждой итерации градиентного метода поиска экстремума соответствующего двойственного функционала Лагранжа, составленного из основного функционала ассоциативной энтропии

$$H(x, \alpha) = 1/2 \sum_n \sum_k [(\eta_{nk}^r)^2 + (\eta_{nk}^i)^2] / \sigma_{nk}^2$$

и линейных ограничений (13), (14).

В качестве альтернативного метода восстановления реликтового фона в «дырах» может быть применена модификация известного алгоритма Фьенапа, реализующего восстановление изображения только по амплитуде спектра источника при полностью неизвестной фазе [13]. В модифицированном нами итерационном алгоритме Фьенапа в ограничениях в спектральной области вместо заданной амплитуды спектра используется лишь информация о его конечной ширине и круговой симметрии (2), а в ограничениях в пространственной области используются известные значения карты.

Алгоритм представляет собой следующую последовательность действий:

1. Формируется начальное приближение карты. В качестве начального приближения мы рекомендуем использовать исходную карту с нулевыми значениями яркости в «дырах».

2. Вычисляется преобразование Фурье начального приближения.

3. Налагаются ограничения в области пространственных частот карты. Для этого полученные на шаге 2 фурье-компоненты модифицируются таким образом, чтобы спектр мощности приобрёл круговую симметрию в соответствии с соотношением (2). Для этого для каждого значения l производится усреднение квадратов амплитуд отсчётов спектра вдоль окружности радиуса l , затем эти отсчёты заменяются на значение, соответствующее полученному среднему квадрату амплитуды, при этом фазы отсчётов не изменяются. Условие пространственной ограниченности спектра выполняется за счёт того, что отсчёты с координатами, удовлетворяющими условию $l > l_d$, приравняются нулю.

4. Выполняется обратное преобразование фурье-спектра, полученного на предыдущем шаге.

5. Налагаются ограничения в пространственной области. Для этого яркость реликтового фона вне «дыр» карты заменяется на известные значения, а в «дырах» не изменяется. Кроме того, налагается ограничение на максимальное значение флуктуации. Если флуктуация яркостной температуры в какой-либо точке превышает заданное значение, флуктуация в этой точке заменяется на априорно известное максимальное возможное значение.

6. Выполняется прямое преобразование Фурье полученной на шаге 5 карты.

7. Выполнение шагов 3–6 до тех пор, пока изображение, получаемое на шаге 5, не будет соответствовать выбранному критерию точности.

Как показывают многочисленные модельные расчёты, видеоизменённый алгоритм Фьенапа довольно быстро сходится к правильному решению и зачастую является более эффективным (благодаря применению быстрого преобразования Фурье), чем алгоритм, использующий ОММЭ.

После восстановления карты флуктуаций реликтового излучения, которую обозначим как CMB_{rec} , можно получить уточнённую карту точечных источников PS_{rec} , если из карты $(\text{CMB} + \text{PS})_{\text{rec}}$, полученной на первом этапе, вычесть карту CMB_{rec} :

$$\text{PS}_{\text{rec}} = (\text{CMB} + \text{PS})_{\text{rec}} - \text{CMB}_{\text{rec}}. \quad (15)$$

Эта операция может быть полезна для решения сопутствующей задачи проекта «Planck» по построению каталога точечных источников [1].

3.3.2. Учёт пиксельного шума. При наличии значительного пиксельного шума измеренная карта реликтового излучения удовлетворяет соотношению (3). Как показывает численное моделирование, при повышении уровня аддитивного шума в данных применение только критерия (7) оказывается

недостаточным для получения удовлетворительного решения. Требуется модификация ОММЭ, обладающая более высокими стабилизирующими свойствами. В данной работе рассматривается одна из возможностей, основанная на модификации подхода, предложенного Фриденем в работе [8]. При этом вводятся новые переменные для гауссового шума с нулевым средним, принимающего как положительные, так и отрицательные значения: $N_{ml} = N_{ml}^+ + N_{ml}^-$. По аналогии с искомым сигналом вводятся соответствующие шумовые члены в функционал ассоциированной энтропии (10) и линейные ограничения (11), (12). Тогда критерий оптимизации принимает вид

$$H(x, \alpha) \equiv - \sum_m \sum_l [x_{ml}^+ \ln(\alpha x_{ml}^+) + x_{ml}^- \ln(\alpha x_{ml}^-)] - \\ - \sum_m \sum_l [N_{ml}^+ \ln(\alpha N_{ml}^+) + N_{ml}^- \ln(\alpha N_{ml}^-)] = \max \quad (16)$$

при условии

$$\sum_m \sum_l (x_{ml}^+ - x_{ml}^-) a_{ml}^{nk} w^{nk} + \sum_m \sum_l (N_{ml}^+ - N_{ml}^-) a_{ml}^{nk} = D_{nk}^r, \quad (17)$$

$$\sum_m \sum_l (x_{ml}^+ - x_{ml}^-) b_{ml}^{nk} w^{nk} + \sum_m \sum_l (N_{ml}^+ - N_{ml}^-) b_{ml}^{nk} = D_{nk}^i, \quad (18)$$

где

$$a_{ml} = \cos[2\pi (nm + kl)/M], \quad b_{ml} = \sin[2\pi (nm + kl)/M].$$

Как видно из (17), (18), сигнал и шум входят в уравнения аддитивно, но в отличие от шума коэффициенты Фурье сигнала домножены на частотную характеристику диаграммы направленности антенны w^{nk} . Если диаграмма направленности антенны имеет форму гауссовой кривой, как принято во многих проектах, в том числе и в проекте «Planck», то её частотная характеристика также будет иметь гауссову форму. Благодаря учёту частотной зависимости диаграммы направленности становится принципиально возможным разделение сигнала и шума, причём такое разделение тем точнее, чем бóльшим числом дискретных отсчётов представлена на карте диаграмма направленности антенны.

4. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

В этом разделе представлены результаты применения предложенного метода к модельным данным. Для создания двумерной карты флуктуаций реликтового фона размером $7,5^\circ \times 7,5^\circ$ была использована функция углового спектра мощности, соответствующая Λ CDM космологической модели Вселенной с параметрами $\Omega_b h^2 = 0,02$; $\Omega_\Lambda = 0,65$; $\Omega_m = 0,3$; $h = 0,65$; $n = 1$ [1]. Карта, удовлетворяющая гауссовой статистике, была численно сгенерирована в соответствии с уравнениями (1) и (2) и нормирована таким образом, чтобы пиковое значение флуктуации реликтового фона было равно единице. Размер дискретных карт был выбран равным 64×64 пикселей, чему соответствует верхняя частота углового спектра мощности $l = 1536$. Предложенный метод разделения компонент был протестирован на двух реализациях фона точечных источников, имеющих сравнимую дисперсию, но разный динамический диапазон яркости. Первая карта PS1 включала сравнительно небольшое число достаточно ярких



Рис. 1

неразрешённых источников. Вторая карта PS2 включала на порядок большее число источников с динамическим диапазоном около 100. Была сгенерирована карта гауссового пиксельного шума с нулевым средним и дисперсией, сравнимой с дисперсией заданной карты реликтового излучения. Дисперсии всех карт приведены в табл. 1. Диаграмма направленности антенны, нормированная таким образом, чтобы её амплитуда была равна единице, представляет собой круговую симметричную гауссову кривую с диаметром сечения 3 пиксела на уровне 0,5 от максимального значения.

Ниже приводятся результаты трёх экспериментов, целью которых было:

Т а б л и ц а 1

1) Исследование качества восстановления карт в зависимости от размера области (l_d, l_f) , или (u, v) -покрытия, используемого в качестве исходных данных для восстановления точечных источников. Этот эксперимент определяет возможность изменения нижней границы l_d частотной области, когда предельная частота углового спектра мощности флуктуаций реликтового фона известна только приблизительно.

Дисперсии заданных карт	
Карта	Дисперсия
СМВ	390,156
PS1	1 692,944
PS2	1 556,411
(СМВ + PS1) * W	5 327,111
N	345,970

2) Сравнение предложенного четырёхэтапного алгоритма с более простым методом, когда оценка карты реликтового фона находится путём вычитания карты точечных источников, восстановленной на втором этапе, из оценки карты $(СМВ + PS)_{гес}$, полученной на первом этапе. Для упрощения этот и рассмотренный ранее четырёхэтапный алгоритмы разделения компонент обозначим как метод 1 и метод 2 соответственно.

3) Исследование метода 2 при наличии в данных значительного пиксельного шума.

Первый эксперимент был выполнен для реализации точечных источников PS1. Пиксельный шум в данных отсутствовал. Карта флуктуаций реликтового фона была сгенерирована в соответствии с выражениями (1) и (2). Исходные карты флуктуаций реликтового излучения, фона точечных источников и измеренная карта $СМВ_{дет} = (СМВ + PS) * W$ показаны на рис. 1а, б и в соответственно. Для выполнения операции обратной свёртки карты $СМВ_{дет}$ диаграммой направленности W использован ОММЭ, описанный в разделе 3.2. В результате получена карта $(СМВ + PS)_{гес}$ с очень высокой точностью благодаря тому, что входные данные свободны от ошибок измерения и для восстановления использовался весь массив фурье-данных, без каких-либо пропусков в интервале $l \leq l_f$.

Результаты эксперимента приведены на рис. 2. На верхних трёх рисунках показаны варианты выбранного (u, v) -покрытия в фурье-области исходных данных, обозначенные буквами а, б и в (в дальнейшем также используем обозначения UV1, UV2 и UV3 соответственно). Размер покрытия увеличивается слева направо, причём покрытия на рис. 2а и б отличаются незначительно, а покрытие на рис. 2в

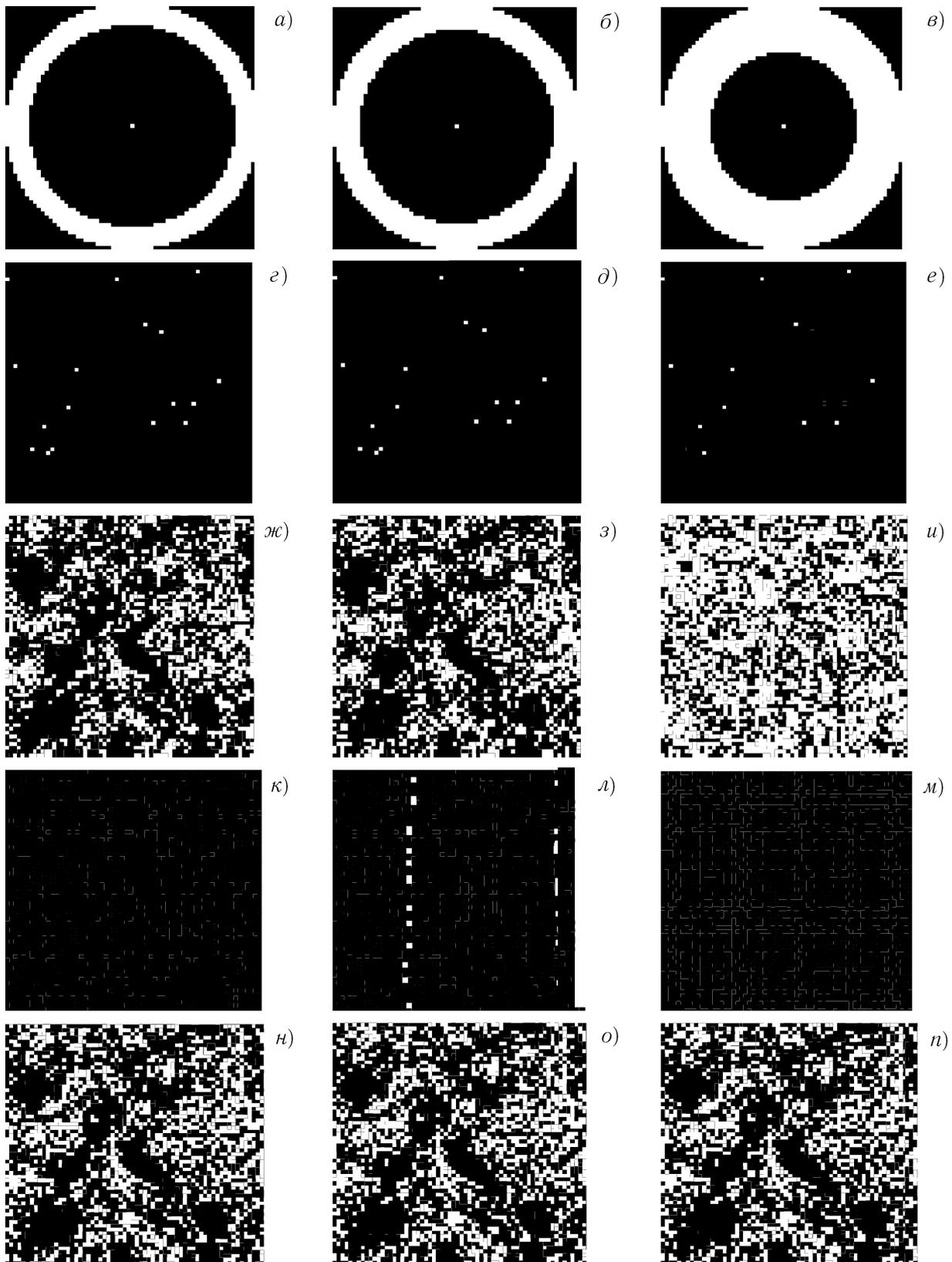


Рис. 2

Т а б л и ц а 2

Дисперсии восстановленных и остаточных карт реликтового излучения

PS-карта	(uv)-покрытие	Наличие пиксельного шума	Метод восстановления	Дисперсия карты CMB_{rec}	Дисперсия остаточной карты
PS1	UV1	—	1	399,250	16,330
PS1	UV1	—	2	390,156	0,000
PS1	UV2	—	1	398,330	16,630
PS1	UV2	—	2	390,156	0,000
PS1	UV3	—	1	1075,000	660,330
PS1	UV3	—	2	390,156	0,000
PS2	UV1	—	1	400,946	19,450
PS2	UV1	—	2	389,360	0,095
PS1	UV1	+	2	366,420	19,330

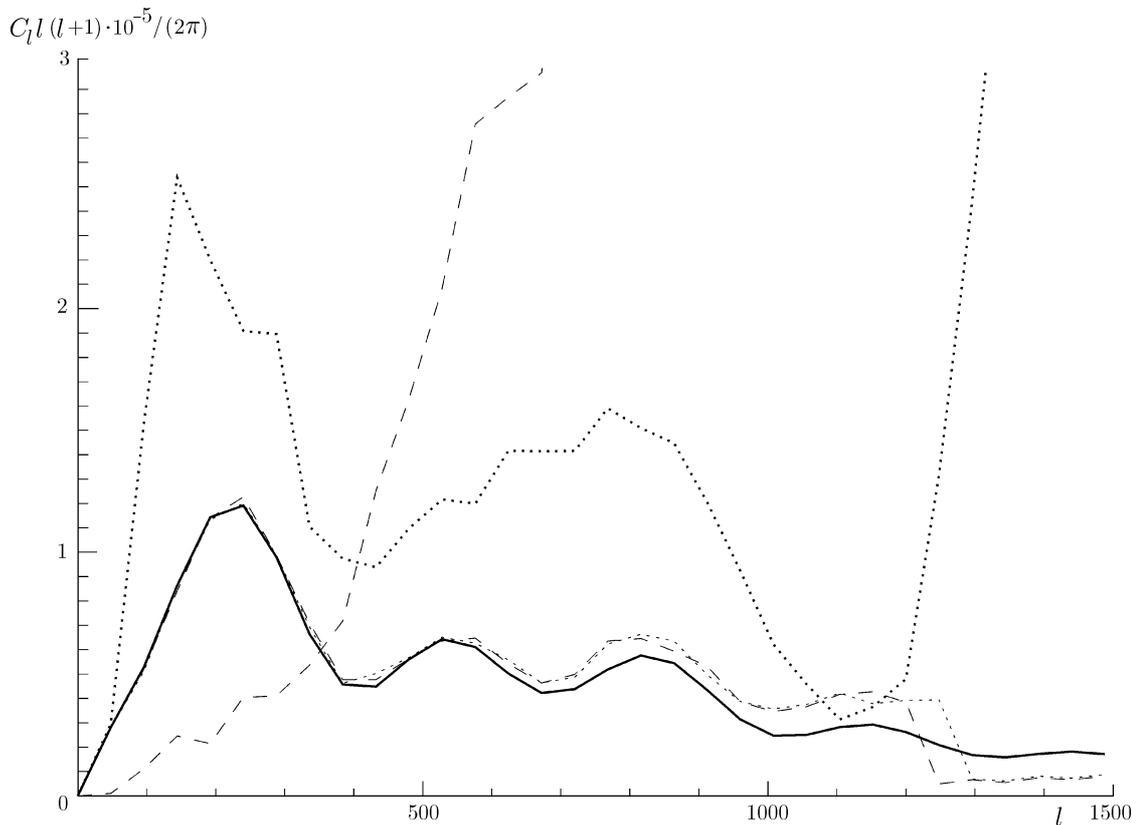


Рис. 3

имеет существенно больший размер. На рисунках под каждым из покрытий приведены соответствующие им результаты восстановления. Во втором ряду приведены изображения фона точечных источников (рис. 2*e–e*), восстановленные с использованием стандартного алгоритма ММЭ. Видно, что с увеличением размера (u, v)-области ухудшается качество восстановления точечных источников из-за увеличения в исходных данных шумовой составляющей, связанной с вкладом реликтового фона. Соответствующие оценки карт реликтового фона, полученные прямым вычитанием восстановленного фона точечных источников из карт $(CMB + PS)_{rec}$ (метод 1), представлены в третьем ряду

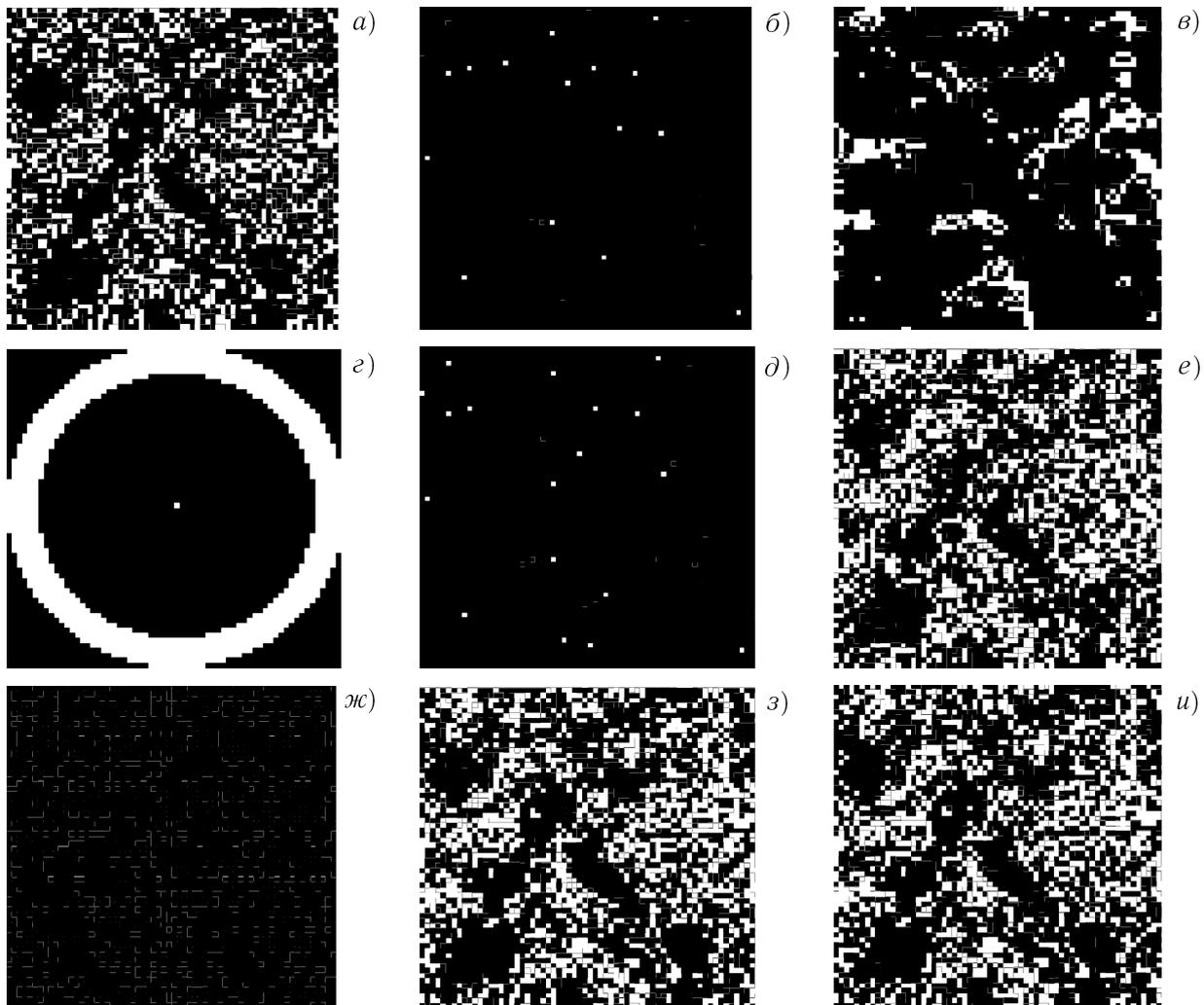


Рис. 4

(рис. 2ж–и). Наихудший результат относится к случаю наибольшего размера (u, v) -покрытия, когда шумовой вклад от реликтового излучения является наибольшим. Очевидно, существует оптимальный размер (u, v) -покрытия, обеспечивающий минимум искажений при достаточном для восстановления фона точечных источников количестве данных.

Для улучшения качества восстановления мы использовали третий и четвертый этапы предложенного алгоритма (метод 2). В четвертом ряду на рис. 2 показаны распределения пикселей, удаленных из карты $(\text{CMB} + \text{PS})_{\text{rec}}$ (рис. 2к–м). Эти распределения получены в результате применения операции амплитудного ограничения к восстановленным картам точечных источников. Порог был выбран равным 10 % от пикового значения флуктуации реликтового фона. Как видно из рис. 2к–м, плотность точек, соответствующих удаленным пикселям, существенно выше плотности распределения точечных источников на исходной карте (см. рис. 1б), и эта плотность тем выше, чем хуже качество восстановления точечных источников из-за появления ложных источников, превосходящих по амплитуде выбранный порог. Ложные источники представляют собой ошибки восстановления, вызванные шумовым вкладом реликтового фона. Увеличивая порог, мы, очевидно, можем уменьшить плотность удаляемых пикселей, соответствующих ошибкам восстановления, но при этом мы рискуем пропустить пиксели,

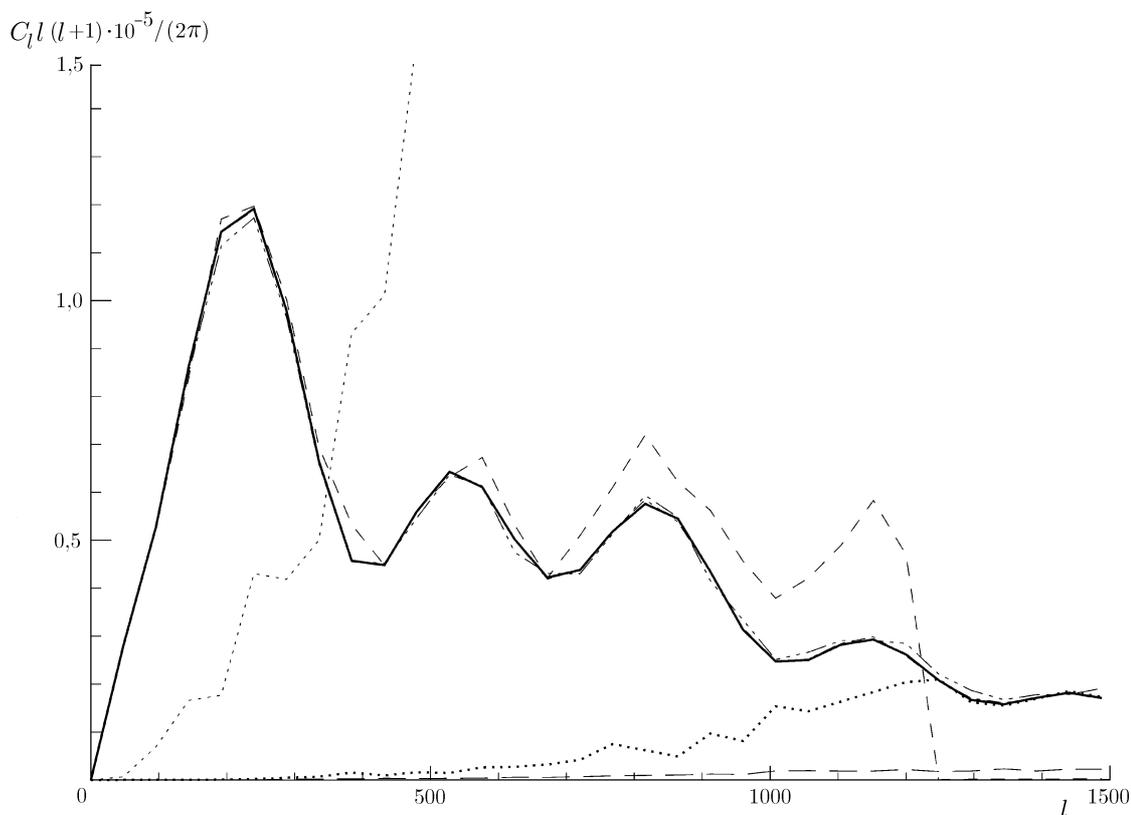


Рис. 5

реально искажённые фоновыми источниками. В каждом конкретном случае существует оптимальный порог, который обеспечивает наилучшее восстановление карты на следующем этапе. Этот порог в каждом случае может быть подобран эмпирически. В результате удаления искажённых пикселей мы получаем карты реликтового излучения с «дырами». Карты с восстановленной яркостью реликтового фона в «дырах» показаны в нижнем ряду (рис. 2*n*–*n*). Анализ полученных изображений показывает, что в этом эксперименте при отсутствии пиксельного шума и использовании достаточно простой реализации карты точечных источников удалось получить предельную точность разделения фоновых компонент. Угловые спектры мощности заданных и восстановленных карт приведены на рис. 3; сравнивая их, можно оценить точность восстановления. Здесь сплошной линией показан угловой спектр мощности заданной карты реликтового излучения; штриховой линией, уходящей вверх на высоких частотах, — спектр заданной карты точечных источников; мелким пунктиром — спектр восстановленной карты реликтового излучения (метод 1) с использованием данных UV1; штриховой линией, проходящей рядом со сплошной, — спектр восстановленной карты реликтового излучения (метод 1) с использованием данных UV2; жирным пунктиром — спектр восстановленной карты реликтового излучения (метод 1) с использованием данных UV3. Угловые спектры мощности восстановленных карт реликтового излучения, полученные с использованием метода 2, на рис. 3 совпадают с угловым спектром заданной карты.

Второй эксперимент был посвящён исследованию качества разделения компонент на примере более сложной реализации фона точечных источников PS2. Здесь также использовались данные, свободные от пиксельного шума. Результаты моделирования показаны на рис. 4, где *a* — заданная карта реликтового излучения, *b* — заданное распределение точечных источников (PS2), *v* — измерен-

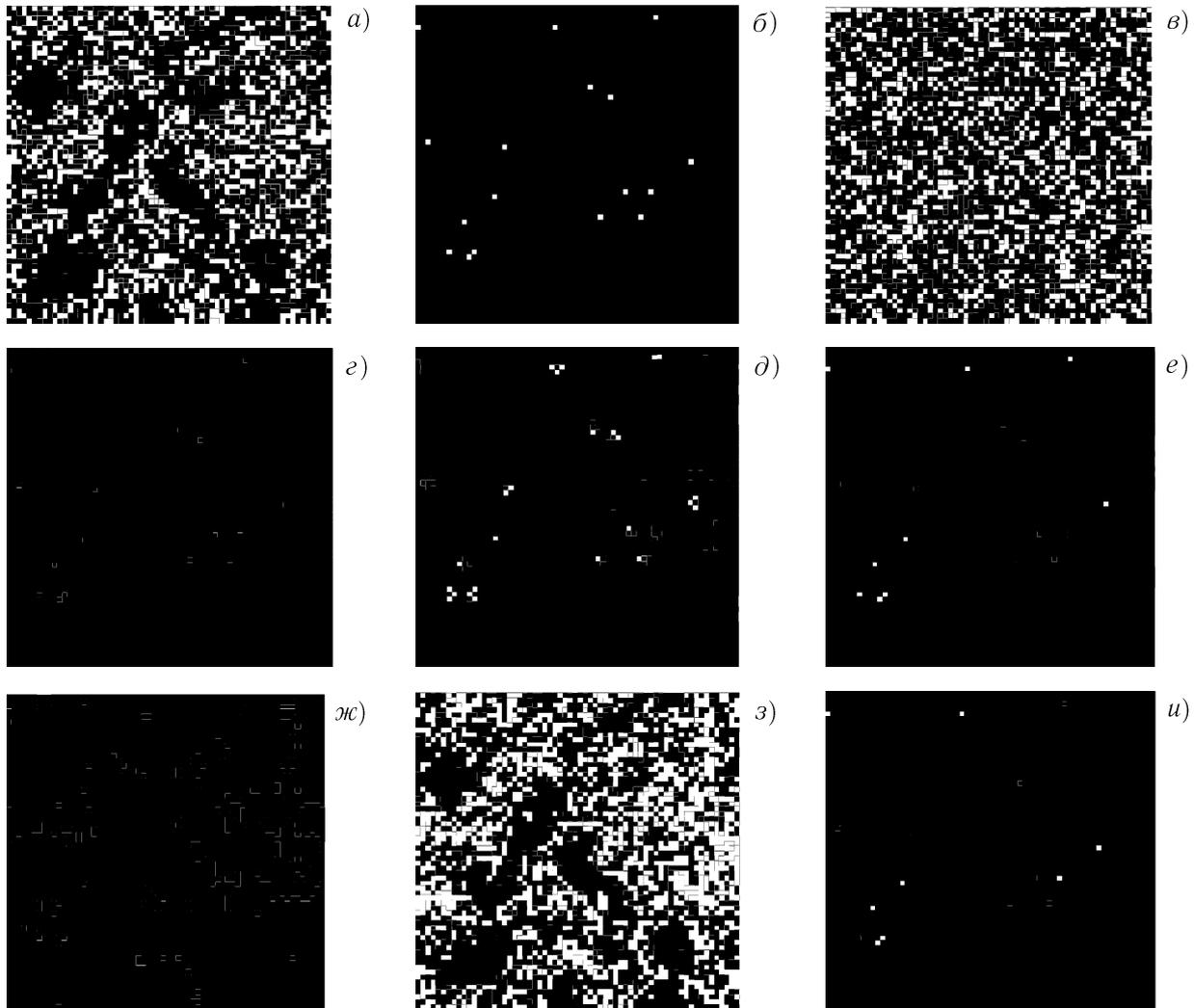


Рис. 6

ная карта $\text{CMB}_{\text{det}} = (\text{CMB} + \text{PS2}) * W$, z — (u, v) -покрытие исходных данных, использованных для восстановления точечных источников, d — карта точечных источников, восстановленная стандартным ММЭ из фурье-данных карты CMB_{det} , e — карта реликтового фона, полученная путём прямого вычитания восстановленного фона точечных источников из карты $(\text{CMB} + \text{PS})_{\text{rec}}$ (метод 1), $ж$ — распределение искажённых пикселей, удалённых из карты, $з$ и $и$ — искомые карты реликтового фона, полученные в результате интерполяции реликтового фона в «дырах» (метод 2) с использованием двух различных нелинейных алгоритмов восстановления — модифицированного ОММЭ и модифицированного алгоритма Фьенапа. Как видно, изображения на рис. 4з и $и$ отличаются незначительно. Угловые спектры мощности заданных и восстановленных карт, а также остаточных карт, полученных в результате вычитания соответствующих восстановленных карт из заданной, приведены на рис. 5. Здесь сплошной линией показан угловой спектр мощности заданной карты реликтового излучения; мелким пунктиром — спектр заданной карты точечных источников; короткими штрихами — спектр восстановленной карты реликтового излучения (метод 1); жирным пунктиром — спектр остаточной карты, соответствующей методу 1; штрих-пунктирной линией — спектр восстановленной карты реликтового

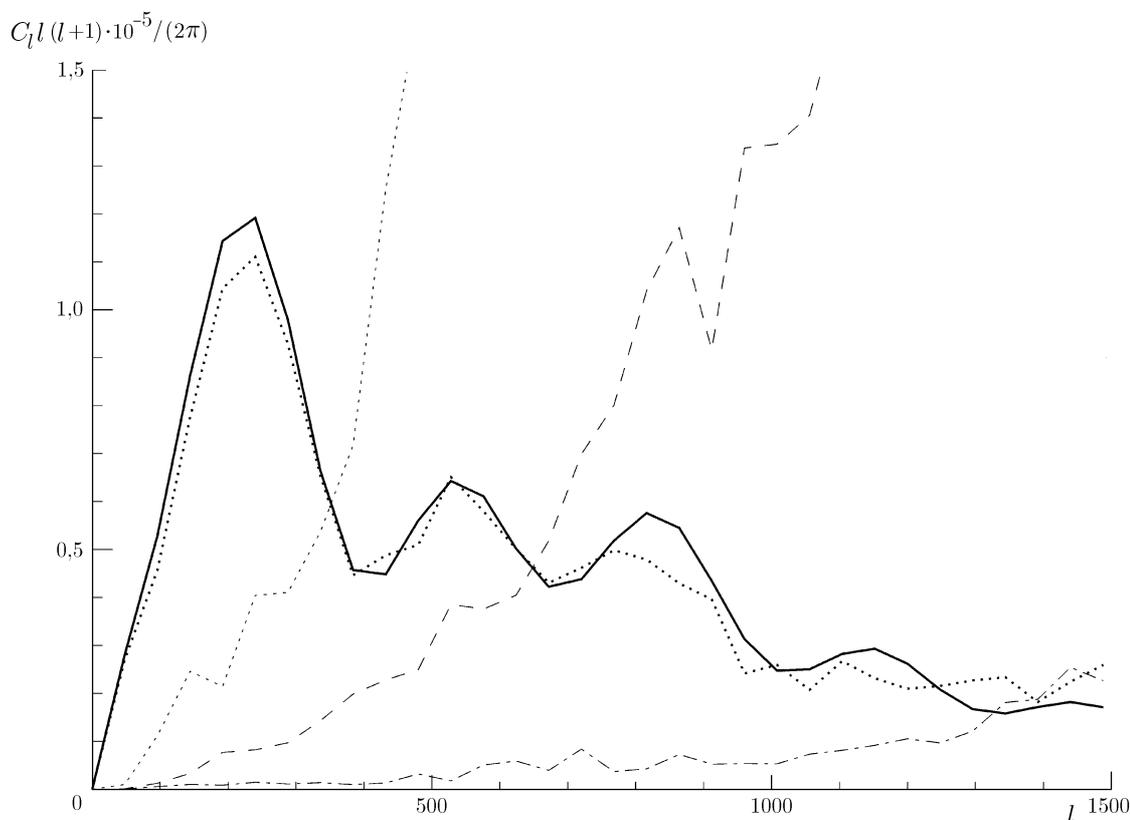


Рис. 7

излучения (метод 2); длинными штрихами — спектр остаточной карты, соответствующей методу 2.

Сравнение полученных результатов с результатами предыдущего эксперимента для одного и того же (u, v) -покрытия показывает некоторое ухудшение качества восстановления, что вызвано наличием в карте PS2 большого количества слабых источников, сравнимых по амплитуде с флуктуациями реликтового фона. Действительно, некоторое число слабых источников, яркость которых ниже выбранного порога, не удалось удалить из карты $(\text{СМВ} + \text{PS})_{\text{rec}}$. Дальнейшее понижение порога оказалось нецелесообразным, поскольку привело к чрезмерному увеличению числа удаляемых пикселей и сделало невозможным дальнейшее восстановление яркости реликтового фона в «дырах». Тем не менее качество восстановления получилось достаточно высоким, гораздо выше, чем при использовании метода 1.

Последний, третий эксперимент был посвящён разделению компонент при наличии в измеренных картах значительного пиксельного шума. В этом случае карта удовлетворяет соотношению (3). В качестве реализации точечных источников мы использовали карту PS1. Для деконволюции исходной карты диаграммой направленности антенны, в отличие от предыдущих экспериментов, мы использовали модификацию ОММЭ, описанную в разделе 3.3, учитывающую как шум, так и сигнал. При этом отношение сигнал/шум до процедуры деконволюции составляет 3,92 (см. табл. 1). Результаты моделирования приведены на рис. 6, где a — заданная карта реликтового излучения, b — заданная карта фона точечных источников, v — заданная карта пиксельного шума, аддитивно входящего в измеренную карту, z — заданная карта $\text{СМВ} + \text{PS}$, d — измеренная карта $\text{СМВ}_{\text{det}} = (\text{СМВ} + \text{PS}) * W + N$, e — карта $(\text{СМВ} + \text{PS})_{\text{rec}}$, восстановленная с помощью ОММЭ в соответствии с уравнениями (16), (17) и (18), $ж$ — распределение удалённых пикселей, полученное путём амплитудного ограничения карты $(\text{СМВ} + \text{PS})_{\text{rec}}$, $з$ — карта СМВ_{rec} , полученная в результате ОММЭ-интерполяции реликтового фона в «дырах», $и$ — уточнённая карта фоновых точечных источников, полученная в результате вычитания

восстановленной карты реликтового фона CMB_{rec} из карты $(CMB + PS)_{\text{rec}}$ в соответствии с уравнением (15). Угловые спектры мощности заданных и восстановленных карт приведены на рис. 7. Здесь сплошной линией показан угловой спектр мощности заданной карты реликтового излучения; мелким пунктиром — спектр заданной карты точечных источников; штриховой линией — спектр заданной карты пиксельного шума; жирным пунктиром — спектр карты реликтового излучения, восстановленной по методу 2 с помощью ОММЭ; штрих-пунктирной линией — спектр остаточной карты.

Из анализа полученных результатов следует, что и в случае сильно зашумлённых данных возможно устранение фоновых точечных источников с приемлемой точностью. Результаты всех трёх экспериментов сведены в табл. 2, где приводятся дисперсии восстановленных и остаточных карт реликтового излучения, характеризующие точность устранения фоновых компонент.

ВЫВОДЫ

В работе предложен и успешно протестирован метод высокоточного устранения фона неразрешённых внегалактических источников из наблюдений реликтового излучения. Результаты моделирования демонстрируют приемлемую точность восстановления углового спектра мощности реликтового излучения даже при наличии в данных наблюдений значительного пиксельного шума.

Предложенный метод отличается от известных тем, что в нём используются свойства аналитичности образов Фурье финитных функций применительно к флуктуациям реликтового фона. Поскольку аналитичность гарантирует возможность восстановления всей функции по известной её части, становится возможным восстановление реликтового излучения в областях искажений фоновыми источниками. Для этого локализуются места расположения фоновых источников, удаляются искажённые участки карты (в этих местах в карте образуются «дыры»), а затем производится восстановление реликтового фона в «дырах» с использованием остальной карты, а также дополнительной информации о конечной ширине и круговой симметрии пространственного спектра мощности флуктуаций реликтового излучения, подчиняющихся гауссовой статистике.

Высокоточная реализация предложенного метода может быть достигнута благодаря применению эффективных нелинейных методов восстановления изображений, обладающих ярко выраженными стабилизирующими и сверхразрешающими свойствами. В качестве таких методов в работе использованы хорошо известный в физике и астрономии метод максимальной энтропии и его модификации, приспособленные к виду восстанавливаемого сигнала.

Автор выражает признательность Ю. Н. Парийскому, И. Д. Новикову и П. Д. Насельскому за плодотворное обсуждение результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Bersanelli M. et al. // COBRAS/SAMBA: ESA Report D/SCI. 1996. P. 3.
2. Tegmark M., Efstathiou G. // MNRAS. 1996. V. 281. P. 1 297.
3. Bond J. R., Efstathiou G. // MNRAS. 1987. V. 226. P. 655.
4. Jones A. W. et al. // astro-ph/9810236. 1998.
5. Hobson M. P. et al. // astro-ph/9810241. 1998.
6. Tegmark M., de Oliveira-Costa A. // astro-ph/9802123. 1998.
7. Högbom J. M. // Astron. Astrophys. Suppl. 1974. No. 15. P. 417.
8. Frieden B. R. // JOSA. 1972. V. 62. P. 511.
9. Skilling J., Gull S. F. // Digital Image Restoration / Ed. by H. C. Andrews, B. R. Hunt. Englewood Cliffs: Prentice-Hall, 1977.
10. Байкова А. Т. // Изв. вузов. Радиофизика. 1991. Т. 34, № 8. С. 919.

11. Bajkova A. T. // Astron. and Astrophys. Transactions. 1992. V. 1. P. 313.
12. Frieden B. R., Bajkova A. T. // Appl. Opt. 1994. V. 33. P. 219.
13. Fienup J. R. // Opt. Lett. 1972. V. 3. P. 27.

Институт прикладной астрономии РАН,
г. Санкт-Петербург, Россия

Поступила в редакцию
6 декабря 2001 г.

**A METHOD FOR SEPARATING THE POINT-SOURCE FOREGROUND
FROM THE COSMIC MICROWAVE BACKGROUND MAPS**

A. T. Bajkova

We propose a new technique for the high-accuracy reconstruction of the cosmic microwave background (CMB) anisotropy from observations contaminated by unresolved extragalactic point sources and pixel noise. The principal difference of our methods from the methods suggested before consists in reconstructing the CMB temperature fluctuations in regions contaminated by point sources with allowance for the *a priori* information on the boundedness and spherical symmetry of the Gaussian power spectrum of the CMB anisotropy, which satisfies the Gaussian statistics. The problems of deconvolving the initial maps, localizing point sources, and interpolating the CMB values in “holes” are solved using effective modifications of the well-known maximum entropy method. The results of simulations demonstrate the possibility of achieving the highest accuracy.

УДК 52.77+523.98+533.951+537.872

ВЛИЯНИЕ НЕОДНОРОДНОСТЕЙ МАГНИТНОГО ПОЛЯ И КОНЦЕНТРАЦИИ ПЛАЗМЫ НА РАСПРОСТРАНЕНИЕ И УСИЛЕНИЕ РАДИОИЗЛУЧЕНИЯ В СОЛНЕЧНОЙ КОРОНЕ

А. А. Кузнецов, В. Г. Власов

Исследуется мазерная генерация миллисекундных всплесков солнечного радиоизлучения (спайков) на циклотронном резонансе быстрой необыкновенной волны в неоднородной среде. Показано, что неоднородность магнитного поля с типичными для солнечной короны параметрами резко ограничивает длительность усиления электромагнитных волн за счёт выхода их из резонанса в пространстве волновых векторов и, таким образом, делает возможным формирование неустойчивого распределения электронов. Эффективная генерация излучения становится возможной только в тех локальных областях, где мелкомасштабная неоднородность концентрации плазмы (с характерными масштабами от десятков до сотен километров) компенсирует влияние неоднородности магнитного поля. Учёт влияния неоднородностей среды позволяет объяснить пространственные и временные характеристики спайков.

ВВЕДЕНИЕ

Миллисекундные всплески солнечного радиоизлучения (радиоспайки) привлекают к себе внимание исследователей уже в течение нескольких десятков лет [1, 2]. Одним из наиболее вероятных механизмов генерации спайков считается циклотронная неустойчивость быстрой необыкновенной волны, обеспечивающей работу мазера на циклотронном резонансе (МЦР) [3]. Обусловлено это высокой яркостной температурой спайков, их малыми временными и спектральными масштабами, а также преобладающей поляризацией спайков, соответствующей X-моду.

Вместе с тем привлечение МЦР для интерпретации спайков приводит к ряду проблем. С точки зрения теоретика весьма важным является вопрос формирования неустойчивого электронного распределения (конуса потерь) и его стабилизации — так называемая проблема начального значения. Дело в том, что благодаря высокой эффективности мазерного механизма характерное время квазилинейной релаксации конуса потерь значительно меньше, чем характерное время формирования указанного распределения. Таким образом, неустойчивое распределение электронов может сформироваться только в случае его стабилизации.

Ряд вопросов возникает при интерпретации наблюдаемых особенностей спайков. Так, спайки возникают большими группами (кластерами). Внутри кластера они распределены, на первый взгляд, случайным образом как по частоте, так и во времени; размеры области, являющейся источником кластера спайков, значительно превосходят размеры источника отдельного всплеска (не превышающие 200 км). Важно отметить, что в течение всего времени существования кластера спайков рентгеновское излучение не прерывается.

Мазерный механизм генерации излучения в солнечной короне в приближении локально однородной среды детально исследуется во многих работах, при этом основное внимание уделяется определению инкрементов электромагнитных волн в МЦР и динамики релаксации неустойчивого распределения электронов [4–7]. В то же время учёт влияния неоднородностей магнитного поля и концентрации плазмы не только делает модель источника спайков более реалистичной, но и позволяет объяснить многие наблюдаемые особенности радиоизлучения. Продольный градиент магнитного поля необходим для формирования конуса потерь, причём существенную роль в этом процессе играют не только

крупномасштабная монотонная неоднородность магнитного поля в корональной петле, но и мелко-масштабные неоднородности (турбулентность) [8]. Мелкомасштабная нерегулярная неоднородность магнитного поля снижает как эффективность усиления радиоизлучения в МЦР, так и степень поглощения этого излучения в гирорезонансных слоях, в результате чего условия выхода радиоизлучения из короны становятся более благоприятными [9]. Влияние мелкомасштабной неоднородности магнитного поля на спектральные характеристики спайков обсуждается в работе [10].

При анализе генерации излучения в неоднородной среде необходимо учитывать ещё одну важную особенность. Как показывают исследования мазерного механизма генерации аврорального километрового радиоизлучения Земли [11–14], основным эффектом, влияющим на данный процесс, является движение электромагнитных волн в пространстве волновых векторов. Под влиянием неоднородностей магнитного поля и концентрации плазмы волновой вектор изменяется как по величине, так и по направлению, в результате чего волна выходит из резонанса с электронным пучком. Именно этот эффект определяет основные особенности аврорального километрового радиоизлучения Земли.

На наш взгляд, учёт влияния неоднородностей среды позволяет естественным образом решить многие из проблем интерпретации спайков, в том числе проблемы формирования кластерной структуры спектров спайков и неустойчивого распределения электронов и его стабилизации.

В данной работе исследуется распространение и усиление первой гармоники необыкновенной волны (X1-моды) в неоднородной среде в типичных для солнечной короны условиях. С помощью уравнений Гамильтона находятся траектории движения электромагнитных волн в обычном пространстве и в пространстве волновых векторов, а также коэффициенты усиления излучения. Временной интервал, в течение которого происходит усиление волн, определяется как с помощью численных расчётов, так и с помощью аналитических оценок. Исследуется зависимость коэффициента усиления волн от параметров неоднородностей магнитного поля и концентрации плазмы и находятся условия, при которых генерация излучения в МЦР появляется и прекращается. Предлагается сценарий генерации спайков при наличии крупномасштабной неоднородности магнитного поля и мелкомасштабной неоднородности концентрации плазмы.

1. РАСПРОСТРАНЕНИЕ ЭЛЕКТРОМАГНИТНЫХ ВОЛН В НЕОДНОРОДНОЙ СРЕДЕ

1.1. Общие уравнения распространения

Дисперсионное уравнение высокочастотных электромагнитных волн имеет вид

$$F = BN^4 + CN^2 + D = 0, \quad (1)$$

где

$$\begin{aligned} B &= -(1 - V - U + UV \cos^2 \theta), & C &= 2(1 - V)^2 + UV \cos^2 \theta - U(2 - V), \\ D &= (1 - V)[U - (1 - V)^2], & U &= (\omega_H/\omega)^2, & V &= (\omega_p/\omega)^2, \end{aligned} \quad (2)$$

ω , ω_H и ω_p — частота излучения, электронные циклотронная и плазменная частоты соответственно, $N = ck/\omega$ — показатель преломления волны, θ — угол между волновым вектором \mathbf{k} и магнитным полем \mathbf{H} , c — скорость света в вакууме.

Распространение волнового пакета, характеризующегося волновым вектором \mathbf{k} и пространственной координатой \mathbf{r} , описывается системой уравнений Гамильтона:

$$\dot{\mathbf{r}} = \frac{\partial \omega}{\partial \mathbf{k}} = -\frac{\partial F}{\partial \mathbf{k}} \bigg/ \frac{\partial F}{\partial \omega}, \quad \dot{\mathbf{k}} = -\frac{\partial \omega}{\partial \mathbf{r}} = \frac{\partial F}{\partial \mathbf{r}} \bigg/ \frac{\partial F}{\partial \omega}, \quad (3)$$

где использовано правило дифференцирования неявной функции. Из дисперсионного уравнения (1) получаем

$$\begin{aligned} \frac{\partial F}{\partial \omega} &= \frac{2}{\omega} \{ N^4 (-V - U + 2UV \cos^2 \theta - 2B) + \\ &+ N^2 [4V(1-V) - 2UV(1 + \cos^2 \theta) + 2U - C] + U(2V - 1) - 3V(1 - V)^2 \}, \\ \frac{\partial F}{\partial k_{\parallel}} &= \frac{2cN_{\parallel}}{\omega} [N^2 (-UV \sin^2 \theta + 2B) + UV \sin^2 \theta + C], \\ \frac{\partial F}{\partial k_{\perp}} &= \frac{2cN_{\perp}}{\omega} [N^2 (UV \cos^2 \theta + 2B) - UV \cos^2 \theta + C], \\ \frac{\partial F}{\partial r_{\parallel, \perp}} &= N^4 \left(\frac{V}{L_{p\parallel, \perp}} + \frac{2U}{L_{H\parallel, \perp}} - \lambda_{\parallel, \perp} UV \cos^2 \theta \right) + \\ &+ N^2 \left[\frac{4V(V-1)}{L_{p\parallel, \perp}} - \frac{4U}{L_{H\parallel, \perp}} + \lambda_{\parallel, \perp} UV(1 + \cos^2 \theta) \right] + \frac{3V(1-V)^2}{L_{p\parallel, \perp}} + \frac{2U}{L_{H\parallel, \perp}} - \lambda_{\parallel, \perp} UV, \\ \lambda_{\parallel} &= \frac{2}{L_{H\parallel}} + \frac{1}{L_{p\parallel}}, \quad \lambda_{\perp} = \frac{2}{L_{H\perp}} + \frac{1}{L_{p\perp}}, \end{aligned} \quad (4)$$

где k_{\parallel} , k_{\perp} , r_{\parallel} , r_{\perp} — компоненты волнового вектора и пространственной координаты волнового пакета вдоль (\parallel) и поперёк (\perp) направления магнитного поля, $N_{\parallel} = ck_{\parallel}/\omega$, $N_{\perp} = ck_{\perp}/\omega$. Масштабы неоднородности концентрации плазмы ($L_{p\parallel}$, $L_{p\perp}$) и магнитного поля ($L_{H\parallel}$, $L_{H\perp}$), входящие в (4), определяются следующим образом:

$$L_{p\parallel} = \frac{n}{\partial n / \partial r_{\parallel}}, \quad L_{p\perp} = \frac{n}{\partial n / \partial r_{\perp}}, \quad L_{H\parallel} = \frac{H}{\partial H / \partial r_{\parallel}}, \quad L_{H\perp} = \frac{H}{\partial H / \partial r_{\perp}}, \quad (5)$$

где n — концентрация плазмы.

1.2. Распространение X1-моды

Для расчёта траекторий движения электромагнитных волн в данной работе используются приведённые выше точные выражения для временных производных координаты и волнового вектора. Для аналитического исследования траекторий движения волн запишем приближённые выражения для указанных производных, а также приближённое дисперсионное уравнение X1-моды, которые следуют из (1)–(4) при $\omega \approx \omega_H$ [14]:

$$N^2 \simeq \frac{\Delta_1 - V}{\Delta_1 - (V/2) \sin^2 \theta}, \quad (6)$$

$$\dot{k}_{\parallel} \simeq -\omega \frac{\delta_{\parallel}}{L_{H\parallel}} \frac{1 - N^2}{1 - N^2 + 2\Delta_1}, \quad \dot{k}_{\perp} \simeq -\omega \frac{\delta_{\perp}}{L_{H\perp}} \frac{1 - N^2}{1 - N^2 + 2\Delta_1}, \quad (7)$$

$$\dot{r}_{\parallel} \simeq \frac{2c\Delta_1 N_{\parallel}}{1 - N^2 + 2\Delta_1}, \quad \dot{r}_{\perp} \simeq \frac{2c\Delta_1 \mu_{\perp} N_{\perp}}{1 - N^2 + 2\Delta_1}, \quad (8)$$

$$\Delta_1 = \frac{\omega - \omega_H}{\omega}, \quad \delta_{\parallel} = 1 + \Delta_1 \frac{L_{H\parallel}}{L_{p\parallel}}, \quad \delta_{\perp} = 1 + \Delta_1 \frac{L_{H\perp}}{L_{p\perp}}, \quad \mu_{\perp} = \frac{1 + N_{\parallel}^2}{2 - N_{\perp}^2}. \quad (9)$$

Вектор $\dot{\mathbf{k}}$, определяющий направление и скорость поворота волны, направлен под углом θ_e к магнитному полю:

$$\operatorname{tg} \theta_e = \frac{\dot{k}_\perp}{\dot{k}_\parallel} \simeq \frac{L_{H\parallel} \delta_\perp}{L_{H\perp} \delta_\parallel}. \quad (10)$$

Если плазма однородна ($L_{p\parallel} = \infty$, $L_{p\perp} = \infty$), то вектор $\dot{\mathbf{k}}$ направлен точно противоположно градиенту магнитного поля. В этом случае электромагнитная волна изменяет своё направление, поворачиваясь в сторону областей с более слабым магнитным полем, а угол θ_e определяет направление наибо́льшего убывания магнитного поля. Аналогично в случае однородного магнитного поля волна поворачивает в сторону областей с менее плотной плазмой. В общем случае направление движения волны определяется совместным влиянием неоднородностей магнитного поля и концентрации плазмы, но влияние неоднородности концентрации ослаблено наличием малого параметра Δ_1 .

Особенно заметно влияние неоднородностей проявляется вблизи точки отсечки (при $N \ll 1$), т. к. при этом длина волнового вектора стремится к нулю, а его временная производная, напротив, максимальна. Поэтому следует ожидать (и численные расчёты это подтверждают), что направление распространения волны будет быстро изменяться и стремиться к θ_e . Например, если магнитное поле однородно в поперечном направлении и неоднородно в продольном, то волна будет распространяться практически параллельно магнитному полю независимо от своего первоначального направления. Другим важным эффектом является изменение величины волнового вектора. Даже если первоначально групповая скорость волны была малой, увеличение длины волнового вектора приводит к соответствующему росту групповой скорости и ускорению пространственного выхода волны из области генерации. По мере увеличения показателя преломления влияние неоднородностей падает (скорость изменения волнового вектора стремится к нулю, а групповая скорость — к скорости света в вакууме), и волна покидает область генерации в направлении, близком к θ_e (тем точнее, чем меньше была первоначальная длина волнового вектора).

1.3. Усиление электромагнитных волн

Мы рассматриваем процесс генерации излучения в линейном приближении, при этом спектральная плотность энергии электромагнитных волн W_f экспоненциально нарастает со временем (или убывает, в зависимости от знака инкремента γ):

$$W_f = W_{f0} \exp \Lambda, \quad \Lambda(t) = 2 \int_0^t \gamma(t') dt', \quad (11)$$

где W_{f0} — начальная спектральная плотность энергии, определяемая тепловыми колебаниями или некогерентными процессами излучения. Таким образом, коэффициент усиления волны определяется интегралом от инкремента вдоль траектории движения волны в обычном пространстве и в пространстве волновых векторов.

2. МОДЕЛЬ ОБЛАСТИ ГЕНЕРАЦИИ СПАЙКОВ

Определим коэффициент усиления электромагнитных волн в МЦР при заданном модельном распределении электронов и заданных неоднородностях концентрации плазмы и магнитного поля. Покажем, что варьирование параметров неоднородностей среды приводит к существенному изменению коэффициента усиления излучения.

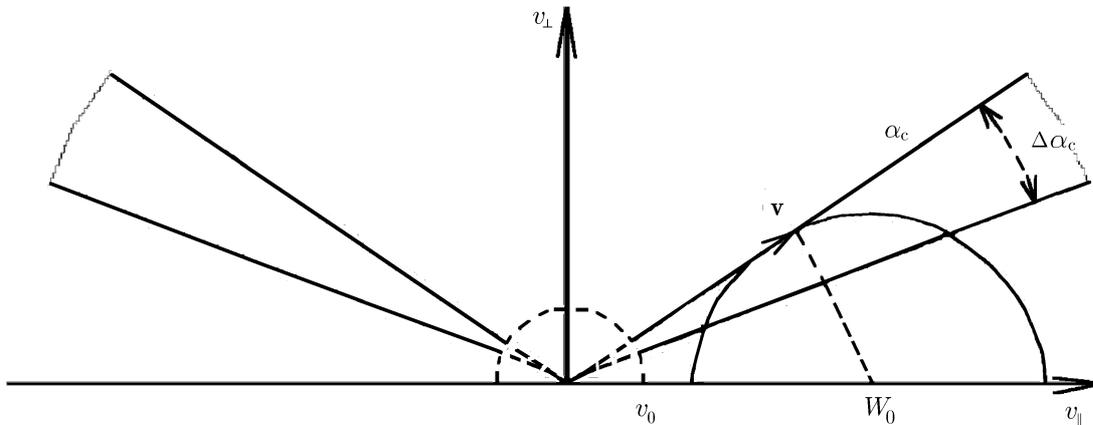


Рис. 1. Модельная функция распределения электронов с конусом потерь. Серым цветом обозначены области, где $\partial f_b / \partial v_{\perp} > 0$. Положение резонансной кривой соответствует максимальному инкременту

2.1. Функция распределения электронов

Для усиления электромагнитных волн в МЦР необходимо неустойчивое распределение электронов с положительным наклоном по поперечной компоненте скорости ($\partial f_b / \partial v_{\perp} > 0$). В данной работе в качестве такого распределения рассматривается конус потерь со степенным распределением частиц по скорости [3]:

$$f_b(v, \alpha) = n_b \frac{\zeta - 3}{4\pi v_0^3 \cos \alpha_c} \begin{cases} (v/v_0)^{-\zeta} \varphi(\alpha), & v \geq v_0; \\ 1, & v < v_0, \end{cases} \tag{12}$$

$$\varphi(\alpha) = \begin{cases} 0, & 0 \leq \alpha < \alpha_c - \Delta\alpha_c; \\ 1 + (\alpha - \alpha_c) / \Delta\alpha_c, & \alpha_c - \Delta\alpha_c \leq \alpha < \alpha_c; \\ 1, & \alpha_c \leq \alpha \leq \pi/2; \\ \varphi(\pi - \alpha), & \pi/2 < \alpha \leq \pi, \end{cases} \tag{13}$$

где v и α — скорость и питч-угол частиц соответственно, n_b — концентрация частиц, имеющих скорость $v \geq v_0$, α_c — критический питч-угол конуса потерь, $\Delta\alpha_c$ — ширина конуса потерь. Данная функция распределения описывает симметричный двусторонний конус потерь, изображённый на рис. 1. Заметим, что для солнечной короны часто используется импульсное представление функции распределения $f_b \propto p^{-\zeta}$, где p — импульс электрона, которое более адекватно описывает распределение электронов с энергией более 100 кэВ. Однако в рассматриваемых нами условиях электроны с такой энергией не вносят заметного вклада в усиление излучения, так что различие между степенными распределениями от импульса и от скорости частиц несущественно.

Процесс усиления излучения в МЦР в однородной среде ранее был детально исследован во многих работах, поэтому мы не будем подробно останавливаться на этом вопросе. Вычисление инкремента электромагнитных волн в МЦР сводится к нахождению интеграла вдоль резонансного эллипса в пространстве скоростей частиц. В данной работе для этого используются формулы, приведённые в [3, 15].

2.2. Физические условия в области генерации

Мы рассматриваем модель, в которой масштабы неоднородности магнитного поля и концентрации плазмы постоянны. В этом случае основные параметры среды экспоненциально зависят от координат:

$$H = H_0 \exp\left(\frac{r_{\parallel}}{L_{H\parallel}} + \frac{r_{\perp}}{L_{H\perp}}\right), \quad n = n_0 \exp\left(\frac{r_{\parallel}}{L_{p\parallel}} + \frac{r_{\perp}}{L_{p\perp}}\right). \tag{14}$$

Относительная концентрация ускоренных частиц n_b/n и остальные параметры конуса потерь считаются постоянными во всей рассматриваемой области.

Во всех расчётах циклотронная частота в начале координат $f_{H0} = 1\,000$ МГц, концентрация плазмы в начале координат $n_0 = 1,25 \cdot 10^8$ см⁻³ (что соответствует отношению $\omega_{p0}/\omega_{H0} = 0,1$), масштаб продольной неоднородности магнитного поля $L_{H\parallel} = -10\,000$ км (поле убывает с высотой), относительная концентрация ускоренных частиц $n_b/n = 10^{-4}$, критический питч-угол конуса потерь $\alpha_c = 30^\circ$, ширина конуса потерь $\Delta\alpha_c = 15^\circ$, степенной показатель функции распределения $\zeta = 5$, минимальная энергия частиц в пучке $E_0 = 20$ кэВ. Переменными параметрами являются масштаб продольной неоднородности концентрации плазмы и масштабы поперечной неоднородности.

3. СТАБИЛИЗАЦИЯ НЕУСТОЙЧИВОГО РАСПРЕДЕЛЕНИЯ ЭЛЕКТРОНОВ КРУПНОМАСШТАБНОЙ НЕОДНОРОДНОСТЬЮ МАГНИТНОГО ПОЛЯ

Для определения траекторий движения и коэффициента усиления электромагнитных волн проводится численное интегрирование уравнений Гамильтона (3) и уравнения (11). В качестве начальных условий при $t = 0$ задаются исходный волновой вектор \mathbf{k}_0 (и соответствующая частота излучения ω), пространственные координаты $r_{\perp 0} = 0$, $r_{\parallel 0} = 0$ и коэффициент усиления $\Lambda_0 = 0$. Расчёт заканчивается, когда волна выходит из области генерации ($|\mathbf{r}|$ или Δ_1 превышают заданную предельную величину). Учитывается, что плотность энергии волн не может быть меньше плотности энергии тепловых колебаний ($\Lambda \geq 0$).

Аналогичный метод расчёта траекторий электромагнитных волн используется в работах [16, 17] в применении к генерации радиоизлучения в магнитосфере Земли.

На рис. 2 изображены траектории движения электромагнитных волн в пространстве волновых векторов (в зависимости от N и θ) и в обычном пространстве (r_{\parallel}, r_{\perp}) при $L_{H\perp} = L_{p\parallel} = L_{p\perp} = -10\,000$ км. В качестве исходных здесь и далее используются волны, распространяющиеся параллельно магнитному полю ($N_{\perp 0} = 0$). Пунктирными линиями обозначены изолинии $t = \text{const}$ (линия $t = 0$ соответствует нижней и верхней границам рис. 2а, т. е. $\theta = 0$ и $\theta = 180^\circ$). Для сравнения на рис. 2а изображены области положительного инкремента, вычисленного в начале отсчёта пространственных координат (т. е. при $\omega_H = \omega_{H0}$, $\omega_p = \omega_{p0}$). По мере того, как волна смещается в область с другими характеристиками магнитного поля и плазмы, инкремент и его зависимость от волнового вектора также будут меняться. Поэтому при $t \neq 0$ изображённое на рис. 2а соответствие волнового вектора и инкремента электромагнитной волны носит приближённый характер.

Тем не менее можно сделать вывод о том, что длительность нахождения электромагнитной волны в области положительного инкремента не превышает нескольких миллисекунд. На рис. 3а и б приведены зависимости инкремента и логарифма коэффициента усиления от времени для двух выбранных траекторий (на рис. 2а они изображены линиями A_1-B_1 и A_2-B_2). Для обеих траекторий инкремент положителен в течение 1,0 мс. Стоит подчеркнуть, что подобное ограничение времени усиления связано именно с движением волны в пространстве волновых векторов, т. к. в обычном пространстве за это же время волна проходит не более 1,5 км.

Как видно из рис. 2а и б, электромагнитные волны с малым первоначальным показателем преломления $N \ll 1$ покидают область генерации под углом к магнитному полю, близким к $\theta = 45^\circ$.

На рис. 4, 5 изображены результаты расчёта траекторий, инкремента и коэффициента усиления электромагнитных волн в случае более резких поперечных градиентов: $L_{H\perp} = L_{p\perp} = -1\,000$ км; $L_{p\parallel} = -10\,000$ км. Скорость движения волны в пространстве волновых векторов увеличилась на порядок по сравнению с предыдущим случаем, соответственно увеличилась и скорость выхода волны из области генерации в обычном пространстве. Волны с малым начальным показателем преломления выходят из области генерации практически перпендикулярно к магнитному полю.

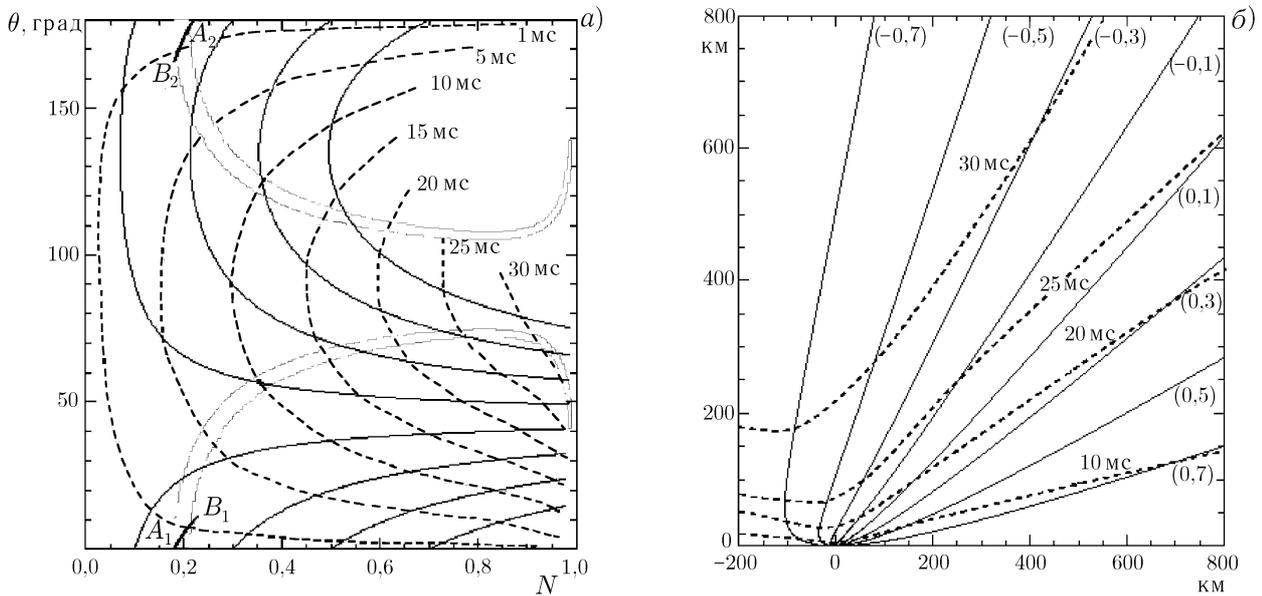


Рис. 2. Траектории движения электромагнитных волн в фазовом (а) и обычном (б) пространствах; $L_{H\perp} = L_{p\parallel} = L_{p\perp} = -10\,000$ км. Пунктирными линиями обозначены изолинии $t = \text{const}$. Серым цветом на рис. 2а обозначены области положительного инкремента (при $r_{\parallel} = r_{\perp} = 0$). На рис. 2б числа в скобках обозначают начальную «продольную компоненту» $N_{\parallel 0}$ показателя преломления для каждой траектории

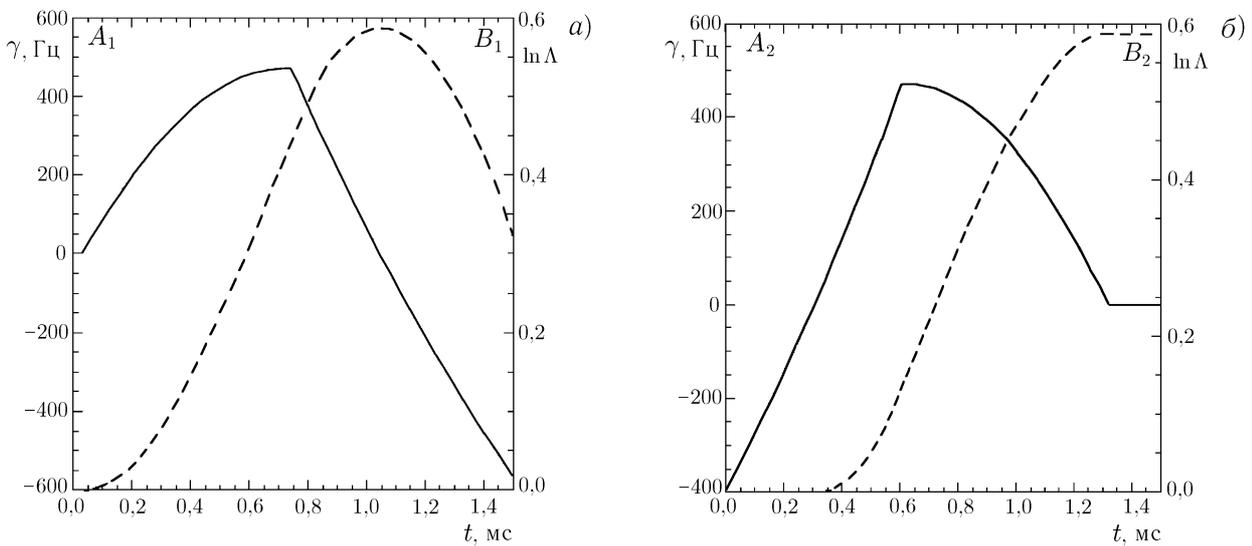


Рис. 3. Зависимость инкремента γ (сплошная линия) и логарифма коэффициента усиления $\ln \Lambda$ электромагнитной волны (пунктир) от времени. Рис. 3а и б соответствуют траекториям волн $A_1—B_1$ и $A_2—B_2$ на рис. 2а

Как показывают приведённые выше результаты расчётов, наличие неоднородностей магнитного поля с масштабами, типичными для солнечной короны, приводит к быстрому выходу электромагнитных волн из области усиления в пространстве волновых векторов. Моделирование эволюции неустойчивого электронного распределения [15] показывает, что функция распределения частиц остаётся практически неизменной в течение времени $\tau_{\text{diff}} \approx 10/\gamma$ после начала генерации. Таким образом, для указан-

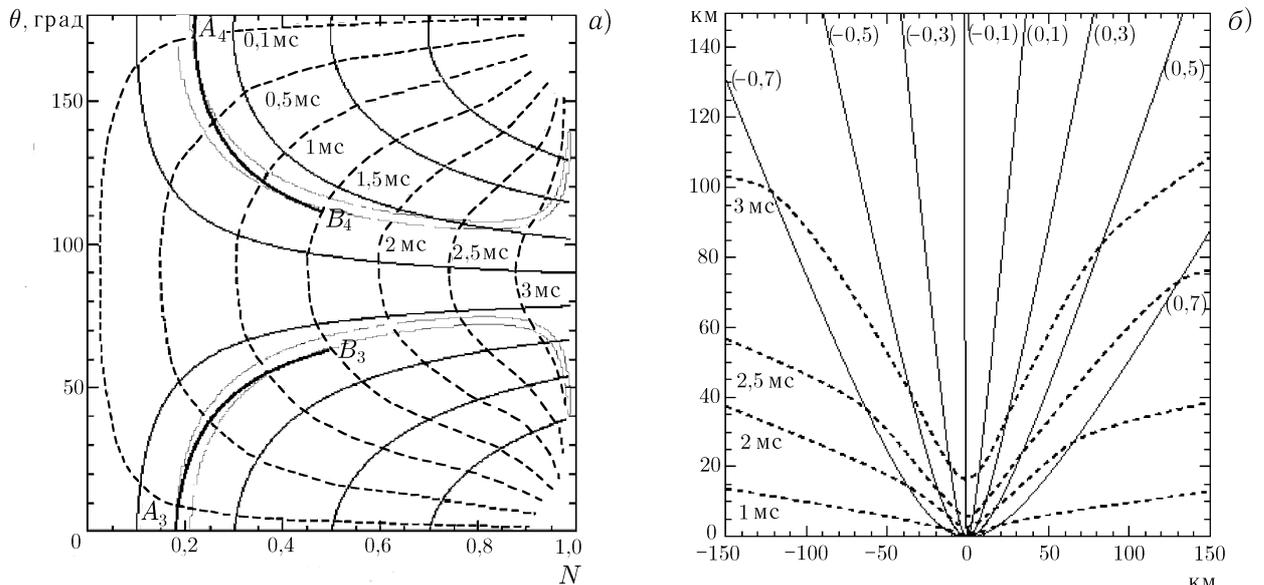


Рис. 4. Траектории движения электромагнитных волн при $L_{H\perp} = L_{p\perp} = -1000$ км; $L_{p\parallel} = -10000$ км (обозначения те же, что и на рис. 2)

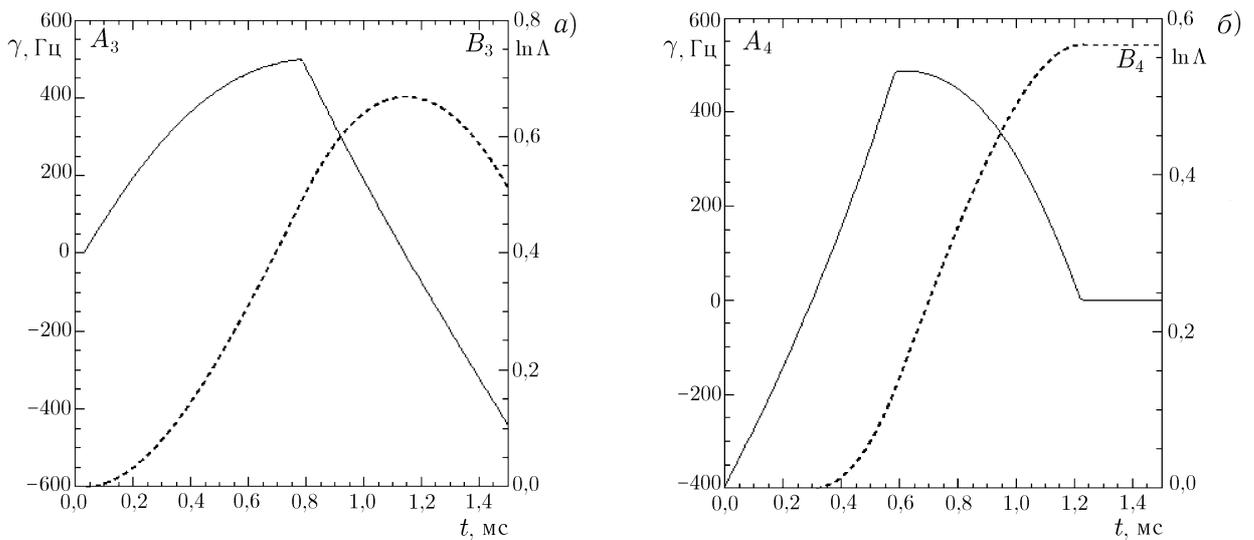


Рис. 5. Зависимость инкремента (сплошная линия) и логарифма коэффициента усиления (пунктир) от времени. Рис. 5а и б соответствуют траекториям волн $A_3—B_3$ и $A_4—B_4$ на рис. 4а

ных выше условий в области генерации спайков при $\gamma < 10^4$ Гц конус потерь оказывается устойчивым (стабилизированным неоднородностями магнитного поля, в первую очередь продольной, как будет показано далее). Соответственно, интенсивность излучения при этом достаточно низка.

4. ГЕНЕРАЦИЯ ИЗЛУЧЕНИЯ ПРИ НАЛИЧИИ МЕЛКОМАСШТАБНЫХ НЕОДНОРОДНОСТЕЙ КОНЦЕНТРАЦИИ ПЛАЗМЫ

4.1. Условия срыва стабилизации

Рассмотрим условия, при которых выход электромагнитных волн из резонанса может замедлиться. Для этого запишем условие циклотронного резонанса:

$$\dot{\psi} = \omega - k_{\parallel}v_{\parallel} - \omega_H/\Gamma = 0, \tag{15}$$

где ψ — относительная фаза волн и частиц, $\Gamma = (1 - v^2/c^2)^{-1/2}$ — лоренц-фактор. Скорость выхода волны из резонанса определяется производной выражения (15) по времени, т. е. второй производной относительной фазы:

$$\ddot{\psi} \approx -\dot{k}_{\parallel} v_{\parallel} - \frac{\omega_H}{\Gamma} \left(\frac{\dot{r}_{\parallel}}{L_{H\parallel}} + \frac{\dot{r}_{\perp}}{L_{H\perp}} \right). \quad (16)$$

Вблизи точки отсечки доминирующим в (16) является слагаемое с \dot{k}_{\parallel} , которое, в свою очередь, определяется продольными градиентами магнитного поля и концентрации плазмы. Влияние поперечных градиентов значительно меньше. Как видно из рис. 3 и 5, даже при увеличении поперечных градиентов на порядок по сравнению с продольными коэффициент усиления волн практически не изменяется.

Скорость выхода волны из резонанса уменьшается, когда вторая производная относительной фазы волн и частиц обращается в нуль: $\ddot{\psi} = 0$. Вблизи точки отсечки это условие может быть выполнено при следующем соотношении масштабов продольной неоднородности концентрации плазмы и магнитного поля:

$$L_{p\parallel} \approx -\Delta_1 L_{H\parallel} \approx -(\omega_p/\omega_H)^2 L_{H\parallel}. \quad (17)$$

При условии (17) влияние неоднородности концентрации плазмы компенсирует влияние неоднородности магнитного поля. При этом неоднородность концентрации должна быть значительно более мелко-масштабной, чем неоднородность магнитного поля ($L_{p\parallel} \approx 100$ км для используемых в расчётах параметров). Кроме того, продольные компоненты градиентов концентрации плазмы и магнитного поля должны быть противоположно направлены.

4.2. Траектории движения электромагнитных волн при условии компенсации влияния неоднородностей

На рис. 6–9 изображены результаты расчёта траекторий, инкремента и коэффициента усиления электромагнитных волн; масштаб продольной неоднородности концентрации плазмы для каждой траектории выбран таким, что в начальный момент ($t = 0$, $r_{\parallel} = r_{\perp} = 0$) вторая производная относительной фазы волн и частиц обращается в нуль. По мере движения волны в обычном пространстве и в пространстве волновых векторов условие $\ddot{\psi} = 0$ нарушается, и волна в конечном итоге выходит из резонанса с электронным пучком. Движение волн рассматривается в области $|r_{\parallel}/L_{p\parallel}| < 1$.

На рис. 6, 7 изображены результаты расчёта траекторий, инкремента и коэффициента усиления электромагнитных волн в случае, когда масштабы поперечной неоднородности равны: $L_{H\perp} = L_{p\perp} = -10\,000$ км. По сравнению со случаем, когда мелко-масштабная неоднородность концентрации плазмы отсутствует (см. рис. 2, 3), длительность усиления волны увеличилась примерно на порядок и составляет 10,4 мс для траектории A_5-B_5 и 14,2 мс для траектории A_6-B_6 ; соответственно, на порядок увеличился логарифм коэффициента усиления Λ . Поскольку интенсивность волны экспоненциально зависит от Λ , компенсация влияния неоднородностей концентрации плазмы и магнитного поля может привести к значительному росту интенсивности излучения. Необходимый масштаб продольной неоднородности концентрации плазмы составляет $L_{p\parallel} \approx 106,7$ км для обеих траекторий.

На рис. 8, 9 изображены результаты расчёта траекторий, инкремента и коэффициента усиления электромагнитных волн в случае $L_{H\perp} = L_{p\perp} = -1\,000$ км. Здесь также наблюдается увеличение длительности и коэффициента усиления волны по сравнению со случаем отсутствия мелко-масштабной неоднородности плазмы (см. рис. 4, 5). Однако теперь проявляется влияние поперечных неоднородностей, и длительность нахождения волны в области положительного инкремента составляет 2,1 мс для траектории A_7-B_7 и 2,7 мс для траектории A_8-B_8 . Необходимый масштаб продольной неоднородности концентрации плазмы составляет $L_{p\parallel} \approx 106,2$ км для обеих траекторий.

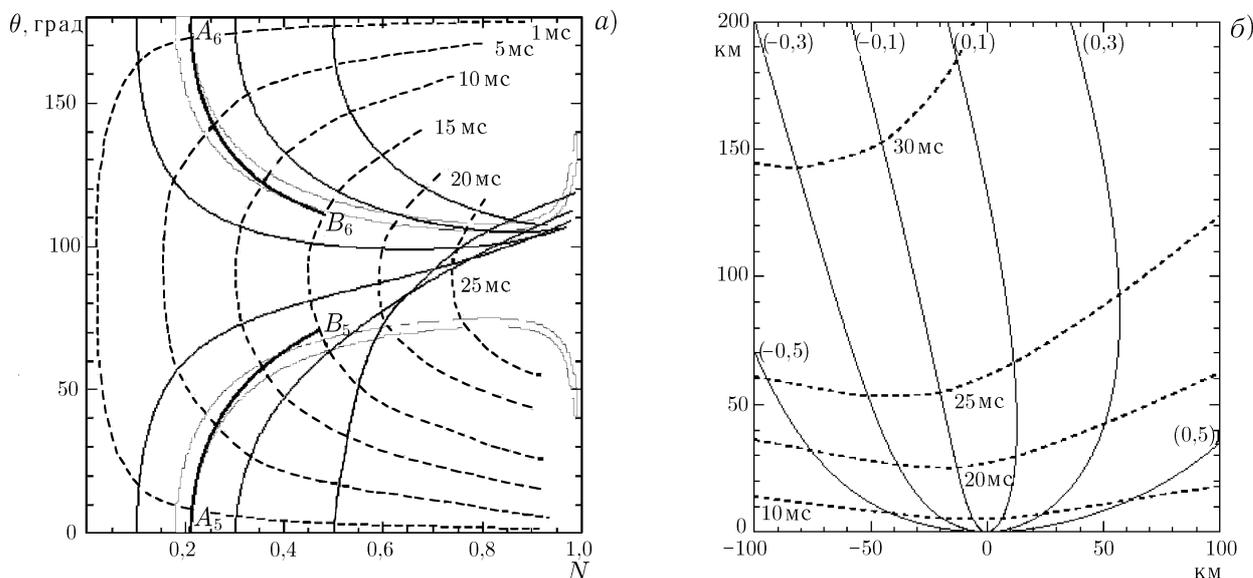


Рис. 6. Траектории движения электромагнитных волн при $L_{H\perp} = L_{P\perp} = -10\,000$ км; масштаб продольной неоднородности концентрации плазмы $L_{P\parallel}$ для каждой траектории выбран таким, что при $r_{\parallel} = r_{\perp} = 0$ вторая производная относительной фазы обращается в нуль (обозначения те же, что и на рис. 2)

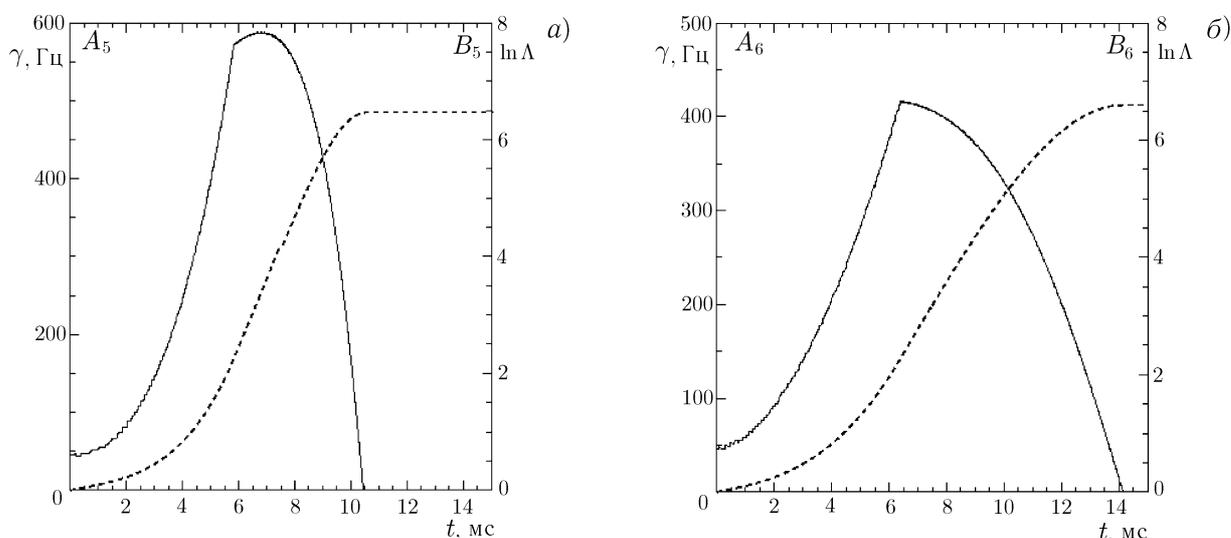


Рис. 7. Зависимость инкремента (сплошная линия) и логарифма коэффициента усиления (пунктир) от времени. Рис. 7а и б соответствуют траекториям волн A_5-B_5 и A_6-B_6 на рис. 6а

5. ОЦЕНКА ВРЕМЕНИ УСИЛЕНИЯ ЭЛЕКТРОМАГНИТНЫХ ВОЛН

Как отмечалось выше, скорость выхода волны из резонанса определяется второй производной относительной фазы волн и частиц (16). Если эта величина стремится к нулю, то необходимо учитывать также производные более высокого порядка. Запишем разложение условия циклотронного резонанса (15) в ряд Тейлора по степеням Δt — времени нахождения волны в области положительного инкремента:

$$\dot{\psi} = \delta\dot{\psi} + \ddot{\psi} \Delta t + \frac{\ddot{\psi}}{2} (\Delta t)^2 + \dots = 0, \tag{18}$$

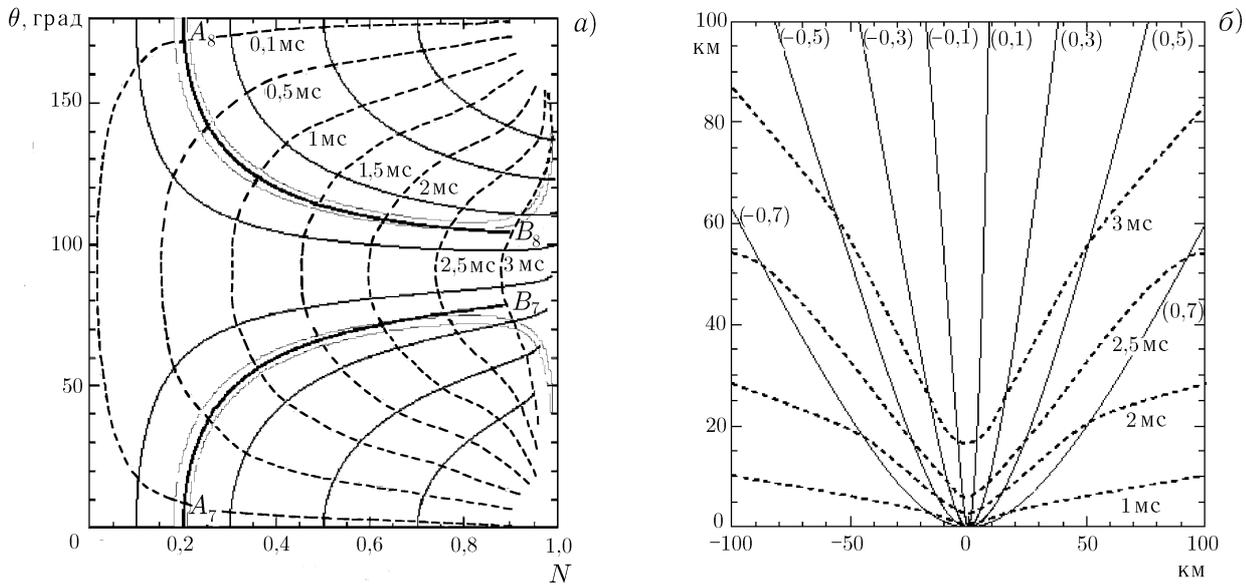


Рис. 8. Траектории движения электромагнитных волн при $L_{H\perp} = L_{P\perp} = -1\,000$ км; масштаб продольной неоднородности концентрации плазмы $L_{P\parallel}$ для каждой траектории выбран таким, что при $r_{\parallel} = r_{\perp} = 0$ вторая производная относительной фазы обращается в нуль (обозначения те же, что и на рис. 2)

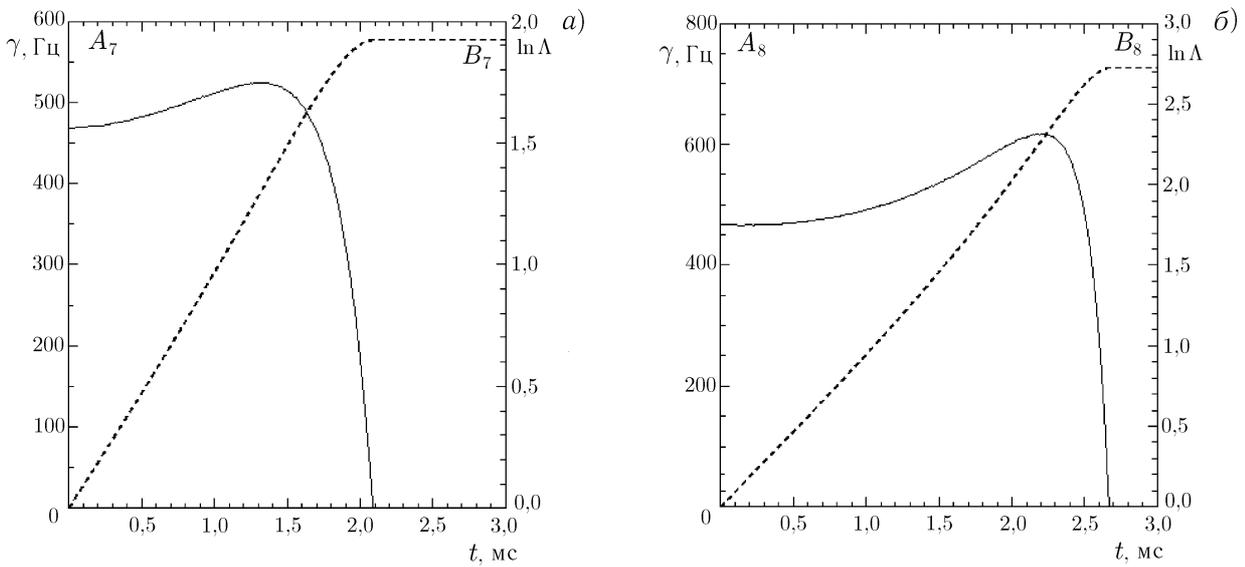


Рис. 9. Зависимость инкремента (сплошная линия) и логарифма коэффициента усиления (пунктир) от времени. Рис. 9а и б соответствуют траекториям волн $A_7—B_7$ и $A_8—B_8$ на рис. 8а

где $\dot{\psi}$ — вариация производной относительной фазы по скоростям электронов в области положительного наклона функции распределения ($\partial f_b / \partial v_{\perp} > 0$):

$$\delta\dot{\psi} = \frac{\partial\dot{\psi}}{\partial v} \Delta v + \frac{\partial\dot{\psi}}{\partial \alpha} \Delta \alpha \approx \left(\frac{\omega_H v}{c^2} - k_{\parallel} \cos \alpha \right) \Delta v + k_{\parallel} v \Delta \alpha \sin \alpha. \tag{19}$$

Для конуса потерь (12), (13) положение и размеры области положительного наклона функции распределения можно оценить следующим образом. В полурелятивистском приближении ($v \ll c$) условие резонанса (15) в пространстве скоростей электронов является уравнением окружности с центром

в точке $v_{\parallel} = W_0 \simeq cN_{\parallel}$, $v_{\perp} = 0$ [3]. Инкремент максимален, когда резонансная окружность касается прямой $\alpha = \alpha_c$ (см. рис. 1), при этом $v = W_0 \cos \alpha_c$; $\Delta v \simeq v - v_0$; $\Delta \alpha = \Delta \alpha_c$. Подставляя эти значения в (19), получаем

$$\delta \dot{\psi} \approx \frac{1}{2} \omega N_{\parallel}^2 \Delta \alpha_c \sin(2\alpha_c). \quad (20)$$

Третья производная относительной фазы волн и частиц $\ddot{\psi}$ включает в себя производные \ddot{k}_{\parallel} , \ddot{r}_{\parallel} , \ddot{r}_{\perp} . Мы приведём её приближённое выражение для случая, когда волна в начальный момент времени распространяется параллельно магнитному полю при $N \ll 1$:

$$\ddot{\psi} \approx \frac{2\omega c^2}{L_{H\parallel}^2} \left\{ \frac{\dot{r}_{\parallel} v_{\parallel}}{c^2} \left[\delta_{\parallel} (\delta_{\parallel} + 1) - \frac{L_{H\parallel}}{2L_{p\parallel}} \right] + \Delta_1 \left(\delta_{\parallel} + \mu_{\perp} \delta_{\perp} \frac{L_{H\parallel}^2}{L_{H\perp}^2} \right) \right\}. \quad (21)$$

При вычислении производных относительной фазы полагаем $v_{\parallel} = W_0 \cos^2 \alpha_c$.

Вместо точного решения уравнения (18) достаточно использовать приближённые корни

$$\Delta t_1 = |\delta \dot{\psi} / \dot{\psi}|, \quad \Delta t_2 = \sqrt{|2\delta \dot{\psi} / \ddot{\psi}|}. \quad (22)$$

В табл. 1 указаны времена резонанса Δt_1 и Δt_2 для различных условий распространения электромагнитных волн. Также указан параметр Δt_{RT} — время нахождения волны в области положительного инкремента, полученное при численном интегрировании уравнений Гамильтона. Время Δt_1 зависит в основном от масштабов продольных неоднородностей $L_{H\parallel}$ и $L_{p\parallel}$ и стремится к бесконечности при условии (17). Поперечные неоднородности с масштабами $L_{H\perp}$ и $L_{p\perp}$ существенным образом ограничивают время резонанса Δt_2 , но их влияние на процесс генерации излучения становится заметным только при условии компенсации влияния продольных неоднородностей. Время усиления волны Δt_{RT} , как видно из табл. 1, с хорошей точностью равно меньшему из времён Δt_1 и Δt_2 .

Таблица 1

Длительность резонанса для различных условий распространения электромагнитных волн

Траектория	$N_{\parallel 0}$	$L_{p\parallel}$, км	$L_{H\perp}$, км	$L_{p\perp}$, км	Δt_1 , мс	Δt_2 , мс	Δt_{RT} , мс
A_1-B_1	0,18	-10^4	-10^4	-10^4	0,942	15,8	1,0
A_2-B_2	-0,22	-10^4	-10^4	-10^4	1,15	18,8	1,0
A_3-B_3	0,18	-10^4	-10^3	-10^3	0,942	2,74	1,1
A_4-B_4	-0,22	-10^4	-10^3	-10^3	1,15	3,30	0,9
A_5-B_5	0,21	106,7	-10^4	-10^4	∞	12,0	10,4
A_6-B_6	-0,21	106,7	-10^4	-10^4	∞	12,0	14,2
A_7-B_7	0,20	106,2	-10^3	-10^3	∞	2,97	2,1
A_8-B_8	-0,20	106,2	-10^3	-10^3	∞	2,97	2,7

6. ЗАКЛЮЧЕНИЕ. ГЕНЕРАЦИЯ СПАЙКОВ В НЕОДНОРОДНОЙ СРЕДЕ

Проведённое исследование показывает, что неоднородности магнитного поля и концентрации плазмы оказывают значительное влияние на формирование и стабилизацию неустойчивого электронного распределения и генерацию излучения на циклотронном резонансе быстрой необыкновенной волны

в солнечной короне. Предлагаемый сценарий генерации миллисекундных всплесков солнечного радиоизлучения выглядит следующим образом:

1) Крупномасштабная продольная неоднородность магнитного поля с характерным масштабом порядка размеров корональной магнитной петли ограничивает эффективность мазерного механизма, тем самым делая возможным формирование неустойчивого конуса потерь. Инкремент электромагнитных волн при этом достаточно высок (до 10^4 Гц для рассматриваемых в данной статье условий); пространственные размеры области с потенциально неустойчивым, но стабилизированным распределением электронов составляют тысячи километров.

2) В некоторой области пространства достаточно быстро возникает мелкомасштабная неоднородность концентрации плазмы с характерным масштабом $L_{p\parallel} \approx -\Delta_1 L_{H\parallel}$. Выход электромагнитных волн из резонанса замедляется, и стабилизация неустойчивого электронного распределения срывается, что приводит к генерации всплеска излучения. Длительность и временной профиль всплеска определяются как эффектами квазилинейной релаксации конуса потерь [8], так и динамикой неоднородности концентрации плазмы (в зависимости от условий один из этих факторов может быть доминирующим).

Очевидно, что условие (17) компенсации влияния неоднородностей магнитного поля и концентрации плазмы может выполняться только в некоторой относительно малой области пространства, которая и будет являться источником отдельного спайка с узкой спектральной полосой. Продольный размер области генерации спайка определяется масштабом $L_{p\parallel}$ и в типичных для солнечной короны условиях составляет от десятков до сотен километров. Оценка поперечных размеров области генерации всплеска требует дополнительных предположений о природе мелкомасштабной неоднородности плазмы, но можно ожидать, что поперечные размеры будут близки к продольным. Таким образом, в области формирования конуса потерь, которая является источником кластера спайков, может возникать большое количество независимых друг от друга локальных областей генерации излучения. При этом спектр излучения в каждый момент времени является отражением пространственного распределения мелкомасштабных неоднородностей плазмы. Движение неоднородности в пространстве приводит к дрейфу излучения по частоте.

В заключение отметим, что данная схема генерации миллисекундных всплесков солнечного радиоизлучения может служить инструментом диагностики не только электронных потоков, но и структуры и динамики плазменной турбулентности в активных областях солнечной короны.

Работа выполнена при поддержке Министерства образования РФ (грант № Е00—8.0—71).

СПИСОК ЛИТЕРАТУРЫ

1. Benz A. O. // Solar Phys. 1986. V. 104. P. 99.
2. Bastian T. S., Benz A. O., Gary D. E. // Annual Rev. Astron. Astrophys. 1998. V. 36. P. 131.
3. Melrose D. B., Dulk G. A. // Astrophys. J. 1982. V. 259. P. 844.
4. Aschwanden M. J. // Astron. Astrophys. 1990. V. 237. P. 512.
5. Robinson P. A. // Solar Phys. 1991. V. 134. P. 299.
6. Fleishman G. D. // Solar Phys. 1994. V. 153. P. 367.
7. Fleishman G. D., Yastrebov S. G. // Solar Phys. 1994. V. 153. P. 389.
8. Флейшман Г. Д., Мельников В. Ф. // УФН. 1998. Т. 168. С. 1 265.
9. Robinson P. A. // Solar Phys. 1991. V. 136. P. 343.
10. Платонов К. Ю., Флейшман Г. Д. // Астрон. журн. 2001. Т. 78. С. 238.
11. Le Quéau D., Pellat R., Roux A. // Ann. Geophys. 1985. V. 3. P. 273.
12. Zarka P., Le Quéau D., Roux A. // J. Geophys. Res. 1986. V. 91. P. 13 542.
13. Vlasov V. G. // Planet. Space Sci. 1991. V. 39. P. 1 223.
14. Власов В. Г. // Физика плазмы. 1991. Т. 17. С. 165.

15. Aschwanden M. J. // *Astron. Astrophys. Suppl. Ser.* 1990. V. 85. P. 1 141.
16. Gaelzer R., Ziebell L. F., Schneider R. S. // *J. Geophys. Res.* 1994. V. 99. P. 8 905.
17. de H. Cavalcanti C. J., Schneider R. S., Ziebell L.F. // *J. Geophys. Res.* 1996. V. 101. P. 24 557.

Иркутский государственный
технический университет, г. Иркутск, Россия

Поступила в редакцию
19 сентября 2001 г.

**THE INFLUENCE OF MAGNETIC-FIELD AND PLASMA-DENSITY NONUNIFORMITIES
ON THE PROPAGATION AND AMPLIFICATION OF RADIO WAVES IN THE SOLAR CORONA**

A. A. Kuznetsov and V. G. Vlasov

We analyze the maser generation of millisecond spikes of the solar radio emission at the cyclotron resonance of the fast extraordinary wave in an inhomogeneous medium. It is shown that the magnetic-field nonuniformity with parameters typical of the solar corona drastically reduces the time of electromagnetic-wave amplification, which is explained by the fact that these waves leave the resonance region in the wave-vector space. As a result, an unstable electron distribution can be formed. The efficient generation of radiation becomes possible only in such local regions where the influence of the magnetic-field nonuniformity is compensated by small-scale inhomogeneities of the plasma density with typical scales ranging from tens to hundreds of kilometers. Taking the effect of inhomogeneous medium into account allows us to explain spatial and temporal characteristics of the spikes.

УДК 535.853.

РЕКОНСТРУКЦИЯ ИЗОБРАЖЕНИЙ СОЛНЦА, ПОЛУЧЕННЫХ НА СИБИРСКОМ СОЛНЕЧНОМ РАДИОТЕЛЕСКОПЕ

С. В. Лесовой

При реконструкции изображений Солнца, полученных на Сибирском солнечном радиотелескопе, возникают две основные проблемы. Во-первых, необходимо с очень высокой точностью знать амплитудно-фазовое распределение в антенно-фидерном тракте, потому что контраст изображений Солнца на длине волны 5,2 см очень велик. Во-вторых, поскольку изображения состоят не только из ярких компактных источников, но содержат также слабоконтрастные протяжённые области, возникает присущая алгоритму CLEAN проблема реконструкции протяжённых областей. Первую проблему мы решаем, определяя амплитудно-фазовые искажения с помощью итерационного анализа самого изображения путём сравнения противоположных боковых лепестков отклика на компактный источник. Для уменьшения влияния других источников на отклик мы анализируем несколько компактных источников. Искажения фазы определяются по асимметрии соответствующих лепестков, искажения амплитуды — по величине лепестков. После каждой итерации изображение корректируется в области пространственного спектра. Проблемы при реконструкции протяжённых областей связаны, с одной стороны, с тем, что алгоритм CLEAN требует весьма значительных вычислительных ресурсов. С другой стороны, при восстановлении изображений протяжённых областей необходимо увеличивать число циклов этого алгоритма. Ещё одна проблема заключается в том, что использование одной и той же «чистой» диаграммы направленности для реконструкции компактных и протяжённых источников ведёт к появлению высокочастотных искажений последних. Основная часть вычислительных ресурсов при использовании алгоритма CLEAN расходуется на сдвиг диаграммы направленности и поиск максимума в исходном изображении. Время, необходимое для сдвига диаграммы направленности, было уменьшено путём исключения из массива значений диаграммы точек, в которых диаграмма не превышает заданный порог, и сортировкой оставшихся точек. Время поиска максимума было уменьшено благодаря использованию плавающего окна поиска. При реконструкции используются несколько уровней отсечки, поиск максимума по всему исходному массиву проводится только после восстановления изображения в окне до текущего уровня отсечки. Для реконструкции ярких компактных источников используется близкая к идеальной диаграмма направленности. Для реконструкции слабоконтрастных протяжённых областей используются модельные диаграммы. Последние представляют собой свёртки диаграммы направленности и изображений протяжённых источников.

ВВЕДЕНИЕ

Рабочая длина волны Сибирского солнечного радиотелескопа (ССРТ), равная 5,2 см, была выбрана из соображений наибольшего контраста изображений активного и спокойного Солнца на этой длине волны [1]. Типичное отношение яркостных температур источников над пятнами к яркостной температуре спокойного Солнца составляет два порядка. Это значит, что диапазон восстановленного («чистого») изображения должен быть не менее 30 дБ, тогда как обычно считается достаточным диапазон восстановленного изображения 20 дБ. Подчеркнём, что речь идёт не об отношении сигнал/шум, а об остаточном уровне боковых лепестков. Такой диапазон «чистого» изображения предъявляет весьма жёсткие требования к амплитудно-фазовому распределению в антенно-фидерном тракте. Как будет показано ниже, технически юстировка ССРТ с требуемой точностью недостижима, поэтому остаётся только возможность коррекции амплитудно-фазового распределения путём обработки данных. Так как выходным сигналом ССРТ является изображение, а не набор пространственных гармоник, то коррекцию амплитудно-фазового распределения можно вести только путём анализа самого изображения.

Применение алгоритма CLEAN для реконструкции изображений всегда связано с некоторыми модификациями самого алгоритма. Вызвано это как тем, что решение некорректной задачи всегда в определённой мере субъективно, так и неконкретностью алгоритма CLEAN. Исходное определение этого

алгоритма, данное Хёгбомом [2], весьма абстрактно. Ясно, что в исходном виде этот алгоритм неприемлем, в частности из-за весьма жестких требований к производительности компьютеров. Конкретизация алгоритма, или адаптация его к определенным данным, всегда приводит к некоторым его изменениям. Хотя эти изменения и не носят принципиального характера, но делают возможным применение этого алгоритма на практике.

1. КОРРЕКЦИЯ АМПЛИТУДНО-ФАЗОВОГО РАСПРЕДЕЛЕНИЯ

Двумерное изображение, получаемое с помощью ССРТ, можно представить в следующем виде:

$$I(x, y) = \left(\iint B(\xi, \eta) K(x - \xi, y - \eta) d\xi d\eta + N \right) F(x, y) + N_q = \\ = (B(x, y) * K(x, y) + N) F(x, y) + N_q,$$

где $K(x, y)$ — диаграмма направленности (ДН) телескопа, N — аддитивный шум, $F(x, y)$ — передаточная функция приёмника ССРТ, N_q — шум квантования, $B(x, y)$ — истинное распределение радиояркости. Представим ДН в виде свёртки идеальной ДН $K_0(x, y)$ и некоторого возмущения $\delta K(x, y)$. Тогда

$$I(x, y) = (B(x, y) * [K_0(x, y) * \delta K(x, y)] + N) F(x, y) + N_q. \quad (1)$$

Отсюда ясно, что возмущение представляет собой набор δ -функций, расположенных в максимумах ДН и имеющих определённые величину и знак. Если возмущение состоит только из одной δ -функции, соответствующей главному максимуму, то мы имеем идеальную ДН. Появление в таком наборе δ -функции, соответствующей любому другому максимуму, связано с гармоническим искажением амплитудно-фазового распределения. Таким образом, задача коррекции амплитудно-фазового распределения сводится к определению $\delta K(x, y)$ и обратной фильтрации исходного изображения с помощью фурье-образа $\delta K(u, v)$ возмущения ДН. Отметим, что даже определив $\delta K(x, y)$ с абсолютной точностью, мы не сможем привести данное изображение к виду $B(x, y) * K_0(x, y) + N$. Этому мешают аддитивный шум и передаточная характеристика приёмника.

Определим требования к точности амплитудно-фазового распределения в антенно-фидерном тракте ССРТ. Будем исходить из желаемого динамического диапазона результирующего изображения и отношения яркостных температур диска Солнца и активных областей. Хорошим динамическим диапазоном для реконструированных изображений традиционно считается 10^2 . В данном случае под динамическим диапазоном мы подразумеваем относительный уровень остающихся после реконструкции боковых лепестков. Типичная яркостная температура компактных областей на рабочей длине волны ССРТ составляет $(1 \div 1,5) \cdot 10^6$ К. Яркостная температура диска Солнца $1,6 \cdot 10^4$ К. Очевидно, что при таком контрасте исходного изображения остаточный уровень боковых лепестков при динамическом диапазоне 10^2 является совершенно неприемлемым, поскольку он сравним с яркостной температурой диска Солнца. Поэтому выберем динамический диапазон равным 10^3 . Для типичной яркостной температуры активных областей остаточный уровень боковых лепестков, составит $(1 \div 1,5) \cdot 10^3$ К. Отметим, что эта величина сравнима с чувствительностью ССРТ, которая составляет $0,8 \cdot 10^3$ К. Так как относительная площадь остаточных боковых лепестков невелика, можно считать, что при динамическом диапазоне реконструированного изображения, равном 10^3 , эти лепестки не оказывают существенного влияния на отношение сигнал/шум. Для определения требуемой точности амплитудно-фазового распределения используем следующую весьма простую процедуру. Будем считать, что нам нужно реконструировать отклик на точечный источник, т. е. для алгоритма CLEAN можно использовать один шаг с фактором усиления, равным 1. При этом мы знаем «грязную» ДН с конечной точностью в окрестности

боковых лепестков с номерами $n, -n$. Другими словами, амплитудно-фазовое распределение известно нам с точностью до гармонической флуктуации в распределении фазы и амплитуды:

$$\varphi(n, k) = \varphi_n \sin \left[2\pi \frac{k(n+1)}{M} \right], \quad A(n, k) = A_n \cos \left[2\pi \frac{k(n+1)}{M} \right], \quad (2)$$

где n — номер бокового лепестка, k — номер антенны, M — число антенн. Очевидно, что $\varphi(n, k)$ будет вносить нечётные, а $A(n, k)$ — чётные искажения в результат алгоритма CLEAN. Набор таких гармонических искажений определяет возмущение $\delta K(x, y)$ в (1). Оценим влияние этих искажений на изображение. Если флуктуацию фазы задавать в радианах, а амплитуды — в долях амплитуды идеального распределения, то нормированная разность между идеальной и возмущённой ДН будет пропорциональна величине возмущений с коэффициентом пропорциональности 0,5; при этом нужно иметь в виду, что возмущения фазы должны быть малы. Для возмущений фазы, не превышающих 0,3 рад (20°), эта пропорциональность сохраняется.

Таким образом, для того, чтобы получить отклонение «грязной» диаграммы направленности от идеальной в окрестности только одного бокового лепестка с номером n , необходимо задать возмущения как амплитуды, так и фазы. При этом величины возмущений должны равняться нормированному отклонению диаграммы направленности. Для динамического диапазона 10^3 отклонение должно быть в пределах 0,1 %. Это значит, что в амплитудно-фазовом распределении не должно быть неизвестных гармонических искажений с точностью 0,1 % по амплитуде и $0,057^\circ$ по фазе. Отметим, что это весьма жёсткая оценка требуемой точности. Реально мы имеем дело, разумеется, не с гармоническим искажением, а со спектром. Однако очевидно, что достигнуть желаемой точности амплитудно-фазового распределения путём юстировки антенно-фидерного тракта не удастся. Единственный практический путь юстировки тракта ССРТ — это подбор геометрических длин волноводов путём так называемых волноводных вставок, при этом 1 мм длины вставки соответствует 6° фазы. Ясно, что таким образом нельзя получить амплитудно-фазовое распределение с желаемой точностью. Более того, обсерватория не в состоянии проводить дорогостоящие геодезические измерения пространственного положения антенн. А это означает, что фазовое распределение будет изменяться как в течение дня наблюдений, так и в течение года. Всё это вынуждает использовать методы коррекции амплитудно-фазового распределения, основанные на анализе конкретного изображения. Отметим, что столь жёсткие требования к амплитудно-фазовому распределению не идут в разрез с проектными параметрами ССРТ. Согласно проекту ССРТ предназначался для наблюдения только активных областей, однако исследовательский интерес к слабоконтрастным структурам побуждает к попыткам выйти за рамки проектных параметров.

Для коррекции амплитудно-фазового распределения с помощью анализа изображения используются два, вообще говоря, независимых способа. В рамках первого способа приближённое распределение получается интегрированием в спектральной области по направлениям, соответствующим проекциям баз интерферометра [3]. Во втором способе используется итерационная процедура коррекции. Суть её заключается в следующем. На изображении выбирается яркий компактный источник в качестве центра системы отсчёта. По заданным диапазонам номеров боковых лепестков измеряются их асимметрия и уровень. На каждом шаге асимметрия и избыточный уровень лепестков корректируются путём внесения соответственно синусных и косинусных возмущений (2) в спектр изображения. Затем выполняется обратная фильтрация изображения с использованием полученных возмущений и переход к следующему шагу. Окончание итераций происходит либо в случае, когда разность уровней боковых лепестков становится равной уровню шума, либо по истечении заданного количества циклов. Практика показывает, что такая процедура является устойчивой при величине возмущений меньше $0,5^\circ$ по фазе и 0,1 % по амплитуде, при этом требуется порядка 100 циклов. Уменьшение указанных величин возмущений, естественно, ведёт к повышению точности распределения и увеличению числа циклов.

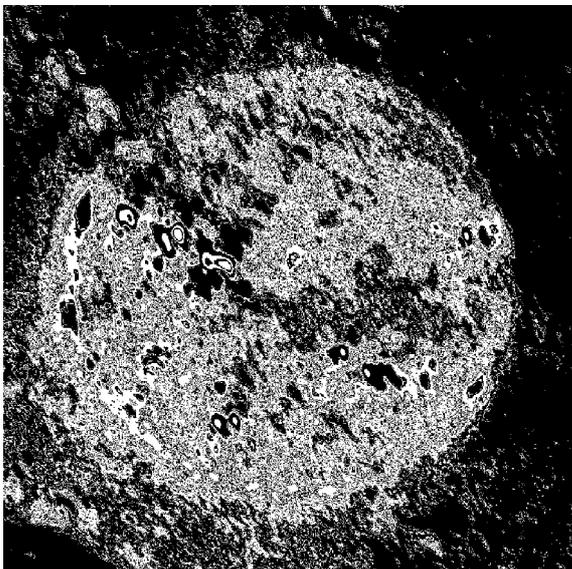


Рис. 1. Исходное изображение Солнца

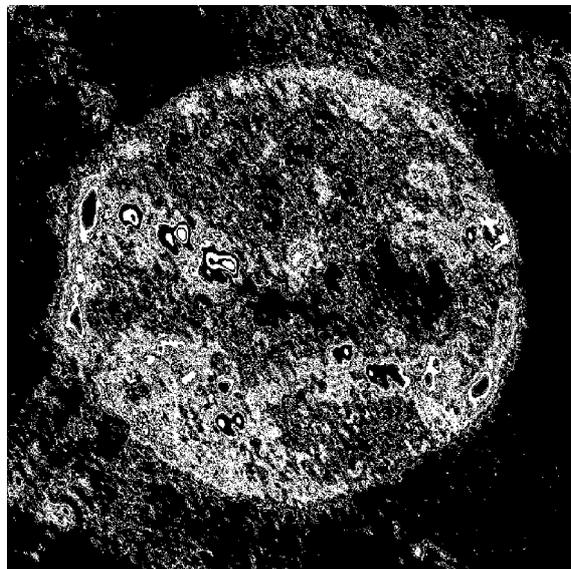


Рис. 2. Изображение Солнца после коррекции амплитудно-фазового распределения

Распределение, полученное первым способом, используется только в качестве первого приближения или не используется вообще. В качестве первого приближения часто используется также распределение, полученное для предыдущих изображений. В случае, если несколько активных областей перекрываются боковыми лепестками, возмущения вносятся вручную. Критерием здесь является визуальное восприятие изображения, что, конечно, требует некоторого опыта от оператора. Рис. 1 и 2 показывают исходное и скорректированное вышеописанным (вторым) способом изображения.

2. ПРИМЕНЕНИЕ АЛГОРИТМА CLEAN

Для того, чтобы сформулировать задачу реконструкции, представим изображение в дискретном виде. Пусть аддитивный шум отсутствует. При этом «грязное» изображение

$$I_0(x, y) = B(x, y) * K(x, y).$$

Исходное изображение $B(x, y)$ можно представить в виде бесконечной суммы модельных источников, имеющих различное местоположение и амплитуду:

$$B(x, y) \approx B_1(x, y) = \sum_{s=0}^{\infty} a_s(x - x_s, y - y_s). \quad (3)$$

Примем довольно слабое ограничение на исходное изображение. Пусть пространственный спектр исходного изображения ограничен, т. е. спадает к краю. Очевидно, что в этом случае представление (3) возможно с любой степенью точности. Необходимо только выбрать достаточно малый пространственный размер модельного источника, соответствующий ширине спектра исходного изображения. В этом случае задачу реконструкции можно определить следующим образом: необходимо аппроксимировать изображение B_1 набором

$$B_2(x, y) = \sum_{s=0}^S a'_s(x - x_s, y - y_s),$$

состоящим из S модельных источников заданного размера, так, чтобы их разность не превышала заданный уровень Tr , т. е.

$$\Delta B(x, y) = B_1(x, y) - B_2(x, y) = \sum_{s=0}^S (a_s - a'_s) + \sum_{s=S}^{\infty} a_s \leq \text{Tr}. \quad (4)$$

В силу линейности (3) мы можем допустить, что индекс s соответствует источникам, отсортированным по мере убывания амплитуды. В этом случае путём надлежащего выбора S остаточный член в (4) может быть сделан сколь угодно малым. Погрешности аппроксимации при этом определится только разностью коэффициентов a_s и a'_s , где $s = 0, 1, \dots, S$. Если исходное изображение составлено из бесконечно узких источников, или, другими словами, имеется бесконечное угловое разрешение, то не видно причины, кроме влияния шума, по которой значения a_s и a'_s , где $s = 0, 1, \dots, S$, будут заметно отличаться. Существенно иная ситуация возникает в случае конечных размеров источников в исходном изображении. Это соответствует случаю, когда ядро $K_0(x, y)$ в (1) отлично от δ -функции, т. е. случаю конечного углового разрешения. В этом случае яркость изображения в каждой точке определяется суперпозицией всех точек изображения. Влияние на данную точку всех остальных точек зависит от яркости изображения в этих точках, протяжённости источников и вида зависимости $K_0(x, y)$. В случае, если $K_0(x, y)$ является неаналитической функцией, что имеет место практически для всех радиотелескопов, на изображении появляются так называемые боковые лепестки. Это не что иное, как боковые интерференционные максимумы. Их происхождение связано с ограниченностью апертуры радиотелескопа и с наличием разрывов на апертуре. Для реконструкции таких изображений применяется алгоритм CLEAN с целью устранения влияния скачков на разрыве идеальной диаграммы направленности $K_0(x, y)$ — боковых лепестков. Очевидно, что для этого необходимо, во-первых, знать точный вид функции $K_0(x, y)$ и, во-вторых, начинать процедуру реконструкции с наибольших значений (в противном случае взаимное влияние частей изображения друг на друга приведёт к большим ошибкам). Наконец, надо иметь в виду, что в случае протяжённых структур этот алгоритм весьма неточен. Действительно, в этом случае яркость изображения в данной точке определяется главным образом остальными точками. Доля этого значения, получаемая на каждом шаге реконструкции, содержит мало информации об истинной яркости источника. При относительно большом факторе усиления (малом количестве шагов) решение становится неустойчивым по отношению к количеству шагов. Решение устойчиво, если после определённого количества шагов дальнейшее их увеличение не даёт заметных изменений в реконструированном изображении. И наоборот, при неустойчивом решении результат реконструкции всегда сильно зависит от числа шагов.

Задача реконструкции, сформулированная в виде (4), может быть решена применительно к изображениям ССРТ путём использования алгоритма CLEAN. Но при этом возникают две серьёзные проблемы. Первая заключается в определении $K_0(x, y)$. О том, как это делается, говорилось выше. Вторая проблема — реконструкция протяжённых областей — заключается в следующем. Так как протяжённые структуры состоят из большего числа точек, то время реконструкции резко возрастает. Затраты вычислительных ресурсов вызваны в основном двумя причинами. Первая — это необходимость сдвига ДН на каждом шаге, вторая — поиск следующего максимума исходного изображения. Практика показывает, что реализация алгоритма CLEAN без каких-либо приёмов ускорения указанных операций приводит к тому, что суммарное время их выполнения составляет около 90 % общего времени реконструкции. Ещё одна трудность при реконструкции протяжённых областей связана с тем, что для устранения высокочастотных искажений результата реконструкции требуется уменьшение фактора усиления. Это ещё больше увеличивает количество шагов и, соответственно, время реконструкции.

Для уменьшения общего времени реконструкции используются следующие приёмы. Процедура сдвига ДН вообще не проводится в явном виде. Дело в том, что даже алгоритм Кларка [4, 5], основанный на применении преобразования Фурье для сдвига ДН, в нашем случае неприемлем с точки

зрения производительности. Реконструированное изображение может содержать до $5 \cdot 10^4$ точек; такое количество преобразований Фурье для массива приемлемой размерности 512×512 требует существенного времени, поэтому используется амплитудная фильтрация ДН. Подавляющая часть значений двумерной ДН крестообразного интерферометра — это или нули, или значения, очень близкие к нулю. Поэтому вводится фильтрация ДН по заданному уровню и используются только значения, превышающие этот уровень. Вместе со значением каждой точки ДН хранятся и её оригинальные координаты. На каждом шаге реконструкции перед вычитанием значения текущей точки ДН вычисляются координаты этой точки в зависимости от её оригинальных координат и координат текущего максимума. Это требует увеличения времени доступа к каждой точке массива ДН, но за счёт резкого сокращения числа точек получается существенный выигрыш. Для ускорения поиска следующего максимума используется процедура разбиения исходного изображения на массивы меньшего размера («плавающие окна»). При этом в алгоритм вводится несколько пороговых уровней окончания реконструкции. Поиск максимума ведётся в текущем окне до тех пор, пока не будет достигнут текущий уровень окончания реконструкции. Затем осуществляется поиск следующего максимума по всему массиву. Пределы окна устанавливаются так, чтобы найденный максимум оказался в его центре. На практике размер окна поиска максимума чаще всего принимается равным 16. Дальнейшее увеличение окна ведёт лишь к увеличению времени реконструкции, не внося заметных улучшений в качество изображения. Напротив, уменьшение размера окна ведёт к ухудшению качества изображения, но не уменьшает время реконструкции. При очень малом размере окна время реконструкции может даже вырасти. Это можно объяснить тем, что уходя от поиска по всему массиву, мы можем прийти к поиску в окне с размером 1×1 , т. е. вернуться к общему поиску с добавлением лишних операций.

Для ускорения реконструкции протяжённых областей целесообразно применять в качестве «чистой» ДН свёртку изображения модельного источника и идеальной ДН. При этом возникает вопрос о зависимости размеров модельного источника от вида реконструируемой области. Можно использовать априорную информацию о размерах реконструируемой области и генерировать для каждой области свою модель, но это сопровождается неоднозначностью процедуры определения размеров, т. к. протяжённые области имеют, как правило, сложную пространственную структуру. Кроме того, многократная генерация моделей ведёт к затратам вычислительных ресурсов. Поэтому в настоящий момент используется простое предположение об увеличении размера активных областей с уменьшением их яркостной температуры. В общем это так — яркие источники, как правило, компактнее слабоконтрастных. Исключения составляют слабоконтрастные компактные источники, например, соответствующие ярким рентгеновским точкам. При реконструкции таких источников данным способом ухудшается пространственное разрешение, но такие источники можно исследовать и без реконструкции.

Таким образом, применение алгоритма CLEAN выглядит следующим образом. Исходное скорректированное изображение калибруется по уровню яркостной температуры неба и диска Солнца, составляющему $1,6 \cdot 10^3$ К. Задаются промежуточные уровни температур, которые используются как для окончания поиска максимума в окне, так и для перехода к следующей модели. Определяются размеры модельных источников, факторы усиления и пороговые значения ДН для каждого уровня. Затем выполняется непосредственно процедура CLEAN.

3. ТЕКУЩИЕ РЕЗУЛЬТАТЫ

В настоящий момент реализованы три режима реконструкции источников. В одном из них пороговые уровни яркостной температуры считаются приращением над уровнем «спокойного» Солнца, при этом модельный солнечный диск добавляется к результату по завершении реконструкции. В другом режиме уровни имеют абсолютное значение; модельный диск при этом, естественно, не добавляется.

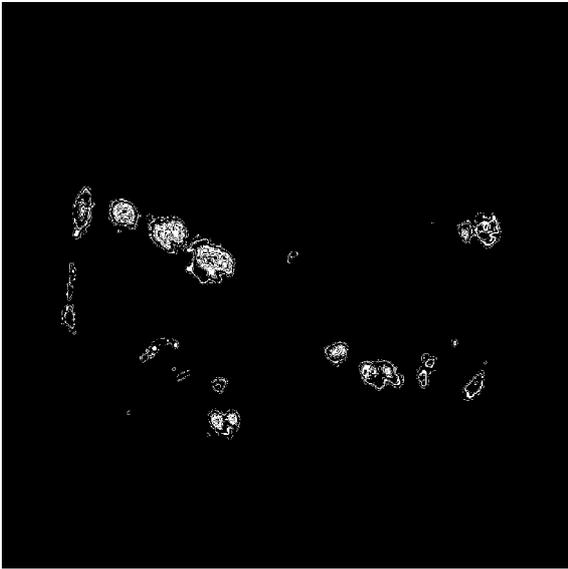


Рис. 3. Восстановленное изображение Солнца (только активные области)

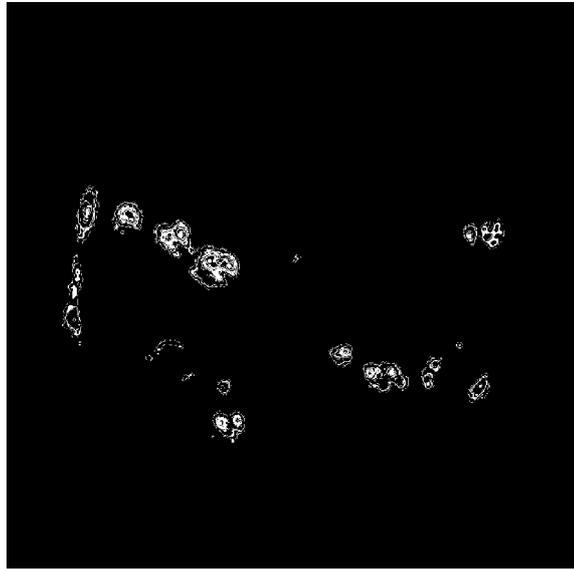


Рис. 4. Восстановленное изображение Солнца (активные области и депрессии)

Второй режим требует большего времени и большей точности амплитудно-фазового распределения (остаточные отрицательные боковые лепестки в первом случае не играют большой роли). Тем не менее второй режим представляется более предпочтительным, т. к. операция добавления модельного диска весьма искусственна. Кроме этого, второй режим не имеет нелинейности (отсечки) при приближении яркостной температуры источников к температуре солнечного диска. В этом режиме естественным образом восстанавливаются депрессии — области с яркостной температурой ниже температуры солнечного диска. Третий режим требует удаления солнечного диска перед реконструкцией. В этом режиме восстанавливаются как активные области, так и депрессии. При этом в отличие от первого режима нет проблем с реконструкцией активных областей, расположенных вне диска спокойного Солнца. Этот режим требует столь же высокой точности амплитудно-фазового распределения, как и второй режим. Затраты вычислительных ресурсов в этом случае меньше, чем для второго режима, но больше, чем для первого. В этом режиме возможна оптимизация затрат времени путём выбора отсечек по активным областям и депрессиям. Депрессии можно реконструировать и первым способом, но это либо сопряжено с отсечками, либо требует искусственной амплитудной фильтрации. На рис. 3 показано изображение, полученное этим способом из изображения, приведённого на рис. 2. Максимальная яркостная температура составляет $1,3 \cdot 10^6$ К, уровень отсечки — 10^3 К над уровнем спокойного Солнца. На рис. 4 показано изображение, полученное третьим способом из того же исходного изображения. Из сравнения рис. 3 и 4 видно, что отличия заключаются только в областях с температурой, меньшей яркостной температуры спокойного Солнца, и на лимбе. Так как эти изображения были получены с использованием одной и той же ДН, то, как и следовало ожидать, на рис. 4 больше заметны искажения, связанные с неточностями ДН. Для такой реконструкции требуется большая точность определения $\delta K(x, y)$, что наряду с восстановлением депрессий требует большого времени вычислений. Поэтому рутинным режимом на сегодняшний день является первый режим, без реконструкции депрессий. Ежедневные изображения Солнца, доступные в интернете по адресу <http://www.eastsib.ru/~ssrt>, получают именно таким способом.

Автор выражает благодарность В. В. Васину, проводившему кропотливую работу по юстировке антенно-фидерного тракта ССРТ, и Н. Н. Кардаполовой за массу полезных советов по практической реализации вышеописанной реконструкции.

Работа выполнена при финансовой поддержке РФФИ (гранты № 00–02–16456 и 00–02–16819) и Совета по поддержке ведущих научных школ (грант № 00–15–96710).

СПИСОК ЛИТЕРАТУРЫ

1. Гельфрейх Г. Б., Корольков Д. В., Смольков Г. Я., Тресков Т. А. Результаты наблюдений и исследований в период МГСС. М.: Наука, 1967. Вып 4. С. 168.
2. Högbom J. // *Astrophys. J. Suppl. Ser.* 1974. V. 15. P. 417.
3. Лесовой С. В., Васин В. И., Занданов В. Г. Определение фазовой характеристики антенно-фидерного тракта ССРТ: Препринт № 5-98 ИСЗФ СО РАН. Иркутск, 1998.
4. Clark B. G. // *Astron. Astrophys.* 1980. V. 89. P. 377.
5. Clark B. G. // *Astron. Astrophys.* 1980. V. 89. P. 417.

Институт солнечно-земной физики
СО РАН, г. Иркутск, Россия

Поступила в редакцию
15 февраля 2001 г.

RECONSTRUCTION OF RADIO IMAGES OF THE SUN OBTAINED BY THE SIBERIAN SOLAR RADIO TELESCOPE

S. V. Lesovoy

Reconstruction of Sun radio images obtained by the Siberian Solar Radio Telescope (SSRT) encounters two problems. First, since the solar radio images at a wavelength of 5.2 cm have high contrast, the amplitude–phase distribution in the antenna–feeder section should be known with the very high accuracy. Second, since such images comprise not only bright compact components but also low-contrast diffuse areas, there is the problem of deconvolution of these diffuse sources, which is inherent to the CLEAN algorithm. To solve first problem, we determine distortions of the amplitude–phase distributions by an iterative analysis of the image itself, in which the opposite sidelobes of the point-source response are compared. To suppress the influence of other sources on the response, we analyze several compact sources. The phase distortions are determined from the asymmetry of the sidelobes, and the amplitude distortions, from their values. The image is corrected in the spatial-spectrum domain after each iteration. On the one hand, the problems of reconstructing extended sources are related to the fact that the CLEAN algorithm requires significant computer resources. On the other hand, reconstructing images of extended areas requires that the number of cycles of this algorithm should be increased. Another problem consists in the fact that the use of the same “clean” beam for reconstructing compact and extended sources results in the appearance of high-frequency distortions of the latter sources. If the CLEAN algorithm is applied, then the computer resources are mainly spent to shift the beam and to find the maximum of the initial image. We decrease the time necessary for the beam shift by excluding the points near the zero value from the beam data set and by sorting the remaining data points. The time of finding the maximum was decreased by using a local search window. In addition, we use a number of cut levels and search the next maximum in the whole image only after reconstructing of the windowed image down to the current cut level. To reconstruct bright compact sources, we use the “clean” beam close to the ideal beam. To reconstruct the low-contrast diffuse areas, we use the model beams obtained by convolving the beam with the images of extended sources.

УДК 621.371.165:528.044.4

НОВАЯ СПЕКТРАЛЬНАЯ ОЦЕНКА ДЛЯ РСА-ИЗОБРАЖЕНИЯ ПОВЕРХНОСТИ ОКЕАНА

М. Б. Каневский

Излагается последовательная теория формирования изображения поверхности океана для радиолокатора с синтезированной апертурой (РСА). Основанная на композиционной модели рассеяния микроволнового излучения водной поверхностью, теория единым образом описывает как собственно изображение волнения, так и спекл-шум, являющийся существенным мешающим фактором при интерпретации РСА-изображений поверхности океана. Предложена новая оценка, позволяющая получить спектр изображения поверхности океана, со статистической точки зрения полностью свободный от спекл-шума, без потери разрешения РСА. Новый метод спектрального оценивания проверен на данных РСА европейских спутников «ERS-1» и «ERS-2».

ВВЕДЕНИЕ

Современная теория формирования изображения поверхности океана радиолокатором с синтезированной апертурой (РСА) базируется на представленной в работах В. Альперса и др. [1–3] интегральной формуле, связывающей интенсивность РСА-сигнала с радиолокационным сечением рассеяния и полем орбитальных скоростей на поверхности. Эта формула явилась отправным пунктом для численных и аналитических исследований механизма формирования изображения, выполненных как в цитированных выше, так и в последующих работах [4–7].

Однако в теории Альперса не учитывается весьма существенный фактор, а именно спекл-шум (шум пятнистости), затрудняющий интерпретацию изображения. В спектре изображения спекл-шум вызывает появление «пьедестала» со случайной формой, который вносит непредсказуемые искажения в спектр волнения.

Стандартным методом борьбы со спекл-шумом является метод сложения нескольких изображений одного и того же участка поверхности, что приводит к понижению уровня шума. Однако число суммируемых изображений не может быть большим во избежание существенного ухудшения разрешения РСА. Таким образом, метод сложения изображений не даёт кардинального решения задачи.

Предлагаются также методы, основанные на введённых априори моделях спекл-шума [8, 9]. Однако результаты их практического применения [8], а также наш теоретический анализ (см. ниже) свидетельствуют о том, что предположения, лежащие в основе этих моделей, далеко не всегда оказываются справедливыми.

В настоящей работе представлена последовательная теория формирования РСА-изображения поверхности океана, описывающая как собственно изображение волнения (соответствующая формула совпадает с формулой Альперса), так и спекл-шум без привлечения какой-либо априорной модели последнего. Единственным предположением теории, базирующейся на общеизвестной двухмасштабной модели водной поверхности, является нормальное распределение мелкой ряби на крупной волне. На основе указанной теории предложена новая оценка спектра изображения, обеспечивающая со статистической точки зрения полное устранение спекл-шумового пьедестала. Новая оценка проверена на экспериментальных данных РСА европейских спутников «ERS-1» и «ERS-2».

1. ОБЩИЕ СООТНОШЕНИЯ

Схема зондирования поверхности океана показана на рис. 1. РСА расположен на носителе, движущемся в плоскости zy параллельно оси y с постоянной скоростью V . Зондирование ведётся с расстояния R при угле падения θ_0 , обеспечивающем выполнение условий брэгговского (резонансного) рассеяния.

Для распределения амплитуды поля $\phi(\mathbf{r}' - \mathbf{r})$ на поверхности примем простую аппроксимацию:

$$\phi(\mathbf{r}' - \mathbf{r}) = \begin{cases} 1, & \mathbf{r}' - \mathbf{r} \in \Delta\mathbf{r}; \\ 0, & \mathbf{r}' - \mathbf{r} \notin \Delta\mathbf{r}, \end{cases} \quad (1)$$

где $\Delta\mathbf{r}(\Delta x, \Delta y)$ — используемая при синтезировании апертуры часть физического элемента разрешения, в пределах которой главный лепесток диаграммы направленности антенны обеспечивает близкое к однородному распределение амплитуды падающего поля.

Определим операцию синтезирования:

$$a_{\text{РСА}}(x, t) = \frac{1}{\Delta t} \int_{t-\Delta t/2}^{t+\Delta t/2} dt' a(x, t') \times \exp\left[-ik \frac{V^2}{R} (t' - t)^2\right], \quad (2)$$

где a и $a_{\text{РСА}}$ — комплексные амплитуды отражённого и синтезированного сигналов, Δt — время синтезирования, $k = 2\pi/\lambda$ — волновое число электромагнитной волны.

Нормируем величину a так, чтобы выполнялось соотношение $\langle aa^* \rangle = \sigma$, где σ — радиолокационное сечение рассеяния поверхности (символ звёздочка и угловые скобки означают комплексное сопряжение и статистическое усреднение соответственно), и в соответствии с теорией брэгговского рассеяния запишем

$$a(x, t) = \frac{2k^2}{\sqrt{\pi}} \exp(2ikR) \int_{\Delta\mathbf{r}} d\mathbf{r}' m(\mathbf{r}', t) \xi(\mathbf{r}', t) \times \exp\{2ik[(x' - x) \sin \theta_0 - \zeta(\mathbf{r}', t) \cos \theta_0 + (y' - Vt)^2/(2R)]\}. \quad (3)$$

Функция $\xi(\mathbf{r}', t)$ описывает статистически однородное поле «стандартной» гравитационно-капиллярной ряби, равномерно покрывающей крупномасштабный рельеф $\zeta(\mathbf{r}', t)$ энергонесущего волнения, а гидродинамическая модуляция реальной ряби описывается функцией $m(\mathbf{r}', t)$, включающей в себя и модуляцию амплитуды отражённого поля наклонами поверхности $\zeta(\mathbf{r}, t)$. Предполагается, что величина $\xi(\mathbf{r}', t)$ имеет нормальное распределение и корреляционную функцию $B_\xi(\mathbf{r}' - \mathbf{r}'', |t' - t''|)$.

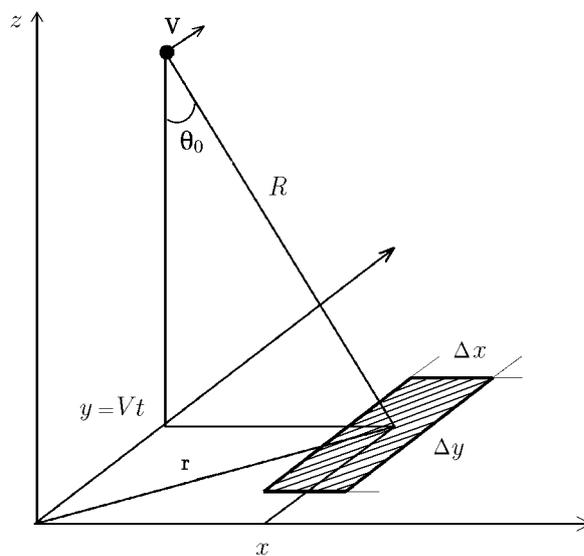


Рис. 1. Схема радиолокационного зондирования

Применив к (3) преобразование (2), получим

$$a_{\text{РСА}}(x, t) = \frac{2k^2}{\Delta t \sqrt{\pi}} \exp(2ikR) \int_{\Delta t} dt' \int_{\Delta \mathbf{r}} d\mathbf{r}' m(\mathbf{r}', t') \xi(\mathbf{r}', t') \exp\left\{2ik \left[(x' - x) \sin \theta_0 - \zeta(\mathbf{r}', t') \cos \theta_0 + (y'^2 - 2Vt'(y' - Vt) - V^2t'^2)/(2R) \right]\right\}. \quad (4)$$

Введём интенсивность РСА-сигнала $I = a_{\text{РСА}} a_{\text{РСА}}^*$ и запишем соответствующую корреляционную функцию:

$$B_I(\boldsymbol{\rho}) = \langle I(\mathbf{r}) I(\mathbf{r} + \boldsymbol{\rho}) \rangle; \quad (5)$$

вектор $\boldsymbol{\rho}$ имеет компоненты ρ_x и $\rho_y = V\tau$, где τ — временной сдвиг. Подставив в (5) выражение для интенсивности с учётом (4), выполним усреднение по реализациям ряби при «замороженном» крупномасштабном волнении, что вполне допустимо, т. к. масштабы временных изменений ряби и крупных волн различаются на два порядка. Воспользовавшись свойством нормально распределённой величины $\xi(\mathbf{r}, t)$

$$\langle \xi(\mathbf{r}_1, t_1) \xi(\mathbf{r}_2, t_2) \xi(\mathbf{r}_3, t_3) \xi(\mathbf{r}_4, t_4) \rangle = B_\xi(\mathbf{r}_1 - \mathbf{r}_2, |t_1 - t_2|) B_\xi(\mathbf{r}_3 - \mathbf{r}_4, |t_3 - t_4|) + B_\xi(\mathbf{r}_1 - \mathbf{r}_4, |t_1 - t_4|) B_\xi(\mathbf{r}_2 - \mathbf{r}_3, |t_2 - t_3|) + B_\xi(\mathbf{r}_1 - \mathbf{r}_3, |t_1 - t_3|) B_\xi(\mathbf{r}_2 - \mathbf{r}_4, |t_2 - t_4|), \quad (6)$$

получим выражение для B_I в виде суммы трёх слагаемых:

$$B_I(\boldsymbol{\rho}) = B_1(\boldsymbol{\rho}) + B_2(\boldsymbol{\rho}) + B_3(\boldsymbol{\rho}), \quad (7)$$

где

$$B_1(\boldsymbol{\rho}) = \langle I_1(\mathbf{r}) I_1(\mathbf{r} + \boldsymbol{\rho}) \rangle, \quad (8)$$

$$B_2(\boldsymbol{\rho}) = \langle I_2(\mathbf{r}, \boldsymbol{\rho}) I_2^*(\mathbf{r}, \boldsymbol{\rho}) \rangle, \quad (9)$$

$$B_3(\boldsymbol{\rho}) = \langle I_3(\mathbf{r}, \boldsymbol{\rho}) I_3^*(\mathbf{r}, \boldsymbol{\rho}) \rangle; \quad (10)$$

$$I_1 = \frac{4k^4}{\pi (\Delta t)^2} \iint_{\Delta t} dt_1 dt_2 \iint_{\Delta \mathbf{r}} d\mathbf{r}_1 d\mathbf{r}_2 m(\mathbf{r}_1, t_1) m^*(\mathbf{r}_2, t_2) B_\xi(\mathbf{r}_1 - \mathbf{r}_2, t_1 - t_2) \times \exp\left\{2ik \left[(x_1 - x_2) \sin \theta_0 - [\zeta(\mathbf{r}_1, t_1) - \zeta(\mathbf{r}_2, t_2)] \cos \theta_0 + [y_1^2 - y_2^2 - 2V(t_1 y_1 - t_2 y_2) + 2V^2 t(t_1 - t_2)] / (2R) \right]\right\}, \quad (11)$$

$$I_2 = \frac{4k^4}{\pi (\Delta t)^2} \int_{\Delta t} dt_1 \int_{\Delta(t+\tau)} dt_2 \int_{\Delta \mathbf{r}} d\mathbf{r}_1 \int_{\Delta(\mathbf{r}+\boldsymbol{\rho})} d\mathbf{r}_2 m(\mathbf{r}_1, t_1) m^*(\mathbf{r}_2, t_2) B_\xi(\mathbf{r}_1 - \mathbf{r}_2, t_1 - t_2) \times \exp\left\{2ik \left[(x_1 - x_2) \sin \theta_0 - [\zeta(\mathbf{r}_1, t_1) - \zeta(\mathbf{r}_2, t_2)] \cos \theta_0 + [y_1^2 - y_2^2 - 2V(t_1 y_1 - t_2 y_2) + 2V^2 t(t_1 - t_2) - 2V^2 t_2 \tau] / (2R) \right]\right\}, \quad (12)$$

$$I_3 = \frac{4k^4}{\pi(\Delta t)^2} \int_{\Delta t} dt_1 \int_{\Delta(t+\tau)} dt_2 \int_{\Delta \mathbf{r}} d\mathbf{r}_1 \int_{\Delta(\mathbf{r}+\boldsymbol{\rho})} d\mathbf{r}_2 m(\mathbf{r}_1, t_1) m^*(\mathbf{r}_2, t_2) B_\xi(\mathbf{r}_1 - \mathbf{r}_2, t_1 - t_2) \times \\ \times \exp \left\{ 2ik \left[(x_1 + x_2) \sin \theta_0 - [\zeta(\mathbf{r}_1, t_1) + \zeta(\mathbf{r}_2, t_2)] \cos \theta_0 + \right. \right. \\ \left. \left. + [y_1^2 + y_2^2 - 2V(t_1 y_1 + t_2 y_2) + 2V^2 t(t_1 + t_2) + 2V^2 t_2 \tau] / (2R) \right] \right\}, \quad (13)$$

Символами $\Delta(t + \tau)$ и $\Delta(\mathbf{r} + \boldsymbol{\rho})$ обозначены области, сдвинутые относительно Δt и $\Delta \mathbf{r}$ на τ и $\boldsymbol{\rho}$ соответственно; в (11)–(13), как и далее, для упрощения записи мы опускаем знак модуля в аргументе $|t_1 - t_2|$ корреляционной функции B_ξ .

В выражениях (8)–(10) усреднение проводится по реализациям крупных волн, т. к. величины I_1, I_2, I_3 , включающие в себя корреляционную функцию ряби, являются случайными функционалами крупномасштабного рельефа поверхности (заметим, что I_1, I_2 и I_3 отнюдь не являются слагаемыми I).

Из последующего будет ясно, что слагаемые B_1 и B_2 корреляционной функции B_I описывают изображение волнения и спекл-шум. Что касается B_3 , то, поскольку в подынтегральную функцию интеграла I_3 входят быстро осциллирующие множители, B_3 оказывается меньше B_2 , по крайней мере, в $(\lambda/\Delta x)^2$ раз, поэтому последним слагаемым в правой части (7) пренебрегаем.

В двух последующих разделах мы рассмотрим интегралы I_1, I_2 и получим для них физически обозримые выражения, для чего введём новые переменные

$$\mathbf{r}_1 - \mathbf{r}_2 = \boldsymbol{\rho}', \quad \mathbf{r}_1 + \mathbf{r}_2 = 2\mathbf{r}'; \quad t_1 - t_2 = \tau', \quad t_1 + t_2 = 2t'.$$

Поскольку функция $B_\xi(\boldsymbol{\rho}', \tau')$ быстро спадает при $\rho' > \rho_\xi, \tau' > \tau_\xi$, где ρ_ξ и τ_ξ — пространственный и временной масштабы ряби, подынтегральные функции в I_1 и I_2 существенно отличны от нуля лишь при $\rho' \leq \rho_\xi, \tau' \leq \tau_\xi$. Поэтому, учитывая малость ρ_ξ и τ_ξ по сравнению с соответствующими масштабами крупномасштабного волнения, положим

$$m(\mathbf{r}_1, t_1) = m(\mathbf{r}_2, t_2) = m(\mathbf{r}', t'), \quad \zeta(\mathbf{r}_1, t_1) - \zeta(\mathbf{r}_2, t_2) = \frac{\partial \zeta}{\partial x}(\mathbf{r}', t') \rho'_x + \frac{\partial \zeta}{\partial y}(\mathbf{r}', t') \rho'_y + \frac{\partial \zeta}{\partial t}(\mathbf{r}', t') \tau',$$

где ρ'_x и ρ'_y — компоненты вектора $\boldsymbol{\rho}'$. Наконец, размер Δx элемента разрешения по дальности, большой по сравнению с ρ_ξ , полагаем малым по сравнению с характерной длиной волны крупномасштабного волнения.

Напомним, что интегралы I_1, I_2 посредством (8) и (9) определяют корреляционную функцию изображения водной поверхности.

2. «ИЗОБРАЖАЮЩАЯ» ЧАСТЬ КОРРЕЛЯЦИОННОЙ ФУНКЦИИ

Обратимся к интегралу I_1 , который с учётом сделанных выше замечаний перепишем в виде

$$I_1 = \frac{4k^4}{\pi(\Delta t)^2} \Delta x \int_{-\infty}^{+\infty} dy' \int_{t-\Delta t/2}^{t+\Delta t/2} dt' |m(\mathbf{r}', t')|^2 \times \\ \times \int_{-\tau'_0(t')}^{\tau'_0(t')} d\tau' \exp \left\{ -2ik \left[\frac{\partial \zeta}{\partial t}(\mathbf{r}', t') \cos \theta_0 - \frac{V}{R} (Vt - y') \right] \tau' \right\} \int_{-\infty}^{+\infty} d\boldsymbol{\rho}' B_\xi(\boldsymbol{\rho}', \tau') \times \\ \times \exp \left\{ 2ik \left[\left(\sin \theta_0 - \frac{\partial \zeta}{\partial x}(\mathbf{r}', t') \cos \theta_0 \right) \rho'_x + \left(\frac{y' - Vt'}{R} - \frac{\partial \zeta}{\partial y}(\mathbf{r}', t') \cos \theta_0 \right) \rho'_y \right] \right\}. \quad (14)$$

Интегрирование по y' формально распространено на бесконечный интервал, поскольку, как будет видно из дальнейшего, основной вклад в интеграл даёт лишь сравнительно узкая область вблизи $y' = Vt$; так же мы поступили и с интегралом по ρ' , где существенной является лишь область $\rho' \leq \rho_\xi$ (в дальнейшем при бесконечных пределах интегрирования символ ∞ опускается). Областью интегрирования по t' и τ' является ромб с вершинами в точках $t' - t = \pm\Delta t/2$ и $\tau' = \pm\Delta t$, т. е.

$$\tau'_0 = \begin{cases} \Delta t + 2(t' - t), & -\Delta t/2 \leq t' - t \leq 0; \\ \Delta t - 2(t' - t), & 0 \leq t' - t \leq \Delta t/2. \end{cases}$$

Внутренний интеграл по ρ' в соответствии с теоремой Винера—Хинчина представим в виде

$$\int d\rho' B_\xi(\rho', \tau') \exp(-i\kappa_{\text{рез}}\rho') = 4\pi^2 \int d\varpi \Psi_\xi(\kappa_{\text{рез}}, \varpi) \exp(-i\varpi\tau'), \tag{15}$$

где $\Psi_\xi(\kappa_{\text{рез}}, \varpi)$ — пространственно-временной спектр ряби на пространственной частоте

$$\kappa_{\text{рез}} = \left\{ -2k \left(\sin \theta_0 - \frac{\partial \zeta}{\partial x}(\mathbf{r}', t') \cos \theta_0 \right); -2k \left(\frac{y' - Vt'}{R} - \frac{\partial \zeta}{\partial y}(\mathbf{r}', t') \cos \theta_0 \right) \right\},$$

которая является резонансной частотой брэгговского рассеяния. Наличие производных $\partial \zeta / \partial x$ и $\partial \zeta / \partial y$ отражает тот факт, что в данном случае $\kappa_{\text{рез}}$ — локальная пространственная частота, зависящая от наклона крупномасштабного рельефа поверхности в текущей точке \mathbf{r}' , а член $(y' - Vt')/R$ учитывает зависимость резонансной частоты от азимутального угла, под которым эта точка видна из места расположения антенны. Ввиду малости этих поправок

$$\kappa_{\text{рез}} \approx \{-2k \sin \theta_0; 0\},$$

поэтому будем считать, что рябь расположена на плоской поверхности, но в каждой точке \mathbf{r}' имеет соответствующую орбитальную скорость $\mathbf{v}_{\text{орб}}$, обусловленную крупномасштабной волной. Тогда

$$\Psi_\xi(\kappa_{\text{рез}}, \varpi) = W_\xi(\kappa_{\text{рез}}) \delta(\varpi - \Omega_{\text{рез}} - \kappa_{\text{рез}} \mathbf{v}_{\text{орб}}), \tag{16}$$

где $W_\xi(\kappa_{\text{рез}})$ — пространственный спектр ряби, Ω — её собственная временная частота, связанная с κ дисперсионным соотношением.

Подставив (16) в (15), после интегрирования по ϖ и подстановки результата в (14) получим

$$I_1 = \frac{16\pi k^4}{(\Delta t)^2} W_\xi(\kappa_{\text{рез}}) \Delta x \int dy' \int_{t-\Delta t/2}^{t+\Delta t/2} dt' |m(x, y'; t')|^2 \int_{-\tau'_0(t')}^{\tau'_0(t')} d\tau' \exp \left\{ 2ik \frac{V}{R} [(Vt - y') - \frac{R}{V} (v_{\text{рад}}(x, y'; t') + v_{\text{рез}}^\phi \sin \theta_0)] \tau' \right\}, \tag{17}$$

где $v_{\text{рад}}$ — радиальная составляющая орбитальной скорости, которая считается положительной в направлении на локатор. При написании (17) было учтено, что

$$2k \frac{\partial \zeta}{\partial t} \cos \theta_0 + \kappa_{\text{рез}} \mathbf{v}_{\text{орб}} = -2\mathbf{k} \mathbf{v}_{\text{орб}} = 2k v_{\text{рад}}.$$

Величина $v_{\text{рез}}^\phi = \Omega_{\text{рез}} / (2k \sin \theta_0)$ — собственная фазовая скорость резонансной ряби. Поскольку фазовая скорость ряби обычно существенно меньше орбитальной, в дальнейшем слагаемое $v_{\text{рез}}^\phi \sin \theta_0$ мы учитывать не будем (оно даёт лишь малый постоянный сдвиг изображения).

Интеграл по τ' в (17) легко берётся. Если предположить, что $\Delta t \ll T_0$, где T_0 — характерный период крупномасштабного волнения, то берётся также и интеграл по t' . В результате получаем

$$I_1 = \Delta x \int dy' \sigma_0(x, y'; t) \left[\frac{\sin w^0(x, y'; t)}{w^0(x, y'; t)} \right]^2, \tag{18}$$

где

$$w^0 = \frac{\pi}{\Delta y_{\text{РСА}}^0} \left[(Vt - y') - \frac{R}{V} v_{\text{рад}}(x, y'; t) \right],$$

$\sigma_0 = 16\pi k^4 |m(x, y'; t)|^2 W_\xi(\kappa_{\text{рез}})$ — удельное сечение рассеяния поверхности, а $\Delta y_{\text{РСА}}^0 = \lambda R / (2V \Delta t)$ — номинальное разрешение РСА по азимуту.

Представим (18) в более удобной для интерпретации форме, для чего условно разделим орбитальную скорость на две компоненты, крупномасштабную и мелкомасштабную по отношению к $\Delta y_{\text{РСА}}^0$:

$$v_{\text{орб}} = \hat{v} + \tilde{v}.$$

Для того, чтобы крупномасштабную компоненту скорости \hat{v} можно было считать мало меняющейся на размере номинального разрешения РСА, соответствующие длины волн в спектре орбитальных скоростей должны, по крайней мере, вдвое превышать $\Delta y_{\text{РСА}}^0$; далее скорость \hat{v} будем называть внешней, а \tilde{v} — внутренней.

Выделим на водной поверхности участок с линейным размером $\Delta y \leq \Delta y_{\text{РСА}}^0$, в пределах которого внешнюю скорость \hat{v} будем считать постоянной. Разделим этот участок на дискретные элементы, так что

$$\sigma_0(y') \propto \sum_n \sigma_{0n} \delta(y' - y_n). \tag{19}$$

Подставив (19) в (18), получим

$$I_1 = \sum_n I_1^{(n)} \propto \sum_n \sigma_{0n} \left(\frac{\sin w_n^0}{w_n^0} \right)^2, \tag{20}$$

где

$$w_n^0 = \frac{\pi}{\Delta y_{\text{РСА}}^0} \left[y - \frac{R}{V} \hat{v}_{\text{рад}} - y_n - \frac{R}{V} \tilde{v}(y_n) \right].$$

Зависимость $I_1(y)$ показана схематически на рис. 2, где $\Delta y_{\text{РСА}} \approx \Delta y_{\text{РСА}}^0 + 2(R/V) \bar{\sigma}_{\text{рад}}$, $\bar{\sigma}_{\text{рад}}$ — среднеквадратическое значение радиальной компоненты внутренней орбитальной скорости. Можно видеть, что реальный элемент разрешения оказывается «размазанным» вследствие наличия внутренних скоростей и сдвинутым как целое благодаря внешней скорости. Таким образом, формирование изображения происходит так, как если бы зависимость $v_{\text{орб}}(y)$ была сглажена, а РСА обладал номинальным разрешением $\Delta y_{\text{РСА}}$.

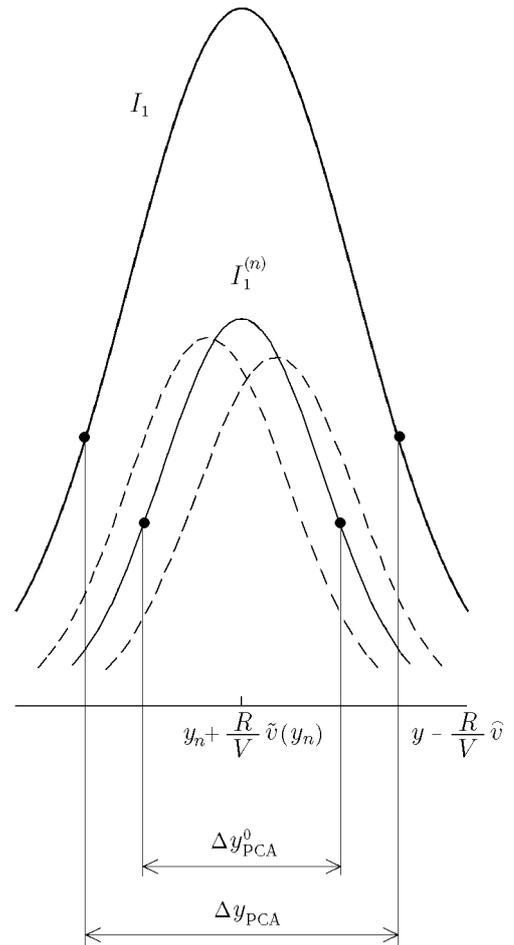


Рис. 2. Формирование реального элемента разрешения РСА

Учитывая сказанное, выражение для I_1 перепишем в следующем виде:

$$I_1 = \Delta x \int dy' \sigma_0(x, y'; t) \left[\frac{\sin w(x, y'; t)}{w(x, y'; t)} \right]^2, \quad (21)$$

где

$$w = \frac{\pi}{\Delta y_{\text{РСА}}} \left[Vt - y' - \frac{R}{V} \widehat{v}_{\text{рад}}(x, y'; t) \right].$$

В выражении (21) в отличие от (18) фигурируют реальное разрешение и внешняя скорость, определяемая крупномасштабной частью спектра с длинами волн, по крайней мере, в два раза превышающими номинальное разрешение. Эта формула практически совпадает с формулой Альперса, в которой под интегралом вместо функции $[\sin(w)/w]^2$ фигурирует $\exp(-w^2)$.

Формирование изображения волнения рассматривалось на основе формулы Альперса в ряде работ (см. введение), поэтому здесь мы не будем останавливаться на этих вопросах.

Поскольку I_1 даёт «изображающую» часть полной корреляционной функции B_I , целесообразно переобозначить I_1 и B_1 как I_i и B_i (от «image»).

3. ШУМОВАЯ ЧАСТЬ КОРРЕЛЯЦИОННОЙ ФУНКЦИИ

Рассмотрим теперь определяемый выражением (12) интеграл I_2 , который с помощью преобразований, аналогичных выполненным в предыдущем разделе, приводится к виду

$$I_2 = \Delta x (1 - |\rho_x|/\Delta x) \int dy' \sigma_0(x, y') K(x, y'; t, \tau), \quad (22)$$

где

$$K = \frac{2}{\beta} \left\{ \int_{t+\tau-\Delta t/2}^{t+\tau/2} dt' \exp(-i\alpha t') \sin[\beta(\Delta t - 2\tau + 2(t' - t))] + \int_{t+\tau/2}^{t+\Delta t/2} dt' \exp(-i\alpha t') \sin[\beta(\Delta t - 2(t' - t))] \right\},$$

$$\alpha = 2k \frac{V^2}{R} \tau, \quad \beta = 2k \frac{V}{R} \left[V(t + \tau/2) - y' - \frac{R}{V} v_{\text{рад}} \right].$$

Интеграл K преобразуем к виду

$$K = \frac{2}{\beta} \exp[-i\alpha(t + \tau/2)] \operatorname{Re} \left\{ \exp \left[i \frac{\alpha}{2} (\Delta t - \tau) \right] \int_0^{\Delta t - \tau} du \exp \left(-i \frac{\alpha}{2} u \right) \sin(\beta u) \right\}, \quad (23)$$

после чего получим

$$K = \exp[-i\alpha(t + \tau/2)] \frac{\sin w_\tau^0}{w_\tau^0} \frac{\sin w_\tau^0}{w_\tau^0}, \quad (24)$$

где

$$w_\tau^0 = \frac{\pi}{\Delta y_{\text{РСА}}^0} \left[V(t + \tau) - y' - \frac{R}{V} v_{\text{рад}} \right].$$

При выводе (24) было учтено, что для фокусированного РСА справедливо соотношение $V \Delta t \gg \Delta y_{\text{РСА}}^0$. Перейдя, как и выше, от $\Delta y_{\text{РСА}}^0$ и $v_{\text{рад}}$ к $\Delta y_{\text{РСА}}$ и $\widehat{v}_{\text{рад}}$, запишем

$$I_2 = \exp \left[-2ik \frac{V^2}{R} \tau (t + \tau/2) \right] I_s, \quad (25)$$

где

$$I_s = \Delta x (1 - |\rho_x|/\Delta x) \int dy' \sigma_0 \frac{\sin w}{w} \frac{\sin w_\tau}{w_\tau}. \quad (26)$$

Как следует из (9), именно I_s (от «speckle») определяет слагаемое B_2 корреляционной функции (7), которое мы переобозначим как B_s :

$$B_s = \langle I_s^2(\mathbf{r}, \boldsymbol{\rho}) \rangle.$$

Рассмотрим пространственные характеристики функции B_s , описывающей пятнистую структуру РСА-изображения поверхности океана.

Очевидно, характерный размер пятен вдоль оси x есть не что иное, как разрешение по дальности Δx . Вопрос об азимутальном размере (вдоль оси y) более сложен.

В подынтегральное выражение в (26) входят функции $\sin(w)/w$ и $\sin(w_\tau)/w_\tau$, которые существенно отличны от нуля вблизи точек пересечения прямых $Vt - y'$ и $V(t + \tau) - y'$ со случайной кривой $(R/V) \widehat{v}_{\text{рад}}$ (см. рис. 3). Очевидно, корни уравнений

$$Vt - y' = \frac{R}{V} \widehat{v}_{\text{рад}}(y'), \quad V(t + \tau) - y' = \frac{R}{V} \widehat{v}_{\text{рад}}(y')$$

сосредоточены в основном внутри областей $Vt \pm (R/V) \widehat{\sigma}_{\text{рад}}$ и $V(t + \tau) \pm (R/V) \widehat{\sigma}_{\text{рад}}$ соответственно. Подынтегральная функция в (26) отлична от нуля лишь в случае, если эти области перекрываются, и, таким образом, характерный масштаб функции $I_s(V\tau)$ не может превысить величину, примерно равную $2(R/V) \widehat{\sigma}_{\text{рад}}$. Заметим, что $2(R/V) \widehat{\sigma}_{\text{рад}} = 100$ м при $R/V = 100$ с и $\widehat{\sigma}_{\text{рад}} = 0,5$ м/с.

Среднее число корней (точек пересечения) $\langle N \rangle$ зависит от характера волнения и параметра

$$\bar{c} = 2\pi \frac{R}{V} \frac{\widehat{\sigma}_{\text{рад}}}{\Lambda_0} |\cos \Phi_0|,$$

где Λ_0 — характерная длина волны на поверхности, а Φ_0 — угол между генеральным направлением волнения и направлением движения носителя РСА. В работах [6, 7] на основе теории нелинейных преобразований случайных процессов получено соотношение между $\langle N \rangle$ и \bar{c} , из которого следует, что $\langle N \rangle \approx 5$ при развитом ветровом волнении, $R/V = 100$ с, $\Lambda_0 = 100$ м, $\widehat{\sigma}_{\text{рад}} = 0,5$ м/с, $\cos \Phi_0 = 1$.

В случае, когда $\langle N \rangle$ существенно превышает единицу, зависимость $B_s(\rho_y)$ принимает вид, подобный показанному на рис. 4, причём с ростом $\langle N \rangle$ крупномасштабный «хвост» функции $B_s(\rho_y)$ становится всё более выраженным. Это означает, что распределение энергии спекл-шума по пространственному спектру регулируется тем же механизмом, который формирует собственно изображение волнения (подробнее см. в [6, 7]). Данное обстоятельство отражает неаддитивный характер спекл-шума, несмотря на аддитивную форму полной корреляционной функции изображения, являющаяся следствием нормального распределения ряби.

Приведённые выше численные оценки свидетельствуют о том, что при определённых условиях значительная часть энергии спекл-шума концентрируется в спектральной области, занимаемой крупномасштабным волнением. Очевидно, в этом случае стандартный метод борьбы со спекл-шумом с помощью сложения изображений одного и того же участка поверхности (см. введение) не может быть эффективным.

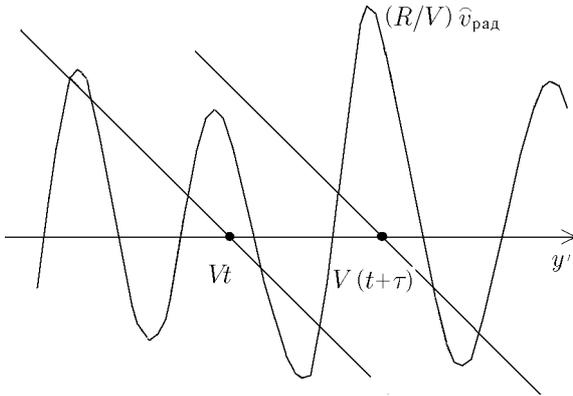


Рис. 3

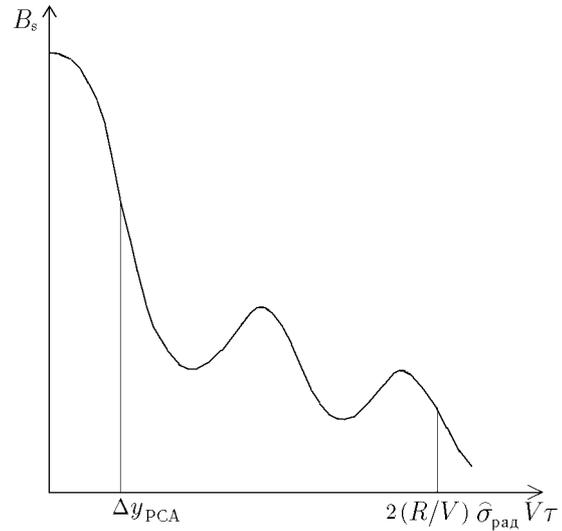


Рис. 4. Шумовая часть корреляционной функции изображения

Суммируя полученные результаты, приведём основные формулы теории:

$$B_I(\rho) = B_i(\rho) + B_s(\rho), \quad B_i = \langle I_i(\mathbf{r})I_i(\mathbf{r} + \rho) \rangle, \quad B_s = \langle I_s^2(\mathbf{r}, \rho) \rangle,$$

$$I_i = \Delta x \int dy' \sigma_0(x, y') \left[\frac{\sin w(x, y')}{w(x, y')} \right]^2,$$

$$I_s = \Delta x (1 - |\rho_x|/\Delta x) \int dy' \sigma_0(x, y') \frac{\sin w(x, y')}{w(x, y')} \frac{\sin w_\tau(x, y')}{w_\tau(x, y')},$$

$$w = \frac{\pi}{\Delta y_{PCA}} \left[Vt - y' - \frac{R}{V} \widehat{v}_{рад}(x, y') \right], \quad w_\tau = \frac{\pi}{\Delta y_{PCA}} \left[V(t + \tau) - y' - \frac{R}{V} \widehat{v}_{рад}(x, y') \right]. \quad (27)$$

Как сказано выше, B_i определяет собственно изображение волнения, а B_s описывает его пятнистую структуру, т. е. спекл-шум. Из этих формул, в частности, следует, что использование интегрального преобразования Альперса обосновано лишь для «чистого», т. е. свободного от спекл-шума, сигнала, однако авторы работы [8] без какого-либо обоснования применили его к величине $n\sigma_0$, где n — белый шум.

Отметим ещё раз, что аддитивная форма корреляционной функции B_I является следствием нормального распределения ряби, в то время как спекл-шум по своей природе является мультипликативным.

4. НОВАЯ СПЕКТРАЛЬНАЯ ОЦЕНКА

На практике из сигнала РСА устраняется фиксированный для данной точки поверхности набег фазы, поэтому в выражении (4) для a_{PCA} опустим множитель $\exp(-2ikx \sin \theta_0)$. Затем введём комплексную величину $I_c = a_{PCA}^2$, которую назовём комплексной интенсивностью, и составим корреляционную

функцию $B_c = \langle I_c(\mathbf{r})I_c^*(\mathbf{r} + \boldsymbol{\rho}) \rangle$:

$$\begin{aligned}
 B_c = & \frac{16k^8}{(\Delta t)^4 \pi^2} \exp \left[4i \frac{kV^2}{R} \tau (t + \tau/2) \right] \iint_{\Delta t} dt_1 dt_2 \iint_{\Delta(t+\tau)} dt_3 dt_4 \iint_{\Delta \mathbf{r}} d\mathbf{r}_1 d\mathbf{r}_2 \times \\
 & \times \iint_{\Delta(\mathbf{r}+\boldsymbol{\rho})} d\mathbf{r}_3 d\mathbf{r}_4 \left\langle m(\mathbf{r}_1, t_1) m(\mathbf{r}_2, t_2) m^*(\mathbf{r}_3, t_3) m^*(\mathbf{r}_4, t_4) \xi(\mathbf{r}_1, t_1) \xi(\mathbf{r}_2, t_2) \xi(\mathbf{r}_3, t_3) \xi(\mathbf{r}_4, t_4) \times \right. \\
 & \times \exp \left\{ 2ik \left[(x_1 + x_2 - x_3 - x_4) \sin \theta_0 - [\zeta(\mathbf{r}_1, t_1) + \zeta(\mathbf{r}_2, t_2) - \zeta(\mathbf{r}_3, t_3) - \zeta(\mathbf{r}_4, t_4)] \cos \theta_0 + \right. \right. \\
 & \quad \left. \left. + \frac{1}{2R} \left[y_1^2 + y_2^2 - y_3^2 - y_4^2 - 2Vt_1 (y_1 - Vt) - \right. \right. \right. \\
 & \quad \left. \left. - 2Vt_2 (y_2 - Vt) + 2Vt_3 (y_3 - V(t + \tau)) + 2Vt_4 (y_4 - V(t + \tau)) \right] \right] \right\} \left. \right\rangle. \quad (28)
 \end{aligned}$$

Используя, как и ранее, свойство (6) нормального поля ряби, перейдём от (28) к сумме трёх слагаемых (два из которых оказываются равными, а третье пренебрежимо мало), после чего с учётом (25) получим

$$B_c = 2 \exp \left[4i \frac{kV^2}{R} \tau (t + \tau/2) \right] \langle I_2^2(\mathbf{r}, \boldsymbol{\rho}) \rangle = 2 \langle I_s^2(\mathbf{r}, \boldsymbol{\rho}) \rangle = 2B_s. \quad (29)$$

Таким образом, шумовая часть B_s полной корреляционной функции изображения есть не что иное, как одна вторая корреляционной функции комплексной интенсивности, откуда следует

$$B_i(\boldsymbol{\rho}) = B_I(\boldsymbol{\rho}) - \frac{B_c(\boldsymbol{\rho})}{2}. \quad (30)$$

Выполнив фурье-преобразование обеих частей (30), запишем

$$W_i(\boldsymbol{\kappa}) = W_I(\boldsymbol{\kappa}) - \frac{W_c(\boldsymbol{\kappa})}{2}; \quad (31)$$

каждый из спектров в (31) соответствует своей корреляционной функции в (30).

Одной из общепринятых оценок спектра является периодограмма [10], которая обычно вычисляется с помощью быстрого преобразования Фурье:

$$\overline{W}_I(\boldsymbol{\kappa}) = \text{FFT}[I(\mathbf{r})],$$

где $I(\mathbf{r})$ — реализация сигнала, FFT — сумма квадратов синус и косинус-преобразований Фурье. Тогда из (31) следует выражение для оценки спектра, свободной от спекл-шума:

$$\begin{aligned}
 \overline{W}_i(\boldsymbol{\kappa}) = & \text{FFT} [\text{Re}^2 a_{\text{PCA}}(\mathbf{r}) + \text{Im}^2 a_{\text{PCA}}(\mathbf{r})] - \\
 & - \frac{1}{2} \text{FFT} [\text{Re}^2 a_{\text{PCA}}(\mathbf{r}) - \text{Im}^2 a_{\text{PCA}}(\mathbf{r})] - 2 \text{FFT} [\text{Re} a_{\text{PCA}}(\mathbf{r}) \text{Im} a_{\text{PCA}}(\mathbf{r})]. \quad (32)
 \end{aligned}$$

Первый член в правой части (32) есть стандартная спектральная оценка, а два остальных предназначены для устранения спекл-шумового пьедестала. Для того, чтобы получить устойчивую оценку спектра, периодограмма должна быть сглажена одним из известных способов [10].

Таким образом, новый метод спектрального оценивания позволяет получить спектр РСА-изображения океана, со статистической точки зрения полностью свободный от спекл-шумового пьедестала, без какой-либо потери разрешения.

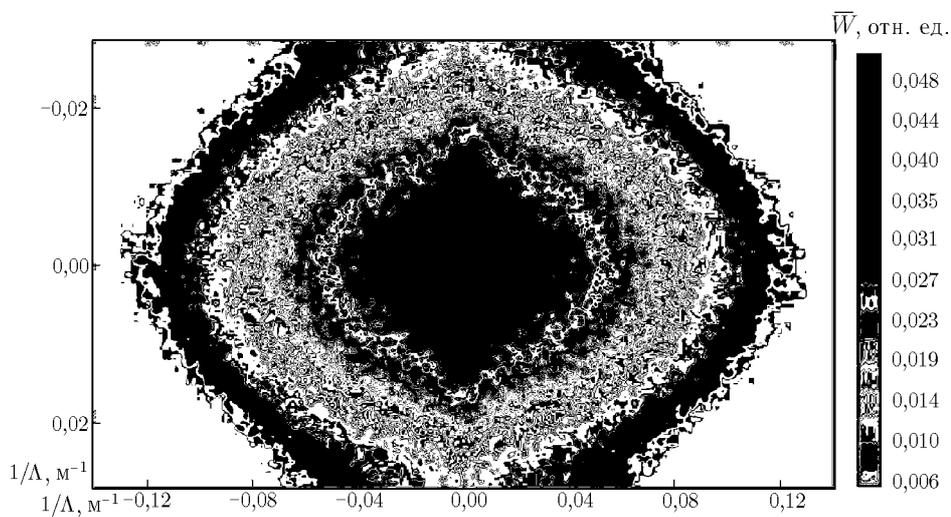


Рис. 5а. Спектр изображения поверхности океана (РСА «ERS-2»), полученный стандартным методом; Λ — длина волны на поверхности

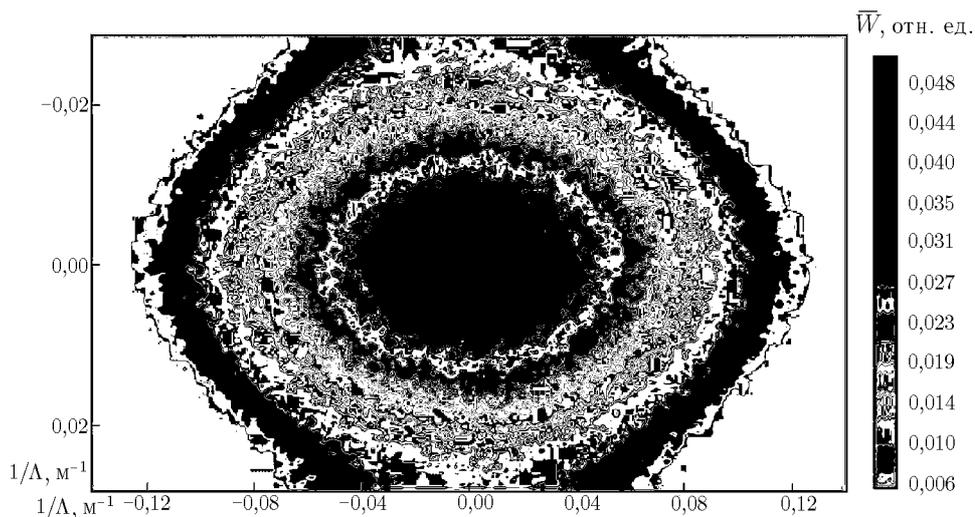


Рис. 5б. Шумовой пьедестал

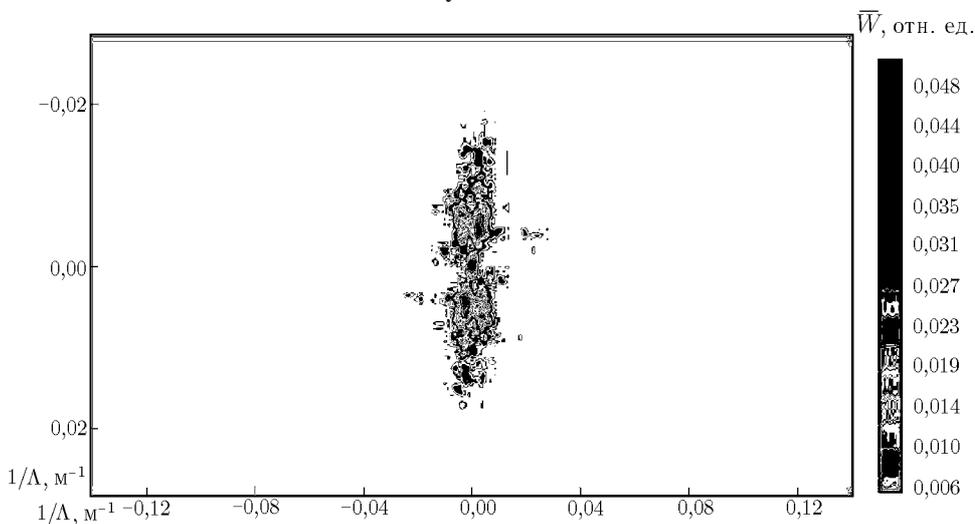


Рис. 5в. Спектр волнения поверхности, полученный новым методом

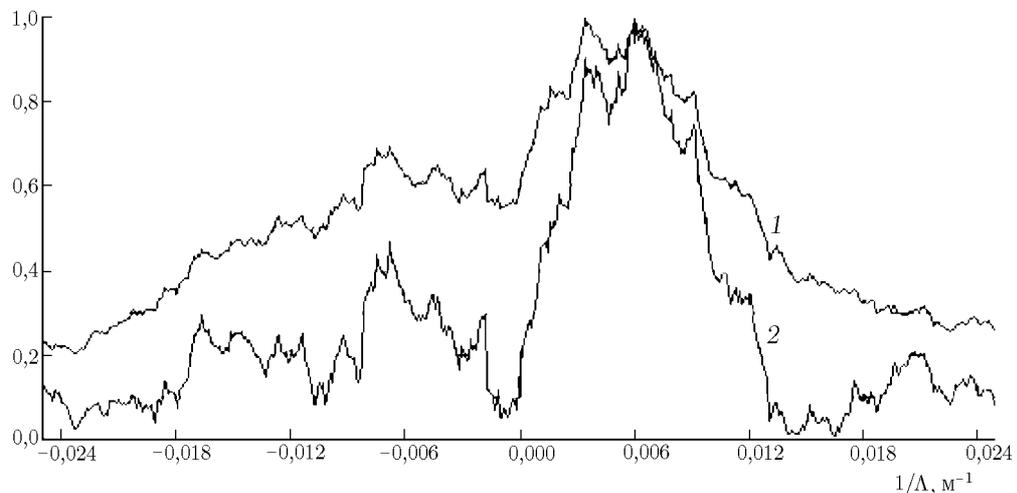


Рис. 6. Сечения нормированных спектров: кривая 1 соответствует спектру на рис. 5а, кривая 2 — спектру на рис. 5б

Новый метод был проверен на данных РСА европейских спутников «ERS-1» и «ERS-2». Эти данные, предоставленные Европейским космическим агентством, включали в себя значения $\text{Re } a_{\text{РСА}}(\mathbf{r})$ и $\text{Im } a_{\text{РСА}}(\mathbf{r})$ для 6 областей размером 50×50 км в Северном море, Мексиканском заливе и вблизи Атлантического побережья США. В этих областях были выделены квадраты 10×10 км, для каждого из которых была проведена обработка данных в соответствии с (32).

На рис. 5 приведён пример результатов обработки. Спектр, показанный на рис. 5а, получен из стандартной спектральной оценки (первый член в правой части (32)) путём сглаживания, т. е. усреднения по матрице 5×5 точек. Присутствующий на рисунке обширный ореол на самом деле есть спекл-шумовой пьедестал под спектром изображения волнения. Как следует из теории, этот пьедестал описывается суммой двух последних слагаемых в правой части (32); эта сумма представлена на рис. 5б. Спектр, свободный от спекл-шума, показан на рис. 5в (спектры, приведённые на рис. 5б, в также сглажены указанным выше способом).

На рис. 6 представлены сечения двух спектров — стандартного и свободного от спекл-шума (каждая кривая нормирована на собственный максимум). Как видно, применение нового метода спектрального оценивания позволило получить более выраженный спектральный максимум.

Подобные результаты были получены для четырёх из шести исследованных квадратов, а в двух случаях шумовой пьедестал воспроизвести не удалось. Для этого мы видим три возможные причины.

1) Развита в настоящей работе теория формирования изображения основана на резонансной теории рассеяния СВЧ радиоволн водной поверхностью. Последняя справедлива для умеренного волнения и определённого интервала углов зондирования. Если какое-либо из условий применимости теории рассеяния нарушается, то, очевидно, все последующие результаты теряют смысл.

2) Возможно, что при каких-то условиях распределение ряби существенно отличается от нормального.

3) Предложенный метод перестаёт работать в случае, когда спекл-шум в изображении не является доминирующим, а представлен в смеси с тепловым шумом. Такая ситуация возникает при очень слабом волнении и, следовательно, слабом отражённом сигнале, что имело место в одном из двух случаев, когда нам не удалось воспроизвести спектр шума. Это естественно, поскольку процедура обработки учитывает именно спекл-шум и никакой иной.

5. ЗАКЛЮЧЕНИЕ

Развитый в настоящей работе новый метод получения спектра РСА-изображения поверхности океана, в принципе, показал свою работоспособность. Представляется весьма желательным проверить метод на более обширной базе данных с тем, чтобы окончательно определить область его применимости.

Экспериментальные данные РСА спутников «ERS-1» и «ERS-2» были предоставлены Европейским космическим агентством (проект С1Р 1005) при финансовой поддержке INTAS-CNES (проект 97–1291). Автор благодарен Г. Куартли (G. Quartly), К. Гомменджинджер (C. Gommenginger) и Ж. Бенвенисте (J. Benveniste) за содействие в получении данных и обсуждение результатов на семинаре в Саутгемптонском океанографическом центре (Великобритания), а также Г. Н. Баландиной, составившей программу обработки данных.

СПИСОК ЛИТЕРАТУРЫ

1. Alpers W. R., Rufenach C. L. // IEEE Trans. Ant. and Prop. 1979. V. 27, No. 5. P. 685.
2. Alpers W. R., Ross D. B., Rufenach C. L. // J. Geophys. Res. C. 1981. V. 86, No. 7. P. 6481.
3. Hasselmann K., Raney R. K., Plant W. R., Alpers W. R. et al. // J. Geophys. Res. C. 1985. V. 90, No. 3. P. 4659.
4. Alpers W. R., Bruening C. // IEEE Trans. Geosci. and Remote Sensing. 1986. V. 24, No. 6. P. 873.
5. Bruening C., Alpers W., Hasselmann K. // Int. J. Remote Sensing. 1990. V. 11, No. 10. P. 1695.
6. Каневский М. Б., Новиков Л. В. // Исслед. Земли из космоса. 1990. № 4. С. 12.
7. Kanevsky M. B. // IEEE Trans. Geosci. and Remote Sensing. 1993. V. 31, No. 5. P. 1031.
8. Cordey R. A., Macklin J. T. // IEEE Trans. Geosci. and Remote Sensing. 1989. V. 27, No. 6. P. 666.
9. Engen G. E., Johnsen H. // IEEE Trans. Geosci. and Remote Sensing. 1995. V. 33, No. 4. P. 1047.
10. Марпл С. Л. (мл). Числовой спектральный анализ и его приложения. М.: Мир, 1990.

Институт прикладной физики РАН,
г. Нижний Новгород, Россия

Поступила в редакцию
26 марта 2002 г.

A NEW SPECTRAL ESTIMATE FOR THE SAR IMAGE OF THE OCEAN

M. B. Kanevsky

We present the consistent theory of synthetic aperture radar (SAR) imaging of the ocean surface. Based on the composite model of microwave backscattering, the theory describes both the image itself and the speckle noise, which is the factor preventing the interpretation of SAR imagery of the ocean. We propose a new spectral estimate, which, from the statistical viewpoint, eliminates completely the speckle noise in the SAR image spectrum and does not deteriorate the SAR resolution. The new technique for estimating the spectrum is tested using the SAR data of the European satellites ERS-1 and ERS-2.

УДК 533.951

О ТРАНСФОРМАЦИИ УГЛОВОГО СПЕКТРА МОЩНОСТИ РАССЕЙАННОГО ИЗЛУЧЕНИЯ В ТУРБУЛЕНТНОЙ ПОГЛОЩАЮЩЕЙ МАГНИТОАКТИВНОЙ ПЛАЗМЕ

В. Г. Гавриленко, А. В. Сорокин

Рассмотрено наклонное падение электромагнитной волны малой амплитуды на плоский слой турбулентной поглощающей плазмы в однородном внешнем магнитном поле. В приближении геометрической оптики получены уравнения для первых двух моментов углового спектра мощности рассеянного излучения. Выяснено, что в некотором направлении влияние двух факторов асимметрии задачи (наклонное падение и анизотропия среды) компенсируется, не давая проявиться эффектам смещения центра тяжести и аномального уширения углового спектра. Найдено условие такого распространения волн. Численно исследована эволюция формы углового спектра с ростом расстояния от границы плазменного слоя вне области применимости малоуглового приближения и подтверждены результаты, полученные в рамках геометрикооптического приближения.

ВВЕДЕНИЕ

В настоящее время известно, что достаточно сильное поглощение в хаотически-неоднородных средах оказывает значительное влияние на статистические характеристики многократно рассеянного излучения. Исследованию распространения волн в такого рода средах посвящены как монографии [1, 2], так и целый ряд оригинальных статей [3–13].

Из [1–5] следует, что при малых углах падения плоской волны на границу раздела однородной прозрачной и хаотически-неоднородной поглощающей сред дисперсия углового спектра мощности рассеянного излучения монотонно возрастает с ростом глубины погружения в поглощающую среду и стремится к некоторому асимптотическому значению (так называемый глубинный режим). В последующих работах было показано, что при наклонном падении излучения под достаточно большими углами на ту же границу раздела возникает целый ряд принципиально новых эффектов. В [6–9] выяснено, что максимум углового спектра мощности рассеянного излучения, как и при малых углах падения, монотонно смещается к нормали к границе. Однако ширина углового спектра меняется при увеличении расстояния от границы немонотонно. При достаточно большом поглощении существует интервал, в котором она значительно превосходит своё глубинное асимптотическое значение. В этом интервале и вблизи освещённой границы среды ширина углового спектра мощности излучения увеличивается с ростом поглощения при прочих равных условиях (эффект аномального уширения). Кроме того, в отмеченном интервале расстояний от границы угловой спектр существенно несимметричен относительно максимума (третий центральный момент значительно отличается от нуля). Эти эффекты являются следствием асимметричной постановки задачи. Такая асимметрия возникает не только в случае наклонного освещения границы раздела, она может быть неотъемлемым свойством самой среды распространения излучения вследствие её анизотропии. Характерным примером хаотически-неоднородной поглощающей анизотропной среды является турбулентная столкновительная магнитоактивная плазма. Многократное рассеяние волн в такой плазме гораздо менее изучено по сравнению с рассеянием волн в изотропных хаотических средах. В работе [10] рассматривалась эволюция статистических моментов углового спектра мощности рассеянного излучения при нормальном падении электромагнитной волны на плоскую границу полубесконечного слоя турбулентной столкновительной магнитоактивной плазмы в наклонном внешнем магнитном поле. Такой подход позволяет исследовать свойства рассеянного излуче-

ния вне рамок малоуглового приближения и получить достоверные результаты о глубинном асимптотическом режиме. Для статистических моментов углового спектра мощности рассеянного излучения до третьего порядка включительно была получена замкнутая система дифференциальных уравнений. Численное решение этой системы привело к результатам, качественно аналогичным получающимся в изотропных мутных средах при наклонном падении. Развитием [10] стала работа [11], где, используя тот же метод, было получено решение системы дифференциальных уравнений для моментов в случае наклонного падения излучения на полубесконечный слой плазмы. Однако использованный в [10, 11] метод не позволяет исследовать эволюцию формы углового спектра мощности рассеянного излучения. Эта задача для случая нормального падения электромагнитной волны на полубесконечный слой турбулентной столкновительной плазмы в наклонном сильном внешнем магнитном поле была решена в работе [12] в рамках численного решения уравнения переноса и статистического моделирования. Оба этих подхода не ограничены областью применимости малоуглового приближения, что позволило проанализировать форму углового спектра мощности излучения для различных глубин погружения в слой. Во всех трёх работах [10–12] индикатриса однократного рассеяния явно или неявно предполагалась гауссовой.

Отметим, что ранее нигде не была рассмотрена эволюция углового спектра мощности рассеянного излучения при одновременном действии различных факторов возникновения асимметрии углового спектра мощности излучения в турбулентной плазме со степенной индикатрисой однократного рассеяния. В первой части настоящей работы статистические характеристики принимаемого в плазме излучения рассчитаны аналитически в приближении комплексной геометрической оптики, позволяющем относительно просто учесть многократное рассеяние волн на малые углы. Сильные искажения углового спектра мощности излучения, вызванные асимметричным влиянием поглощения, численно исследуются во второй части статьи.

1. ПОСТАНОВКА ЗАДАЧИ

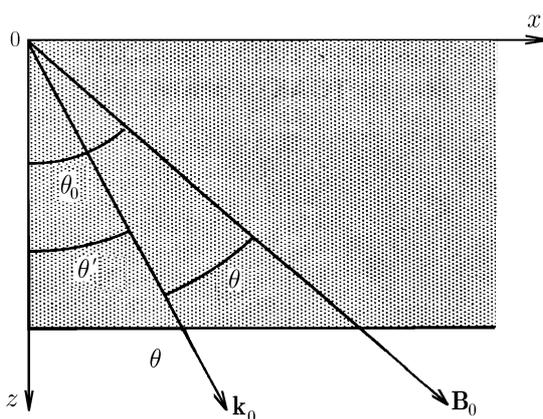


Рис. 1

Пусть на полубесконечный слой столкновительной магнитоактивной турбулентной плазмы из вакуума падает плоская электромагнитная волна. Выберем декартову систему координат таким образом, чтобы плоскость xu являлась границей раздела двух сред, а ось z была направлена вглубь плазмы. Координатную плоскость xz расположим так, чтобы она совпадала с плоскостью, образованной вектором внешнего магнитного поля \mathbf{B}_0 и волновым вектором преломлённой волны \mathbf{k}_0 . Введём также следующие обозначения (см. рис. 1): θ_0 — угол между направлением магнитного поля и осью z , θ — угол между векторами \mathbf{B}_0 и \mathbf{k}_0 , θ' — угол преломления волны относительно нормали к границе раздела,

θ_i — угол падения волны относительно нормали к границе раздела.

Электронная концентрация в плазменном слое может быть представлена в виде $p(\mathbf{r}) = p_0 + p_1(\mathbf{r})$, где p_0 — постоянная составляющая, $p_1(\mathbf{r})$ — случайная функция пространственных координат, описывающая флуктуации электронной концентрации. Характеристики рассеянного излучения фиксируются на заданной глубине z .

В наиболее общем случае магнитоактивная плазма является анизотропной средой, компоненты её тензора диэлектрической проницаемости $\hat{\epsilon}$ определяются циклической частотой падающей волны ω ,

эффективной частотой соударений электронов с другими частицами в плазме ν_{eff} , плазменной частотой $\omega_p = e(4\pi p/m)^{1/2}$ и циклической гирочастотой $\omega_B = eB_0/(mc)$. Приведём эти компоненты в системе координат, где ось z сонаправлена с магнитным полем:

$$\begin{aligned}\varepsilon_{xx} = \varepsilon_{yy} &= 1 - \frac{\omega_p^2}{2} \left(\frac{1}{\omega^2 - \omega\omega_B - i\omega\nu_{\text{eff}}} + \frac{1}{\omega^2 + \omega\omega_B - i\omega\nu_{\text{eff}}} \right), \\ \varepsilon_{xy} = -\varepsilon_{yx} &= \frac{-i\omega_p^2\omega_B}{\omega(\omega + \omega_B - i\nu_{\text{eff}})(\omega - \omega_B - i\nu_{\text{eff}})}, \\ \varepsilon_{xz} = \varepsilon_{zx} = \varepsilon_{yz} = \varepsilon_{zy} &= 0, \quad \varepsilon_{zz} = 1 - \frac{\omega_p^2}{\omega^2 - i\omega\nu_{\text{eff}}},\end{aligned}\quad (1)$$

где e и m — элементарный заряд и масса электрона соответственно, c — скорость света в вакууме, B_0 — индукция внешнего магнитного поля.

Также предположим, что характерный размер неоднородностей много больше длины волны, что позволяет использовать метод геометрической оптики для нахождения характеристик рассеянного поля.

2. РАСЧЁТ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПРИНИМАЕМОГО ИЗЛУЧЕНИЯ МЕТОДОМ КОМПЛЕКСНОЙ ГЕОМЕТРИЧЕСКОЙ ОПТИКИ

Рассчитаем электрическое поле E в плоскости xz . Распространяющаяся в плазме электромагнитная волна удовлетворяет дисперсионному уравнению

$$\mathbf{k}^2 = \frac{\omega^2}{c^2} n^2, \quad (2)$$

где $\mathbf{k}(\mathbf{r})$ — комплексный волновой вектор в данной точке пространства, n — комплексный показатель преломления. В холодной турбулентной магнитоактивной плазме этот показатель преломления определяется формулой [14]

$$n^2 = 1 + \frac{2v(1-v-is)}{2(1-is)(1-v-is) - u\sin^2\theta \pm \sqrt{u^2\sin^4\theta + 4u(1-v-is)^2\cos^2\theta}}, \quad (3)$$

где u , v и s — безразмерные параметры, отражающие вклад в показатель преломления магнитного поля, электронной концентрации и поглощения соответственно:

$$u = \omega_B^2/\omega^2, \quad v = \omega_p^2/\omega^2, \quad s = \nu_{\text{eff}}/\omega. \quad (4)$$

В нулевом приближении без учёта случайных неоднородностей ($p_1 = 0$), вклад в результирующее поле даёт только исходная падающая из вакуума на слой плоская волна. Проекции её комплексного волнового вектора на координатные оси x и z являются чисто действительными величинами:

$$k_x = \frac{\omega}{c} \sin\theta_i, \quad k_y = \frac{\omega}{c} \cos\theta_i = \sqrt{\frac{\omega^2}{c^2} - k_x^2}. \quad (5)$$

В частности, из этого следует, что производная $\partial k_{z1}/\partial k_x$ тоже является действительной величиной:

$$\frac{\partial k_{z1}}{\partial k_x} = -\frac{k_x}{k_{z1}} = -\text{tg}\theta_i. \quad (6)$$

Преломляясь на границе раздела, эта волна становится неоднородной, т. к., подчиняясь граничным условиям, проекция k_x её волнового вектора остаётся действительной, а проекция k_{z2} становится комплексной:

$$k_{z2} = \sqrt{\frac{\omega^2}{c^2} n^2 - k_x^2}. \quad (7)$$

Хаотические неоднородности электронной концентрации в плазменном слое вызывают флуктуации волнового поля в точке наблюдения. Статистические характеристики этого поля в основном определяются флуктуациями комплексной фазы φ неоднородной плоской волны в случае малоуглового рассеяния [15]. Фазовые характеристики конкретной нормальной волны в приближении геометрической оптики задаются уравнением эйконала:

$$\mathbf{k}(\mathbf{r}) = -\nabla\varphi. \quad (8)$$

Показатель преломления в анизотропной среде зависит от направления волнового вектора: $n^2 = n^2(p(\mathbf{r}), \omega, k_x, k_y)$, поэтому, действуя на обе части дисперсионного уравнения (2) оператором градиента, получим

$$\mathbf{k} \nabla \mathbf{k} - \frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n^2)}{\partial \mathbf{k}_\perp} \nabla \mathbf{k} = \frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n^2)}{\partial p} \nabla p, \quad (9)$$

где $\mathbf{k}_\perp = \mathbf{k}_\perp(k_x, k_y)$.

Полагая флуктуации электронной концентрации достаточно малыми:

$$p_1 \ll p_0, \quad (10)$$

представим фазу и волновой вектор в виде рядов:

$$\mathbf{k} = \mathbf{k}^0 + \mathbf{k}_1(\mathbf{r}) + \dots, \quad \varphi = \varphi_0 + \varphi_1 + \dots. \quad (11)$$

Формулой (5) определяется проекция на ось x невозмущённого волнового вектора \mathbf{k}^0 . Флуктуирующие слагаемые \mathbf{k}_1 и φ_1 пропорциональны малому безразмерному параметру p_1/p_0 . Учитывая слагаемые только первого порядка малости, из уравнений (9)–(11) получаем

$$k_z^0 \frac{\partial \varphi_1}{\partial z} + \left[k_x^0 - \frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n_0^2)}{\partial k_x^0} \Big|_{k_y^0=0} \right] \frac{\partial \varphi_1}{\partial x} = -\frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n_0^2)}{\partial p_0} p_1, \quad (12)$$

где $k_z^0 = k_{z2}$, а n_0 — невозмущённый показатель преломления. В этом уравнении уже учтено равенство $\partial(n_0^2)/\partial k_y^0|_{k_y^0=0} = 0$. Для краткости в дальнейшем индекс нуль везде опускается.

Интегрируя уравнение (12) вдоль его комплексных характеристик с учётом того, что

$$\frac{\partial k_{z2}}{\partial k_x} = - \left(k_x - \frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n^2)}{\partial k_x} \right) / k_{z2}, \quad (13)$$

получим

$$\varphi_1 = \frac{\alpha}{k_{z2}} \iint_{-\infty}^{+\infty} \exp(ik_x x + ik_y y) dk_x dk_y \int_0^z p_1(k_x, k_y, \xi) \exp \left[-ik_x \left(\frac{\partial k_{z2}}{\partial k_x} \right)_0 (z - \xi) \right] d\xi, \quad (14)$$

где флуктуации электронной концентрации выражены через их двумерный фурье-образ

$$p_1(x, y, \xi) = \iint_{-\infty}^{+\infty} p_1(k_x, k_y, \xi) \exp(ik_x x + ik_y y) dk_x dk_y \quad (15)$$

и введено обозначение

$$\alpha = -\frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n^2)}{\partial p}. \quad (16)$$

Рассчитаем поперечную корреляционную функцию фазы:

$$\begin{aligned} R_\varphi &= \langle \varphi_1(x, y, z) \varphi_1^*(x + \rho_x, y + \rho_y, z) \rangle = \\ &= \frac{\alpha^2}{k_{z2}^2} \iint_{-\infty}^{+\infty} \exp[ik_x(x + \rho_x) - ik'_x x + ik_y(y + \rho_y) - ik'_y y] \langle p_1(k_x, k_y, \xi) p_1^*(k'_x, k'_y, \xi') \rangle \times \\ &\quad \times \exp[-ik_x(\beta + i\gamma)(z - \xi) + ik'_x(\beta - i\gamma)(z - \xi')] dk_x dk_y dk'_x dk'_y d\xi d\xi'. \end{aligned} \quad (17)$$

Производя замену переменных

$$\rho_z = \xi - \xi', \quad \eta = (\xi + \xi')/2 \quad (18)$$

и интегрируя по этим переменным с учётом того, что масштаб корреляции параметров среды значительно меньше глубины z регистрации излучения, в итоге получаем

$$\begin{aligned} R_\varphi(\rho_x, \rho_y, z) &= 2\pi \frac{\alpha^2}{k_{z2}^2} \iint_{-\infty}^{+\infty} \exp[ik_x \rho_x + ik_y \rho_y] \frac{\exp(2k_x \gamma z)}{2k_x \gamma} [1 - \exp(2k_x \gamma z)] \times \\ &\quad \times \Phi_p(k_x, k_y, -\beta k_x) dk_x dk_y, \end{aligned} \quad (19)$$

где $\Phi_p(k_x, k_y, k_z)$ — трёхмерный спектр мощности статистически однородных флуктуаций электронной концентрации, $\partial k_{z2}/\partial k_x = \beta + i\gamma$.

Функцию корреляции комплексного поля (функцию пространственной когерентности) можно записать по определению в виде

$$\begin{aligned} R_E &= \langle E(x, y, z) E^*(x + \rho_x, y + \rho_y, z) \rangle = \\ &= E_0^2 \exp[-2z \operatorname{Im}(k_{z2}) + ik_x^0 \rho_x] \langle \exp[i\varphi_1(x, y, z) - i\varphi_1(x + \rho_x, y + \rho_y, z)] \rangle. \end{aligned} \quad (20)$$

В наиболее интересном случае сильных флуктуаций фазы ($\langle \varphi_1 \varphi_1^* \rangle \gg 1$) можно предположить, как обычно [15], что они имеют нормальное распределение. Так как корреляционная функция быстро убывает с ростом ρ_x и ρ_y , то в выражении (20) можно разложить в ряд аргумент во втором экспоненциальном множителе:

$$R_E(\rho_x, \rho_y) = E_0^2 \exp[-2z \operatorname{Im}(k_{z2}) + ik_x^0 \rho_x] \exp\left(\frac{\partial R_\varphi}{\partial \rho_x} \rho_x + \frac{1}{2} \frac{\partial^2 R_\varphi}{\partial \rho_x^2} \rho_x^2 + \frac{1}{2} \frac{\partial^2 R_\varphi}{\partial \rho_y^2} \rho_y^2\right),$$

где R_φ определяется выражением (19). Производные от корреляционной функции фазы берутся в точке $\rho_x = \rho_y = 0$.

Пространственный (угловой) спектр мощности рассеянного излучения, который имеет большое практическое значение, является фурье-образом корреляционной функции поля [15]:

$$S(k_x, k_y, z) = \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} d\rho_x d\rho_y R_E(\rho_x, \rho_y) \exp[-ik_x \rho_x - ik_y \rho_y]. \quad (21)$$

Для сильных флуктуаций фазы пространственный спектр мощности имеет гауссову форму:

$$S(k_x, k_y, z) = S_0 \exp \left[-\frac{(k_x - k_x^0 - \Delta k_x)^2}{2 \langle k_x^2 \rangle} - \frac{k_y^2}{2 \langle k_y^2 \rangle} \right], \tag{22}$$

где Δk_x определяет смещение максимума пространственного спектра мощности рассеянного излучения, вызванное случайными неоднородностями, а $\langle k_x^2 \rangle$ и $\langle k_y^2 \rangle$ определяют ширину этого спектра в плоскостях xz и yz соответственно. Выражения для Δk_x , $\langle k_x^2 \rangle$ и $\langle k_y^2 \rangle$ можно получить дифференцированием непосредственно из уравнения (19):

$$\Delta k_x = \frac{2}{i} \frac{\partial R_\varphi}{\partial \rho_x} \Big|_{\substack{\rho_x=0, \\ \rho_y=0}}, \quad \langle k_x^2 \rangle = -\frac{\partial^2 R_\varphi}{\partial \rho_x^2} \Big|_{\substack{\rho_x=0, \\ \rho_y=0}}, \quad \langle k_y^2 \rangle = -\frac{\partial^2 R_\varphi}{\partial \rho_y^2} \Big|_{\substack{\rho_x=0, \\ \rho_y=0}}. \tag{23}$$

В результате имеем

$$\Delta k_x = \frac{2\pi}{\gamma} \frac{\alpha^2}{k_{z2}^2} \iint_{-\infty}^{+\infty} \Phi_p(k_x, k_y, -\beta k_x) [\exp(2k_x \gamma z) - 1] dk_x dk_y, \tag{24}$$

$$\langle k_x^2 \rangle = \frac{\pi}{\gamma} \frac{\alpha^2}{k_{z2}^2} \iint_{-\infty}^{+\infty} \Phi_p(k_x, k_y, -\beta k_x) k_x [\exp(2k_x \gamma z) - 1] dk_x dk_y, \tag{25}$$

$$\langle k_y^2 \rangle = \frac{\pi}{\gamma} \frac{\alpha^2}{k_{z2}^2} \iint_{-\infty}^{+\infty} \Phi_p(k_x, k_y, -\beta k_x) \frac{k_y^2}{k_x} [\exp(2k_x \gamma z) - 1] dk_x dk_y. \tag{26}$$

Из формул (24)–(26) следует, что аномальные эффекты уширения и смещения центра тяжести углового спектра мощности рассеянного излучения имеют место при отличной от нуля мнимой части производной $\partial k_{z2} / \partial k_x$, причём, как отмечалось выше, на эту мнимую часть оказывает влияние как наклонное падение излучения, так и анизотропия самой среды. Приведём значение этой производной для некоторых частных случаев.

а) Наклонное падение плоской волны на изотропную поглощающую среду:

$$\frac{\partial k_{z2}}{\partial k_x} = -\frac{k_x}{k_{z2}}, \quad \text{Im} \frac{\partial k_{z2}}{\partial k_x} = 0 \Leftrightarrow k_x = 0. \tag{27}$$

б) Нормальное падение плоской волны на слой сильно замагниченной плазмы ($u \gg 1; v < 1$) [13]:

$$\frac{\partial k_{z2}}{\partial k_x} = -\frac{k_x (\cos^2 \theta_0 + \varepsilon_{zz} \sin^2 \theta_0) + k_{z2} \sin \theta_0 \cos \theta_0 (\varepsilon_{zz} - 1)}{k_{z2} (\sin^2 \theta_0 + \varepsilon_{zz} \cos^2 \theta_0) + k_x \sin \theta_0 \cos \theta_0 (\varepsilon_{zz} - 1)}, \quad \text{Im} \frac{\partial k_{z2}}{\partial k_x} = 0 \Leftrightarrow \mathbf{k} \parallel \mathbf{B}_0. \tag{28}$$

в) Наклонное падение плоской волны на слой магнитоактивной плазмы в конечном магнитном поле:

$$\frac{\partial k_{z2}}{\partial k_x} = -\frac{k_x - \frac{1}{2} \frac{\omega^2}{c^2} \frac{\partial(n^2)}{\partial k_x}}{k_{z2}}; \quad \frac{\partial(n^2)}{\partial k_x} = \frac{\partial(n^2)}{\partial \theta'} \frac{\partial \theta'}{\partial k_x} = -\frac{1}{\partial k_x / \partial \theta'} \frac{\partial(n^2)}{\partial \theta} = -\frac{1}{k \cos \theta'} \frac{\partial(n^2)}{\partial \theta}, \tag{29}$$

т. к. $\theta = \theta_0 - \theta'$, $k_x = k \sin \theta'$. Производная $\partial(n^2) / \partial \theta$ находится путём прямого дифференцирования формулы (3).

Случаи а) и б) были детально проанализированы в работах [6–9] и [12, 13] соответственно. Главным результатом этих исследований было обнаружение аномальных эффектов уширения и смещения максимума углового спектра мощности рассеянного излучения при соответствующей постановке задачи. Эти эффекты проявляются, если $k_{xm}\gamma z > 1$, где k_{xm} определяет ширину пространственного спектра флуктуаций электронной концентрации Φ_p . Как следует из формул (27) и (28), в случаях а) и б) могут иметь место аномальные эффекты. Исследование случая в) показывает, что существует некоторое выделенное направление, определяемое соотношением $\gamma = \text{Im}(\partial k_{z2}/\partial k_x) = 0$, при распространении волны вдоль которого нет смещения центра тяжести и аномального уширения углового спектра вследствие взаимной компенсации влияния наклонного падения и анизотропии среды. В дальнейшем это направление будем называть направлением компенсации. При распространении волны под другим углом аномальное уширение будет иметь место, а центр тяжести углового спектра мощности рассеянного излучения с увеличением глубины z будет смещаться к направлению компенсации.

3. ЧИСЛЕННЫЙ АНАЛИЗ УГЛОВОГО РАСПРЕДЕЛЕНИЯ ПЛОТНОСТИ ПОТОКА РАССЕЯННОГО ИЗЛУЧЕНИЯ

Полученные методом геометрической оптики особенности поведения углового спектра мощности не претендуют на полное описание процессов многократного рассеяния излучения в плазме. В первую очередь это связано с тем, что уравнение (9) решено методом возмущений. Такой подход к его решению означает, что рассеяние должно быть малоугловым. На начальном этапе распространения волны это ограничение не играет значительной роли, но уже в переходной области к глубинной асимптотической форме углового спектра (так называемый переходный режим) геометрооптическое решение отражает только некоторые качественные стороны процесса трансформации углового спектра. Не существует какого-либо универсального или частного аналитического метода исследования трансформации углового спектра из приповерхностного режима в переходный и далее в глубинный. Поэтому информацию о поведении этого спектра и зависимости его от параметров задачи можно получить только численно.

Известно, что влияние поглощения на статистические характеристики волн в хаотической среде существенно зависит от формы пространственного спектра мощности случайных неоднородностей электронной концентрации. Поэтому в качестве модели такого спектра была выбрана наиболее реалистичная степенная зависимость (характерная, например, для ионосферной плазмы). Кроме того, эта модель для простоты полагалась статистически изотропной:

$$\Phi_p(k) = \begin{cases} C(\sqrt{2}/90)^{-7/2}, & k \in [0, (\sqrt{2}/90)\omega/c]; \\ Ck^{-7/2}, & k \in [(\sqrt{2}/90)\omega/c, \sqrt{2}\omega/c]; \\ 0, & k > \sqrt{2}\omega/c, \end{cases} \quad (30)$$

где константа C определяется через условие нормировки:

$$4\pi \int_0^{\sqrt{2}\omega/c} k^2 \Phi_p(k_x, k_y, k_z = 0) dk = \langle p_1^2 \rangle. \quad (31)$$

Одним из наиболее удобных подходов для численного решения поставленной задачи является статистическое моделирование процесса распространения излучения (метод Монте-Карло). В отличие от ранее выполненных авторами расчётов [12], реализация этого метода для магнитоактивной турбулентной плазмы в данной работе учитывает зависимость дифференциального сечения рассеяния (а следовательно, и коэффициента экстинкции по рассеянию) от угла между магнитным полем и волновым вектором.

Дифференциальное сечение однократного рассеяния в магнитоактивной турбулентной плазме задаётся выражением [16]

$$\frac{\partial \sigma_s}{\partial \Omega} = \frac{1}{2\pi} \frac{e^4}{m^2 c^2} \frac{\omega^4}{\omega_p^4} \frac{n'^3 |\mathbf{q}^* (\hat{\varepsilon} - 1) \mathbf{q}'|^2}{n (|\mathbf{q}|^2 - |\mathbf{k}\mathbf{q}|^2/k^2) (\mathbf{q}'^* \hat{\varepsilon} \mathbf{q})} \langle \delta p^2 \rangle, \quad (32)$$

где \mathbf{q} — вектор поляризации нерассеянной волны, \mathbf{q}' — вектор поляризации рассеянной волны, $\langle \delta p^2 \rangle$ — спектральное распределение мощности флуктуаций электронной концентрации. Интегрируя это выражение по телесному углу Ω , можно получить явную зависимость полного сечения рассеяния на единицу объёма рассеивающей среды (коэффициента экстинкции по рассеянию) σ_s от угла θ между волновым вектором и магнитным полем. В некоторых частных случаях эту операцию можно сделать аналитически, например в случае малоуглового рассеяния в замагниченной плазме с малым на длине волны поглощением. Довольно громоздкие вычисления приводят к следующей формуле:

$$\sigma_s(v, \theta) = \frac{\sigma_{s0} v^2 \sin^4 \theta}{(1 - v \cos^2 \theta)^4}. \quad (33)$$

В наиболее общем случае эту зависимость аналитически явно выделить не удаётся, но её можно получить путём прямого численного расчёта по формуле (32).

Применительно к процессу моделирования это означает, что среднее расстояние между двумя актами рассеяния зависит от угла между магнитным полем и направлением распространения плоской волны.

Численный эксперимент проводился при различных значениях параметров u , v , s и угла наклона магнитного поля θ_0 . Для каждого набора параметров на ЭВМ проводилась серия расчётов углового спектра мощности рассеянного излучения для различных начальных углов преломления падающей волны. По окончании серии из результатов моделирования находился угол, при котором наблюдается эффект компенсации, и его значение сравнивалось с корнем уравнения $\text{Im}(\partial k_{z_2}/\partial k_x) = 0$. Во всех рассмотренных нами случаях с точностью до ошибок численного счёта эти две величины совпадали.

В качестве иллюстрации на рис. 2–5 приведены результаты моделирования для $u = 1,25$; $v = 0,2$; $s = 0,02$ и угла наклона магнитного поля $\theta_0 = 20^\circ$. Величина $\text{Im}(\partial k_{z_2}/\partial k_x)|_{\theta_0=20^\circ}$ при этих параметрах обращается в нуль при $\theta = 15,9^\circ$; её зависимость от угла θ в окрестности указанного направления приведена на рис. 2. На рис. 3–5 показаны графики зависимостей первых трёх статистических моментов углового спектра мощности от глубины погружения в слой. На всех графиках глубина отложена в безразмерных единицах $\sigma_s(\theta_0 - \theta') z$; кривая 1 соответствует направлению компенсации, кривая 2 — углу преломления $\theta = 30^\circ$, кривая 3 — нормальному падению на слой.

На рис. 3 изображён график зависимости центра тяжести углового спектра мощности от глубины погружения в слой. Очевидно, что результаты численного моделирования полностью совпадают с предсказаниями теории: в случае преломления в направлении компенсации смещение центра тяжести отсутствует. При преломлении под другим углом центр тяжести асимптотически стремится к направлению компенсации с увеличением толщины слоя. Отметим, что в формуле (24) при малой глубине z экспоненту под интегралом можно разложить в ряд. Поскольку спектр мощности флуктуаций Φ_p является чётной функцией аргумента k_x , то интеграл от линейного члена разложения будет равен нулю. Следовательно, первым ненулевым слагаемым будет интеграл от квадратичного члена, что даёт $\Delta k_x \propto z^2$. Этот результат неплохо подтверждается в численном эксперименте.

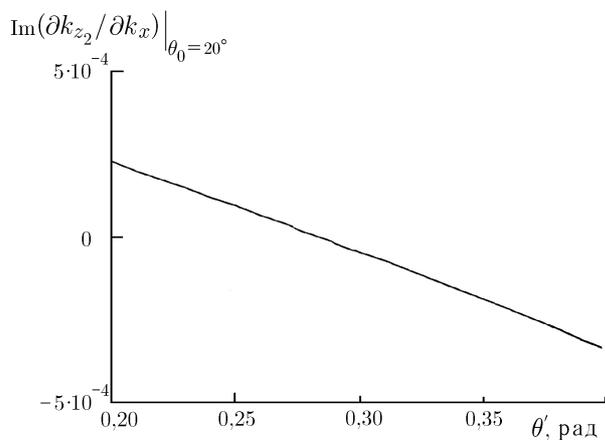


Рис. 2

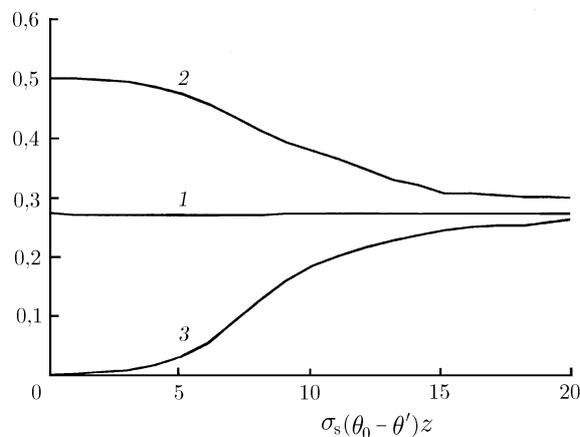


Рис. 3

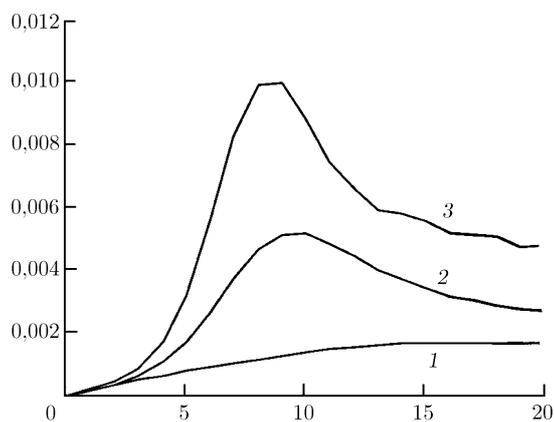


Рис. 4

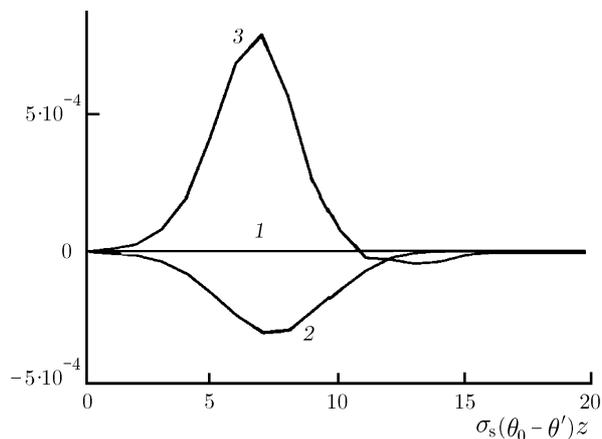


Рис. 5

График зависимости второго центрального момента углового спектра (дисперсии в плоскости падения) представлен на рис. 4. Как и следует из геометрических соображений, в направлении компенсации эта зависимость очень похожа на аналогичную для случая нормального падения плоской волны на слой изотропной среды [2]. Это сходство подчёркивается монотонным характером обеих зависимостей. Два фактора асимметрии задачи (анизотропия среды и наклонное падение на слой) полностью компенсируют друг друга, не давая проявиться аномальному уширению спектра и немонотонному характеру роста дисперсии. Вне направления компенсации (кривые 2 и 3) действие одной из причин асимметрии преобладает, и имеют место немонотонная зависимость дисперсии от толщины слоя и эффект аномального уширения. В формуле (25) при малой глубине z экспоненту под интегралом можно разложить в ряд, что в итоге приводит к известной [2] линейной зависимости $\langle k_x^2 \rangle \propto z$, хорошо согласующейся с результатами численного моделирования.

На рис. 5 изображена зависимость третьего центрального момента углового спектра (коэффициента асимметрии) от глубины погружения в слой. Как и ожидалось, в направлении компенсации пространственный спектр симметричен (с точностью до ошибок моделирования коэффициент асимметрии равен нулю), а вне его знак третьего момента отражает стремление (слева или справа) центра тяжести углового спектра мощности излучения к направлению компенсации.

Отметим далее ряд особенностей формы углового спектра, которые удалось численно обнаружить даже для относительно простого случая сильно замагниченной плазмы. Хотя центр тяжести углового

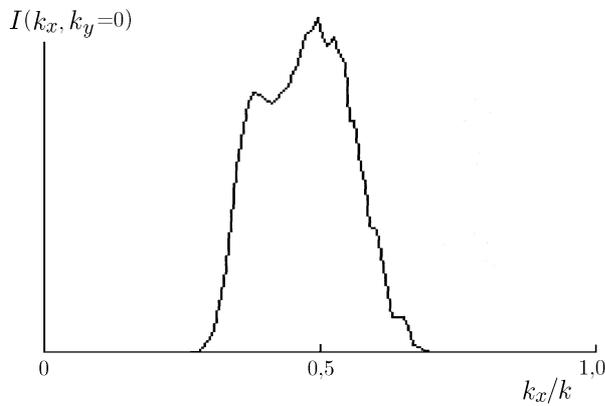


Рис. 6

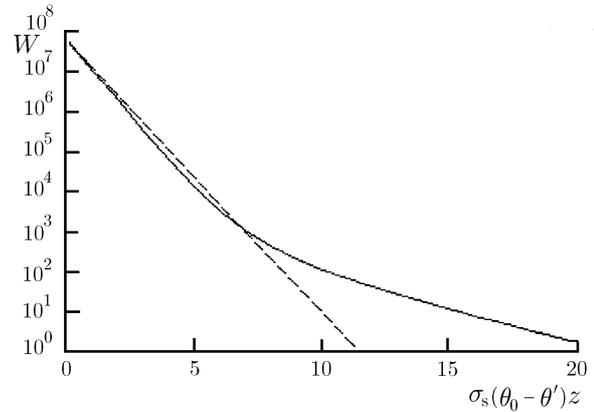


Рис. 7

спектра мощности рассеянного излучения с ростом толщины слоя стремится к направлению магнитного поля при любом начальном угле преломления (эффект компенсации отсутствует), в такой среде существуют другие, не менее интересные эффекты трансформации углового спектра. Характерным свойством замагниченной плазмы является сильное ослабление поглощения и рассеяния излучения вдоль направления магнитного поля. Численный эксперимент показывает, что даже для случая нормального падения волны на границу плазмы в наклонном внешнем магнитном поле ($u = 10$; $v = 0,7$; $s = 0,001$; $\theta_0 = 30^\circ$) это обстоятельство приводит к «двугорбой» форме углового спектра мощности рассеянного излучения в переходном режиме (см. рис. 6; по оси абсцисс отложена величина k_x/k , по оси ординат — нормированный угловой спектр мощности $I(k_x, k_y = 0)$ на глубине $\sigma_s(\theta_0 - \theta')z = 8,5$). Левый максимум на этом графике соответствует рассеянными на относительно малые углы волнам, в которых изначально была сосредоточена большая часть энергии всего рассеянного излучения. Правый максимум соответствует изначально ничтожной доле многократно рассеянного излучения, попавшего на направление, близкое к направлению магнитного поля. С увеличением толщины слоя совокупная энергия первой группы волн уменьшается из-за значительного поглощения в направлении её распространения, а потери энергии второй группы волн гораздо меньше, что приводит к выравниванию энергий обеих групп волн на определённой глубине. Это обстоятельство способствует росту дисперсии углового спектра мощности рассеянного излучения в переходном режиме, что обеспечивает его аномальное уширение.

Другим результатом того же численного эксперимента является тот факт, что при приближении к направлению внешнего магнитного поля рассеянное излучение медленнее теряет свою энергию (эффект статистического просветления плазмы). Зависимость полной энергии рассеянного излучения от толщины слоя при нормальном падении по результатам моделирования показана на рис. 7 (по оси абсцисс отложена нормированная толщина слоя, а по оси ординат — полная энергия W рассеянного излучения в логарифмическом масштабе). В отсутствие рассеяния аналогичная зависимость имела бы вид прямой линии, обозначенной на рис. 7 пунктиром. Ввиду того, что на малых глубинах смещение центра тяжести углового спектра ещё незначительно, основная часть рассеянного излучения проходит немного больший путь, чем излучение, идущее строго по нормали. Как следствие, рассеянное излучение поглощается больше, и его полная энергия несколько меньше энергии излучения в отсутствие рассеяния. Затем начинает проявляться эффект статистического просветления плазмы за счёт приближения центра тяжести углового спектра к направлению магнитного поля. Рис. 7 служит наглядной иллюстрацией сказанного.

4. ЗАКЛЮЧЕНИЕ

Полученные в данной статье результаты показывают, что угловой спектр мощности электромагнитного излучения, многократно рассеянного в слое столкновительной турбулентной магнитоактивной плазмы, существенно зависит от двух факторов, обуславливающих асимметрию задачи: наклонного падения волны на слой и анизотропии среды. Каждый из этих факторов, действуя отдельно, приводит к искажениям углового спектра мощности рассеянного излучения по сравнению с полностью симметричным случаем при нормальном падении плоской волны на изотропную хаотическую среду. Но их взаимное действие может усиливать или ослаблять отмеченные выше эффекты аномального влияния поглощения на угловой спектр мощности излучения. Наиболее интересным, на наш взгляд, является вывод о том, что в определённом направлении действие этих факторов взаимно компенсируется, так что вдоль него излучение распространяется как при нормальном падении на изотропную среду. Методом геометрической оптики было найдено условие такого распространения и с помощью численного эксперимента подтверждена правильность этого расчёта.

Другими результатами, достойными внимания, являются «двугорбый» характер углового спектра мощности и просветление плазмы в случае сильного магнитного поля. Эти эффекты не были ранее отмечены в литературе. Дело в том, что при аналогичной с рассматриваемой в данной работе постановке задачи в [13] расчёты ограничены малоугловым приближением, а в [10, 11] применялся метод статистических моментов, не позволяющий исследовать форму углового спектра мощности излучения. Важно также отметить, что в предыдущих работах по многократному рассеянию в магнитоактивной столкновительной турбулентной плазме [10–12] предполагалось для простоты, что индикатриса однократного рассеяния на плавных неоднородностях электронной концентрации плазмы быстро спадает (по гауссовому закону) при больших углах рассеяния. Между тем влияние поглощения на форму спектра и полную интенсивность излучения наиболее сильно проявляется на тех расстояниях от границы, где центр тяжести углового спектра значительно смещается. Поэтому в более близком к реальности случае степенной модели индикатрисы, использованной в данной работе, отмеченные выше эффекты проявляются значительно сильнее и приобретают качественно новые черты. Это объясняется тем, что степенная индикатриса, одинаковая с гауссовой на уровне 0,5 от максимального значения, даёт более плавное спадание мощности рассеянного излучения при дальнейшем увеличении угла рассеяния. Что же касается зависимостей первых трёх статистических моментов от расстояния до границы слоя, то в данной работе и предыдущих статьях они получились качественно аналогичными.

Полученные результаты могут представлять интерес для специалистов, разрабатывающих системы дальней радиосвязи, а также системы зондирования искусственных и природных плазменных образований (ионосферы, солнечной короны, полупроводниковой плазмы и т. п.) методом просвечивания.

Работа выполнена при частичной финансовой поддержке Конкурсного центра фундаментального естествознания (грант № Е00–3.5–469) и Российского фонда фундаментальных исследований (программа поддержки ведущих научных школ, грант № 00–15–96741).

СПИСОК ЛИТЕРАТУРЫ

1. Оптика океана. Т. 1. Физическая оптика океана / Под ред. А. С. Моница. М.: Наука, 1983.
2. Долин Л. С., Левин И. М. Справочник по теории подводного видения. Л.: Гидрометеиздат, 1991.
3. Долин Л. С. // ДАН СССР. 1981. Т. 260, № 6. С. 1344.
4. Ремизович В. С., Шехмаметьев Ш. А. // Изв. вузов. Радиофизика. 1990. Т. 33, № 2. С. 202.
5. Зеге Э. П., Полонский И. Н., Чайковская Л. И. // Изв. АН СССР. Физика атмосферы и океана. 1987. Т. 23, № 5. С. 486.

6. Гавриленко В. Г., Петров С. С. // Изв. вузов. Радиофизика. 1985. Т. 28, № 11. С. 1 408.
7. Аистов А. В., Гавриленко В. Г., Петров С. С. // Изв. АН СССР. Физика атмосферы и океана. 1991. Т. 27, № 3. С. 297.
8. Gavrilenko V. G., Petrov S. S. // Waves in Random Media. 1992. V. 2. P. 237.
9. Аистов А. В., Гавриленко В. Г. // Оптика и спектроскопия. 1995. Т. 72, № 4. С. 672.
10. Аистов А. В., Гавриленко В. Г. // Физика плазмы. 1996. Т. 22, № 8. С. 712.
11. Аистов А. В., Гавриленко В. Г., Джандиери Г. В. // Изв. вузов. Радиофизика. 1999. Т. 42, № 12. С. 1 165.
12. Гавриленко В. Г., Петров С. С., Семериков А. А., Сорокин А. В. // Изв. вузов. Радиофизика. 2000. Т. 43, № 1. С. 28.
13. Gavrilenko V. G., Semerikov A. A., Jandieri G. V. // Waves in Random Media. 1999. V. 9. P. 427.
14. Гинзбург В. Л. Распространение электромагнитных волн в плазме. М.: Наука, 1967.
15. Рытов С. М., Кравцов Ю. А. Татарский В. И. Введение в статистическую радиофизику. Ч. 1 и 2. М.: Наука, 1978.
16. Электродинамика плазмы / Под ред. А. М. Ахиезера. М.: Наука, 1974.

Нижегородский государственный
университет им. Н. И. Лобачевского,
г. Нижний Новгород, Россия

Поступила в редакцию
28 января 2002 г.

TRANSFORMATION OF THE ANGULAR POWER SPECTRUM OF SCATTERED RADIATION IN A TURBULENT ABSORBING MAGNETOACTIVE PLASMA

V. G. Gavrilenko and A. V. Sorokin

We consider oblique incidence of a small-amplitude plane electromagnetic wave on a layer of turbulent absorbing plasma in a uniform external magnetic field. The equations for the first two statistical momenta of the angular power spectrum of scattered radiation are derived in the geometrical-optics approximation. We show that two asymmetry factors of the problem, i.e., the oblique incidence and the influence of the anisotropic medium, compensate each other along a certain direction and impede the appearance of the effect of displacement of the power-spectrum maximum and the effect of anomalous spectrum broadening. We find the condition at which such a case of wave propagation is realized. The dependence of the angular-spectrum shape on the distance from the plasma-layer boundary is studied numerically without using the small-angle approximation. The calculations confirm the results obtained using the geometrical-optics approximation.

УДК 535.2+535.853

О ВЛИЯНИИ НЕКОМПЛАНАРНОЙ НАМОТКИ СВЕТОВОДА В ВОЛОКОННОМ КОЛЬЦЕВОМ ИНТЕРФЕРОМЕТРЕ НА РАЗНОСТЬ ФАЗ ВСТРЕЧНЫХ ВОЛН

Е. И. Якубович, Г. Б. Малыкин

Рассмотрен ряд физических явлений (фотоупругих и геометрооптических), возникающих в волоконном кольцевом интерферометре с некомпланарной намоткой волоконного контура. Показано, что ни одно из этих явлений не может привести к увеличению чувствительности такого интерферометра к вращению.

ВВЕДЕНИЕ

Эффект Саньяка [1–3] (см. также обзоры [4–11]) заключается в том, что волна во вращающемся кольцевом интерферометре приобретает фазовый сдвиг относительно другой, встречной волны, который пропорционален угловой скорости вращения, частоте волны, а также площади проекции поперечного сечения интерферометра на плоскость, ортогональную вектору угловой скорости. Этот кинематический эффект специальной теории относительности (СТО) [11–13], как показано в [10, 11, 13], является следствием релятивистского закона сложения скоростей — линейной скорости вращения интерферометра и фазовой скорости каждой из встречных волн. В настоящее время кроме оптического диапазона эффект Саньяка зарегистрирован для радиоволн [14], рентгеновских лучей [15], а также для волн не электромагнитной природы — волн де Бройля материальных частиц: электронов [7, 16], нейтронов [17, 18], атомов кальция [19], натрия [20] и цезия [21]. Теоретически показано, что данный эффект существует для γ -лучей [22], акустических, поверхностных акустических и поверхностных магнитостатических волн (так называемых медленных волн) [10, 23] и для волн де Бройля π -мезонов [24].

Следует отметить, что разность фаз встречных волн в оптических кольцевых интерферометрах и, в частности, в волоконных кольцевых интерферометрах (ВКИ) [25, 26] определяется не только эффектом Саньяка, но и рядом других эффектов, которые не связаны с вращением. Эти эффекты подробно рассмотрены в обзоре [26]. В случае, если ВКИ используется в качестве волоконного гироскопа — датчика угловой скорости вращения (следует иметь в виду, что ВКИ находит также ряд других применений [25, 26]), то все эффекты, приводящие к возникновению не связанной с вращением разности фаз встречных волн в ВКИ, являются паразитными. Разность фаз встречных волн, обусловленная паразитными эффектами, аддитивно складывается с разностью фаз встречных волн, обусловленной эффектом Саньяка, и, соответственно, приводит к возникновению аддитивной ошибки в измерении угловой скорости вращения. При этом рассматриваемые эффекты никоим образом не могут привести к мультипликативной ошибке в измерении угловой скорости вращения, а также не могут увеличить обусловленную эффектом Саньяка разность фаз встречных волн в ВКИ, т. е. не могут увеличить чувствительность ВКИ к вращению.

Однако в последнее время появился ряд экспериментальных [27–29] и теоретических работ [30–32], в которых утверждается, что величина эффекта Саньяка в ВКИ зависит от характера намотки на катушку одномодового волоконного световода (ОВС), из которого изготовлен контур ВКИ. В работах [27–29] наблюдалось изменение не связанной с вращением разности фаз встречных волн в ВКИ при изменении шага тороидальной намотки ОВС контура ВКИ на цилиндр. При этом в [27–29] утверждалось, что в эксперименте регистрировалось изменение геометрооптической фазы [33–43] (именуемой иногда фазой Берри [37–39, 44]) встречных волн, которая связана с неплоской (некомпланарной)

намоткой световода, имеет различный знак для встречных волн и, следовательно, приводит к появлению фазовой невязности ОВС контура ВКИ для встречных волн. В работах [30–32] обсуждается возможность существенного увеличения обусловленной вращением разности фаз встречных волн для случая неплоской, например тороидальной, намотки ОВС.

Цель данной работы — провести всесторонний анализ различных эффектов, связанных с неплоской намоткой световода в контуре ВКИ, и наглядно продемонстрировать ошибочность утверждений, высказанных в [27–32]. В частности, будет показано, что:

1) геометрооптический эффект, связанный с неплоской намоткой ОВС, — в действительности это хорошо известный эффект Рытова [33–35] (см. также [36, 39, 40, 45]) — не может привести к появлению фазовой невязности ОВС для встречных волн;

2) величина эффекта Саньяка не зависит от формы намотки ОВС контура ВКИ, а определяется только площадью проекции контура на плоскость, ортогональную вектору угловой скорости.

Рассмотрим явления, которые могут приводить к различного рода фазовым набегам для каждой из встречных волн в световоде произвольной конфигурации.

1. ПОГОННАЯ ФАЗА В ОВС

В поляризационно изотропном световоде всегда имеет место фазовый набег $\varphi = 2\pi Ln/\lambda$ (так называемая погонная фаза), который пропорционален оптической длине световода Ln , где L — длина световода, n — эффективный коэффициент преломления ОВС, λ — длина волны света в вакууме. Погонная фаза не зависит от типа намотки световода при условии, что связанные с намоткой деформации не влияют на величину n . Поскольку встречные волны в контуре ВКИ проходят один и тот же оптический путь, то даже в том случае, когда связанные с намоткой деформации влияют на величину n , погонные фазы встречных волн равны. Следовательно, погонная фаза никоим образом не влияет на результат интерференции встречных волн.

2. АНИЗОТРОПИЯ ОВС, ОБУСЛОВЛЕННАЯ НЕПЛОСКОЙ НАМОТКОЙ СВЕТОВОДА

Выше мы рассмотрели так называемый изотропный ОВС, для которого имеет место вырождение ортогональных поляризационных мод. Следует иметь в виду, что в реальном ОВС всегда имеет место та или иная поляризационная анизотропия, и эффективный коэффициент преломления для ортогональных поляризационных мод несколько отличается: от 10^{-3} для ОВС с сильным линейным двулучепреломлением до $10^{-9} \div 10^{-8}$ для так называемых связных ОВС, в которых линейное двулучепреломление в значительной мере подавлено. Неплоская намотка световода обязательно предполагает некоторый его изгиб, который приводит к возникновению дополнительного линейного двулучепреломления [46, 47]:

$$\beta[\text{рад/м}] = 1,34 \cdot 10^6 (d/D)^2, \quad (1)$$

где $\beta = 2\pi \Delta n/\lambda$, Δn — разность показателей преломления медленной и быстрой мод ОВС, d — диаметр световода, D — диаметр намотки. Обусловленное изгибом линейное двулучепреломление возбуждает собственное (невозмущённое) линейное двулучепреломление ОВС. В случае, если быстрые и медленные оси собственного и наведённого двулучепреломления совпадают, происходит их сложение, а если быстрая ось собственного двулучепреломления совпадает с медленной осью наведённого двулучепреломления — вычитание. В общем случае происходит их векторное сложение, и оси суммарного двулучепреломления оказываются повернутыми относительно осей собственного двулучепреломления.

Кроме того, при некомпланарной намотке, в принципе, могут возникнуть кручения световода, которые вызовут торсионные деформации и приведут к появлению наведённого циркулярного двулучепреломления, величина которого определяется выражением [46–49]

$$\beta_c = 2(1 - g)t, \quad (2)$$

где g — коэффициент фотоупругости, который для кварцевых ОВС составляет $0,065 \div 0,08$ [46–49], t — погонное кручение волокна (в радианах на метр). В случае, если ОВС имеет собственное (невозмущённое) линейное двулучепреломление, собственные поляризационные моды скрученного ОВС становятся эллиптическими [50, 51]. В принципе, можно избежать торсионных деформаций световода при неплоской намотке. Для этого волокно наматывается на некоторый каркас (катушку или поверхность цилиндра) не непосредственно, а предварительно уложенное в тефлоновую трубочку [27–29, 39, 52–55], обладающую малым коэффициентом трения. При этом тефлоновая трубочка испытывает торсионные деформации, а волокно не испытывает, поскольку имеет возможность свободно раскручиваться в трубочке таким образом, чтобы торсионные напряжения отсутствовали.

Пусть укладка световода на цилиндрическую катушку осуществляется таким образом, что оси собственного линейного двулучепреломления сохраняют свою ориентацию от одного витка к другому, как это показано на рис. 1а. (Отметим, что это условие выполняется автоматически при укладке так называемого ленточного ОВС [56–58] (см. рис. 1а), поперечное сечение которого имеет форму, близкую к прямоугольной, а оси двулучепреломления совпадают с нормальными к сторонам прямоугольника.) Тогда в волокне возникнет торсионное кручение. Погонное кручение t волокна и угол T поворота осей линейного двулучепреломления на одном витке волокна составят [39]

$$t = 2\pi h / [(\pi D)^2 + h^2], \quad (3)$$

$$T = 2\pi h / \sqrt{(\pi D)^2 + h^2}, \quad (4)$$

где D — диаметр цилиндра, h — шаг намотки ОВС.

3. ЭФФЕКТ РЫТОВА В ОВС ПРИ НЕКОМПЛАНАРНОЙ НАМОТКЕ

Эффект Рытова [33, 34] (см. также [35, 36, 39, 40, 45]) заключается в том, что при распространении линейно поляризованного света по некомпланарной траектории возникает поворот плоскости поляризации на угол, численно равный телесному углу, который описала касательная к траектории луча [35]. Вследствие эффекта Рытова также возникает дополнительный фазовый набег, равный углу поворота плоскости поляризации [39, 40]. Следует особо отметить, что это взаимный набег фазы, который одинаков для прямого и обратного направления распространения луча, т. е. эффект Рытова не приводит к фазовой невзаимности. Некомпланарная лучевая траектория имеет место, в частности, в случае распространения луча в световоде, имеющем неплоскую намотку. Поскольку контур ВКИ состоит из большого числа витков ОВС, уложенного на катушку, то конец каждого витка смещён относительно начала витка на расстояние не менее поперечного диаметра световода. Величина эффекта Рытова строго определена в частном случае циклической эволюции естественного трёхгранника [35], что характерно для геометрической фазы (фазы Берри) [39, 44].

Следует также отметить, что эффект Рытова приводит не только к повороту плоскости поляризации луча, но и к соответствующему повороту поперечной структуры луча в случае, если таковая имеется. Так, в работе [59] экспериментально наблюдался поворот спекл-структуры света в многомодовом све-

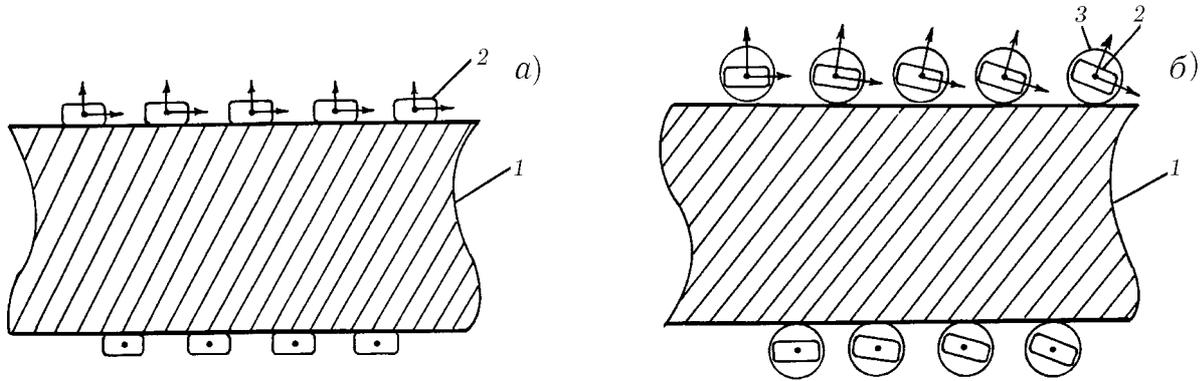


Рис. 1. Продольное сечение катушки с намотанным на неё ленточным ОВС: *a* — световод намотан непосредственно на катушку; *б* — световод уложен внутри намотанной на катушку тефлоновой трубочки. 1 — катушка; 2 — ленточный ОВС; 3 — тефлоновая трубочка. Стрелками обозначены направления осей линейного двулучепреломления ОВС

товоде при его намотке на катушку, численно соответствующий пространственному углу, который описала касательная к световоду.

Угол поворота плоскости поляризации излучения, обусловленного эффектом Рытова, на одном витке ОВС на цилиндрической катушке составляет [39]:

$$\gamma = 2\pi \left(1 - \frac{h}{\sqrt{(\pi D)^2 + h^2}} \right) = 2\pi (1 - \cos \theta), \quad (5)$$

где θ — угол между осью световода и образующей цилиндра. Поворот плоскости поляризации света, обусловленный эффектом Рытова, экспериментально регистрируют следующим образом. Чтобы избежать торсионных деформаций световода при намотке (см. выше), волокно наматывается на катушку уложенным в тефлоновую трубочку [27–29, 39, 52–55]. В этом случае реализуется некомпланарная структура витков ОВС, однако без кручения волокна.

На рис. 1б показано расположение осей линейного двулучепреломления и поперечного сечения ленточного ОВС, уложенного на катушку в тефлоновой трубочке. Рис. 1б делает совершенно прозрачным физический смысл эффекта Рытова в ОВС при отсутствии торсионных кручений: это поворот поперечного сечения и, следовательно, осей линейного двулучепреломления световода от одного витка к последующему, обусловленный некомпланарной структурой витков, т. е. геометрическими свойствами намотки световода на катушку.

Отметим здесь, что величина эффекта Рытова не зависит от наличия или отсутствия торсионных кручений волокна — это два совершенно независимых эффекта. Поскольку эффект Рытова вызывает поворот плоскости поляризации, его можно формально свести к появлению некоторой дополнительной оптической активности, т. е. эффективного циркулярного двулучепреломления, которое в лабораторной системе отсчёта составляет

$$\beta_c = 2t_{\text{эфф}}, \quad (6)$$

где $t_{\text{эфф}} = \gamma / \sqrt{(\pi D)^2 + h^2}$ — эффективное погонное кручение волокна, обусловленное геометрическими свойствами его намотки. Вместе с тем в сопровождающей такое кручение винтовой системе координат [50, 51, 60] собственные поляризационные моды ОВС остаются эллиптическими. В частности, если собственное (невозмущённое) двулучепреломление ОВС было линейным, то в винтовой системе координат поляризационные моды будут оставаться линейно поляризованными.

4. ПОЛЯРИЗАЦИОННАЯ НЕВЗАИМНОСТЬ ВКИ, СВЯЗАННАЯ С НЕКОМПЛАНАРНОЙ НАМОТКОЙ ОВС

Явление поляризационной невязимности в ВКИ с немонахроматическим источником излучения [26, 61–75] связано с неодинаковыми условиями возбуждения встречных волн на противоположных входах контура ВКИ и имеет место при отсутствии «настоящих» невязимных эффектов, таких, как эффекты Фарадея и Физо, и даже при отсутствии вращения, т. е. при отсутствии эффекта Саньяка. Простое и наглядное объяснение этого эффекта можно привести для случая, когда в ОВС контура ВКИ имеет место только линейное двулучепреломление [26, 74, 75]. В этом случае, если оси линейного двулучепреломления ОВС на входе контура развёрнуты относительно друг друга, а возбуждение осуществляется циркулярно поляризованным излучением, то возбуждение встречных волн происходит со сдвигом во времени. Таким образом, встречные волны приобретают разность фаз уже на входе контура ВКИ. Когда встречные волны после обхода контура интерферируют на выходе ВКИ, они имеют такой же относительный фазовый сдвиг, как и на входе контура, который не зависит от наличия или отсутствия вращения и в общем случае аддитивно складывается с фазовым сдвигом, обусловленным эффектом Саньяка. Это приводит к возникновению ошибки при измерении угловой скорости вращения.

Следует отметить, что поляризационная невязимность является геометрооптическим эффектом и может рассматриваться как проявление геометрической фазы. Как показано в наших работах [71, 72], фазовая невязимность может быть определена как площадь сферического треугольника на сфере Пуанкаре [40, 76], вершины которого задаются тремя точками: первая соответствует состоянию поляризации излучения на входе ВКИ, а две другие — состояниям поляризации излучения на обоих выходах контура ВКИ (рассмотрение проводится по отдельности для медленной и быстрой поляризационных мод ОВС контура). Таким образом, поляризационная невязимность относится к тому же классу явлений, к которому принадлежат эффекты Рытова [33–35] и Панчаратнама [77] (см. также [37–40]), и может рассматриваться как проявление фазы Берри в оптике.

Как показано в наших работах [71, 73], в случае, когда в ОВС контура ВКИ имеет место только циркулярное двулучепреломление, поляризационная невязимность отсутствует. В общем случае, когда собственные поляризационные моды ОВС контура ВКИ являются эллиптическими, простого объяснения механизма поляризационной невязимности нет, однако в [73] получено выражение, которое связывает разность фаз на выходе ВКИ с поляризационными параметрами ОВС контура и состоянием поляризации излучения на входе ВКИ:

$$\Delta\varphi = \arctg \left[\frac{S_L \sin(\alpha_1 + \alpha_2) \sin(4\chi) \cos(\alpha_1 - \alpha_2 - 2\eta) + S_3 \sin[2(\alpha_1 + \alpha_2)] \cos(2\chi)}{\cos^2(\alpha_1 + \alpha_2) + 1/2 \sin[2(\alpha_1 + \alpha_2)] \sin(2\chi)} \right], \quad (7)$$

где $\chi = \arctg(b/a)$, a и b — большая и малая оси эллипса поляризации [35, 76], $b/a = \beta_c/(\beta_e + \beta)$, $\beta_e = \sqrt{\beta^2 + \beta_c^2}$ [78, 79]; параметр b/a и углы α_1 , α_2 характеризуют эллиптичность собственных поляризационных мод ОВС контура ВКИ и ориентацию больших осей эллипсов поляризации на концах контура соответственно, $S_L = \sqrt{S_1^2 + S_2^2}$ — линейная составляющая поляризации излучения, η — азимут плоскости линейной составляющей поляризации излучения на входе ВКИ, S_1 , S_2 и S_3 — компоненты вектора Стокса [35, 76] поляризации излучения на входе ВКИ (S_3 — круговая компонента поляризации).

Из выражения (7) видно, что величина $\Delta\varphi$ — обусловленная поляризационной невязимностью разность фаз встречных волн (сдвиг нуля ВКИ) — будет меняться как при изменении эллиптичности χ поляризационных мод ОВС контура, так и при изменении ориентации α_1 и α_2 больших осей их эллипсов поляризации на концах контура.

Отметим, что выражение (7) записано для ВКИ без поляризатора. При наличии поляризатора (так называемая минимальная схема ВКИ [65]; такие ВКИ применяются для целей навигации) сдвиг нуля

в главном порядке величины уменьшается в число раз, равное коэффициенту экстинкции поляризатора по электрическому полю [66]. Однако в ВКИ с поляризатором практически невозможно наблюдать изменение сдвига нуля при изменении конфигурации намотки. Вследствие этого все известные эксперименты по данному вопросу [27–29] проводились в ВКИ без поляризатора, чему и соответствует выражение (7).

Очевидно, что изменения как линейного (см. выражение (1)), так и циркулярного (см. выражение (2)) двулучепреломления в ОВС, которые связаны с некомпланарной намоткой световода и подробно рассмотрены в разделе 2, приведут к изменению эллиптичности χ его собственных поляризационных мод. В этом случае в соответствии с выражением (7) изменится обусловленная поляризационной невзаимностью разность фаз встречных волн в контуре ВКИ. Как было отмечено выше, эта разность фаз аддитивно складывается с разностью фаз, обусловленной эффектом Саньяка, и никоим образом не может привести к увеличению чувствительности ВКИ к вращению.

В работах [27–29] утверждается, что их авторы экспериментально наблюдали изменение геометро-оптической фазы Берри (в действительности — проявление эффекта Рытова). Для этой цели использовался ВКИ без поляризатора с контуром из слабоанизотропного ОВС круглого сечения, часть которого была уложена в трубочку из полимера и наматывалась на цилиндрическую катушку. В экспериментах [27] на вход ВКИ подавалось циркулярно поляризованное излучение, в [28, 29] — линейно поляризованное излучение. При изменении шага намотки менялась разность фаз встречных волн на выходе ВКИ. Объяснение этого явления приведено в начале данного раздела. ОВС не испытывает кручения внутри полимерной трубочки только в том случае, когда его концы остаются свободными [27–29, 52–55]. В экспериментах же [27–29] азимуты осей двулучепреломления ОВС на входе контура ВКИ были зафиксированы в светоделителе [28, 29] или в светозапитывающем устройстве [27]. При изменении шага намотки происходит торсионное кручение волокна и, следовательно, изменение эллиптичности собственных поляризационных мод ОВС контура ВКИ. В этом случае разность фаз встречных волн в ВКИ изменяется вследствие изменения поляризационной невзаимности контура ВКИ [61–75]. Таким образом, сам эффект Рытова не может вызвать разность фаз встречных волн в ВКИ, поскольку это взаимный эффект, но сопровождающие эксперименты [27–29] торсионные напряжения волокна приводят к изменению эллиптичности и азимутов собственных поляризационных мод ОВС и, соответственно, к изменению фаз встречных волн.

5. ЭФФЕКТ САНЬЯКА В ВОЛОКОННОМ КОЛЬЦЕВОМ ИНТЕРФЕРОМЕТРЕ С НЕКОМПЛАНАРНОЙ НАМОТКОЙ ОВС

Рассмотрим теперь влияние некомпланарности намотки ОВС на величину эффекта Саньяка, т. е. на чувствительность ВКИ к вращению.

Прежде всего отметим, что ещё в 1904 г. А. Майкельсон с точностью до малых релятивистских поправок порядка v^2/c^2 , где v — линейная скорость вращения интерферометра, c — скорость света, показал для частного случая плоского оптического контура кольцевого интерферометра произвольной формы, произвольно расположенного относительно центра вращения, что величина эффекта Саньяка прямо пропорциональна площади контура и угловой скорости вращения и не зависит от формы контура [80].

Как известно [10–12, 26], эффект Саньяка является эффектом СТО первого порядка по v/c . Однако, если пренебречь весьма малыми поправками порядка v^2/c^2 , то рассмотрение можно проводить с помощью метода геометрической оптики.

Приведём элементарный вывод выражения, описывающего эффект Саньяка для произвольной трассы распространения света. Анализ проведём в лабораторной системе отсчёта. Забегая вперед, отметим, что полученное нами приращение эйконала (фазы) в лабораторной системе отсчёта будет

равно его приращению в движущейся системе, т. к. эйконал является истинным скаляром (см., например, [45]).

Приращение эйконала при распространении по бесконечно малому участку трассы из-за вращения последнего с угловой скоростью Ω равно ¹

$$d\Phi = \frac{\partial\Phi}{\partial l} \frac{dl}{dt} dt = (1 - \mu) \mathbf{k} \mathbf{v} dt, \quad (8)$$

где $\mathbf{k} = k\mathbf{a}$ — волновой вектор (в вакууме), $k = 2\pi/\lambda$, λ — длина волны света в вакууме, \mathbf{a} — единичный вектор, касательный к элементу трассы, $\mathbf{v} = [\Omega, \mathbf{l}]$ — скорость изменения бесконечно малого участка оптической длины трассы из-за вращения, $l = nr$ — оптическая длина элемента трассы, r — геометрическая длина этого же элемента трассы, n — коэффициент преломления среды, μ — коэффициент увлечения света, dt — время прохождения светом этого участка. Как показано в работе Эйнштейна [81] в случае, когда источник света, кольцевой интерферометр и заполняющая его оптическая среда вращаются как единое целое, $\mu = 1 - 1/n^2$. Отсюда

$$d\Phi = n^2 (1 - \mu) k [\Omega, \mathbf{r}] \mathbf{a} n dt. \quad (9)$$

Поскольку $n^2 (1 - \mu) \equiv 1$, то приращение эйконала равно

$$d\Phi = k [\Omega, \mathbf{r}] d\mathbf{r}/c \quad (10)$$

и, как известно [9–11, 14, 82], не зависит от коэффициента преломления среды.

Для замкнутой трассы (не обязательно плоской) приращение эйконала для волны, бегущей в одну сторону, равно

$$\Delta\Phi = \frac{k}{c} \Omega \oint [\mathbf{r}, d\mathbf{r}] = \frac{2k}{c} \Omega \mathbf{S}, \quad (11)$$

а для встречных волн разность фаз в два раза больше и определяет величину эффекта Саньяка Φ_S для случая неплоской трассы:

$$\Phi_S = 2 \Delta\Phi = \frac{4k}{c} \Omega \mathbf{S}. \quad (12)$$

Здесь $\mathbf{S} = \frac{1}{2} \oint [\mathbf{r}, d\mathbf{r}]$. Вектор \mathbf{S} не зависит от положения системы координат и полностью определяется видом и направлением контура интегрирования. Действительно, перенесём начало координат в точку \mathbf{r}_0 . При этом вид интеграла \mathbf{S} изменится: $\mathbf{S} = \frac{1}{2} \oint [\mathbf{r} - \mathbf{r}_0, d\mathbf{r}]$. Однако, как легко видеть, интеграл $\oint [\mathbf{r}_0, d\mathbf{r}]$ равен нулю, поэтому \mathbf{S} не изменится.

Покажем, что вектор \mathbf{S} имеет следующий геометрический смысл: площадь проекции контура трассы на произвольную плоскость (с учётом направления обхода) равна проекции вектора \mathbf{S} на нормаль к данной плоскости (см. рис. 2). Таким образом, вектор \mathbf{S} может быть назван «вектором площади». Для

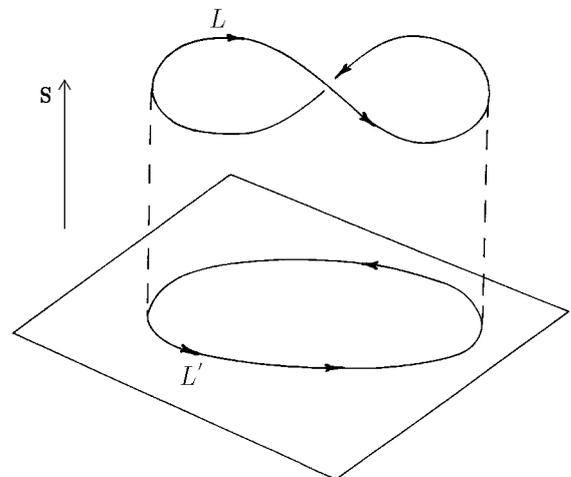


Рис. 2. Некомпланарная замкнутая трасса светового луча и её проекция на плоскость. \mathbf{S} — «вектор площади» контура L ; L' — проекция контура L на плоскость. Стрелками обозначено направление обхода трассы

¹ Поскольку цель настоящей работы — изучить влияние геометрии трассы на эффект Саньяка, мы не будем учитывать его зависимость от других факторов. Последние не зависят от геометрии трассы, и, если будут существенны, то приведут лишь к изменению скалярного множителя перед выражением для набег фазы.

этого умножим вектор \mathbf{S} скалярно на произвольный единичный вектор \mathbf{e} . После несложных преобразований вида $\mathbf{S}\mathbf{e} = \frac{1}{2}\mathbf{e} \oint [\mathbf{r}, d\mathbf{r}] = \frac{1}{2} \oint [\mathbf{e}, \mathbf{r}] d\mathbf{r}$ ясно, что интеграл берётся по плоскости, ортогональной к единичному вектору \mathbf{e} , и, следовательно, равен площади проекции контура трассы на эту плоскость. В частности, для плоской кривой модуль вектора \mathbf{S} равен площади, ограниченной этой кривой. Отсюда следует также, что ракурсов, под которыми заданная замкнутая кривая видна с наибольшей площадью, всего два и их направления параллельны \mathbf{S} .

Итак, эффект Саньяка максимален при параллельности «вектора площади» \mathbf{S} вектору угловой скорости и не может быть увеличен за счёт некомпланарности контура трассы.

6. ЗАКЛЮЧЕНИЕ

Основной результат, полученный в данной работе, заключается в том, что разность фаз встречных волн во вращающемся волоконном кольцевом интерферометре прямо пропорциональна площади проекции волоконного контура на плоскость, ортогональную вектору угловой скорости. Таким образом, при заданной длине волоконного контура ВКИ оптимальной формой намотки волокна является кольцо, плоскость которого ортогональна вектору угловой скорости: в этом случае площадь контура будет максимальной и, следовательно, разность фаз встречных волн, обусловленная эффектом Саньяка, также будет максимальной.

Поскольку габариты реальных датчиков угловой скорости вращения на основе ВКИ не превышают $10 \div 20$ см, то оптимальной формой намотки волокна является многовитковое кольцо. Полученный вывод с точки зрения специалистов по волоконной гироскопии является в высшей степени банальным — именно таким образом и изготовлены волоконные контуры всех современных ВКИ, так же был намотан и контур из радиокабеля кольцевого интерферометра в работе Берштейна [14], предназначенный для работы на длине волны 10 м.

Полученный нами в самом общем виде результат заключается в том, что неплоская намотка контура ВКИ, вопреки утверждению работ [30–32], никоим образом не может привести к росту его чувствительности к вращению. Последняя определяется исключительно площадью проекции контура на плоскость, ортогональную вектору угловой скорости.

Следует отметить, что в работах [30–32] содержится также утверждение о якобы неголономном характере эффекта Саньяка (подобного рода утверждения приводились и ранее [83, 84]), т. е. о том, что его величина не может однозначно определяться площадью проекции контура на плоскость, ортогональную вектору угловой скорости. Из полученных в разделе 5 результатов следует, что данное утверждение является неверным.

Применение некомпланарной намотки контура ВКИ повлечёт исключительно ухудшение характеристик ВКИ, что связано с необходимостью увеличивать длину ОВС контура по сравнению со случаем компланарной намотки для сохранения чувствительности ВКИ к вращению (так называемого масштабного коэффициента ВКИ):

1) уменьшится соотношение сигнал/шум на фотоприёмнике, поскольку увеличение длины ОВС контура приведёт к росту потерь света, что, в свою очередь, приведёт к уменьшению реальной чувствительности ВКИ, которая ограничена уровнем дробовых шумов излучения [14, 85];

2) возрастут габариты прибора (многослойная волоконная катушка большей длины может занимать много места в малогабаритном датчике вращения), что крайне нежелательно;

3) увеличение длины ОВС контура заметно увеличит стоимость прибора (цена высококачественного сильноанизотропного ОВС, который характеризуется весьма малым уровнем связи поляризационных мод на случайных неоднородностях ОВС и, соответственно, позволяет достигнуть малой величины дрейфа нуля ВКИ, составляет десятки долларов США за 1 м).

Если же сохранить длину ОВС контура ВКИ при некомпланарной намотке такой же, как и при компланарной намотке, то это приведёт к уменьшению масштабного коэффициента, поскольку при той же длине волокна суммарная площадь проекции витков контура на плоскость, ортогональную вектору угловой скорости, будет всегда меньше, чем при плоской намотке.

Что касается содержащегося в работах [27–29] утверждения об экспериментальном наблюдении изменения разности фаз встречных волн при изменении шага неплюской намотки волокна контура ВКИ, то, как было показано в разделе 4, объяснение этого явления заключается в том, что при фиксации концов волокна изменение шага намотки приводит к торсионному кручению волокна и, следовательно, к изменению эллиптичности собственных поляризационных мод ОВС контура ВКИ. В этом случае происходит изменение разности фаз встречных волн вследствие изменения поляризационной невязимности контура ВКИ [61–75].

Авторы выражают благодарность Вл. В. Кочаровскому и Ю. А. Кравцову за обсуждение вопросов, связанных с проявлением эффекта Рытова в ОВС. Работа частично поддержана РФФИ (гранты № 00–15–96732, 00–02–17344 и 00–15–96772) и Советом по государственной поддержке ведущих научных школ (грант № 00–15–96772).

СПИСОК ЛИТЕРАТУРЫ

1. Sagnac M. G. // *Compt. Rend.* 1913. V. 157, No. 17. P. 708.
2. Sagnac M. G. // *Compt. Rend.* 1913. V. 157, No. 25. P. 1 410.
3. Sagnac M. G. // *J. Phys. (Paris). Ser. 5.* 1914. V. 4, No. 3. P. 177.
4. Вавилов С. И. Собрание сочинений М.: Изд-во АН СССР, 1956. Т. 4. С. 9.
5. Post E. J. // *Rev. Mod. Phys.* 1967. V. 39, No. 2. P. 475.
6. Франкфурт У. И., Френк А. М. Оптика движущихся тел. М.: Наука, 1972. 210 с.
7. Hasselbach F., Nicklaus M. // *Phys. Rev. A.* 1993. V. 48, No. 1. P. 143.
8. Anderson R., Bilger H. R., Stedman G. E. // *Am. J. Phys.* 1994. V. 62, No. 11. P. 975.
9. Малыкин Г. Б. // *УФН.* 1997. Т. 167, № 3. С. 337.
10. Вугальтер Г. А., Малыкин Г. Б. // *Изв. вузов. Радиофизика.* 1999. Т. 42, № 4. С. 373.
11. Малыкин Г. Б. // *УФН.* 2000. Т. 170, № 12. С. 1 325.
12. Логунов А. А., Чугреев Ю. В. // *УФН.* 1988. Т. 156, № 1. С. 137.
13. Малыкин Г. Б. // *УФН.* 2002. Т. 172, № 8. С. 969.
14. Берштейн И. Л. // *ДАН СССР.* 1950. Т. 75, № 5. С. 635.
15. Высоцкий В. И., Воронцов В. И., Кузьмин Р. Н., Безирганян П. А., Ростомян А. Г. // *УФН.* 1994. Т. 164, № 3. С. 309.
16. Neutze R., Hasselbach F. // *Phys. Rev. A.* 1998. V. 58, No. 1. P. 557.
17. Werner S. A., Staudenmann J. L., Colella R. // *Phys. Rev. Lett.* 1979. V. 42, No. 17. P. 1 103.
18. Dresden M., Yang C. N. // *Phys. Rev. D.* 1979. V. 20, No. 8. P. 1 846.
19. Riehle F., Kisters T., Witte A., Helmcke J., Borde C. H. // *Phys. Rev. Lett.* 1991. V. 67, No. 2. P. 177.
20. Lene F., Hammond T. D., Smith E. T., Chapman M. S., Rubinstein R. A., Pritchard D. E. // *Phys. Rev. Lett.* 1997. V. 78, No. 5. P. 760.
21. Gustavson T. L., Bouyer P., Kasevich M. A. // *Phys. Rev. Lett.* 1997. V. 78, No. 11. P. 2 046.
22. Пат. 3102953 США / Wallace A.; Заявл. 09.07.58; Оpubл. 03.09.63.
23. Newburgh R. G., Blacksmith P., Budreau A. J., Sethares J. C. // *Proc. IEEE.* 1974. V. 62, No. 12. P. 1 621.
24. Малыкин Г. Б. // *Изв. вузов. Радиофизика.* 1998. Т. 41, № 6. С. 767.
25. Lefevre H. *The fiber-optic gyroscope.* Boston, London: Artech House, 1993. 313 p.

26. Андропова И. А., Малыкин Г. Б. // УФН. 2002. Т. 172, № 7. С. 849.
27. Frins E. M., Dultz W. // Opt. Commun. 1997. V. 136, No. 5–6. P. 354.
28. Senthilkumaran P., Culshaw B., Thursby G. // J. Opt. Soc. Am. B. 2000. V. 17, No. 11. P. 1914.
29. Senthilkumaran P., Thursby G., Culshaw B. // Opt. Lett. 2000. V. 25, No. 8. P. 533.
30. Staroverova M. E. // Proc. Int. Conf. "Geometrization of Physics III", Kazan State University, Kazan, 1–5 October 1997. P. 153.
31. Bashkov V. I., Sintsova Yu. V. // Gravitation and Cosmology. 1999. V. 5, No. 4 (20). P. 319.
32. Bashkov V. I., Sintsova Yu. V. // Reports Math. Phys. 2001. V. 48, No. 3. P. 353.
33. Рытов С. М. // ДАН СССР. 1938. Т. 18, № 4–5. С. 263.
34. Рытов С. М. // Труды физического института. М., Л.: Изд-во АН СССР, 1940. Т. 2, вып. 1. С. 41.
35. Владимирский В. В. // ДАН СССР. 1941. Т. 31, № 3. С. 222.
36. Кравцов Ю. А., Орлов Ю. И. Геометрическая оптика неоднородных сред. М.: Наука, 1980. 304 с.
37. Berry M. // Nature. 1987. V. 326, No. 6110. P. 277.
38. Berry M. // Phys. Today. 1990. V. 43, No. 12. P. 34.
39. Виноцкий С. И., Дербов В. Л., Дубовик В. М., Марковски Б. Л., Степановский Ю. П. // УФН. 1990. Т. 160, № 6. С. 1.
40. Малыкин Г. Б. // Изв. вузов. Радиофизика. 1997. Т. 40, № 3. С. 265.
41. Anandan J. // Nature. 1992. V. 360, No. 6402. P. 307.
42. Клышко Д. Н. // УФН. 1993. Т. 163, № 11. С. 1.
43. Бондарчук В. И., Давтян Л. С., Корнеев Д. А. // УФН. 1996. Т. 166, № 2. С. 185.
44. Berry M. V. // Proc. R. Soc. (London). Ser. A. 1984. V. 392, No. 1802. P. 45.
45. Ландау Л. Д., Лифшиц Е. М. Теоретическая физика. Т. 2. Теория поля. М.: Наука, 1967. 460 с.
46. Rashleigh S. C. // Opt. Lett. 1978. V. 3, No. 2. P. 60.
47. Ulrich R. // Fiber-Optics rotation sensors / Ed by S. Ezekiel, H. J. Arditty. Berlin: Springer Verlag, 1982. P. 52.
48. Ulrich R., Simon A. // Appl. Opt. 1979. V. 18, No. 13. P. 2241.
49. Barlow A. J., Ramskov-Hansen J. J., Payne D. N. // Appl. Opt. 1981. V. 20, No. 17. P. 2962.
50. Железняков В. В., Кочаровский В. В., Кочаровский Вл. В. // УФН. 1983. Т. 141, № 2. С. 257.
51. Малыкин Г. Б., Позднякова В. И., Шерешевский И. А. // Оптика и спектроскопия. 1997. Т. 83, № 5. С. 843.
52. Ross J. N. // Optical and Quantum Electron. 1984. V. 16, No. 5. P. 455.
53. Varnham M. P., Birch R. D., Payne D. N. // Proc. of the Fifth Intern. Conf. on Integrated Optics and Optical Fiber Communication and Eleventh Europ. Conf. on Optical Communication (Istituto Internazionale delle Comunicazioni, Genova, Italy, 1985). P. 135.
54. Tomita A., Chiao R. Y. // Phys. Rev. Lett. 1986. V. 57, No. 8. P. 937.
55. Frins E. M., Dultz W. // J. Lightwave. Techn. 1997. V. 15, No. 1. P. 144.
56. Simson J. K., Stolen R. H., Sears F. M., Pleibel J. B. // J. Lightwave. Techn. 1983. V. 1, No. 2. P. 370.
57. Бычков А. В., Гусовский Д. Д., Дианов Е. М., Зайцев Ю. И., Кофанов С. В., Малыкин Г. Б., Смирнов О. Б., Степанов Д. П., Хрулёв В. П. // Тезисы V Всесоюз. конф. «Волоконно-оптические системы передачи». Секция 2. М., 1988. С. 42.
58. Арутюнян З. Ю., Грудинин А. Б., Гурьянов А. Н., Гусовский Д. Д., Дианов Е. М., Зайцев Ю. И., Игнатъев С. В., Малыкин Г. Б., Смирнов О. Б., Степанов Д. П. // Тезисы V Всесоюз. конф. «Волоконно-оптические системы передачи». Секция 3. М., 1988, С. 7.
59. Катаевская И. В., Кундикова Н. Д. // Квантовая электроника. 1995. Т. 22, № 9. С. 959.
60. Гинзбург В. Л. // ЖТФ. 1944. Т. 14, № 3. С. 181.
61. Schiffrer G., Leeb W. K., Krammer H., Wittmann J. // Appl. Opt. 1979. V. 18, No. 13. P. 2096.
62. Ulrich R., Johnson M. // Opt. Lett. 1979. V. 4, No. 5. P. 152.

63. Логозинский В. Н. Исследование волоконного интерферометра как датчика угловой скорости: Дипломная работа. МФТИ, 1979. 30 с.
64. Гордон С. А., Логозинский В. Н., Новиков А. Г. // Квантовая электроника. 1980. Т. 7, № 10. С. 2352.
65. Ulrich R. // Opt. Lett. 1980. V. 5, No. 5. P. 173.
66. Kintner E. // Opt. Lett. 1981. V. 6, No. 3. P. 154.
67. Pavlath G. A., Shaw H. J. // Appl. Opt. 1982. V. 21, No. 8. P. 1400.
68. Burns W. K., Moeller R. P. // J. Lightwave Techn. 1984. V. 2, No. 4. P. 430.
69. Козел С. М., Листвин В. Н., Шаталин С. В., Юшкайтис Р. В. // Оптика и спектроскопия. 1986. Т. 60, № 6. С. 1295.
70. Малыкин Г. Б. // Изв. вузов. Радиофизика. 1991. Т. 34, № 7. С. 817.
71. Малыкин Г. Б. // Оптика и спектроскопия. 1996. Т. 81, № 3. С. 474.
72. Малыкин Г. Б. // Оптика и спектроскопия. 1998. Т. 84, № 3. С. 515.
73. Малыкин Г. Б. // Оптика и спектроскопия. 1997. Т. 83, № 6. С. 1013.
74. Андропова И. А., Геликонов Г. В., Малыкин Г. Б. // Квантовая электроника. 1999. Т. 26, № 3. С. 271.
75. Andropova I. A., Gelikonov G. V., Malykin G. B. // Proc. SPIE. 1999. V. 3736. P. 423.
76. Шерклифф У. Поляризованный свет. М.: Мир, 1965. 264 с.
77. Pancharatnam S. // Proc. Ind. Acad. Sci. A. 1956. V. 44, No. 5. P. 247.
78. Малыкин Г. Б., Позднякова В. И., Шерешевский И. А. Связь эллиптических винтовых поляризационных мод в одномодовых волоконных световодах с линейным двулучепреломлением и регулярным кручением осей анизотропии при наличии случайного кручения осей: Препринт ИПФ РАН № 489. Нижний Новгород, 1999. 32 с.
79. Малыкин Г. Б., Позднякова В. И., Шерешевский И. А. // Оптика и спектроскопия. 2000. Т. 88, № 3. С. 477.
80. Michelson A. A. // Philos. Mag. Ser. 6. 1904. V. 8, No. 48. P. 716.
81. Einstein A. // Astron. Nach. 1914. V. 199. P. 8
82. Pogany B. // Ann. der Phys. 1928. V. 85, No. 2. P. 244.
83. Corum J. F. // J. Math. Phys. 1977. V. 18, No. 4. P. 770.
84. Winterberg F. // Z. Naturforsch. A. 1989. V. 44. P. 1145.
85. Алексеев Э. И., Базаров Е. Н., Израелян В. Г., Кухта А. В. // Квантовая электроника. 1987. Т. 4, № 1. С. 192.

Институт прикладной физики РАН,
г. Нижний Новгород, Россия

Поступила в редакцию
18 июня 2002 г.

THE EFFECT OF NONCOPLANAR WINDING OF THE OPTICAL FIBER IN A FIBER RING INTERFEROMETER ON THE PHASE DIFFERENCE OF COUNTER-PROPAGATING WAVES

E. I. Yakubovich and G. B. Malykin

We consider a number of photoelastic and geometric-optical physical phenomena arising in a fiber ring interferometer with noncoplanar winding of the fiber ring. It is shown that the sensitivity of such an interferometer to rotation cannot be improved using any of these phenomena.

УДК 621.391.1

ПРОСТРАНСТВЕННОЕ РАЗДЕЛЕНИЕ ПОЛЬЗОВАТЕЛЕЙ В МІМО-СИСТЕМАХ, ИСПОЛЬЗУЮЩИХ ПАРАЛЛЕЛЬНУЮ ПЕРЕДАЧУ ДАННЫХ

А. Г. Флакман

Рассматриваются системы сотовой связи с антенными решётками на обоих концах линии связи (так называемые МІМО-системы (multiple-input multiple-output systems)), в которых передача данных осуществляется по параллельным каналам. Эти каналы формируются с помощью адаптивных передающей и приёмной диаграммообразующих схем (ДОС) и согласованы с пространственным каналом связи. Для увеличения пропускной способности МІМО-систем предложен способ пространственного разделения пользователей, не требующий оценки направлений прихода сигналов и основанный на ортогонализации всех параллельных каналов всех пользователей. Найдены отношение мощности сигнала к мощности шума на выходе собственных каналов и полная пропускная способность МІМО-системы при одновременном обслуживании произвольного числа пользователей. Приведены результаты численного моделирования в случае релейских замираний сигналов, подтверждающие высокую эффективность предложенного способа пространственного разделения пользователей.

ВВЕДЕНИЕ

Главной проблемой в создании и развитии перспективных систем сотовой связи является значительное увеличение скорости передачи информации при высоком качестве обслуживания абонентов. Основной причиной, ограничивающей пропускную способность и надёжность передачи данных в таких системах, являются глубокие замирания (фединги) сигналов, которые обусловлены многолучевым распространением сигналов и характерны для больших городов. Пропускную способность можно увеличить с помощью разнесённой передачи и/или разнесённого приёма, т. е. за счёт использования антенных решёток на одном или на обоих концах линии связи. Такие системы связи часто называют МІМО-системами. В [1–3] рассматривается формирование независимых (ортогональных) параллельных каналов в МІМО-системах с помощью адаптивных ДОС на обоих концах линии связи. Такие каналы называются собственными каналами, т. к. ДОС построены на собственных векторах матрицы коэффициентов передачи рассеивающей среды. Формирование собственных каналов предполагает оценку комплексных коэффициентов передачи между всеми антеннами на приёмном конце линии связи и передачу данной информации по обратному каналу на передающий конец линии. При этом обеспечивается согласование со случайным каналом связи не только приёмника, но и передатчика и тем самым значительное увеличение пропускной способности МІМО-системы [1–3].

В настоящее время широко используются временной, частотный и кодовый способы разделения пользователей. Системы с кодовым разделением обладают возможностью одновременного обслуживания пользователей, работающих в одном частотном диапазоне, и поэтому имеют наибольшую пропускную способность. Известен также способ пространственного разделения пользователей, который называется SDMA (Space Division Multiple Access) [4–6], основанный на возможности адаптивного формирования системы ортогональных лучей, каждый из которых обеспечивает передачу информации только одному пользователю. Эффективность SDMA в сотовых сетях связи рассматривалась в [4–6], где для реализации SDMA рекомендуется предварительно оценивать направления прихода сигналов от пользователей, а затем формировать систему ортогональных лучей, каждый из которых обеспечивает передачу сигналов только соответствующему пользователю. Некоторые экспериментальные результаты для SDMA приведены в [7].

Однако оценка азимута направления прихода сигналов от пользователей является сложной проблемой в сотовых системах связи, часто работающих в условиях случайной среды распространения сигналов. В таких условиях пользователь окружён отражателями, которые рассеивают его сигнал. Поэтому для базовой станции пользователь представляет собой распределённый источник, причём угловые размеры такого источника могут составлять десятки градусов и превышать ширину луча антенны базовой станции [8]. Более того, «центр тяжести» излучения такого источника может смещаться в достаточно больших пределах, т. к. число рассеивателей, их угловое положение и эффективная поверхность рассеяния являются случайными величинами. Отметим также, что для ММО-систем, использующих передачу данных по независимым параллельным собственным каналам, пространственное разделение пользователей ранее не рассматривалось.

В настоящей работе предложен способ пространственного разделения пользователей для ММО-систем, не требующий оценки направлений прихода сигналов. Предполагается, что передача информации осуществляется по параллельным собственным каналам. SDMA обеспечивается за счёт дополнительной обработки сигналов, на основе ортогонализации пространственных каналов всех пользователей. Показано, что существует оптимальное число пространственно разделяемых пользователей, обеспечивающее максимальную пропускную способность ММО-системы. Это объясняется тем, что число собственных каналов (а следовательно, и пропускная способность) возрастает с увеличением числа пользователей. Однако при этом потери в отношении мощности сигнала к мощности шума (ОСШ) на выходе собственных каналов из-за разделения пользователей также возрастают, что приводит к уменьшению пропускной способности. Проведён сравнительный анализ пропускной способности систем, использующих SDMA, по сравнению с аналогичными системами без SDMA в случае точно известных матриц комплексных коэффициентов передачи различных пользователей.

1. ПЕРЕДАЧА ДАННЫХ ОДНОМУ ПОЛЬЗОВАТЕЛЮ

Рассмотрим сначала адаптивное формирование независимых параллельных каналов для передачи данных в ММО-системе, состоящей из M передающих и N приёмных антенн. Такую систему будем обозначать как ММО- $\{M; N\}$. Если пространственный канал связи с M входами и N выходами является частотно неселективным, то он может быть описан матрицей \mathbf{H} размерности $N \times M$ комплексных коэффициентов передачи h_{nm} сигналов из m -й передающей антенны в n -ную приёмную антенну. В случае распространения сигналов в случайной однородной рассеивающей среде коэффициенты h_{nm} имеют релеевское распределение амплитуд и равномерное распределение фаз в интервале $[0, 2\pi]$ [9]. Без ограничения общности можно считать, что коэффициенты передачи в среднем нормированы к единице. Тогда коэффициенты корреляции $\langle h_{nm} h_{pq}^* \rangle = \delta_{np} \delta_{mq}$, где δ_{nm} — символ Кронекера, угловые скобки означают статистическое среднее, а индекс звёздочка — комплексное сопряжение.

Модулированные сигналы разделяются на L параллельных потоков, которые передаются независимо друг от друга. Число таких потоков должно быть не больше минимального числа передающих или приёмных антенн, т. е. $L \leq \min\{M, N\}$. В передающей ДОС сигналы из каждого потока умножаются на соответствующие весовые коэффициенты. Обозначим $\mathbf{C}(t) = [c_1(t), c_2(t), \dots, c_L(t)]^T$ L -мерный вектор-столбец сигналов на входе ДОС, где индекс T обозначает операцию транспонирования. Общая схема такой системы показана на рис. 1.

Для формирования L независимых параллельных потоков приёмник должен непрерывно оценивать канальную матрицу \mathbf{H} , затем эта информация должна сообщаться на передающий конец линии. Сигналы передаются в виде последовательности фреймов (пакетов импульсов), каждый из которых имеет длительность τ_{fr} и состоит из двух частей. Первая часть является служебной и предназначена для оценивания матрицы \mathbf{H} , передачи информации об этой матрице на передающий конец линии и для

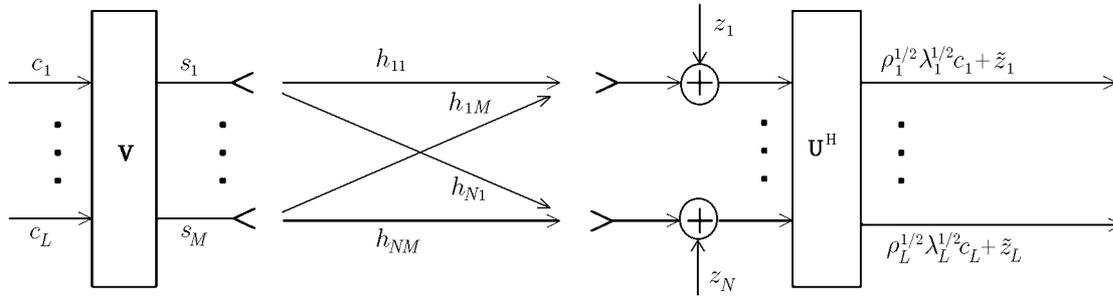


Рис. 1

вычисления весовых векторов обеих ДОС. Вторая часть фрейма служит непосредственно для передачи символов $c_i(t)$. Таким образом, адаптация передающей и приёмной ДОС заключается в оценивании матрицы \mathbf{H} и в нахождении весовых векторов в служебной части каждого фрейма. Далее эти векторы используются до конца соответствующего фрейма.

Длительность фрейма τ_{fr} должна быть значительно меньше времени корреляции τ_{cor} замираний сигнала для эффективного отслеживания состояния канала. В релейском канале с изотропным распределением рассеивателей относительно передающей и приёмной антенн функция корреляции флуктуаций сигналов имеет вид функции Бесселя первого рода нулевого порядка: $f(\tau) = J_0(2\pi f_d \tau)$, где $f_d = v\lambda^{-1}$ — максимальная доплеровская частота, v — скорость движения абонента, λ — длина волны [9]. Время корреляции замираний по уровню 0,5 составляет $\tau_{cor} = 0,2f_d^{-1}$. Например, для хаотически перемещающихся пользователей-пешеходов ($v \leq 5$ км/ч) при $\lambda \approx 33$ см (несущая частота приблизительно 900 МГц) имеем $\tau_{fr} \ll 0,05$ с. Для движущегося со скоростью 50 км/ч автомобиля длительность фрейма должна удовлетворять более жёсткому условию $\tau_{fr} \ll 0,005$ с.

Оценка канальной матрицы \mathbf{H} производится с помощью псевдошумовых обучающих (тренирующих) последовательностей, состоящих из конечного числа известных символов. В случае простейшей модуляции, когда одному сигналу соответствует один бит, символ равен +1 или -1. Обычно используются максимально правдоподобные оценки матрицы \mathbf{H} или оценки, основанные на поиске минимума среднеквадратической ошибки [10]. Возможны параллельный и последовательный способы оценивания матрицы \mathbf{H} в MIMO-системе. В первом из них число обучающих последовательностей должно быть равно числу передающих антенн, все они излучаются одновременно, но имеют разные кодовые последовательности. Это даёт возможность в каждой приёмной антенне разделить сигналы, принимаемые от всех антенн, т. е. идентифицировать обучающие сигналы каждой из передающих антенн. Второй способ заключается в использовании одной обучающей последовательности, которая должна излучаться поочерёдно разными антеннами. На практике матрица \mathbf{H} всегда оценивается с некоторой ошибкой, обусловленной влиянием собственных шумов приёмных устройств. В условиях релейских замираний сигналов эта ошибка может приводить к значительному увеличению вероятности битовой ошибки при передаче информации [11]. Другой составляющей ошибки является изменение состояния канала за время между двумя его последовательными оценками.

Будем рассматривать потенциальные характеристики MIMO-системы, реализуемые при точно известной на обоих концах линии матрице \mathbf{H} . То есть мы считаем, что состояние пространственного канала (матрица \mathbf{H}) остаётся постоянным в течение каждого фрейма и изменяется случайным образом по релейскому закону от фрейма к фрейму. В этом случае обеспечивается полное согласование приёмника и передатчика со случайным каналом передачи и приёма сигналов.

M -мерный вектор-столбец сигналов $\mathbf{S}(t) = (s_1(t), s_2(t), \dots, s_M(t))^T$ на выходе передающей ДОС

может быть представлен в виде

$$\mathbf{S}(t) = \mathbf{V}\mathbf{R}^{1/2}\mathbf{C}(t), \quad (1)$$

где матрица $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_L)$ размерности $M \times L$ состоит из весовых векторов \mathbf{V}_i передающей ДОС, $\mathbf{R} = \text{diag}\{\rho_1, \rho_2, \dots, \rho_L\}$ — диагональная матрица, составленная из чисел ρ_i , которые дают распределение мощности между параллельными каналами.

Вектор $\mathbf{X}(t) = (x_1(t), x_2(t), \dots, x_N(t))^T$ сигналов в элементах приёмной решётки равен

$$\mathbf{X}(t) = \mathbf{H}\mathbf{S}(t) + \mathbf{Z}(t), \quad (2)$$

где $\mathbf{Z}(t) = (z_1(t), z_2(t), \dots, z_N(t))^T$ — вектор собственных шумов, которые будем считать гауссовскими некоррелированными во времени и в приёмных каналах случайными процессами с нулевыми средним и единичной дисперсией, т. е. $\langle z_i \rangle = 0$ и $\langle z_i(t_1)z_j^*(t_2) \rangle = \delta_{ij}\delta_{t_1t_2}$.

В приёмной ДОС происходит преобразование сигнала, которое может быть описано матрицей $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_L)$ размерности $N \times L$. На выходе ДОС вектор сигнала имеет вид $\mathbf{Y}(t) = \mathbf{U}^H\mathbf{X}(t)$, где индекс H обозначает эрмитовое сопряжение. С учётом (2) имеем $\mathbf{Y}(t) = \mathbf{U}^H\mathbf{H}\mathbf{S}(t) + \tilde{\mathbf{Z}}(t)$, где $\tilde{\mathbf{Z}}(t) = \mathbf{U}^H\mathbf{Z}(t)$ — вектор шумов на выходе ДОС. Теперь с помощью (1) получим

$$\mathbf{Y}(t) = \mathbf{U}^H\mathbf{H}\mathbf{V}\mathbf{R}^{1/2}\mathbf{C}(t) + \tilde{\mathbf{Z}}(t). \quad (3)$$

Из (3) следует, что для формирования независимых ортогональных пространственных каналов необходимо, чтобы матрица $\mathbf{U}^H\mathbf{H}\mathbf{V}$ была диагональной, а матрицы \mathbf{U} и \mathbf{V} для приёмной и передающей ДОС должны быть унитарными. Задача нахождения матриц \mathbf{U} и \mathbf{V} решается методом сингулярного разложения матрицы коэффициентов передачи \mathbf{H} , которое имеет вид [12]

$$\mathbf{H} = \sum_{i=1}^L \sqrt{\lambda_i} \mathbf{F}_i \mathbf{G}_i^H, \quad (4)$$

где λ_i — собственные числа матрицы $\mathbf{H}\mathbf{H}^H$ при $M > N$ или матрицы $\mathbf{H}^H\mathbf{H}$ при $M < N$, векторы \mathbf{F}_i и \mathbf{G}_i являются собственными векторами матриц $\mathbf{H}\mathbf{H}^H$ и $\mathbf{H}^H\mathbf{H}$ соответственно и образуют ортонормированные векторные системы.

Подставляя (4) в (3), можно прийти к выводу, что весовые векторы приёмной и передающей ДОС должны совпадать с векторами сингулярного разложения (4) матрицы \mathbf{H} , т. е. $\mathbf{U}_i = \mathbf{F}_i$ и $\mathbf{V}_i = \mathbf{G}_i$. При этом матрица $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^H$, где $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_L\}$ — диагональная матрица, составленная из собственных чисел λ_i . Сформированные таким образом каналы называются собственными. Вектор сигнала на выходе собственных каналов равен

$$\mathbf{Y}(t) = \mathbf{\Lambda}^{1/2}\mathbf{R}^{1/2}\mathbf{C}(t) + \tilde{\mathbf{Z}}(t). \quad (5)$$

Приёмная ДОС не изменяет статистические свойства собственных шумов, которые являются независимыми в разных выходных каналах ДОС. Таким образом, ММО-система с помощью линейного преобразования сигналов в передающей и приёмной ДОС может быть представлена в виде $L \leq \min\{M, N\}$ независимых параллельных каналов (одноканальных систем).

Эффективность передачи данных в системах связи определяется вероятностью битовой ошибки, которая зависит как от статистических свойств флуктуаций сигналов в пространственном канале и излучаемой мощности, так и от выбранного способа кодирования и модуляции передаваемой информации. Существует большое число методов кодирования и модуляции [10, 13], которые существенно влияют на битовую ошибку. Поэтому часто систему характеризуют шенноновской пропускной способностью, равной максимальному числу бит, которые можно передать без ошибки через пространственный канал за одну секунду в полосе частот 1 Гц (бит·с⁻¹·Гц⁻¹). Шенноновская пропускная способность

является удобным параметром, т. к. не зависит от способа кодирования и модуляции и определяется статистическими свойствами замираний, излучаемой мощностью и методом распределения этой мощности между собственными каналами.

Из (5) следует, что ОСШ на выходе i -го собственного канала равно $\eta_i = \rho_i \lambda_i$. Поэтому пропускная способность этого канала имеет вид $C_i = \log_2(1 + \rho_i \lambda_i)$ [10, 13]. Поскольку собственные каналы являются независимыми, то пропускная способность всей системы определяется суммой пропускных способностей отдельных каналов:

$$C = \sum_{i=1}^L \log_2(1 + \rho_i \lambda_i). \quad (6)$$

Пропускная способность ММО-системы зависит от способа распределения полной мощности P_0 между собственными каналами. Максимальная пропускная способность обеспечивается при оптимальном распределении мощности в соответствии с так называемым правилом «наполнения водой» («water pouring»)[10]. Мощность в i -м собственном канале равна $\rho_i = \mu - \lambda_i^{-1}$, где константа μ находится из условия ограничения полной мощности ($\rho_1 + \rho_2 + \dots + \rho_L = P_0$). Величину $\mu - \lambda_i^{-1}$ можно интерпретировать как глубину дна чаши, разную для разных собственных каналов. Если в эту чашу налить определённое количество воды, соответствующее мощности P_0 , то вода распределится по чаше так, чтобы достигался максимум пропускной способности. Чем глубже дно (больше коэффициент передачи λ_i соответствующего собственного канала), тем большая доля мощности распределяется в этот канал. Более того, если для некоторого j -го канала $\lambda_j^{-1} > \mu$, то $\rho_j = 0$, т. е. мощность в данный канал не распределяется, и этот канал не будет использоваться для передачи данных.

2. ПРОСТРАНСТВЕННОЕ РАЗДЕЛЕНИЕ ПОЛЬЗОВАТЕЛЕЙ

Рассмотрим теперь одновременную передачу данных к нескольким пользователям, основанную на их пространственном разделении. Пусть число одновременно обслуживаемых пользователей равно Q . Каждый из них может иметь антенную решётку с разным числом элементов. Сделаем принципиальное предположение, что антенны всех пользователей состоят из одинакового числа N элементов. Передающая антенна базовой станции должна иметь большее число M элементов ($M > N$). Максимальное число пространственно разделяемых пользователей удовлетворяет условию $Q \leq [M/N]$, где $[x]$ обозначает целую часть x , а число собственных каналов L каждого пользователя не может превышать числа элементов приёмной антенны ($L \leq N$). Будем считать, что формируется максимальное число собственных каналов, т. е. $L = N$.

Коэффициенты передачи $h_{nm}^{(q)}$ сигналов из m -й передающей антенны в n -ную приёмную антенну q -го пользователя образуют канальную матрицу $\mathbf{H}^{(q)}$. Свойства канала связи для всех пользователей могут быть описаны Q матрицами $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots, \mathbf{H}^{(Q)}$, которые имеют одинаковую размерность $N \times M$ и состоят из элементов $h_{nm}^{(q)}$ с одинаковыми статистическими свойствами. Передающую и приёмную ДОС, формирующие собственные каналы для q -го пользователя, будем описывать матрицами $\mathbf{V}^{(q)}$ и $\mathbf{U}^{(q)}$. Эти каналы обеспечивают параллельную передачу L символов q -му пользователю. Таким образом, одновременно может быть передано LQ символов для Q пользователей.

Обозначим $\mathbf{C}^{(q)}$ L -мерный вектор сигналов, одновременно передаваемых q -му пользователю. Тогда M -мерный вектор сигналов на выходе передающей ДОС имеет вид

$$\mathbf{S}(t) = \sum_{q=1}^Q \mathbf{V}^{(q)} (\mathbf{R}^{(q)})^{1/2} \mathbf{C}^{(q)}(t), \quad (7)$$

где $\mathbf{R}^{(q)}$ — диагональная матрица, элементы $\rho_i^{(q)}$ которой определяют распределение мощности, предназначенной q -му пользователю, по его собственным каналам; $i = 1, 2, \dots, L$. При этом естественно предполагать, что мощность, предназначенная каждому пользователю, остаётся равной P_0 , т. е. при одновременном обслуживании Q пользователей общая мощность увеличивается в Q раз и составляет QP_0 .

Вектор сигнала, принимаемого q -м пользователем, имеет вид $\mathbf{X}^{(q)}(t) = \mathbf{H}^{(q)}\mathbf{S}(t) + \mathbf{Z}^{(q)}(t)$. Представим матрицу $\mathbf{H}^{(q)}$ канальных коэффициентов в виде сингулярного разложения: $\mathbf{H}^{(q)} = \mathbf{U}^{(q)} [\mathbf{\Lambda}^{(q)}]^{1/2} \times \mathbf{V}^{(q)\text{H}}$. Учитывая (7), получим

$$\mathbf{X}^{(q)}(t) = \sum_{p=1}^Q \mathbf{U}^{(q)} [\mathbf{\Lambda}^{(q)} \mathbf{R}^{(p)}]^{1/2} (\mathbf{V}^{(q)\text{H}} \mathbf{V}^{(p)}) \mathbf{C}^{(p)}(t) + \mathbf{Z}^{(q)}(t). \quad (8)$$

Система векторов $\mathbf{V}_i^{(q)}$ является ортогональной только для соответствующего (q -го) пользователя, т. е. $\mathbf{V}_i^{(q)}$ и $\mathbf{V}_i^{(p)}$ — ортогональные векторы только при $q = p$. Следовательно, в (8) матрица $\mathbf{V}_i^{(q)\text{H}} \mathbf{V}_i^{(p)}$ не диагональная, и q -й пользователь принимает информацию, предназначенную и другим пользователям. В самом деле, для вектора сигнала q -го пользователя на выходе собственных каналов имеем

$$\mathbf{Y}^{(q)}(t) = [\mathbf{\Lambda}^{(q)} \mathbf{R}^{(q)}]^{1/2} \mathbf{C}^{(q)}(t) + \sum_{\substack{p=1 \\ p \neq q}}^Q [\mathbf{\Lambda}^{(q)} \mathbf{R}^{(p)}]^{1/2} (\mathbf{V}^{(q)\text{H}} \mathbf{V}^{(p)}) \mathbf{C}^{(p)}(t) + \tilde{\mathbf{Z}}^{(q)}(t), \quad (9)$$

где $\tilde{\mathbf{Z}}^{(q)}(t) = \mathbf{U}^{(q)\text{H}} \mathbf{Z}^{(q)}(t)$ — вектор собственных шумов. Первое слагаемое в (9) представляет собой полезный сигнал, передаваемый q -му пользователю. Второе слагаемое является помехой для этого пользователя. Таким образом, неортогональность системы векторов приёмных ДОС разных пользователей приводит к тому, что пользователи оказываются пространственно неразделёнными. Поэтому независимая параллельная передача всех LQ символов оказывается невозможной.

Для пространственного разделения пользователей все LQ собственных каналов должны быть ортогонализированы между собой. Процедура ортогонализации представляет собой дополнительное преобразование сигналов и может быть выполнена с помощью матриц-проекторов $\mathbf{P}_i^{(q)}$ на подпространство, ортогональное всем векторам, кроме вектора $\mathbf{V}_i^{(q)}$.

Будем использовать следующую процедуру ортогонализации полной системы векторов $\mathbf{V}_i^{(q)}$, где $q = 1, \dots, Q$; $i = 1, \dots, L$. Пусть $\mathbf{A} = (\mathbf{V}_1^{(1)}, \dots, \mathbf{V}_L^{(1)}, \mathbf{V}_1^{(2)}, \dots, \mathbf{V}_L^{(2)}, \dots, \mathbf{V}_1^{(Q)}, \dots, \mathbf{V}_L^{(Q)})$ — матрица размерности $M \times QL$, составленная из всех векторов $\mathbf{V}_i^{(q)}$, $\mathbf{A}_i^{(q)}$ — матрица, образованная из матрицы \mathbf{A} исключением вектора $\mathbf{V}_i^{(q)}$, т. е. в матрицу $\mathbf{A}_i^{(q)}$ не входит вектор-столбец, формирующий i -й собственный канал q -го пользователя.

Найдём проекцию вектора $\mathbf{V}_i^{(q)}$ на подпространство, ортогональное всем остальным векторам. Соответствующая матрица-проектор равна $\mathbf{A}_i^{(q)} (\mathbf{A}_i^{(q)\text{H}} \mathbf{A}_i^{(q)})^{-1} \mathbf{A}_i^{(q)\text{H}}$. Учтём, что матрица-проектор на полное пространство совпадает с единичной матрицей \mathbf{I} . Тогда матрица-проектор $\mathbf{P}_i^{(q)}$ на подпространство, ортогональное всем векторам, кроме вектора $\mathbf{V}_i^{(q)}$, имеет вид

$$\mathbf{P}_i^{(q)} = \mathbf{I} - \mathbf{A}_i^{(q)} (\mathbf{A}_i^{(q)\text{H}} \mathbf{A}_i^{(q)})^{-1} \mathbf{A}_i^{(q)\text{H}}, \quad (10)$$

а искомая проекция равна $\mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)}$. Нормируем этот вектор. При этом учтём, что матрица $\mathbf{P}_i^{(q)}$ является эрмитовой и не изменяется при возведении в квадрат, т. е. $\mathbf{P}_i^{(q)} = \mathbf{P}_i^{(q)\text{H}}$ и $\mathbf{P}_i^{(q)} = (\mathbf{P}_i^{(q)})^2$. В ре-

зультате получим, что нормированный вектор равен

$$\tilde{\mathbf{V}}_i^{(q)} = \frac{\mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)}}{\sqrt{\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)}}}. \quad (11)$$

Теперь вектор $\mathbf{S}(t)$ сигналов на выходе передающей ДОС и вектор сигналов $\mathbf{X}^{(q)}(t)$ на входе приёмной ДОС q -го пользователя будут определяться с помощью (7) и (8), если в этих формулах заменить $\mathbf{V}_i^{(q)}$ на $\tilde{\mathbf{V}}_i^{(q)}$, а вектор сигнала на выходе ДОС q -го пользователя будет равен

$$\mathbf{Y}^{(q)}(t) = \sum_{p=1}^Q [\boldsymbol{\Lambda}^{(q)} \mathbf{R}^{(p)}]^{1/2} (\mathbf{V}^{(q)H} \tilde{\mathbf{V}}^{(p)}) \mathbf{C}^{(p)}(t) + \tilde{\mathbf{Z}}^{(q)}(t). \quad (12)$$

Однако в отличие от (9) матрица $(\mathbf{V}^{(q)H} \tilde{\mathbf{V}}^{(p)})$ является диагональной. В самом деле, из свойств матриц-проекторов следует, что элемент такой матрицы равен

$$(\mathbf{V}^{(q)H} \tilde{\mathbf{V}}^{(p)})_{lm} = \mathbf{V}_l^{(q)H} \tilde{\mathbf{V}}_m^{(p)} = \sqrt{\mathbf{V}_l^{(q)H} \mathbf{P}_m^{(p)} \mathbf{V}_m^{(p)}} = \delta_{pq} \delta_{lm}. \quad (13)$$

Подставляя (13) в (12) для вектора сигнала, принимаемого q -м пользователем, получаем следующее выражение:

$$\mathbf{Y}^{(q)}(t) = (\boldsymbol{\Lambda}^{(q)})^{1/2} (\mathbf{W}^{(q)})^{1/2} (\mathbf{R}^{(q)})^{1/2} \mathbf{C}^{(q)}(t) + \tilde{\mathbf{Z}}^{(q)}(t), \quad (14)$$

где $\mathbf{W}^{(q)}$ — диагональная матрица, составленная из чисел $\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)}$ ($i = 1, 2, \dots, L$), которые дают распределение мощности между параллельными каналами после их ортогонализации. Из (14) следует, что пространственное разделение пользователей обеспечивает параллельную передачу LQ символов. Общая схема системы с SDMA показана на рис. 2.

Выше рассматривалась передача информации от базовой станции к нескольким пользователям. Покажем, что предложенная процедура ортогонализации собственных каналов обеспечивает также разделение пользователей при передаче в обратном направлении, от пользователей к базовой станции. В этом случае все каналные матрицы являются эрмитово сопряжёнными по отношению к матрицам, соответствующим передаче в прямом направлении. Следовательно, канал связи в обратном направлении определяется Q матрицами $\mathbf{H}^{(1)H}, \mathbf{H}^{(2)H}, \dots, \mathbf{H}^{(Q)H}$. Кроме этого, приёмные антенны пользователей становятся передающими, а передающие антенны базовой станции — приёмными, т. е. $\mathbf{U}^{(q)}$ является матрицей передающей ДОС q -го пользователя, а матрица $\mathbf{V}^{(q)H}$ играет роль матрицы приёмной ДОС этого пользователя. Аналогичным образом для сигнала, принимаемого q -м пользователем на выходе i -го собственного канала, можно получить выражение, совпадающее с (14). Следовательно, собственные каналы являются независимыми при передаче информации в любом направлении при использовании процедуры ортогонализации.

Найдём пропускную способность системы при пространственном разделении Q пользователей. Из (14) получим, что ОСШ на выходе i -го собственного канала q -го пользователя равно

$$\eta_i^{(q)} = \rho_i^{(q)} \lambda_i^{(q)} \left(\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)} \right). \quad (15)$$

Сравним полученное выражение с ОСШ при отсутствии пространственного разделения пользователей, которое также определяется с помощью выражения (15), если в нём матрицу-проектор $\mathbf{P}_i^{(q)}$ заменить единичной матрицей \mathbf{I} . Учтём, что для любой системы векторов справедливо неравенство $\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)} \leq \mathbf{V}_i^{(q)H} \mathbf{V}_i^{(q)}$. Отсюда следует, что пространственное разделение пользователей приводит к уменьшению ОСШ на выходе всех собственных каналов. С ростом числа Q пользователей отношение мощности сигнала к мощности шума будет уменьшаться в большей степени.

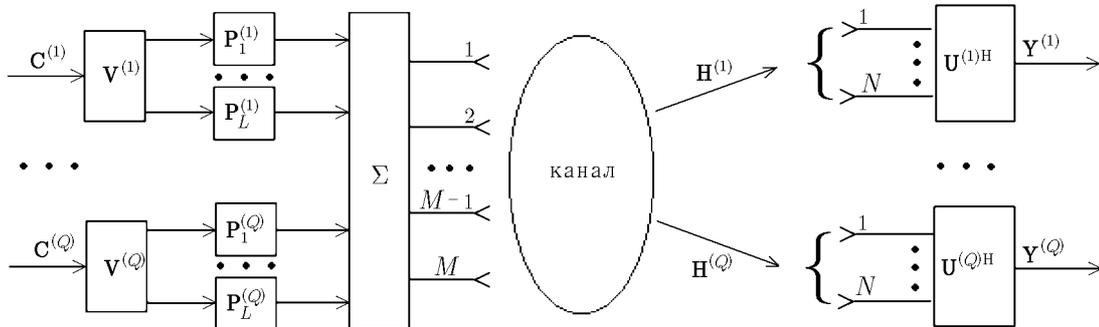


Рис. 2

Сначала найдём удельную пропускную способность $C^{(q)}$ ММО-системы, приходящуюся на одного пользователя. С помощью (15) получаем

$$C^{(q)} = \sum_{i=1}^L \log_2 \left[1 + \rho_i^{(q)} \lambda_i^{(q)} \left(\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)} \right) \right]. \quad (16)$$

Поскольку все пользователи являются пространственно разделёнными, полную пропускную способность системы можно получить, суммируя удельные пропускные способности:

$$C = \sum_{q=1}^Q \sum_{i=1}^L \log_2 \left[1 + \rho_i^{(q)} \lambda_i^{(q)} \left(\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)} \right) \right]. \quad (17)$$

Если используется более простое, равномерное распределение мощности между всеми Q пользователями и всеми собственными каналами каждого из них, то $\rho_i^{(q)} = P_0/L$. Тогда из (17) для полной пропускной способности ММО-системы получаем

$$C = \sum_{q=1}^Q \sum_{i=1}^L \log_2 \left[1 + \frac{P_0}{L} \lambda_i^{(q)} \left(\mathbf{V}_i^{(q)H} \mathbf{P}_i^{(q)} \mathbf{V}_i^{(q)} \right) \right]. \quad (18)$$

3. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Приведём результаты моделирования для случая распространения сигналов в случайной однородной рассеивающей среде, когда коэффициенты $h_{nm}^{(q)}$ имеют релейское распределение амплитуд и равномерное распределение фаз и некоррелированы между собой [9]. Пропускная способность ММО-системы — случайная величина, полной характеристикой которой является функция плотности вероятности. Однако часто ограничиваются нахождением усреднённой по релейским замираниям пропускной способности C_{mean} .

Вначале формировалась случайная релейская выборка коэффициентов передачи $h_{nm}^{(q)}$ сигналов из m -й передающей антенны в n -ную приёмную антенну для каждого пользователя, т. е. создавалась одна реализация случайных канальных матриц $\mathbf{H}^{(q)}$ для всех Q пользователей. Для этой реализации матриц $\mathbf{H}^{(q)}$, где $q = 1, 2, \dots, Q$, вычислялись векторы матриц $\mathbf{V}^{(q)}$ и $\mathbf{U}^{(q)}$ передающей и приёмной ДОС q -го пользователя как векторы сингулярного разложения матрицы $\mathbf{H}^{(q)}$, а также собственные числа

$\lambda_i^{(q)}$. Далее осуществлялось пространственное разделение пользователей. С этой целью с помощью (10) находились проекционные матрицы $\mathbf{P}_i^{(q)}$ для всех пользователей, и на основе (18) вычислялась пропускная способность ММО-системы, соответствующая заданной реализации матриц $\mathbf{H}^{(q)}$. Затем формировалась следующая выборка коэффициентов передачи релейского канала, и аналогично находилась соответствующая пропускная способность ММО-системы. Всего использовалось 1000 реализаций каждой из матриц $\mathbf{H}^{(q)}$. Моделирование завершалось нахождением интегральной функции плотности вероятности пропускной способности и её среднего значения C_{mean} .

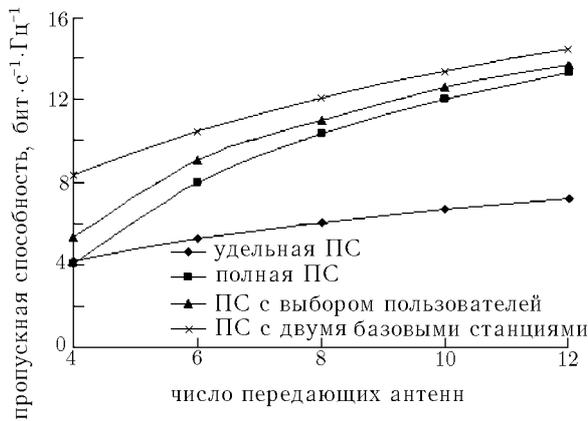


Рис. 3

Рассмотрим вначале наиболее простой случай пространственного разделения двух пользователей ($Q = 2$), каждый из которых имеет по две антенны ($N = 2$), т. е. число собственных каналов $L = 2$. Число антенн на базовой станции будем считать произвольным, но не меньшим четырёх ($M \geq 4$). На рис. 3 показана средняя удельная пропускная способность (ПС), приходящаяся на одного пользователя, и полная пропускная способность ММО-системы в зависимости от числа M антенн. Верхняя кривая соответствует случаю обслуживания двух пользователей двумя базовыми станциями. Четвёртая кривая будет пояснена ниже. Мощность, выделяемая каждому пользователю, равна $P_0 = 2$, т. е. ОСШ в каждой приёмной антенне равно единице ($\rho_i^{(q)} = 1; i = 1, 2; q = 1, 2$). Видно, что

пространственное разделение пользователей увеличивает пропускную способность тем значительнее, чем больше антенн используется на базовой станции. С ростом M коэффициент увеличения пропускной способности системы стремится к 2, т. е. пропускная способность ММО-системы с разделением двух пользователей приближается к пропускной способности при обслуживании двух пользователей двумя базовыми станциями.

Отметим, что возможно дальнейшее увеличение пропускной способности, когда имеется большее число пользователей-кандидатов ($Q' > 2$), из которых можно оптимальным образом выбрать два пользователя для их пространственного разделения. Сформируем $Q' - 1$ матриц $\mathbf{B}^{(q)} = (\mathbf{V}_1^{(q)}, \mathbf{V}_2^{(q)}, \mathbf{V}_1^{(q)}, \mathbf{V}_2^{(q)})$, где $q = 2, 3, \dots, Q'$, и вычислим определители $\det(\mathbf{B}^{(q)\text{H}}\mathbf{B}^{(q)})$. Если векторы $\mathbf{V}_1^{(1)}, \mathbf{V}_2^{(1)}, \mathbf{V}_1^{(q)}, \mathbf{V}_2^{(q)}$, имеющие единичную длину, ортогональны между собой, то $\det(\mathbf{B}^{(q)\text{H}}\mathbf{B}^{(q)}) = 4$. Это означает, что q -й пользователь уже является пространственно разделённым с первым пользователем и, следовательно, наиболее предпочтительным для одновременного обслуживания. Если векторы $\mathbf{V}_1^{(1)}, \mathbf{V}_2^{(1)}, \mathbf{V}_1^{(q)}, \mathbf{V}_2^{(q)}$ не ортогональны друг с другом при всех q , то $\det(\mathbf{B}^{(q)\text{H}}\mathbf{B}^{(q)}) < 4$. Однако наиболее предпочтительным является тот пользователь, для которого определитель наиболее близок к 4. При этом потери в ОСШ при пространственном разделении будут минимальными, а пропускная способность — максимальной. Усреднённая по релейским замираниям пропускная способность системы с оптимальным выбором двух пользователей из четырёх показана на рис. 3. Видно, что пропускная способность увеличивается более всего при малом числе M передающих антенн, т. е. в области наибольших потерь в ОСШ за счёт пространственного разделения пользователей. Коэффициент увеличения пропускной способности при $M = 4$ составляет приблизительно 1,3. Отметим, что оптимальный выбор пользователей становится значительно более сложным при большем числе разделяемых пользователей ($Q > 2$).

Рассмотрим теперь пространственное разделение произвольного числа ($Q > 2$) пользователей,

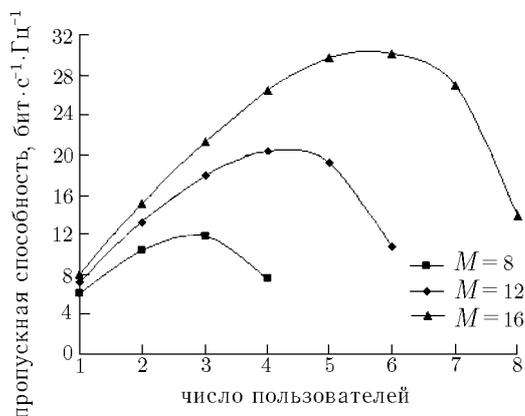


Рис. 4

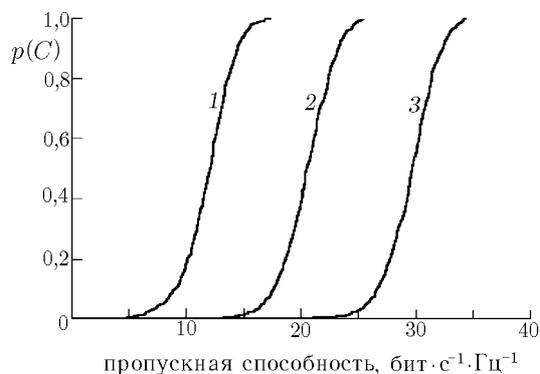


Рис. 5

каждый из которых по-прежнему имеет по две антенны ($N = 2$). Мощность, выделяемую каждому пользователю, будем считать прежней ($P_0 = 2$). На рис. 4 показана средняя пропускная способность случайного релейского канала в зависимости от числа Q пользователей для разного числа M антенн на базовой станции. Видно, что полная пропускная способность MIMO-системы при увеличении числа пользователей сначала возрастает, а затем начинает убывать из-за потерь в ОСШ на выходе собственных каналов. Максимальное число Q_{\max} разделяемых пользователей равно $[M/N]$, что составляет $Q_{\max} = 4; 6$ и 8 при $M = 8; 12$ и 16 соответственно. Однако из рис. 4 видно, что максимальная пропускная способность достигается при меньшем числе пользователей. Например, при $M = 16$ максимальная пропускная способность имеет место при $Q = 6$ и равна $C_{\text{mean}} = 30,1$ бит·с⁻¹·Гц⁻¹. Отметим, что при $Q = 5$ пропускная способность равна $C_{\text{mean}} = 29,7$ бит·с⁻¹·Гц⁻¹, т. е. оптимальным является разделение $Q = 5$ пользователей, а добавление шестого пользователя практически не увеличивает пропускную способность. Коэффициенты увеличения пропускной способности MIMO-системы за счёт SDMA при различном числе передающих M и приёмных N антенн приведены в табл. 1. Видно, что оптимальным для SDMA является число пользователей Q_{opt} , составляющее 50 ÷ 60 % от максимально возможного, т. е. $Q_{\text{opt}} \approx [(0,5 \div 0,6)Q_{\max}] = [(0,5 \div 0,6)M/N]$.

На рис. 5 показаны интегральные функции плотности вероятности $p(C)$ пропускной способности MIMO-системы при оптимальном числе пользователей и разном числе антенн базовой станции: $M = 8, Q = 3; M = 12, Q = 4; M = 16, Q = 5$ (кривые 1, 2, 3 соответственно). Видно, что флуктуации пропускной способности являются небольшими по сравнению со средним значением. Например, при $Q = 5$ пользователей и $M = 16$ антенн на базовой станции в 80 % реализаций пропускная способность изменялась в пределах от 26,6 до 32,3 бит·с⁻¹·Гц⁻¹ при среднем значении, равном $C_{\text{mean}} = 29,7$ бит·с⁻¹·Гц⁻¹.

Таблица 1

$\{M; N\}$	Число пользователей						
	2	3	4	5	6	7	8
$\{4; 2\}$	1,27	—	—	—	—	—	—
$\{8; 2\}$	1,72	1,96	1,26	—	—	—	—
$\{12; 2\}$	1,85	2,50	2,84	2,67	1,49	-	-
$\{16; 2\}$	1,90	2,68	3,31	3,71	3,77	3,37	1,74
$\{16; 4\}$	1,69	1,85	0,78	—	—	—	—
$\{12; 3\}$	1,70	1,89	0,95	—	—	—	—
$\{12; 4\}$	1,47	0,70	—	—	—	—	—

4. ЗАКЛЮЧЕНИЕ

В настоящей работе рассмотрены ММО-системы, в которых передача данных осуществляется по параллельным собственным каналам, каждый из которых является адаптивно согласованным со случайным пространственным каналом связи. Формирование собственных каналов предполагает оценку на приёмном конце линии комплексных коэффициентов передачи между всеми антеннами и передачу этой информации по обратному каналу передатчику. Для увеличения пропускной способности ММО-систем предложен способ пространственного разделения (SDMA) пользователей, не требующий оценки направлений прихода сигналов. Это упрощает реализацию SDMA в условиях случайной рассеивающей среды, когда оценка направлений является сложной проблемой из-за пространственного распределения источника и флуктуаций «центра тяжести» его излучения. Разделение пользователей обеспечивается за счёт дополнительной обработки сигналов в ММО-системе, на основе ортогонализации всех собственных каналов всех пользователей.

Исследованы потенциальные характеристики ММО-системы с пространственным разделением пользователей, реализуемые при точно известной на обоих концах линии связи матрице комплексных коэффициентов передачи между всеми антеннами. Найдено ОСШ на выходе собственных каналов и полная пропускная способность ММО-системы при одновременном обслуживании произвольного числа пользователей. Показано, что пространственное разделение пользователей приводит к уменьшению ОСШ, поэтому полная пропускная способность ММО-системы с ростом числа пользователей сначала увеличивается, а затем начинает убывать из-за потерь в ОСШ. Приведены результаты численного моделирования в случае релеевских замираний сигналов при различном числе передающих и приёмных антенн, подтверждающие высокую эффективность предложенного способа пространственного разделения пользователей. Установлено, что максимальная пропускная способность достигается в случае, когда число разделяемых пользователей составляет $(50 \div 60)$ % от максимального, равного отношению числа антенн базовой станции к числу антенн отдельного пользователя.

Работа выполнена при поддержке программы «Ведущие научные школы» (грант № 00–15–96620) и РФФИ (грант № 00–02–17602).

СПИСОК ЛИТЕРАТУРЫ

1. Andersen J. B. // IEEE Antennas and Propagation Magazine. 2000. V. 42, No. 2. P. 12.
2. Ермолаев В. Т., Маврычев Е. А., Флакман А. Г. // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники. 2001. № 9. С. 50.
3. Maltsev A. A., Rubtsov A. E., Tiraspol'sky S. A. // Proc. of the 5th Sci. Conf. on Radio Phys. devoted to the 100th Anniversary of A. A. Andronov's Birth. Nizhny Novgorod, Russia, May 7, 2001. P. 191.
4. Kohno R. // IEEE Personal Commun. 1998. February. P. 28.
5. Vornefeld U., Walk C., Walk B. // IEEE Commun. Magaz. 1999. V. 37, No. 11. P. 52.
6. Sheikh K., Gesbert D., Gore D., Paulraj A. // IEEE Commun. Magaz. 1999. V. 37, No. 11. P. 100.
7. Tsoulos G., McGeehan J., Beach M. // IEE Proc. Radar, Sonar and Navigation. 1998. V. 145, No. 1. P. 73.
8. Pedersen K. I., Mogensen P. E., Fleury B. H. // IEEE Trans. on Vehicular Technology. 2000. V. 49, No. 2. P. 437.
9. Parsons J. D. The Mobile Radio Propagation Channel. London: Pentech Press Publisher, 1994.
10. Прокис Дж. Цифровая связь. М.: Радио и связь, 2000.
11. Bevan D. D. N., Ermolayev V. T., Flaksman A. G. // IEE Proc. Radar, Sonar and Navigation. 1998. V. 145, No. 1. P. 63.
12. Воеводин В. В. Линейная алгебра. М.: Наука, 1980.

13. Финк Л. М. Теория передачи дискретных сообщений. М.: Советское радио, 1970.

Нижегородский госуниверситет им. Н. И. Лобачевского,
г. Нижний Новгород, Россия

Поступила в редакцию
26 марта 2002 г.

SPACE DIVISION OF MULTIPLE ACCESS IN MIMO SYSTEMS WITH PARALLEL DATA TRANSMISSION

A. G. Flaksman

The cellular communication multiple-input multiple-output (MIMO) systems using antenna arrays at both link ends and the parallel channels for the data transmission are considered. These channels (the so-called eigenchannels) are formed using an adaptive transmitting and receiving beamformer processors and matched with the random fading environment. We propose a technique for the space division of multiple access (SDMA) of many users aimed at increasing the MIMO system capacity. The proposed method is based on the orthogonalization of the parallel channels of all users and does not require the estimation of the direction of arrival of the signals. The signal-to-noise ratio at the eigenchannel outputs and the full MIMO system capacity are found for an arbitrary user number. The simulation results for the case of the Rayleigh fading are given, which confirm the high efficiency of the proposed SDMA technique.