

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Известия высших учебных заведений

РАДИОФИЗИКА

ежемесячный научно-технический журнал

Издаётся с апреля 1958 г.

Том XLV № 10

Нижний Новгород

2002

Содержание

- Афраймович Э. Л., Водяников В. В., Воейков С. В., Первалова Н. П., Яковец А. Ф.**
Динамические характеристики перемещающихся волновых пакетов возмущений пол-
ного электронного содержания.....809
- Урядов В. П., Понятов А. А., Розанов С. В., Поддельский И. Н.** Особенности распро-
странения верхнего луча на трассах, пересекающих ионосферный провал828
- Заборонкова Т. М., Кудрин А. В., Лях М. Ю., Попова Л. Л.** Несимметричные свистовые
волны, направляемые цилиндрическими дактами с повышенной плотностью плазмы.....835
- Шульга С. Н.** Эффективный импеданс слабошероховатой границы произвольно анизо-
тропной среды.....858
- Костюков И. Ю., Крячко А. Ю., Токман М. Д.** О некоторых особенностях ЭЦР-нагрева
на половинной и фундаментальной гармониках в магнитной ловушке.....869
- Короновский А. А., Трубецков Д. И., Храмов А. Е., Храмова А. Е.** Универсальные зако-
номерности переходных процессов.....880
- Ханкина С. И., Яковенко В. М., Яковенко И. В.** Поверхностные плазменные волны на
неровной границе твёрдого тела887
- Турчин В. И.** Оценки параметров сигнала с минимальной ковариацией.....894
- Моисеев С. Н.** Предельное время предсказуемости процессов авторегрессии—скользя-
щего среднего900

УДК 550.388.2

ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПЕРЕМЕЩАЮЩИХСЯ ВОЛНОВЫХ ПАКЕТОВ ВОЗМУЩЕНИЙ ПОЛНОГО ЭЛЕКТРОННОГО СОДЕРЖАНИЯ

Э. Л. Афраймович¹, В. В. Водяников², С. В. Воейков¹,
Н. П. Первалова¹, А. Ф. Яковец²

С применением метода SOPHASE [1] и метода GPS-интерферометрии перемещающихся ионосферных возмущений [2, 3] дан детальный анализ пространственно-временных свойств перемещающихся волновых пакетов возмущений полного электронного содержания. Анализ проведён с использованием разработанной в ИСЗФ СО РАН технологии GLOBDET глобального детектирования и мониторинга ионосферных возмущений естественного и техногенного происхождения по данным измерений вариаций полного электронного содержания с помощью глобальной сети приёмников навигационной системы GPS на примере наиболее яркого проявления перемещающихся волновых пакетов 18 октября 2001 г. в Калифорнии (США) [3]. Во временной области перемещающиеся волновые пакеты представляют собой квазипериодические колебания полного электронного содержания длительностью порядка 1 часа с периодом колебаний $10 \div 20$ минут и амплитудой, превышающей амплитуду «фоновых» флуктуаций полного электронного содержания, как минимум, на порядок. Скорость и направление перемещения волновых пакетов соответствуют аналогичным параметрам среднеширотных среднемасштабных перемещающихся ионосферных возмущений, полученным ранее при анализе фазовых характеристик КВ радиосигналов, а также сигналов геостационарных ИСЗ и дискретных космических радиостанций.

ВВЕДЕНИЕ

Неослабевающий интерес к исследованиям атмосферных акустико-гравитационных волн (АГВ) на протяжении уже более 40 лет, начиная с пионерских работ [4, 5], обусловлен той важной ролью, которую играют эти волны в энергетике и динамике атмосферы Земли. Этим исследованиям посвящено большое количество публикаций, включая ряд обстоятельных обзоров [6, 7].

Обычно АГВ проявляются в ионосфере в форме перемещающихся ионосферных возмущений (ПИВ), детектируемых различными радиофизическими методами. ПИВ подразделяются на крупно- и среднемасштабные, различающиеся горизонтальной фазовой скоростью, которая больше (в случае крупного масштаба) или меньше (для среднего масштаба) скорости звука в нижней термосфере (порядка 300 м/с), и периодами в пределах $0,5 \div 3$ ч и $10 \div 40$ мин соответственно. Среднемасштабные перемещающиеся ионосферные возмущения (СМ ПИВ) наблюдаются преимущественно в дневные часы и ассоциируются с АГВ, которые генерируются в нижней атмосфере. Крупномасштабные ПИВ преобладают в ночные часы и тесно связаны с геомагнитной и авроральной активностью.

Известно, что источниками среднемасштабных АГВ могут быть как естественные процессы различной природы (магнитные бури, авроральные явления, погодные фронты, тропосферная турбулентность и струйные течения, солнечный терминатор, сильные землетрясения, извержения вулканов), так и антропогенные воздействия (запуски ракет, взрывы, ядерные испытания). Вследствие этого наблюдаемая картина возмущения электронной концентрации, по существу, является суммарным интерференционным волновым полем АГВ различного происхождения. Выделение из этого поля АГВ определённого типа представляет собой весьма трудную и в общем случае практически неразрешимую задачу. Поэтому наиболее достоверные измерения основных параметров среднемасштабных АГВ (параметры волнового вектора АГВ, спектральные и дисперсионные характеристики и т. д.) можно выполнить

только для весьма редкого и экзотического типа СМ ПИВ — квазипериодических (монохроматических) колебаний. Такие колебания впервые были зарегистрированы в виде соответствующих вариаций доплеровского смещения частоты F_D отражённого от ионосферы КВ радиосигнала [8–11].

Данные эксперименты позволили исследовать пространственно-временные характеристики СМ ПИВ, поскольку подобные записи легко визуально идентифицировать как отдельные монохроматические АГВ. К сожалению, такие исследования удалось осуществить для весьма ограниченного объёма экспериментальных данных. Так, например, авторы [10] смогли найти только несколько монохроматических ПИВ в данных, полученных за более чем 100 часов наблюдений.

В работе [11] также зарегистрировано лишь несколько реализаций монохроматических ПИВ за два периода наблюдений в зимние месяцы 1989 и 1990 гг. Авторы [11], по-видимому, впервые применили для обозначения квазимонохроматических вариаций частоты F_D термин «волновые пакеты» и сделали попытку объяснения их природы на основе изучения фазовой структуры колебаний. В [11] наблюдались два типа вариаций F_D : квазистохастические ПИВ и монохроматические ПИВ в виде волновых пакетов. Авторы [11] пришли к заключению, что квазистохастические ПИВ характеризуются случайным поведением фазы, малой длиной когерентности и большой вертикальной фазовой скоростью. Волновые пакеты показывают квазимонохроматические осцилляции частоты F_D , постоянное значение фазы, большую длину когерентности и меньшую вертикальную фазовую скорость.

Вслед за [11] мы остановились на использовании термина «волновые пакеты», расширив его до «перемещающиеся волновые пакеты» (ПВП). Проведённое в настоящей статье исследование достаточно убедительно показало, что это обозначение наиболее адекватно исследуемому явлению.

Несмотря на многолетние экспериментальные и теоретические исследования, до сих пор нет чёткого представления не только о физической природе квазимонохроматических СМ ПИВ, но даже об их морфологии (о частоте появления СМ ПИВ как функции географического положения, времени, уровня геомагнитной и метеоактивности и т. д.).

Для решения этих вопросов требуется получение статистически значимых наборов экспериментальных данных с хорошим пространственно-временным разрешением, которые позволили бы проследить как морфологические, так и динамические характеристики квазимонохроматических СМ ПИВ: направление перемещения, скорость распространения, положение возможного источника возмущения. Таковую возможность впервые предоставляет использование международной наземной сети двухчастотных приёмников навигационной системы GPS, насчитывающей к началу 2002 г. не менее 1 000 пунктов и поставляющей данные в сеть Интернет.

Использование разработанной в ИСЗФ СО РАН технологии глобального детектирования и определения параметров ионосферных возмущений различных классов позволило автоматизировать анализ и выявление ПВП. В работе [12] впервые представлены данные, характеризующие морфологию ПВП, наблюдавшихся в течение 30 суток в 1998–2001 гг. при различном уровне геомагнитной активности и числе станций глобальной сети GPS от 100 до 300. Общее число используемых для анализа радиотрасс составило около 300 000. Оказалось, что ПВП наблюдаются не более чем на $0,1 \div 0,4$ % от общего числа радиотрасс, причём чаще всего при низкой геомагнитной активности в дневное время зимой и осенью.

Целью настоящей работы является изучение пространственно-временных свойств ПВП по данным глобальной сети приёмников GPS с применением метода SOPHASE [1] и метода GPS-интерферометрии ПИВ [2, 3] на примере наиболее яркого проявления ПВП 18 октября 2001 г. в Калифорнии (США).

1. МЕТОД ОБРАБОТКИ ДАННЫХ ГЛОБАЛЬНОЙ СЕТИ ПРИЁМНИКОВ GPS. ОТБОР ПВП

Технология GPS предоставляет возможность детектирования волновых возмущений в ионосфере

на основе фазовых измерений полного электронного содержания (ПЭС) I_0 [13]:

$$I_0 [\text{м}^{-2}] = \frac{1}{40,308} \frac{f_1^2 f_2^2}{f_1^2 - f_2^2} [(L_1 \lambda_1 - L_2 \lambda_2) + \text{const} + nL], \quad (1)$$

где $L_1 \lambda_1$ и $L_2 \lambda_2$ — приращения фазового пути радиосигнала, вызванные задержкой фазы в ионосфере, L_1 и L_2 — число полных оборотов фазы, λ_1 и λ_2 — длины волн (в метрах), соответствующие частотам f_1 и f_2 (в герцах), const — некоторый неизвестный начальный фазовый путь (в метрах), nL — ошибка в определении фазового пути (в метрах).

Измерения фазы в системе GPS проводятся с высокой степенью точности, при которой ошибка в определении ПЭС при 30-секундных интервалах усреднения не превышает 10^{14} м^{-2} , хотя начальное значение ПЭС и остаётся неизвестным [13]. Это позволяет детектировать неоднородности ионизации и волновые процессы в ионосфере в широком диапазоне амплитуд (до 10^{-4} суточного изменения ПЭС) и периодов (от 5 мин до суток). Ниже мы будем использовать общепринятую в литературе единицу измерения полного электронного содержания TECU, равную 10^{16} м^{-2} .

Первичными данными являются ряды измерений ПЭС вдоль наклонного луча на ИСЗ («наклонного» ПЭС) $I(t)$, а также соответствующие им ряды угла места $\theta_s(t)$ и азимута $\alpha_s(t)$ луча на ИСЗ, рассчитанные по разработанной нами программе CONVTEC, которая преобразует полученные по сети Интернет стандартные для системы GPS RINEX-файлы. Для определения характеристик ПВП выбираются непрерывные ряды измерений $I(t)$ длиной не менее 2,5 часа.

Для нормировки амплитуды возмущений ПЭС используется преобразование «наклонного» ПЭС в эквивалентное «вертикальное» значение [14]

$$I = I_0 \cos \left[\arcsin \left(\frac{R_z}{R_z + h_{\max}} \cos \theta_s \right) \right], \quad (2)$$

где R_z — радиус Земли, h_{\max} — высота максимума слоя F_2 ионосферы.

Наиболее достоверные результаты определения параметров ПИВ соответствуют большим углам места $\theta_s(t)$ луча на ИСЗ, поскольку при этом эффекты сферичности становятся достаточно малыми. В нашем случае все результаты получены для углов места $\theta_s(t) > 30^\circ$.

С целью исключения вариаций регулярной ионосферы, а также трендов, обусловленных движением спутника, используется процедура удаления тренда с подходящим для конкретного эксперимента временным окном.

Разработанная в ИСЗФ СО РАН технология глобального детектирования ПВП позволяет в автоматическом режиме из большого количества экспериментального материала отбирать возмущения ПЭС, которые можно отнести к классу ПВП. Отбор ПВП проводился по двум критериям. Прежде всего, выбираются вариации ПЭС, среднеквадратичное отклонение которых превышает заданный порог ε . Обсуждаемая в этом разделе статистика ПВП получена для $\varepsilon = 0,1$ TECU. Далее, для каждого отфильтрованного ряда проверяется выполнение условия «квазимонохроматичности» колебаний ПЭС, при котором отношение R суммарной спектральной амплитуды сигнала в полосе частот δF заданной ширины в окрестности максимального значения S_{\max} амплитудного спектра к суммарной спектральной амплитуде сигнала вне полосы δF превышает заданное число R_{\min} . Все приводимые в настоящей статье данные получены при $R_{\min} = 2$.

2. ОБЩИЕ СВЕДЕНИЯ ОБ ЭКСПЕРИМЕНТЕ

18 октября 2001 г. между 15:00 и 18:00 UT по сигналам нескольких ИСЗ на многих станциях GPS, расположенных в Калифорнии, были зарегистрированы многочисленные перемещающиеся ионосферные возмущения типа ПВП. Район Калифорнии в пределах $220 \div 260^\circ$ в. д., $28 \div 42^\circ$ с. ш. удобен для

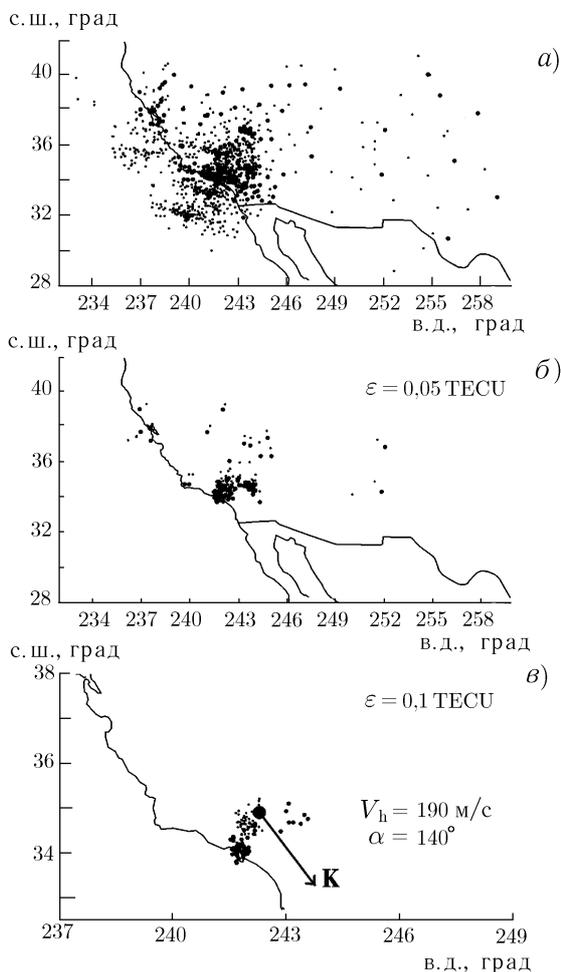


Рис. 1

В качестве элементов GPS-решёток для расчёта параметров ПВП использовались станции, отмеченные на рис. 1в.

Геомагнитную обстановку 18 октября 2001 г. можно охарактеризовать как слабо возмущённую. Это должно привести к некоторому увеличению фоновых флуктуаций ПЭС, однако не может вызвать крупномасштабных изменений электронной концентрации, характерных для геомагнитно-возмущённой ионосферы. D_{st} -вариации геомагнитного поля 18 октября 2001 г. представлены на рис. 2а.

Для анализа геомагнитной обстановки мы использовали также данные магнитной обсерватории Виктория ($48,52^\circ$ с. ш., $236,58^\circ$ в. д.), где для интересующего нас интервала времени 15:00–18:00 UT было зарегистрировано слабое геомагнитное возмущение, которое выразилось в уменьшении горизонтальной компоненты $H(t)$ магнитного поля на 60 нТл (рис. 2б). Одновременно наблюдалось небольшое увеличение амплитуды флуктуаций $dH(t)$ магнитного поля в диапазоне периодов $2 \div 20$ мин (рис. 2в). Размах колебаний геомагнитного индекса D_{st} для выбранного интервала времени был также сравнительно небольшим (не более 20 нТл), но в период 15:30–18:00 UT наблюдается хорошо выраженное падение индекса D_{st} , совпадающее с периодом уменьшения горизонтальной компоненты $H(t)$ магнитного поля (рис. 2а).

На рис. 2г дано распределение $N(t)$ числа ПВП, обнаруженных в этот день на всех анализируемых станциях глобальной сети GPS, со среднеквадратичным отклонением выше $\varepsilon = 0,1 \text{ TECU}$. Рис. 2д

наших исследований тем, что там расположено большое число станций GPS. Это позволяет получить целый набор GPS-решёток различной конфигурации для определения параметров ПИВ и обеспечивает возможность проверки достоверности расчётных данных. Важно также, что для указанного интервала времени (15:00–18:00 UT) и выбранного долготного диапазона местное время менялось от 08:00 до 11:00, что снижает уровень фоновых флуктуаций ПЭС, характерных для ночной ионосферы.

Рис. 1 иллюстрирует геометрию эксперимента 18 октября 2001 г. Жирными точками изображены станции GPS, меньшими — положение подионосферных точек для лучей приёмник GPS—ИСЗ. Поскольку в каждом приёмном пункте одновременно наблюдается несколько (не менее четырёх) спутников GPS, число радиотрасс значительно превышает число станций, что расширяет возможности анализа. На рис. 1а представлен весь набор станций GPS, использовавшийся в эксперименте. На рис. 1б и в показаны те станции и подионосферные точки, где в вариациях ПЭС были обнаружены ПВП с амплитудой, превышающей заданный порог ($\varepsilon = 0,05 \text{ TECU}$ и $\varepsilon = 0,1 \text{ TECU}$ соответственно). За единичным исключением, ПВП зарегистрированы на трассах, проходящих над сушей преимущественно в северо-восточном направлении. Как видно из рис. 1, увеличение порога регистрации в два раза сократило число зарегистрированных событий.

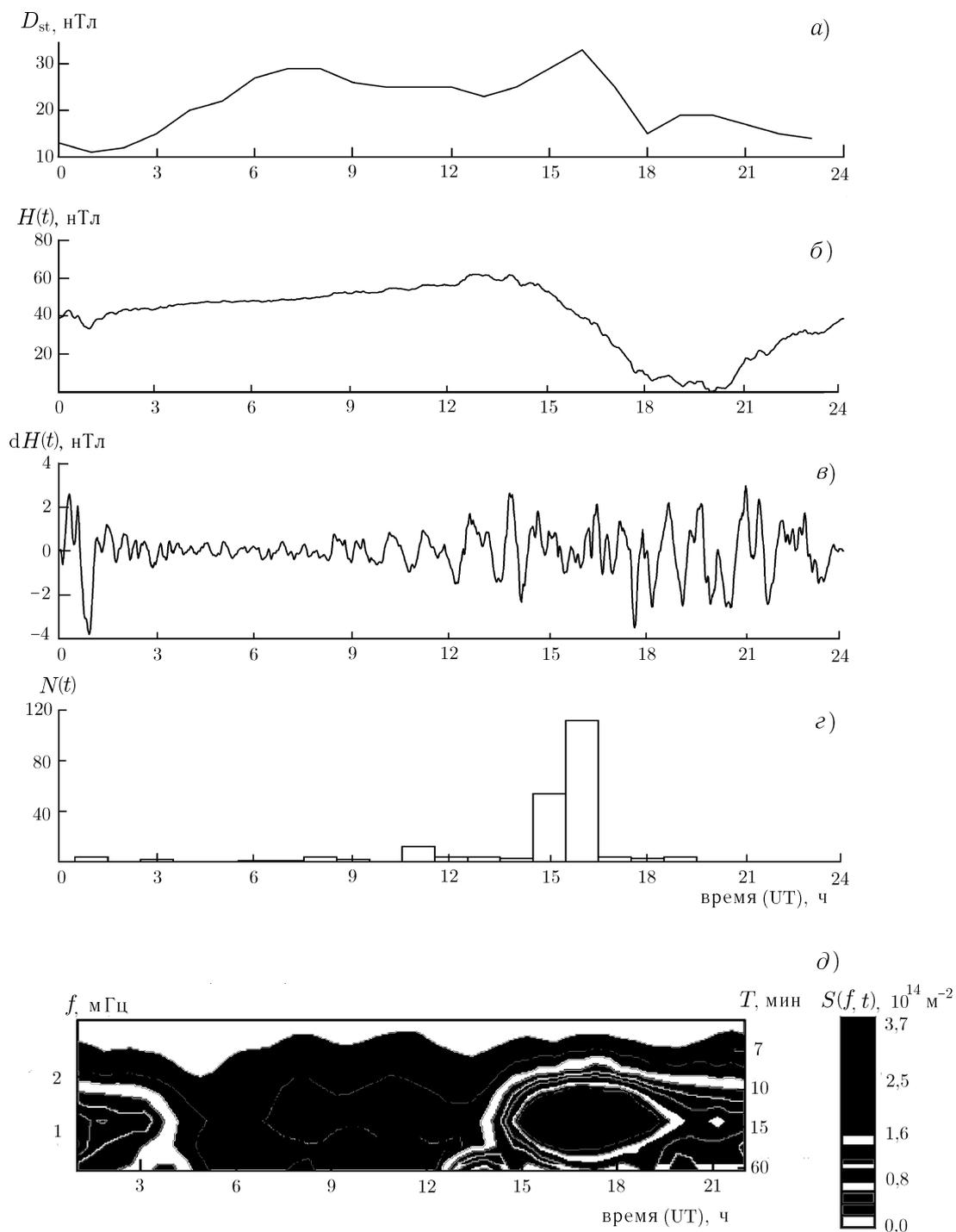


Рис. 2

демонстрирует динамический амплитудный спектр $S(f, t)$ вариаций ПЭС в диапазоне периодов $T = 5 \div 60$ мин, полученный с использованием метода пространственного усреднения спектров для всего района Калифорнии [15].

В целом вариации ПЭС коррелируют с изменениями геомагнитного поля. Между 15:00 и 19:00 UT одновременно с усилением колебаний H -компоненты магнитного поля, наблюдается расширение спек-

тра и увеличение амплитуды флуктуаций ПЭС. Наибольшую интенсивность имеют колебания ПЭС с периодами 12÷17 мин в интервале 15:30—17:00 UT. В этот же период времени зарегистрировано наибольшее количество ПВП (см. рис. 2г).

3. ОПРЕДЕЛЕНИЕ ФОРМЫ И ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ПВП МЕТОДОМ SADM-GPS

Метод SADM-GPS является развитием метода [16, 17] и подробно описан и протестирован в работах [2, 18, 19]. Ниже даётся краткое изложение метода, необходимое для понимания полученных результатов.

Мы определяем скорость и направление перемещения фазовой интерференционной картины (фазового фронта) в рамках какой-либо модели фазового фронта, адекватный выбор которой имеет решающее значение. В простейшей форме пространственно-временные изменения ПЭС в ионосфере $I(t, x, y)$ в каждый данный момент времени t можно представить в виде фазового фронта, перемещающегося без изменения формы (отсутствие дисперсии):

$$I(t, x, y) = F(t - x/u_x - y/u_y), \quad (3)$$

где $u_x(t)$ и $u_y(t)$ — скорости перемещения фазового фронта (линии равного значения ПЭС) вдоль осей x (направление на восток) и y (направление на север) соответственно.

Наиболее часто для описания движений фазового фронта используется частный случай формы (3), представляющий собой монохроматический волновой пакет:

$$I(t, x, y) = \delta \sin(\Omega t - K_x x - K_y y + \varphi_0), \quad (4)$$

где $I(t, x, y)$ — пространственно-временные вариации ПЭС, $\delta(t) = \exp[-(t - t_{\max})^2/t_d^2]$ — амплитуда, K_x , K_y и Ω — x - и y -проекции волнового вектора \mathbf{K} и угловая частота возмущения соответственно, φ_0 — начальная фаза возмущения, t_{\max} — момент времени, соответствующий максимальной амплитуде возмущения, t_d — полуширина «волнового пакета».

В действительности в чистом виде идеальные модели (3) и (4) не реализуются. Это обусловлено тем, что вызывающие ПИВ акустико-гравитационные волны распространяются в атмосфере в виде диспергирующего волнового пакета с конечной шириной углового спектра. Но в первом приближении для небольших времён усреднения, сравнимых с периодом отфильтрованных вариаций ПЭС, можно считать, что фазовая интерференционная картина перемещается без существенного изменения формы.

Автор работы [16] предложил статистический метод для анализа фазовой интерференционной картины. В этом методе первичные данные представляют собой временные зависимости пространственных производных ПЭС $I'_x(t)$ и $I'_y(t)$ вдоль направлений x и y . Суть метода заключается в определении ряда мгновенных значений азимута

$$\alpha(t) = \arctg[I'_x(t)/I'_y(t)] \quad (5)$$

с последующим построением на выбранном интервале времени функции распределения азимута $P(\alpha)$. Центральное значение α используется как оценка азимута преимущественного распространения ПИВ (с точностью до 180°).

Другой метод основан на анализе анизотропии фазовой интерференционной картины путём определения контраста C интерференционной картины [16]. В этом случае определяется отношение C_{xy} :

$$C_{xy} = \begin{cases} \sigma_X/\sigma_Y, & \sigma_X > \sigma_Y; \\ \sigma_Y/\sigma_X, & \sigma_Y > \sigma_X, \end{cases} \quad (6)$$

где σ_X и σ_Y — среднеквадратичные отклонения соответствующих рядов $X(t)$ и $Y(t)$, которые получены преобразованием исходных рядов $I'_x(t)$ и $I'_y(t)$ путём вращения первоначальной системы координат (x, y) на угол β :

$$X(t) = I'_x(t) \sin \beta + I'_y(t) \cos \beta, \quad Y(t) = -I'_x(t) \cos \beta + I'_y(t) \sin \beta. \quad (7)$$

В [16] показано, что существует такой угол β_0 , при котором отношение C_{xy} , становится максимальным и равным контрасту C . Этот параметр характеризует степень анизотропии фазовой интерференционной картины. Угол β_0 в этом случае показывает направление вытянутости фазового фронта, а угол $\alpha_c = \beta_0 + \pi/2$ — направление волнового вектора \mathbf{K} , совпадающее (с точностью до 180°) с направлением перемещения фазового фронта. Метод [16], по существу, позволяет определить только анизотропию и направление вытянутости фазовой интерференционной картины (с точностью до 180°), но не скорость перемещения фазового фронта. Ниже метод [16] используется для определения характеристик формы ПИВ.

Автор работы [18] разработал статистический угломерно-доплеровский метод SADM (Statistical Angle-of-arrival and Doppler Method) определения характеристик динамики ПИВ при трансионосферном зондировании на основе измерений не только пространственных производных $I'_x(t)$ и $I'_y(t)$, но и производной ПЭС по времени $I'_t(t)$. Это даёт возможность однозначно установить ориентацию $\alpha(t)$ волнового вектора \mathbf{K} в диапазоне $0 \div 360^\circ$, а также определить модуль горизонтальной компоненты скорости $V_h(t)$ в каждый момент времени.

В работе [2] описана модификация метода SADM для GPS-решёток (метод SADM-GPS), которая основана на простой модели перемещения фазовой интерференционной картины, движущейся без изменения формы. Этот метод использует текущую информацию об угле места $\theta_s(t)$ и азимуте $\alpha_s(t)$ луча на ИСЗ для коррекции изменения модуля горизонтальной скорости $V_h(t)$, обусловленного движением ИСЗ. Метод SADM-GPS позволяет определить в каждый момент времени модуль горизонтальной компоненты скорости $V_h(t)$ и азимут $\alpha(t)$ перемещения ПИВ в неподвижной системе координат (x, y) :

$$\begin{aligned} \alpha(t) &= \operatorname{arctg}[u_y(t)/u_x(t)]; \\ u_x(t) &= I'_t(t)/I'_x(t) = u(t)/\cos \alpha(t), \quad u_y(t) = I'_t(t)/I'_y(t) = u(t)/\sin \alpha(t), \\ u(t) &= |u_x(t)u_y(t)|/[u_x^2(t) + u_y^2(t)]^{1/2}; \\ V_x(t) &= u(t) \sin \alpha(t) + w_x(t), \quad V_y(t) = u(t) \cos \alpha(t) + w_y(t), \quad V_h(t) = [V_x^2(t) + V_y^2(t)]^{1/2}, \end{aligned} \quad (8)$$

где u_x и u_y — скорости перемещения фазового фронта вдоль осей x и y в плоскости GPS-решётки, w_x и w_y — x - и y -проекции скорости \mathbf{w} перемещения подыоносферной точки (для учёта движения ИСЗ).

Скорость \mathbf{w} в приближении достаточно больших углов места $\theta_s(t)$ определяется следующим образом. Координаты подыоносферной точки $x_s(t)$ и $y_s(t)$ на высоте h_{\max} в топоцентрической системе координат записываются в виде

$$x_s(t) = h_{\max} \sin[\alpha_s(t)] \operatorname{ctg}[\theta_s(t)], \quad y_s(t) = h_{\max} \cos[\alpha_s(t)] \operatorname{ctg}[\theta_s(t)], \quad (9)$$

а x - и y -компоненты скорости перемещения \mathbf{w} в виде

$$w_x(t) = \frac{d}{dt} x_s(t), \quad w_y(t) = \frac{d}{dt} y_s(t), \quad w(t) = [w_x^2(t) + w_y^2(t)]^{1/2}. \quad (10)$$

Изложим кратко последовательность процедур обработки данных. Из большого числа станций GPS выбираются три (A, B, C) пункта, расстояния между которыми не превышают величины порядка

половины ожидаемой длины волны Λ возмущения. Пункт B принимается за центр топоцентрической системы отсчёта, ось x которой направлена на восток, ось y — на север. Приёмные пункты в данной системе отсчёта имеют координаты $A(x_A, y_A)$, $B(0, 0)$, $C(x_C, y_C)$. Такая конфигурация приёмников GPS представляет собой GPS-решётку с минимальным необходимым количеством элементов. В районах с густой сетью GPS-приёмников мы можем получить целый набор GPS-решёток различной конфигурации, что даёт возможность проверить достоверность получаемых данных; в настоящей работе мы использовали эту возможность.

Поскольку расстояние между элементами GPS-решётки (от десятков до нескольких сотен километров) много меньше расстояния до ИСЗ (более 20 000 километров), геометрия решётки на высоте $h_{\max} = 300$ км подобна геометрии решётки на уровне земной поверхности, так что относительные координаты подионосферных точек известны с заведомо высокой точностью (единицы метров).

Ряды линейных преобразований разности значений отфильтрованного ПЭС $I_B(t) - I_A(t)$ и $I_B(t) - I_C(t)$ в точках приёма A , B и C используются далее для определения компонент градиента ПЭС I'_x и I'_y [2]. Временная производная I'_t определяется дифференцированием $I_B(t)$ в точке B .

Полученные ряды используются для определения в каждый момент времени модуля горизонтальной скорости $V_h(t)$ и азимута $\alpha(t)$ перемещения ПИВ с учётом движения ИСЗ на орбите (формула (8)). Далее ряды $V_h(t)$ и $\alpha(t)$ подвергаются статистической обработке. Она включает в себя построение распределений скорости $P(V_h)$ и направления $P(\alpha)$, которые анализируются с целью проверки гипотезы о существовании преимущественного направления распространения. Если такое направление существует, то картину можно считать перемещающейся без изменения формы, а среднюю горизонтальную скорость её перемещения и преимущественное направление перемещения можно определить из распределений $P(V_h)$ и $P(\alpha)$.

При известном азимуте $\alpha(t)$ можно оценить также угол места $\theta(t)$ волнового вектора \mathbf{K} ПИВ и модуль фазовой скорости $V(t)$, поскольку амплитуда волновых возмущений ПЭС имеет хорошо выраженную ракурсную зависимость [20]. Максимум амплитуды наблюдается для тех ПИВ, волновой вектор \mathbf{K} которых перпендикулярен направлению луча \mathbf{r} приёмник—ИСЗ, т. е. когда выполнено условие

$$\operatorname{tg} \theta = -\cos(\alpha_s - \alpha) / \operatorname{tg} \theta_s. \quad (11)$$

Модуль фазовой скорости V определяется как

$$V = V_h \cos \theta. \quad (12)$$

Ракурсная зависимость амплитуды возмущения ПЭС имеет важное значение при исследовании волновых возмущений. Условие (11) ограничивает число траекторий, для которых возможно уверенное детектирование ПИВ на фоне шумов. Ракурсный эффект смещает положение максимума возмущения на временной оси, что может приводить к ошибкам в определении скорости перемещения ПИВ, если скорость рассчитывается по временным задержкам. Кроме того, как будет показано ниже, за счёт ракурсного эффекта в вариациях ПЭС могут наблюдаться структуры типа волновых пакетов, которые не существуют в действительности.

Для распределения ионизации, близкого к гауссовскому, с возмущениями электронной концентрации в виде плоской монохроматической волны коэффициент M изменения амплитуды возмущения ПЭС определяется углом γ между направлением распространения возмущения \mathbf{K} и лучом зрения \mathbf{r} на ИСЗ, углом места θ_s вектора \mathbf{r} , а также отношением длины волны возмущения $\Lambda = 2\pi/|K|$ к полутолщине максимума ионизации h_d [20]:

$$M(\gamma) \propto \frac{h_d}{\sin \theta_s} \exp \left[- \left(\frac{\pi h_d \cos \gamma}{\Lambda \sin \theta_s} \right)^2 \right]. \quad (13)$$

Исследование влияния ракурсных эффектов проводилось с помощью моделирования волновых возмущений электронной концентрации (см. раздел 5).

Таким образом, в результате использования описанных в данном разделе преобразований для каждой из выбранных для анализа GPS-решёток мы получаем следующие средние за интервал времени $1 \div 2$ часа параметры ПВП: $\langle A_I \rangle$ — амплитуда возмущения ПЭС, $\langle \alpha \rangle$ и $\langle \theta \rangle$ — азимут и угол места волнового вектора \mathbf{K} , $\langle V_h \rangle$ и $\langle V \rangle$ — горизонтальная компонента и модуль фазовой скорости, а также контраст C и азимут нормали к фронту волны α_c по методу [16].

4. ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПВП, ПОЛУЧЕННЫЕ С ПОМОЩЬЮ МЕТОДА SADM-GPS

На рис. 3а приведены характерные временные зависимости ПЭС $I(t)$ для спутника GPS PRN14, полученные на трёх GPS-станциях BRAN, CHMS и DUPS в Калифорнии. Три указанные пункта GPS представляют собой типичную GPS-решётку, обработка данных которой проводилась по методике, описанной в предыдущем разделе. На рис. 3б для тех же станций даны вариации ПЭС $dI(t)$, отфильтрованные из исходных рядов $I(t)$ полосовым фильтром, выделяющим вариации с периодами от 5 до 20 мин.

Отфильтрованные ряды в период 15:00–16:00 UT демонстрируют наличие значительных колебаний ПЭС типа уединённого волнового пакета продолжительностью около одного часа и амплитудой $A = 0,5$ TECU. Размах отфильтрованных колебаний $dI(t)$ не только намного превышает погрешность фазовых измерений (10^{-2} TECU), но и почти на порядок превосходит уровень фоновых вариаций ПЭС [15]. Колебания ПЭС на трёх пространственно разнесённых станциях GPS показывают высокую степень подобия и имеют небольшой временной сдвиг. Это свидетельствует о том, что мы имеем дело с одним и тем же перемещающимся возмущением.

Результаты расчёта по алгоритму SADM-GPS средних значений азимута $\langle \alpha \rangle$ и горизонтальной скорости $\langle V_h \rangle$ перемещения возмущения для каждого 30-секундного интервала времени представлены на рис. 3в и г. Как видно из рисунка, данный волновой пакет двигался преимущественно в юго-восточном направлении со средней скоростью около 180 м/с. Разброс отсчётов обусловлен неполным соответствием реальной картины и идеальной модели ПВП в виде монохроматического пакета (4), в частности, из-за наличия фоновых некоррелированных флуктуаций ПЭС [2].

Обработка данных других GPS-решёток в том же районе (см. рис. 1в) с использованием метода SADM-GPS позволила получить распределения основных параметров ПВП, зарегистрированных 18 октября 2001 г. в Калифорнии. Общее число различных комбинаций GPS-решёток для интервала времени 15:00–16:00 UT составило 231. Статистические данные показывают совпадение средних значений рассчитываемых параметров в пределах их среднеквадратичных отклонений, что свидетельствует о хорошей устойчивости получаемых данных независимо от конфигурации GPS-решётки.

Распределения средних значений параметров ПВП, рассчитанных для каждой из 231 GPS-решётки, приведены на рис. 4а–в. Горизонтальная скорость V_h распространения ПВП (рис. 4а) согласно нашим данным меняется в среднем от 60 до 270 м/с, наиболее вероятное значение V_h составляет 190 м/с. Длина волны Λ ПВП при среднем периоде колебаний порядка 10^3 с оказывается порядка $150 \div 200$ км.

Анализ распределения азимутов $P(\alpha)$ (рис. 4б) показывает хорошо выраженное юго-восточное направление перемещения ПВП: $\alpha \approx 140^\circ \pm 20^\circ$. Среднее по 231 GPS-решёткам значение контраста $C = 7$ свидетельствует о сильной анизотропии ПВП. На рис. 4б отображено также распределение азимута нормали к фронту волны $P(\alpha_c)$, определённого по методу [16]. Это распределение практически совпадает с $P(\alpha)$, что указывает на то, что ПВП перемещаются поперёк направления своей вытянутости. Таким образом, характерный размер всего волнового пакета в направлении распространения

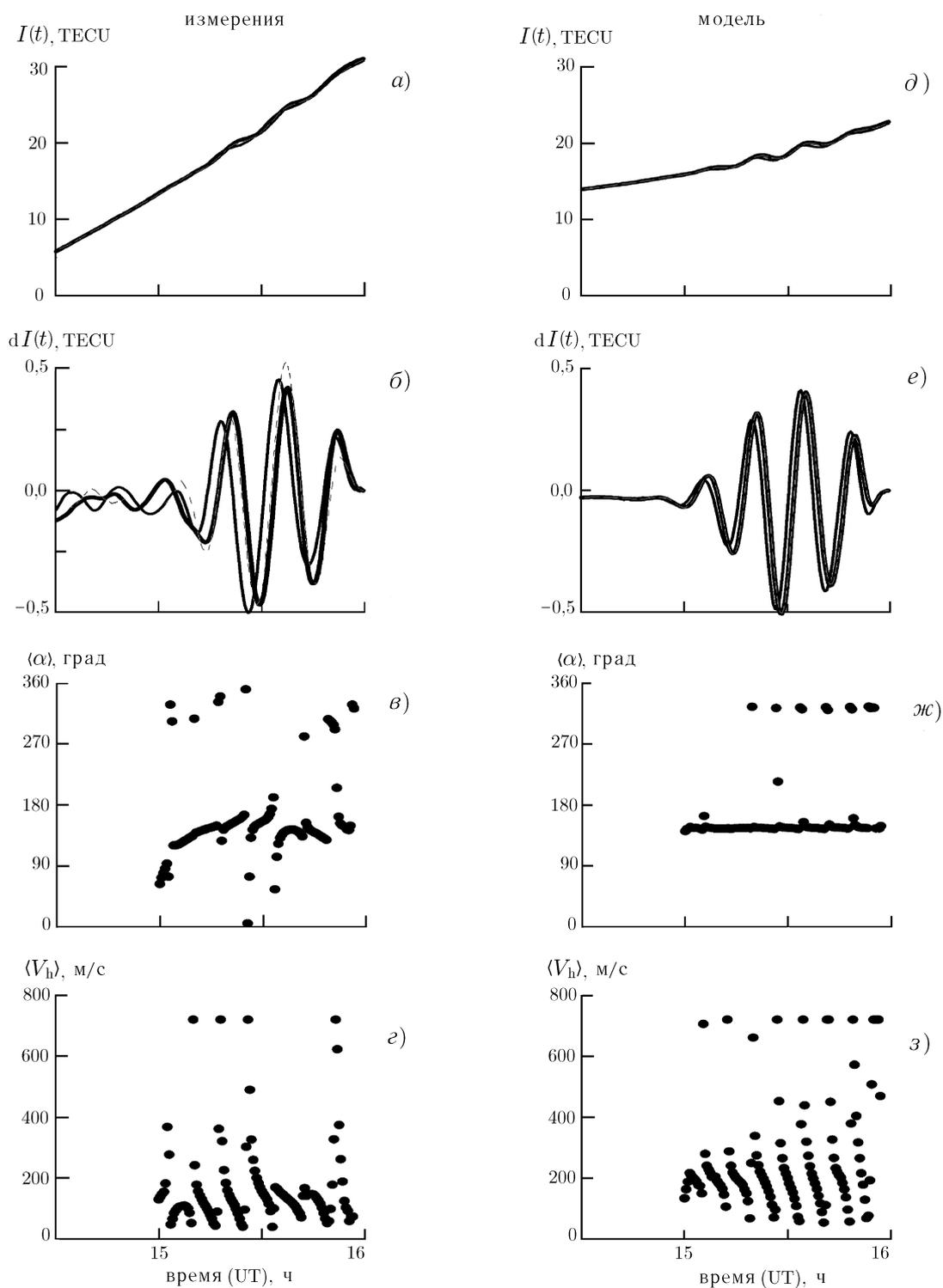


Рис. 3

составляет около $300 \div 500$ км, а вдоль фронта волны — до 10^3 км.

На рис. 1в стрелочкой схематически показан волновой вектор \mathbf{K} ПВП. Представленные на рис. 1в значения α и V_h соответствуют наиболее вероятным значениям азимута распространения и модуля

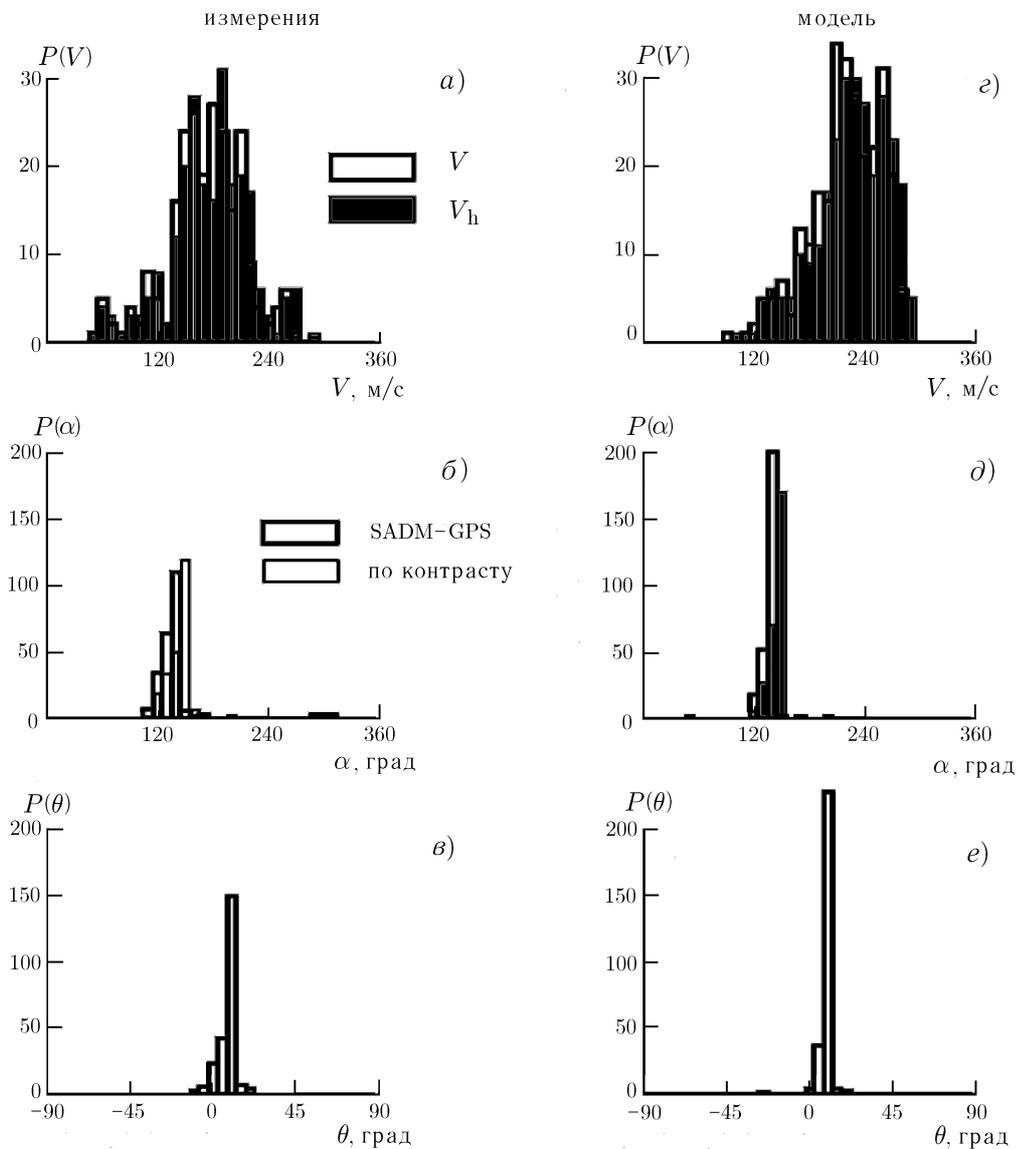


Рис. 4

горизонтальной скорости ПВП.

Угол места волнового вектора ПВП, определённый по ракурсному условию (11), в большинстве случаев имеет небольшое положительное значение (рис. 4в), т. е. фаза колебаний распространяется вверх. Поскольку средний угол места составляет всего 10° , можно считать, что ПВП представляют собой практически горизонтальную волну. Соответственно, оценки модуля фазовой скорости V (рис. 4а) дают значения, близкие к её горизонтальной проекции V_h : $V \sim 50 \div 270$ м/с, наиболее вероятное значение V составляет 180 м/с.

5. МОДЕЛИРОВАНИЕ

В настоящей работе исследование влияния ракурсных эффектов на характер изменения ПЭС и точность рассчитываемых параметров ПВП, а также проверка достоверности определения характеристик ПВП проводились с помощью моделирования волновых возмущений электронной концентрации для условий наблюдения 18 октября 2001 г.

Разработанная нами модель измерений ПЭС на GPS-интерферометре позволяет рассчитывать пространственное и временное распределение локальной электронной концентрации N_e в ионосфере, приближённое к реальному, а затем, используя координаты пунктов приёма и спутников, интегрировать N_e вдоль луча приёмник—ИСЗ с заданным шагом по времени [2]. В результате получаются временные ряды ПЭС, аналогичные экспериментальным данным, обрабатывать которые можно теми же методами, что и экспериментальные.

Модель ионизации учитывает высотное распределение концентрации N_e , суточные и сезонные изменения концентрации, определяемые зенитным углом Солнца, а также нерегулярные возмущения N_e с меньшими амплитудами и пространственными масштабами в виде дискретной суперпозиции плоских бегущих волн. Подробное описание использовавшейся модели дано в работе [2].

В настоящей работе с целью выяснения природы ПВП моделировались три типа возмущений:

1) Возмущение в форме одиночной плоской волны с амплитудой A_1 , составляющей 3 % от значения N_e в максимуме ионизации, периодом $T_1 = 15$ мин и длиной волны $\Lambda_1 = 172$ км. Угол места направления распространения волны $\theta_1 = 10^\circ$, азимут $\alpha_1 = 146^\circ$.

2) Возмущение в виде суперпозиции двух плоских волн с периодами $T_1 = 15$ мин, $T_2 = 12$ мин и длинами волн $\Lambda_1 = 172$ км, $\Lambda_2 = 138$ км. Углы места и азимуты направления распространения волн задавались равными: $\theta_1 = \theta_2 = 10^\circ$, $\alpha_1 = \alpha_2 = 146^\circ$; амплитуды волн $A_1 = A_2$ составляли 3 % от значения N_e в максимуме ионизации.

3) Возмущение в форме одиночного волнового пакета с полушириной $t_d = 20$ мин и максимальной амплитудой в момент $t_{\max} = 15:30$ UT. Осцилляции внутри пакета имели период $T_1 = 15$ мин и длину волны $\Lambda_1 = 172$ км.

Параметры возмущений выбирались близкими к полученным из экспериментальных данных по методике [2].

Моделирование ПВП преследовало две цели: проверку достоверности расчёта характеристик ПВП и выяснение природы ПВП.

На рис. 3д—з представлены результаты расчётов по модели ПВП в форме одиночного волнового пакета для GPS-станций BRAN, CHMS, DUPB 18 октября 2001 г. Параметры волнового пакета выбирались близкими к экспериментальным данным (рис. 3а—с).

Рис. 3 иллюстрирует хорошее подобие экспериментальных и модельных вариаций ПЭС $I(t)$, $dI(t)$ и зависимостей $V_h(t)$, $\alpha(t)$. Отметим более слабый наклон модельного ПЭС $I(t)$ (рис. 3д) по сравнению с экспериментальным (рис. 3а). Это обусловлено тем, что в модели с целью простоты и наглядности суточные изменения ионизации выбраны пропорциональными косинусу зенитного угла Солнца. В реальных условиях зависимость более сложная, что даёт более быстрый временной рост ПЭС. Поскольку в дальнейшем из рядов удаляется тренд и рассматриваются только относительные вариации ПЭС $dI(t)$, указанное несоответствие в поведении модельного и экспериментального ПЭС не сказывается на полученных результатах.

Сравнение заданных в модели параметров возмущения ПЭС с соответствующими значениями, полученными после обработки экспериментальных данных методом SADM-GPS, показывает достаточную близость этих величин (рис. 4з—е). Азимут волнового вектора ПВП полагался $\alpha = 146^\circ$, рассчитанное по методу SADM-GPS среднее значение азимута составляет $146,2^\circ$ (рис. 3ж и 4д); горизонтальная скорость задавалась равной 180 м/с, рассчитанное среднее значение скорости — 207,5 м/с (рис. 3з и 4г); угол места в модели $\theta = 10^\circ$, наиболее вероятный рассчитанный угол места также составлял 10° (рис. 4е). Определённый с помощью метода SADM-GPS азимут распространения ПВП близок также к азимуту, вычисленному по анализу контраста фазовой интерференционной картины [16] (рис. 4б, д). Всё это доказывает адекватность методики SADM-GPS и подтверждает достоверность полученных с её помощью параметров ПВП.

Рассмотрим теперь возможные механизмы образования структур типа ПВП в наблюдаемых вари-

ациях ПЭС. Регистрируемый ПВП представляет собой одиночный волновой пакет с длительностью порядка 60 мин, периодом осцилляций внутри пакета $12 \div 15$ мин и амплитудой, превышающей уровень фоновых флуктуаций в $2 \div 5$ раз (рис. 3б). Структуры такого типа могут возникать в вариациях ПЭС различными способами.

Если в ионосфере распространяется монохроматическая волна с периодом около 15 мин (период наблюдаемых вариаций ПЭС), то осцилляции ПЭС типа ПВП могли образоваться за счёт ракурсного эффекта. Рис. 5а, г иллюстрируют такую ситуацию: представлены отфильтрованные ряды ПЭС, полученные моделированием одиночной плоской волны при разных условиях её наблюдения. В приёмном пункте BRAN (рис. 5а) углы места ИСЗ PRN14 близки к 90° , а волновой вектор возмущения перпендикулярен лучу зрения приёмник—ИСЗ в течение всего периода наблюдений. Об этом свидетельствует характер теоретической зависимости $M(\gamma)$, рассчитанной по формуле (13) для данного луча: $M(\gamma)$ близко к 1 на всём интервале наблюдения (рис. 5в). Поэтому на луче BRAN—PRN14 наблюдаются волновые возмущения ПЭС, близкие к заданной монохроматической волне.

В пункте BRU1 для ИСЗ PRN05 условия детектирования возмущений значительно хуже. Первые полчаса волна в этом случае распространяется практически вдоль луча зрения ($M(\gamma) < 0,3$, рис. 5е), а в конце интервала наблюдения луч зрения BRU1—PRN05 перпендикулярен направлению распространения волны, но имеет очень низкие углы места ($M(\gamma) > 1$, рис. 5е). За счёт этого в вариациях ПЭС наблюдается рост амплитуды колебаний от 0 до 0,7 TECU — возникает образование, похожее на волновой пакет, которого не было в исходной заданной модели волны. Однако, как видно из рис. 5, в данном случае рост амплитуды колебаний ПЭС имеет почти линейный характер, в то время как в эксперименте (см. рис. 3б) наблюдается нелинейная модуляция амплитуды.

Исследование характера ракурсной зависимости для всех использовавшихся GPS-решёток, показало, что в 97 % случаев (в том числе и в тех, которые приведены на рис. 3б) реализуется ситуация, отображённая на рис. 5в, т. е. $M(\gamma) \sim 1$ на протяжении всего интервала наблюдений. Учитывая всё вышесказанное, можно сделать вывод, что регистрируемые ПВП обусловлены распространением реального волнового пакета возмущения локальной электронной концентрации в ионосфере, а не являются следствием условий регистрации.

Другая возможность образования волнового пакета ПВП — сложение двух или нескольких монохроматических волн с близкими периодами, распространяющихся в ионосфере [11]. С целью анализа такой ситуации моделировалось распространение двух квазигоризонтальных волн с периодами $T_1 = 15$ мин и $T_2 = 12$ мин, равными амплитудами, скоростями и направлением распространения. Полученные вариации ПЭС на двух лучах BRAN—PRN14 и BRU1—PRN05 после удаления тренда представлены на рис. 5б и д соответственно.

Как и следовало ожидать, в этом случае возникают амплитудно-модулированные осцилляции ПЭС с периодом модуляции около $T = T_1 T_2 / (T_2 - T_1) = 60$ мин. Период T близок к длительности регистрируемого в эксперименте волнового пакета, однако простая суперпозиция двух волн даёт не один,

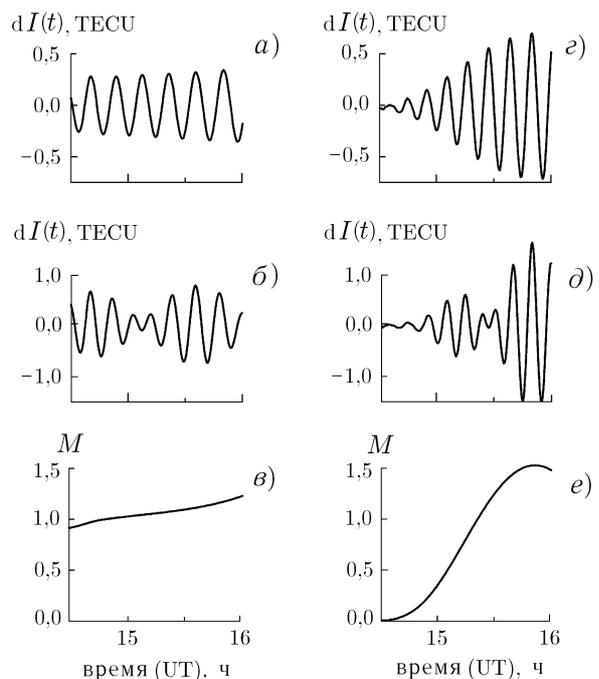


Рис. 5

а целую цепочку волновых пакетов. В случае, когда ракурсный эффект мало сказывается на характере вариаций ПЭС (рис. 5в), эти пакеты имеют, кроме того, и равные амплитуды (рис. 5б). Изменение фазы суммируемых волн мало сказывается на представленной картине, лишь незначительно меняя форму и положение минимума амплитуды. Ракурсная же зависимость (рис. 5е) вносит дополнительную модуляцию в регистрируемый сигнал: происходит существенное усиление одного из пакетов цепочки (рис. 5д), а картина приближается к наблюдаемой в эксперименте (рис. 3б).

Тем не менее нельзя считать, что наблюдаемые одиночные волновые пакеты ПВП являются следствием ракурсного эффекта. Во-первых, при ракурсной модуляции в силу медленного (почти линейного) изменения амплитуды прослеживается не один, а, по крайней мере, два волновых пакета. Во-вторых, как уже отмечалось выше, характер ракурсной зависимости в большинстве случаев регистрации ПВП оказался таким, что он мало влияет на амплитуду возмущений. Чтобы получить осцилляции ПЭС, наиболее близкие к наблюдаемым в эксперименте, нам пришлось ввести искусственную модуляцию, т. е. амплитуда исходной монохроматической волны возмущения концентрации N_e задавалась не постоянной, а имела зависимость от времени в виде функции Гаусса. Именно такая модель вариаций ПЭС показана на рис. 3е.

Сопоставляя обнаруженные нами ПВП с данными работ [4, 11, 21, 22], а также с результатами проведённого моделирования, можно предположить, что в атмосфере существует дополнительный механизм модуляции амплитуды волновых процессов, который позволяет получить осцилляции ПЭС в виде одиночного волнового пакета. Наиболее вероятно, что такой механизм обусловлен дисперсионными свойствами атмосферы.

6. ФАЗОВАЯ СТРУКТУРА ПАКЕТА И ОПРЕДЕЛЕНИЕ СКОРОСТИ И НАПРАВЛЕНИЯ ПЕРЕМЕЩЕНИЯ ПВП МЕТОДОМ SOPHASE

Для расчёта параметров ПВП применялся также метод SOPHASE [1, 23, 24]. Ранее этот метод был с успехом применён при анализе сейсмических волн, акустических сигналов, генерируемых при запуске ракеты «Saturn—Apollo» [23], а также при исследовании атмосферных волн над Казахстаном [24].

Анализ проводился по 12 рядам измерений ПЭС $I(t)$, полученным 18 октября 2001 года на станциях GPS в Калифорнии и содержащим хорошо выраженные ПВП.

На рис. 6 представлены временные зависимости амплитуды $A(t)$ (слева) и фазы $\varphi(t)$ (справа) отфильтрованных рядов ПЭС $dI(t)$ для отдельных станций и спутников GPS (отмечены у каждой кривой), рассчитанные методом комплексной демодуляции [11, 25]. Отмечены два интервала времени, на которых наблюдаются колебания амплитуды $A(t)$ при постоянной или медленно изменяющейся фазе: интервал A (14:06–14:20 UT) и интервал B (15:15–16:06 UT). На интервале A амплитуда $A(t)$ мала (не более 0,1 TECU), что не позволяет с уверенностью выделить на данном временном отрезке волновой пакет. На интервале B амплитуда колебаний $A(t)$ в 4–5 раз превышает амплитуду колебаний на интервале A . Соответственно, на отрезке B хорошо различим волновой пакет. Для дальнейшего анализа мы использовали волновой пакет, наблюдаемый на интервале B .

Данные и использованные методы не позволяют разделить основной пакет, наблюдаемый на интервале времени 15:15–16:06 UT, на какие-либо составляющие. Поэтому приходится считать этот пакет монохроматическим, что вполне согласуется с результатами моделирования (см. раздел 5).

Статистическая характеристика SOPHASE, выражающая мощность принимаемого сигнала как функцию горизонтальной скорости и направления распространения волны, для данных с решётки де-

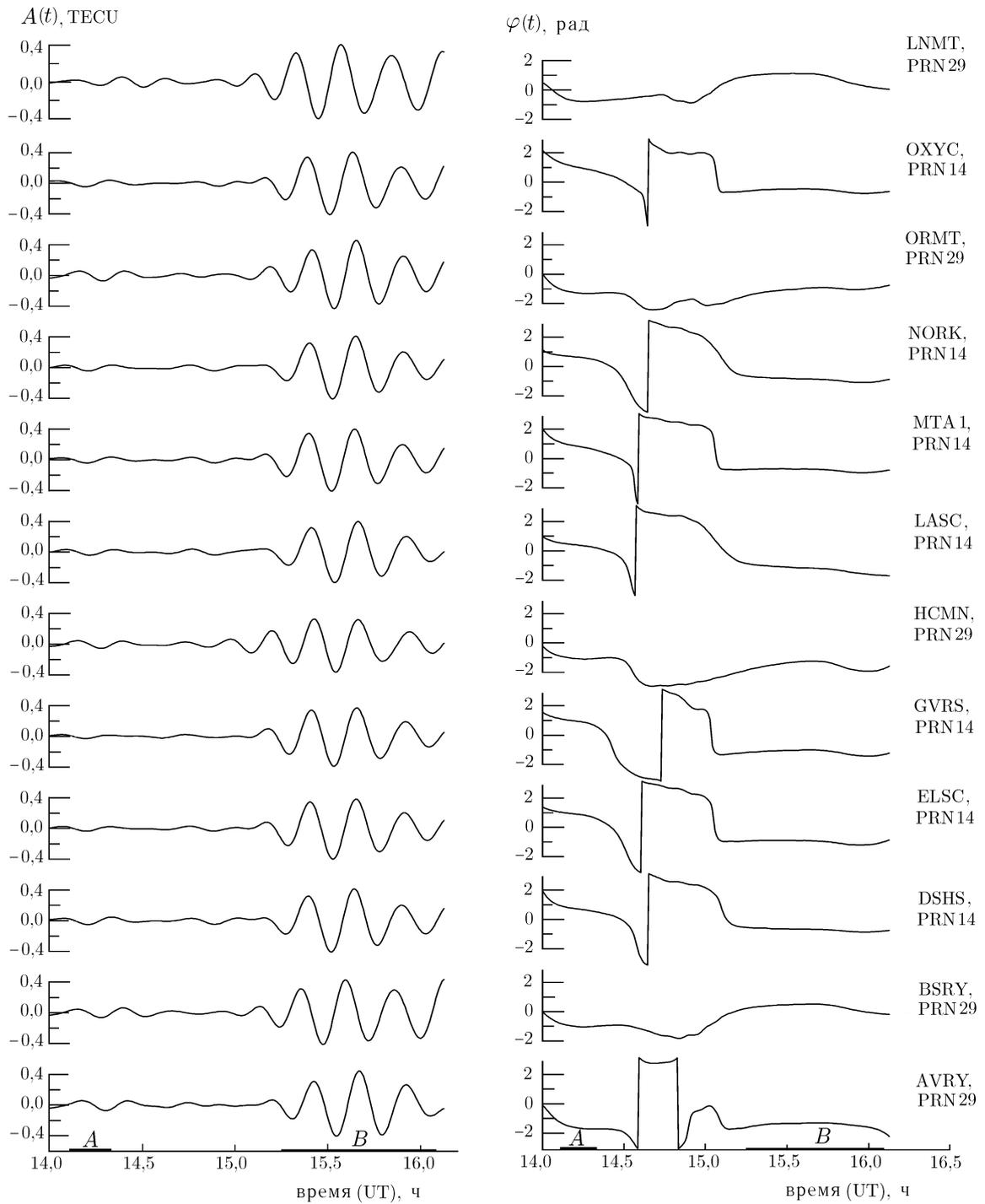


Рис. 6

текторов, принимающих сигнал и некоррелированный шум, определяется следующим образом:

$$C(V_h, \alpha) = \frac{\sum_{n=N_1}^{N_2} \sum_{k=1}^K \sum_{j=k+1}^{K-1} [A_j(t_n) + A_k(t_n)] \cos \psi}{\sum_{n=N_1}^{N_2} \sum_{k=1}^K \sum_{j=k+1}^{K-1} [A_j(t_n) + A_k(t_n)]}, \quad (14)$$

где $\cos \psi = \cos[(\varphi_j^F - \varphi_k^F) - (\Delta\varphi_j^V + \Delta\varphi_k^V)/2]$, V — фазовая скорость возмущения, α — азимут направления распространения волны, n — порядковый номер точки отсчёта (в интервале от N_1 до N_2 , что соответствует началу и концу выбранного временного интервала), j, k — номер записи соответствующего детектора, K — общее число записей, $A_j(t_n)$ — амплитуда j -й записи в момент времени t_n , φ_j^F — фаза j -й записи в момент времени t_n , $\Delta\varphi_j^V = -\omega_j r_{jk} \cos(\gamma - \alpha)/V$, $\omega_j = 2\pi f + d\varphi_j/dt$, r_{jk} — расстояние между детекторами j и k в момент времени t_n , γ — азимут радиус-вектора, направленного от детектора k к детектору j .

В общем случае характеристика $C(V_h, \alpha)$ является результатом расчёта нормированной взвешенной суммы по всем $K(K - 1)/2$ парам детекторов и по всем моментам времени косинусов разностей между разностью фаз записей и предполагаемой разностью фаз. Отметим, что $C(V_h, \alpha) = 0$, если записи детекторов не коррелированы, и $C(V_h, \alpha) = 1$, если записи детекторов воспроизводят недиспергирующий распространяющийся сигнал с предполагаемыми значениями V_h и α , соответствующими фактическим значениям скорости и азимута сигнала.

В нашем случае в качестве детекторов используются подыоносферные точки, поэтому, чтобы использовать метод SOPHASE для обработки GPS-данных, необходимо учесть движение спутников GPS. При учёте движения спутников аналогично методу SADM-GPS (раздел 3), в формуле (14) изменяется выражение для $\Delta\varphi_j^V$:

$$\Delta\varphi_j^V = -\frac{\omega_j r_{jk} \cos(\gamma - \alpha)}{V_h - w_j \cos(\alpha_{sj} - \alpha)}. \tag{15}$$

В нашем случае количество записей $K = 12$, центральная частота $f = 1,1028 \cdot 10^{-3}$ Гц, w_j — скорость перемещения подыоносферной точки j , α_{sj} — азимут вектора скорости перемещения подыоносферной точки j .

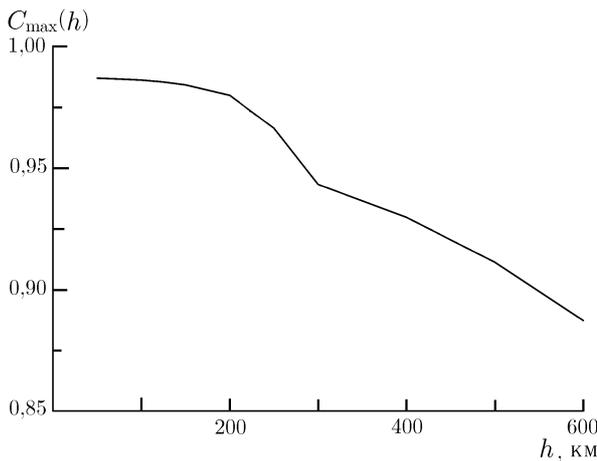


Рис. 7

Контраст $C(V_h, \alpha)$ интерференционной картины рассчитывался при разной высоте подыоносферной точки h . Подыоносферные точки определялись как точки пересечения луча приёмник GPS—ИСЗ со сферой, имеющей радиус $(6370 + h)$ км. На рис. 7 представлена зависимость значения C_{\max} от h для интервала 15:25–16:10 UT. Как видно, C_{\max} имеет высокое и почти постоянное значение (около 0,98) для высот $50 \div 200$ км, затем наблюдается небольшое уменьшение C_{\max} с высотой. При $h = 600$ км C_{\max} достигает 0,88. Исходя из того соображения, что функция C_{\max} показывает наибольшее максимальное значение при правильном выборе высоты h подыоносферных точек, можно предположить, что ПВП в данном случае распространяется на высоте не более 200 км.

На рис. 8 представлен график функции $C(V_h, \alpha)$ в виде линий равного уровня в полярной системе координат, рассчитанный для интервала времени 15:25–16:10 UT. Максимальное значение $C_{\max} = 0,98$ при $h = 100$ км соответствует модулю горизонтальной скорости $V_h \approx (155 \pm 5)$ м/с и азимуту направления распространения $\alpha \approx 150^\circ \pm 5^\circ$.

Эти данные с хорошей для геофизических измерений точностью ($10 \div 20$ м/с и 10° соответственно) согласуются с результатами применения метода SADM-GPS.

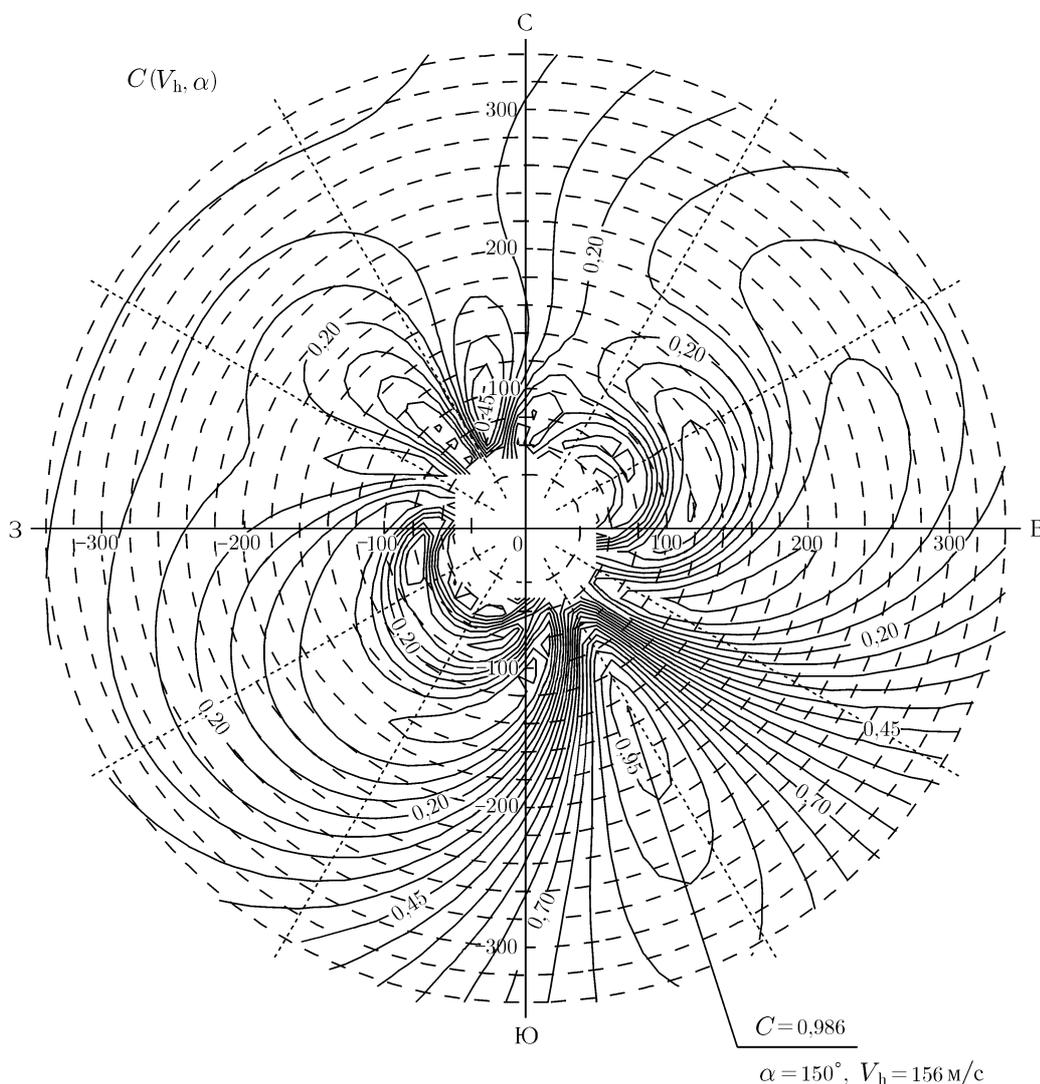


Рис. 8

7. ЗАКЛЮЧЕНИЕ

В работе получены следующие основные результаты:

1) Во временной области ПВП представляют собой квазипериодические колебания ПЭС длительностью около 1 часа с периодом колебаний в диапазоне $10 \div 20$ мин и амплитудой, превышающей амплитуду фоновых флуктуаций ПЭС, как минимум, на порядок.

2) Параметры ПВП, зарегистрированных 18 октября 2001 года в Калифорнии, определялись двумя различными методами: SADM-GPS и SOPHASE. Оба метода дали почти совпадающие результаты. Скорость и азимут перемещения ПВП, рассчитанные методом SOPHASE ($V_h = 160$ м/с, $\alpha = 150^\circ$), с хорошей для геофизических измерений точностью ($10 \div 20$ м/с и 10° соответственно) согласуются с результатами применения метода SADM-GPS ($V_h = 180$ м/с, $\alpha = 140^\circ$).

3) Скорость и направление перемещения ПВП соответствуют аналогичным параметрам среднеширотных среднесаботных перемещающихся ионосферных возмущений, полученным ранее при анализе фазовых характеристик КВ радиосигналов [9, 10, 21, 26–29], а также сигналов навигационных

ИСЗ первого поколения [30], геостационарных ИСЗ [19, 31] и дискретных космических радиоисточников [16, 17, 32].

4) Регистрируемые ПВП не являются следствием условий регистрации, а представляют собой квазимонохроматические волны в атмосфере, которые благодаря дисперсионным свойствам среды распространения приобретают форму волновых пакетов.

Представление о дисперсионных свойствах среды как о наиболее вероятной причине образования одиночных волновых пакетов в вариациях ПЭС согласуются с результатами работ [4, 21, 22]. Автор [22], рассматривая авроральный электроджет как источник ПИВ, показал, что после распространения через атмосферу в F -область отражённые от Земли волны приобретают свойства волнового пакета. На основе дисперсионного соотношения в работах [4, 21] показано, что ПИВ, связанные с тропосферным струйным течением, появляются в F -области ионосферы в виде волнового пакета с квазимонохроматическими осцилляциями, период которых является функцией угла места волнового вектора при распространении волны от источника к месту наблюдения в F -области.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 01–05–65374 и 00–05–72026), а также Совета по поддержке ведущих научных школ Российской Федерации (грант № 00–15–98509) и INTAS (грант No. 99–1186).

СПИСОК ЛИТЕРАТУРЫ

1. Posmentier E. S., Herrman R. W. // *J. Geophys. Res.* 1971. V. 76. P. 2194.
2. Afraimovich E. L., Palamartchouk K. S., Perevalova N. P. // *J. Atmos. and Solar-Terr. Phys.* 1998. V. 60. P. 1205.
3. Afraimovich E. L., Palamartchouk K. S., Perevalova N. P. // *Adv. of Space Res.* 2000. V. 26, No. 6. P. 1001.
4. Hines C. O. // *Can. J. Phys.* 1960. V. 38. P. 1441.
5. Hines C. O., Reddy C. A. // *J. Geophys. Res.* 1967. V. 72. P. 1015.
6. Hocke K., Schlegel K. // *Ann. Geophys.* 1996. V. 14. P. 917.
7. Oliver W. L., Otsuka Y., Sato M., Takami T., Fukao S. // *J. Geophys. Res.* 1997. V. 102. P. 14449.
8. Davies K., Jones J. E. // *J. Atmos. Terr. Phys.* 1971. V. 33. P. 39.
9. Waldock J. A., Jones T. B. // *J. Atmos. Terr. Phys.* 1987. V. 49. P. 195.
10. Jacobson A. R., Carlos R. C. // *J. Atm. Terr. Phys.* 1991. V. 53. P. 53.
11. Yakovets A. F., Kaliev M. Z., Vodyannikov V. V. // *J. Atmos. Terr. Phys.* 1999. V. 61. P. 629.
12. Афраймович Э. Л., Воейков С. В., Липко Ю. В. // Труды XX Всероссийской конференции по распространению радиоволн. Нижний Новгород, 2002. С. 86.
13. Hofmann-Wellenhof B., Lichtenegger H., Collins J. *Global Positioning System: Theory and Practice.* New York: Springer-Verlag Wien, 1992. 327 p.
14. Klobuchar J. A. // *IEEE Transactions on Aerospace and Electronics Systems.* 1986. V. 23, No. 3. P. 325.
15. Афраймович Э. Л., Косогоров Е. А., Лесюта О. С., Ушаков И. И. // *Изв. вузов. Радиофизика.* 2001. Т. 44, № 10. С. 828.
16. Mercier C. // *J. Atmos. Terr. Phys.* 1986. V. 48. P. 605.
17. Mercier C. // *Ann. Geophys.* 1996. V. 14. P. 42.
18. Afraimovich E. L. // *Acta Geod. Geophys. Hung.* 1997. V. 32, No. 3–4. P. 461.
19. Афраймович Э. Л., Бойтман О. Н., Жовтый Е. И., Калихман А. Д., Пирог Т. Г. // *Геомагнетизм и аэрономия.* 1997. Т. 37, № 4. С. 86.
20. Afraimovich E. L., Terekhov A. I., Udodov M. Yu., Fridman S. V. // *J. Atmos. Terr. Phys.* 1992. V. 54. P. 1013.

21. Waldock J. A., Jones T. B. // *J. Atmos. Terr. Phys.* 1984. V. 46. P. 217.
22. Francis S. H. // *J. Geophys. Res.* 1974. V. 79. P. 5 245.
23. Montes H. A., Posmentier E. S. // *Geophys. J. R. Astr. Soc.* 1971. V. 26. P. 271.
24. Aushev V. M., Pogoreltsev A. I., Vodyannikov V. V., Wiens R. H., Shepherd G. G. // *Phys. Chem. Earth B.* 2000. V. 25. P. 409.
25. Waters C. L., Menk F. W., Fraser B. J., Ostwald P. M. // *Planetary and Space Science.* 1991. V. 39. P. 569.
26. Yakovets A. F., Drobjev V. I., Litvinov Yu. G. // *J. Atmos. Terr. Phys.* 1995. V. 57. P. 25.
27. Kalikhman A. D. // *J. Atmos. Sol. Terr. Phys.* 1980. V. 42. P. 697.
28. Афраймович Э. Л. Интерференционные методы радиозондирования ионосферы. М.: Наука, 1982. 198 с.
29. Waldock J. A., Jones T. B. // *J. Atmos. Sol. Terr. Phys.* 1986. V. 48. P. 245.
30. Spoelstra T. A. Th. // *J. Atmos. Terr. Phys.* 1992. V. 54. P. 1 185.
31. Jacobson A. R., Carlos R. C., Massey R. S., Wu G. // *J. Geophys. Res.* 1995. V. 100. P. 1 653.
32. Velthoven P. F. J., Mercier C., Kelder H. // *J. Atmos. Terr. Phys.* 1990. V. 52. P. 305.

¹ Институт солнечно-земной физики СО РАН, г. Иркутск,
Россия,

Поступила в редакцию
17 июля 2002 г.

² Институт ионосферы МОН РК, г. Алматы, Казахстан

DYNAMICAL CHARACTERISTICS OF TRAVELING WAVE PACKETS OF TOTAL ELECTRON CONTENT DISTURBANCES

*E. L. Afraimovich, V. V. Vodyannikov, S. V. Voyeikov,
N. P. Perevalova, and A. F. Yakovets*

Using the CPHASE method [1] and the GPS interferometry method for traveling ionospheric disturbances [2, 3], we analyze in detail spatio-temporal properties of traveling wave packets (TWPs) of total electron content (TEC) disturbances. The analysis is performed on the example of the clearest TWP manifestation observed in California, USA, on October 18, 2001 [3], using the GLOBDET technique developed at the Institute of Solar-Terrestrial Physics of the Siberian Branch of RAS for global detection and monitoring of natural and technogenic ionospheric disturbances on the basis of TEC variations retrieved from the global network of GPS receivers. In the time domain, TWPs are quasi-periodic TEC oscillations lasting about 1 h with the period of 10–20 min and amplitudes exceeding those of the background TEC fluctuations by at least an order of magnitude. The velocity and direction of TWP motion are similar to those of mid-latitude mesoscale traveling ionospheric disturbances, as obtained earlier from the analysis of phase parameters of radio signals from HF transmitters, geostationary satellites, and discrete space radio sources.

УДК 550.388.2

ОСОБЕННОСТИ РАСПРОСТРАНЕНИЯ ВЕРХНЕГО ЛУЧА НА ТРАССАХ, ПЕРЕСЕКАЮЩИХ ИОНОСФЕРНЫЙ ПРОВАЛ

В. П. Урядов¹, А. А. Понятов¹, С. В. Розанов², И. Н. Поддельский²

Представлены результаты экспериментальных исследований особенностей распространения верхнего луча на субавроральной (Магадан—Нижний Новгород) и среднеширотной (Хабаровск—Нижний Новгород) трассах ЛЧМ-зондирования, пересекающих ионосферный провал вблизи его северной и южной границ соответственно. В течение 1998–2000 гг. исследовалась сезонно-суточная зависимость частоты появления верхнего луча и его частотный диапазон. Установлено, что на среднеширотной трассе имеет место максимум частоты появления верхнего луча в весенне-летний период; во все сезоны за исключением периода летнего солнцестояния наблюдается максимум частоты появления верхнего луча в вечерние и ночные часы (19:00–03:00 LT). В июне имеет место практически равномерное суточное распределение частоты появления. На субавроральной трассе наблюдается более сглаженная сезонно-суточная зависимость частоты появления верхнего луча. В июне частотные диапазоны верхнего луча на обеих трассах близки. Зимой и весной частотный диапазон верхнего луча на субавроральной трассе больше, чем на среднеширотной трассе. Высказывается предположение, что различные сезонные и суточные зависимости частоты появления верхнего луча, а также его частотные диапазоны на двух трассах связаны с особенностями формирования поля коротковолновых сигналов в условиях воздействия двух факторов: процесса фотоионизации и мелкомасштабного расслоения электронной концентрации в окрестности ионосферного провала, которое способно привести к формированию флуктуационного волновода.

ВВЕДЕНИЕ

Отличительной особенностью F -области ночной субавроральной ионосферы является провал электронной концентрации. Это область пониженной концентрации, примыкающая к южной границе аврорального овала [1]. В исследованиях провала ионизации много внимания уделяется вопросам формирования провала, его динамике, генерации неоднородностей в окрестности провала и установлению их связи с геофизическими явлениями. Провал характеризуется наличием градиентов электронной концентрации, особенно сильно выраженных на его северной границе, где концентрация увеличивается очень быстро с широтой, возрастая в некоторых случаях до 10 раз при изменении широты на 1° . Провал имеет суточную и сезонную зависимости, что отражается в статистике появления эффектов, связанных с неоднородной структурой провала. Согласно радарным измерениям [2, 3] интенсивные дециметровые неоднородности, ответственные за обратное рассеяние коротких волн в субавроральной ионосфере, локализованы вблизи северной границы провала. В этой же области локализованы субкилометровые неоднородности, ответственные за мерцания сигналов ИСЗ [4, 5]. Указанные мерцания слабы или отсутствуют в центре провала [6].

Хотя за последние годы получено много новых результатов, уточняющих и расширяющих существующие представления о механизмах образования ионосферных неоднородностей в окрестности провала, многие вопросы остаются ещё открытыми и требуют дополнительных исследований. Например, неоднородности на субавроральных широтах встречаются как в спокойных, так и в возмущённых условиях, что делает статистическое моделирование неоднородностей F -слоя на этих широтах весьма трудной задачей. Поэтому использование новых методов диагностики, проведение исследований неоднородной структуры провала в различных геофизических условиях, набор статистических данных в сочетании с моделированием имеют важное значение для понимания физических процессов формирования тонкой структуры субавроральной ионосферы и создания прогностической модели провала.

Новые возможности для изучения динамики ионосферного провала и его тонкой структуры открылись благодаря организации российской сети трасс наклонного зондирования, пересекающих субавроральную и среднеширотную ионосферу в азиатском долготном секторе. Радиолинии оснащены современной аппаратурой с высокой степенью автоматизации регистрации и обработки ионограмм при зондировании сигналом с линейной частотной модуляцией (ЛЧМ).

Для исследования неоднородной структуры ионосферы особый интерес представляет анализ характеристик верхнего луча (моды Педерсена) на трассах, пересекающих провал, поскольку мода Педерсена является чувствительным индикатором магнитосферно-ионосферных возмущений, во время которых увеличивается интенсивность неоднородностей [7, 8].

В работе представлены результаты измерений характеристик верхнего луча на двух протяжённых трассах российской сети ЛЧМ-иозондов, пересекающих ионосферный провал в азиатском долготном секторе.

1. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Исследования проводились на субавроральной (Магадан—Нижний Новгород) и среднеширотной (Хабаровск—Нижний Новгород) трассах в ноябре и декабре 1998 г., феврале, марте, мае и декабре 1999 г., апреле, июне и октябре 2000 г. Трасса Магадан—Нижний Новгород проходит вблизи северной границы провала, а трасса Хабаровск—Нижний Новгород — вблизи южной границы [9]. Такая геометрия эксперимента при одновременных измерениях характеристик КВ сигналов позволяет исследовать динамику границ провала и особенности проявления неоднородной структуры субавроральной ионосферы в окрестностях северной и южной стенок провала.

Зондирование осуществлялось непрерывным ЛЧМ сигналом в диапазоне частот $4 \div 30$ МГц. Скорость изменения частоты составляла 100 кГц/с. Передатчики работали круглосуточно с интервалом 15 мин. Измерялись дистанционно-частотные и амплитудно-частотные характеристики. Измерения проходили на фоне умеренной магнитной активности, индекс K_p изменялся в диапазоне от 1 до 4. Этих данных недостаточно, чтобы выделить зависимость характеристик КВ сигнала от магнитной активности, поэтому при анализе результатов наблюдений данные для различных значений K_p были объединены. Анализировался сезонно-суточный ход частоты появления верхнего луча различных мод сигнала и его частотный диапазон $\Delta f_{\text{п}} = \text{МНЧ}_{\text{п}} - \text{ННЧ}_{\text{п}}$, где $\text{МНЧ}_{\text{п}}$ и $\text{ННЧ}_{\text{п}}$ — максимальная наблюдаемая и наименьшая наблюдаемая частоты моды Педерсена.

Рис. 1а и б иллюстрируют сезонный ход появляемости верхнего луча (моды Педерсена) на трассах Магадан—Нижний Новгород и Хабаровск—Нижний Новгород соответственно. Видно некоторое отличие частоты появления моды Педерсена на этих трассах. Так, на среднеширотной трассе наибольшая частота появления имеет место в весенне-летний период (апрель, май, июнь). На субавроральной трассе частота появления в целом несколько выше, чем на среднеширотной трассе, и имеет более равномерное распределение.

Рис. 2а–е иллюстрируют суточную зависимость частоты появления моды Педерсена на двух трассах в различные сезоны. Для увеличения статистики наблюдательного материала месяцы условно разделены на «зимние» (октябрь, ноябрь, декабрь, февраль, март; рис. 2а, з) и весенние (апрель, май; рис. 2б, д). Данные за июнь приведены отдельно (рис. 2в, е), поскольку они схожи для обеих трасс. Частота появления построена в зависимости от местного времени суток, соответствующего средней точке трасс и отличающегося на три часа от времени суток в точке приёма ($LT = MSK + 3$ ч). Это время соответствует меридиану, где обе трассы наиболее глубоко заходят в область высоких широт. Из рис. 2 видно, что имеется суточная зависимость вероятности появления моды Педерсена, причём на субавроральной и среднеширотной трассах она выражена по-разному для различных сезонов (кроме июня). На трассе Хабаровск—Нижний Новгород в зимние и весенние месяцы наблюдается максимум

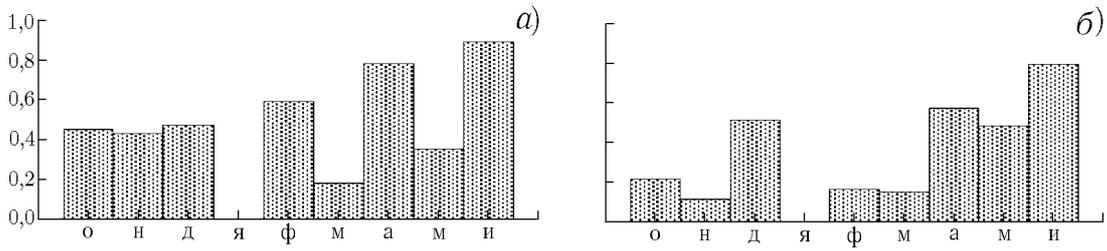


Рис. 1. Сезонные вариации частоты появления верхнего луча на субавроральной (а) и среднеширотной (б) трассах. Обозначения на горизонтальной оси соответствуют месяцам, в которые проводились наблюдения (с октября (о) по июнь (и))

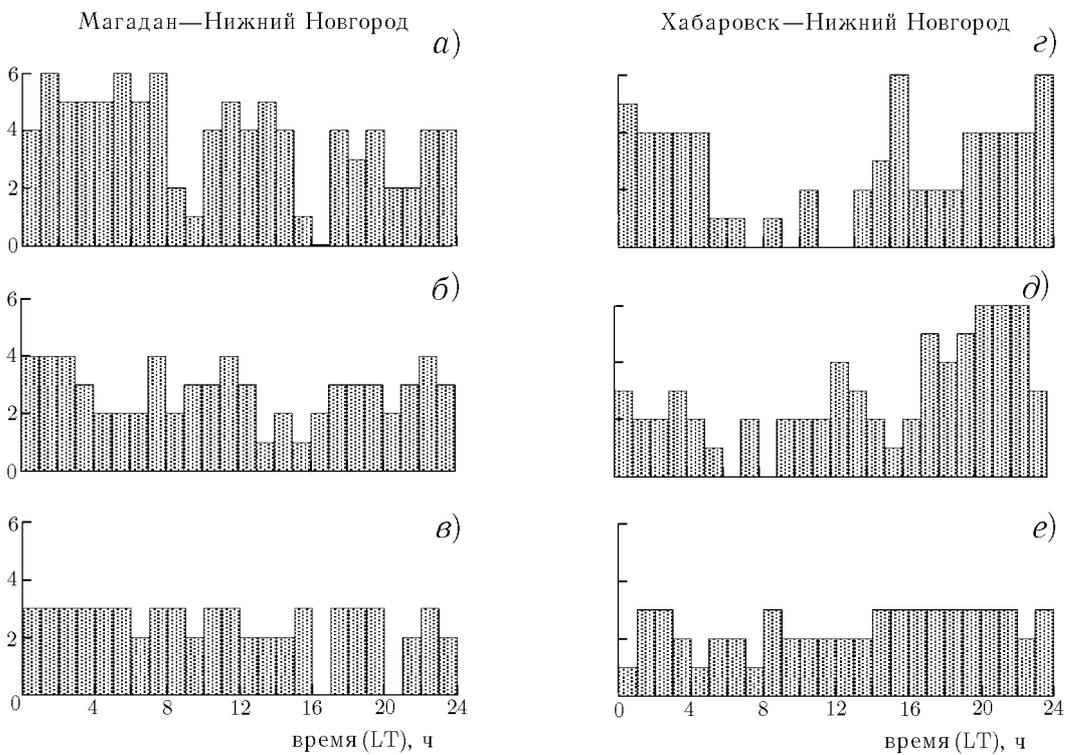


Рис. 2. Суточные вариации частоты появления верхнего луча на субавроральной (а, б, в) и среднеширотной (г, д, е) трассах

частоты появления верхнего луча в вечерние и ночные часы 19:00–03:00 ЛТ. В июне наблюдается практически равномерное суточное распределение частоты появления моды Педерсена (см. рис. 2е). На субавроральной трассе суточные зависимости частоты появления в различные сезоны более равномерны.

На рис. 3а–е представлены гистограммы распределения частотного диапазона $\Delta f_{\text{п}}$ верхнего луча для различных сезонов для трасс Магадан–Нижний Новгород и Хабаровск–Нижний Новгород. Согласно полученным данным в июне средние частотные диапазоны верхнего луча на обеих трассах близки: $\langle \Delta f_{\text{п}} \rangle \approx 2,05$ МГц на среднеширотной трассе и $\langle \Delta f_{\text{п}} \rangle \approx 1,95$ МГц на субавроральной трассе. Зимой и весной частотный диапазон верхнего луча на субавроральной трассе больше, чем на среднеширотной.

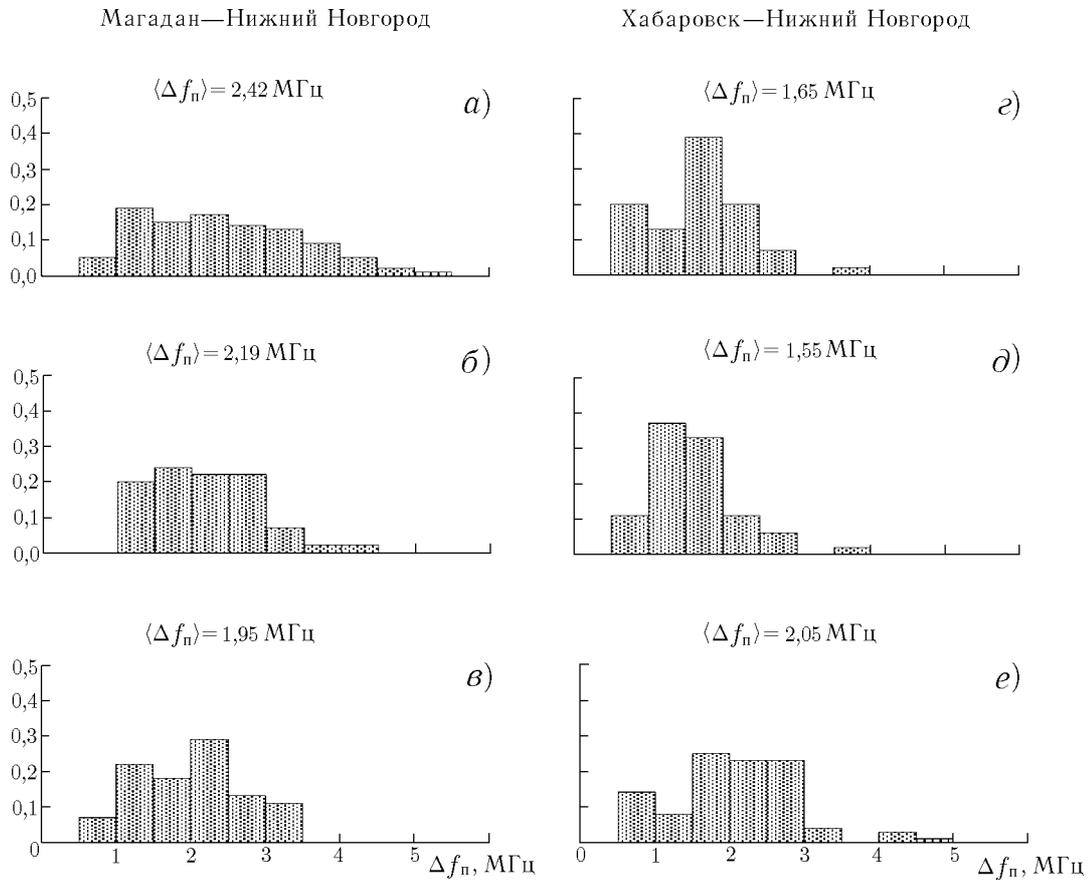


Рис. 3. Гистограммы распределения частотного диапазона верхнего луча $\Delta f_n = \text{МНЧ}_n - \text{ННЧ}_n$ на субавроральной (а, б, в) и среднеширотной (з, д, е) трассах зимой (а, з), весной (б, д) и в июне (в, е)

2. ОБСУЖДЕНИЕ

Различные сезонные и суточные зависимости частоты появления моды Педерсена на среднеширотной и субавроральной трассах, на наш взгляд, связаны с особенностями формирования поля КВ сигналов с учётом ионизации ионосферы и наличия развитой неоднородной структуры высокоширотной ионосферы.

Тот факт, что летом суточные зависимости частоты появления моды Педерсена практически одинаковы на обеих трассах и равномерны, по-видимому, свидетельствует об определяющем влиянии процессов фотоионизации на формирование ионосферных слоёв и о существенном влиянии слоя F_1 на распространение верхнего луча. Дело в том, что наличие слоя F_1 (более чётко выраженного в летний период) приводит к уплощению высотного профиля электронной концентрации, а этот фактор способствует уменьшению утечки радиоволн, распространяющихся верхним лучом, посредством стандартного скользящего механизма [10].

В другие сезоны вероятность появления верхнего луча максимальна в вечерние и ночные часы, причём более чётко выраженная зависимость частоты появления от времени суток имеет место для среднеширотной трассы. В вечерние и ночные часы начинает проявляться неоднородная структура ионосферы. В это время наряду со скользящим механизмом распространения верхнего луча заметный вклад в частоту появления моды Педерсена может вносить механизм квазиволнового распро-

странения, связанный с образованием в окрестности максимума F -слоя флуктуационного волновода, формируемого развитой неоднородной структурой электронной концентрации. При этом, как показано в [8], определяющую роль в каналировании верхнего луча играет мелкомасштабная стратификация ионосферы вдоль трассы распространения. Относительную роль этого механизма на двух трассах можно проследить, анализируя суточные зависимости частоты появления и гистограммы распределения частотного диапазона моды Педерсена. Как уже отмечалось, для зимних месяцев, когда влияние слоя F_1 на распространение моды Педерсена посредством скользящего механизма сказывается в меньшей степени, частотный диапазон моды Педерсена на субавроральной трассе больше, чем на среднеширотной. Это указывает на более развитую стратификацию ионосферы в окрестности северной границы провала (вблизи которой проходит субавроральная трасса), способную формировать флуктуационный волновод и каналировать радиоволны в окрестности максимума слоя.

О более развитой ионосферной турбулентности в окрестности северной границы провала свидетельствуют и результаты измерений диффузности сигналов на двух трассах [9]. Это подтверждается также сравнением суточной зависимости частоты появления моды Педерсена на двух трассах: на субавроральной трассе суточная зависимость частоты появления более равномерная. Здесь следует отметить, что по данным о мерцаниях сигналов спутников «NN S S» с ростом широты увеличивается временной интервал появления мерцаний [11], т. е. на более высоких широтах неоднородности, ответственные за мерцания (с масштабами порядка километра), существуют большее время в течение суток. На среднеширотной трассе в суточной зависимости частоты появления имеется максимум в предполуденные и ночные часы (за исключением июня), когда согласно многочисленным наблюдениям (см., например, работу [1] и библиографию в ней) провал смещается на более низкие широты и часть среднеширотной трассы попадает в область провала с неоднородностями, способными формировать флуктуационный волновод.

Отметим, что на среднеширотной трассе в весенние месяцы (см. рис. 2*d*) наблюдается некоторая асимметрия в суточной зависимости частоты появления луча Педерсена. Как видно из рис. 2*d*, частота появления выше в послеполуденном секторе до середины ночи и меньше в ранние утренние часы. Аналогичная зависимость имеет место в поведении неоднородностей, ответственных за авроральные мерцания сигналов ИСЗ [12], и в прямых зондовых измерениях флуктуаций электронной концентрации [13]. Как отмечается в [14], этот эффект, возможно, связан с уменьшением широтного градиента концентрации плазмы в послеполуденном секторе и ослаблением механизма генерации ионосферных неоднородностей, связанного с градиентно-дрейфовой неустойчивостью.

Совокупность полученных результатов позволяет сделать вывод, что в период умеренной магнитной активности летом на среднеширотной трассе, когда преобладающую роль играют процессы фотоионизации, превалирует стандартный скользящий механизм распространения моды Педерсена в окрестности максимума F -слоя. Этот же механизм, по-видимому, даёт основной вклад в вероятность появления моды Педерсена летом и на субавроральной трассе. Зимой и весной вероятность появления моды Педерсена на субавроральной трассе в значительной мере контролируется развитой неоднородной структурой высокоширотной ионосферы, способной формировать флуктуационный волновод и каналировать в нём радиоволны на значительное расстояние. Этот же механизм каналирования становится достаточно эффективным в вечерние и ночные часы зимой, а также весной и на среднеширотной трассе, когда она попадает в область южной границы провала, которая характеризуется менее развитой турбулентностью, чем окрестность северной границы.

Таким образом, сравнение характеристик моды Педерсена (весьма чувствительной к ионосферным неоднородностям) при одновременном наблюдении на субавроральной и среднеширотной трассах в широкой полосе частот зондирования позволяет контролировать динамику движения провала и оценивать пространственно-временное распределение неоднородной структуры электронной концентрации в окрестности провала.

ВЫВОДЫ

Основные результаты работы состоят в следующем.

Исследованы сезонно-суточные зависимости вероятности появления верхнего луча и его частотный диапазон на субавроральной (Магадан—Нижний Новгород) и среднеширотной (Хабаровск—Нижний Новгород) трассах ЛЧМ-зондирования, проходящих вблизи северной и южной границ ионосферного провала. Установлено, что на среднеширотной трассе максимум частоты появления верхнего луча наблюдается в весенне-летний период. На среднеширотной трассе максимум частоты появления верхнего луча наблюдается в вечерние и ночные часы (19:00—03:00 LT) во все сезоны за исключением периода летнего солнцестояния. В июне наблюдается практически равномерное суточное распределение частоты появления верхнего луча. На субавроральной трассе наблюдается более равномерная сезонно-суточная зависимость частоты появления верхнего луча. В июне частотные диапазоны верхнего луча на обеих трассах близки, зимой и весной частотный диапазон верхнего луча на субавроральной трассе больше, чем на среднеширотной трассе.

Выявленные особенности поведения верхнего луча на среднеширотной и субавроральной трассах могут быть объяснены при учёте двух факторов, оказывающих существенное влияние на формирование поля КВ сигналов: процесса фотоионизации и возможности каналирования верхнего луча во флуктуационном волноводе, который образуется вследствие мелкомасштабного расслоения ионосферы в окрестности провала, более чётко выраженного вблизи его северной границы.

Работа выполнена при поддержке РФФИ (грант № 02–05–64383).

СПИСОК ЛИТЕРАТУРЫ

1. Moffet R. J., Quegan S. // J. Atm. Terr. Phys. 1983. V. 45. P. 315.
2. Moller H. G., Tauriainen A. // J. Atm. Terr. Phys. 1975. V. 37. P. 161.
3. Ruohoniemi J. M., Greenwald R. A., Villain J. P. et al. // J. Geophys. Res. A. 1988. V. 93, No. 11. P. 12871.
4. Weber E. J., Tsunoda R. T., Buchau J. et al. // J. Geophys. Res. 1985. V. 90. P. 6497.
5. Basu S., Basu S., Senior C. et al. // Geophys. Res. Lett. 1986. V. 13. P. 101.
6. Fremouw E. J., Lansinger J. M. // J. Geophys. Res. 1981. V. 86. P. 10087.
7. Иванов В. А., Рябова Н. В., Урядов В. П., Шумаев В. В. // Геомагнетизм и аэрономия. 1995. Т. 35, № 5. С. 131.
8. Erukhimov L. M., Uryadov V. P., Cherkashin Yu. N. et al. // Waves in Random Media. 1997. V. 7, No. 4. P. 531.
9. Урядов В. П., Понятов А. А., Розанов С. В. и др. // Изв. вузов. Радиофизика. 2001. Т. 44, № 3. С. 255.
10. Сажин В. И., Тинин М. В. // Геомагнетизм и аэрономия. 1975. Т. 15. С. 748.
11. Pryse S. E., Kersley L., Russell C. D. // Radio Sci. 1991. V. 26, No. 4. P. 1105.
12. Basu S., Aarons J. // Radio Sci. 1980. V. 15, No. 1. P. 59.
13. Clark D. H., Raitt W. J. // Planet. Space Sci. 1975. V. 23. P. 1623.
14. Tsunoda R. T. // Rev. Geophys. 1988. V. 26, No. 4. P. 719.

¹ Научно-исследовательский радиофизический институт,
г. Нижний Новгород,

Поступила в редакцию
18 апреля 2002 г.

² Институт космофизических исследований и распространения
радиоволн ДВО РАН,
пос. Паратунка Камчатской обл., Россия

FEATURES OF HIGH-ANGLE RAY PROPAGATION ON PATHS CROSSING THE IONOSPHERIC TROUGH

V. P. Uryadov, A. A. Ponyatov, S. V. Rozanov, and I. N. Poddel'sky

We present the results of experimental study of the features of high-angle ray propagation on subauroral (Magadan — N. Novgorod) and mid-latitude (Khabarovsk — N. Novgorod) chirp-sounding paths crossing the ionospheric trough near its north and south boundaries, respectively. Seasonal and daily occurrence of the high-angle ray and its frequency range are studied in 1998–2000. It is found that the frequency of occurrence of the high-angle ray on the mid-latitude path is maximum during spring and summer. This frequency is also maximum in evening and nighttime hours (19:00–03:00 LT) during all seasons excluding the summer solstice. The daily distribution of the frequency of occurrence in June is almost uniform. The seasonal-daily behavior of the frequency of occurrence of the high-angle ray, observed on the subauroral path, is more smooth. The frequency ranges of high-angle ray on both paths are similar in June, while in winter and spring the frequency range of the high-angle ray on the subauroral path is larger than on the mid-latitude one. We put forward a hypothesis that different seasonal and daily behavior of occurrence of the high-angle ray and frequency range of this ray on two paths are stipulated by features of the formation of HF-signal field under the influence of two factors: photoionization and small-scale stratification of the electron density in the vicinity of the ionospheric trough, which can form a fluctuation waveguide.

УДК 533.951

НЕСИММЕТРИЧНЫЕ СВИСТОВЫЕ ВОЛНЫ, НАПРАВЛЯЕМЫЕ ЦИЛИНДРИЧЕСКИМИ ДАКТАМИ С ПОВЫШЕННОЙ ПЛОТНОСТЬЮ ПЛАЗМЫ

*Т. М. Заборонкова*¹, *А. В. Кудрин*², *М. Ю. Лях*², *Л. Л. Попова*²

Исследуется распространение волн свистового диапазона, поля которых зависят от азимутального угла, в цилиндрических плазменно-волноводных каналах (дактах плотности), ориентированных вдоль внешнего магнитного поля и окружённых однородной фоновой плазмой. Основное внимание уделено дактам с повышенной плотностью плазмы. Показано, что такие дакты при определённых условиях могут поддерживать собственные моды и слабывтекающие несобственные моды. Приведены результаты анализа дисперсионных свойств и структуры полей несимметричных мод свистового диапазона, направляемых цилиндрическими дактами.

ВВЕДЕНИЕ

В последние годы значительное внимание уделяется исследованию возбуждения и распространения электромагнитных волн в плазменных каналах (дактах плотности), расположенных в магнитоактивной плазме [1–5]. Повышенный интерес вызывают волны, частота ω которых принадлежит нижнегибридному (свистовому) диапазону:

$$\omega_{\text{ЛН}} < \omega < \omega_{\text{Н}} \ll \omega_{\text{р}}, \quad (1)$$

где $\omega_{\text{ЛН}}$ — нижняя гибридная частота, $\omega_{\text{Н}}$ и $\omega_{\text{р}}$ — гирочастота и плазменная частота электронов соответственно. Это объясняется, с одной стороны, важной ролью, которую свистовые волны (вистлеры) играют во многих фундаментальных физических процессах в магнитосфере и ионосфере Земли [6, 7]. С другой стороны, использование направляющих свойств дактов плотности (как естественного, так и искусственного происхождения) открывает дополнительные возможности для решения ряда прикладных задач, связанных с ОНЧ диагностикой околоземного космического пространства [8] и разработкой новых методов эффективного возбуждения вистлеров [9, 10] в магнитоактивной плазме. Такие методы представляют особый интерес в связи с постановкой соответствующих натуральных ионосферных [11, 12] и модельных лабораторных [2–5, 13] экспериментов, а также для применения в различных плазменных установках (например, в геликонных источниках плазмы [14]). Для некоторых из указанных выше приложений весьма перспективным представляется использование несимметричных свистовых волн, направляемых цилиндрическими дактами плотности.

Следует отметить, что большинство теоретических работ, посвящённых распространению несимметричных свистовых волн в замагниченных цилиндрических плазменных волноводах, относится к случаю ограниченной плазмы, когда плазменный цилиндр окружён металлическим экраном или диэлектриком [15–20]. Применительно к цилиндрическим дактам плотности, расположенным в неограниченной магнитоактивной плазме, достаточно подробно рассматривались лишь свойства азимутально-симметричных волн [1, 5, 21, 22]. Исследованию же несимметричных мод, поддерживаемых такими дактами, посвящено сравнительно немного работ (см., например, [23, 24]). В них обсуждаются преимущественно слабонеоднородные каналы, параметры которых отвечают условиям земной магнитосферы. Каналирование несимметричных свистовых волн в дактах плотности, параметры которых соответствуют условиям упоминавшихся выше экспериментов, рассматривалось лишь в [1] для некоторых простейших частных случаев.

В настоящей работе проведено систематическое исследование дисперсионных свойств и структуры полей собственных (локализованных) и несобственных вытекающих (квазилокализованных) несимметричных мод свистового диапазона (1), направляемых цилиндрическими дактами с повышенной относительно фона плотностью плазмы. Такие структуры могут возникать в замагниченной плазме в результате нелинейного взаимодействия поля источника с окружающей (фоновой) плазменной средой [2, 5, 11–13] и представляют значительный интерес для управления электродинамическими характеристиками источника [1, 25, 26].

Данное исследование является непосредственным продолжением работы [22], в которой изучались азимутально-симметричные свистовые моды, поддерживаемые дактами с повышенной плотностью. Так же, как и в [22], мы ограничимся рассмотрением однородного дакта с резкой границей, чтобы выяснить наиболее важные особенности каналированного распространения несимметричных мод. Оговоримся сразу же, что, хотя при некоторых условиях учёт столкновений может иметь принципиальное значение для свистовых мод [1, 27], мы с целью упрощения анализа остановимся лишь на случае бесстолкновительной плазмы.

1. ОСНОВНЫЕ СООТНОШЕНИЯ

Рассмотрим однородный цилиндрический дакт плотности, ориентированный вдоль внешнего магнитного поля $\mathbf{B}_0 = B_0 \mathbf{z}_0$ и окружённый однородной фоновой плазмой. Плотность плазмы при наличии такого дакта может быть представлена в виде

$$N(\rho) = N_a + (\tilde{N} - N_a) [1 - U(\rho - a)], \quad (2)$$

где a — радиус дакта, U — единичная функция Хевисайда, ρ , ϕ и z — цилиндрические координаты, \tilde{N} и N_a — плотность плазмы внутри и вне дакта соответственно (предполагается, что $\tilde{N} > N_a$).

Напомним, что тензор диэлектрической проницаемости холодной магнитоактивной плазмы записывается следующим образом:

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon & -ig & 0 \\ ig & \varepsilon & 0 \\ 0 & 0 & \eta \end{pmatrix}. \quad (3)$$

Величины ε , g , η определяются параметрами среды и приведены, например, в [28]. Далее компоненты тензора (3) внутри и вне дакта будем обозначать $\tilde{\varepsilon}$, \tilde{g} , $\tilde{\eta}$ и ε_a , g_a , η_a соответственно.

Поля мод, направляемых дактом, представим в виде

$$\mathbf{E}(\rho, \phi, z) = \mathbf{E}(\rho) \exp[-im\phi - ik_0pz], \quad \mathbf{H}(\rho, \phi, z) = \mathbf{H}(\rho) \exp[-im\phi - ik_0pz], \quad (4)$$

где m — азимутальный индекс ($m = 0, \pm 1, \pm 2, \dots$), p — нормированная постоянная распространения, $k_0 = \omega/c$ — волновое число в свободном пространстве. В выражениях для поля опущен временной множитель $\exp(i\omega t)$.

Векторные функции $\mathbf{E}(\rho)$, $\mathbf{H}(\rho)$ могут быть описаны с помощью двух скалярных функций $E_z(\rho)$, $H_z(\rho)$, которые являются решениями системы уравнений [1]

$$\begin{aligned} (p^2 - \varepsilon) \hat{L}_m H_z + k_0^2 [g^2 - (p^2 - \varepsilon)^2] H_z &= ipg \hat{L}_m E_z, \\ [g^2 + \varepsilon (p^2 - \varepsilon)] \hat{L}_m E_z + k_0^2 \eta [g^2 - (p^2 - \varepsilon)^2] E_z &= -ipg \hat{L}_m H_z, \end{aligned} \quad (5)$$

где

$$\hat{L}_m = \frac{d^2}{d\rho^2} + \frac{1}{\rho} \frac{d}{d\rho} - \frac{m^2}{\rho^2}.$$

Решая уравнения (5) и используя связь поперечных компонент поля $E_\rho(\rho)$, $E_\phi(\rho)$, $H_\rho(\rho)$, $H_\phi(\rho)$ с продольными компонентами $E_z(\rho)$, $H_z(\rho)$, приходим к следующим выражениям:

а) при $\rho < a$

$$\begin{aligned}
 E_\rho &= - \sum_{k=1}^2 B_k \left[\frac{\tilde{n}_k p + \tilde{g}}{\tilde{\varepsilon}} J_{m+1}(k_0 \tilde{q}_k \rho) + \tilde{\alpha}_k m \frac{J_m(k_0 \tilde{q}_k \rho)}{k_0 \tilde{q}_k \rho} \right], \\
 E_\phi &= i \sum_{k=1}^2 B_k \left[J_{m+1}(k_0 \tilde{q}_k \rho) + \tilde{\alpha}_k m \frac{J_m(k_0 \tilde{q}_k \rho)}{k_0 \tilde{q}_k \rho} \right], \\
 E_z &= \frac{i}{\tilde{\eta}} \sum_{k=1}^2 B_k \tilde{n}_k \tilde{q}_k J_m(k_0 \tilde{q}_k \rho), \\
 H_\rho &= -i \sum_{k=1}^2 B_k \left[p J_{m+1}(k_0 \tilde{q}_k \rho) - \tilde{n}_k \tilde{\beta}_k m \frac{J_m(k_0 \tilde{q}_k \rho)}{k_0 \tilde{q}_k \rho} \right], \\
 H_\phi &= - \sum_{k=1}^2 B_k \tilde{n}_k \left[J_{m+1}(k_0 \tilde{q}_k \rho) - \tilde{\beta}_k m \frac{J_m(k_0 \tilde{q}_k \rho)}{k_0 \tilde{q}_k \rho} \right], \\
 H_z &= - \sum_{k=1}^2 B_k \tilde{q}_k J_m(k_0 \tilde{q}_k \rho);
 \end{aligned} \tag{6}$$

б) при $\rho > a$

$$\begin{aligned}
 E_\rho &= - \sum_{k=1}^2 C_k \left[\frac{n_k p + g_a}{\varepsilon_a} K_{m+1}(k_0 s_k \rho) + \alpha_k m \frac{K_m(k_0 s_k \rho)}{k_0 s_k \rho} \right], \\
 E_\phi &= i \sum_{k=1}^2 C_k \left[K_{m+1}(k_0 s_k \rho) + \alpha_k m \frac{K_m(k_0 s_k \rho)}{k_0 s_k \rho} \right], \\
 E_z &= - \frac{i}{\eta_a} \sum_{k=1}^2 C_k n_k s_k K_m(k_0 s_k \rho), \\
 H_\rho &= -i \sum_{k=1}^2 C_k \left[p K_{m+1}(k_0 s_k \rho) - n_k \beta_k m \frac{K_m(k_0 s_k \rho)}{k_0 s_k \rho} \right], \\
 H_\phi &= - \sum_{k=1}^2 C_k n_k \left[K_{m+1}(k_0 s_k \rho) - \beta_k m \frac{K_m(k_0 s_k \rho)}{k_0 s_k \rho} \right], \\
 H_z &= \sum_{k=1}^2 C_k s_k K_m(k_0 s_k \rho).
 \end{aligned} \tag{7}$$

Здесь

$$\begin{aligned}
 \tilde{n}_k &= n_k(p, \tilde{N}), & \tilde{\alpha}_k &= \alpha_k(p, \tilde{N}), & \tilde{\beta}_k &= \beta_k(p, \tilde{N}), & \tilde{q}_k^2 &= q_k^2(p, \tilde{N}), \\
 n_k &= n_k(p, N_a), & \alpha_k &= \alpha_k(p, N_a), & \beta_k &= \beta_k(p, N_a), & s_k^2 &= -q_k^2(p, N_a),
 \end{aligned} \tag{8}$$

J_m — функции Бесселя порядка m , K_m — функции Макдональда порядка m , B_k и C_k — некоторые константы. Функции, входящие в формулы (8), определяются следующим образом:

$$n_k(p, N) = -\frac{\varepsilon}{pg} \left(p^2 + q_k^2(p, N) + \frac{g^2}{\varepsilon} - \varepsilon \right),$$

$$\alpha_k(p, N) = q_k^2(p, N) \frac{1 + \eta^{-1} p n_k(p, N)}{p^2 - \mathcal{P}^2(N)},$$

$$\beta_k(p, N) = -q_k^2(p, N) \frac{p m_k^{-1}(p, N) + \eta^{-1} \mathcal{P}^2(N)}{p^2 - \mathcal{P}^2(N)},$$

$$q_k(p, N) = \frac{1}{\sqrt{2}} \left\{ \varepsilon - \frac{g^2}{\varepsilon} + \eta - \left(\frac{\eta}{\varepsilon} + 1 \right) p^2 - \left(\frac{\eta}{\varepsilon} - 1 \right) (-1)^k [(p^2 - \mathcal{P}_b^2(N)) (p^2 - \mathcal{P}_c^2(N))]^{1/2} \right\}^{1/2}, \quad (9)$$

где

$$\mathcal{P}(N) = (\varepsilon - g)^{1/2},$$

$$\mathcal{P}_{b,c}(N) = \left\{ \varepsilon - (\eta + \varepsilon) \frac{g^2}{(\eta - \varepsilon)^2} + \frac{2\chi_{b,c}}{(\eta - \varepsilon)^2} [\varepsilon g^2 \eta (g^2 - (\eta - \varepsilon)^2)]^{1/2} \right\}^{1/2},$$

$$\chi_b = -\chi_c = -1. \quad (10)$$

При $N = \tilde{N}$ и $N = N_a$ в правые части формул (9), (10) в качестве компонент тензора диэлектрической проницаемости подставляются величины $\tilde{\varepsilon}$, \tilde{g} , $\tilde{\eta}$ и ε_a , g_a , η_a соответственно. При выводе формул для α_k , β_k в (9) учтено тождество $n_1 n_2 = -\eta$.

Следует отметить, что выражения (6)–(10), соответствующие магнитоактивной плазме с тензором диэлектрической проницаемости общего вида (3), справедливы для дактов как с повышенной, так и с пониженной плотностью плазмы и ни в коей мере не ограничиваются рамками частотного диапазона (1). Обратим внимание на то, что наличие членов, отвечающих двум ветвям функций \tilde{q}_k^2 , s_k^2 , в выражениях (6), (7) для поля связано с анизотропией магнитоактивной плазменной среды, что кардинально отличает полученные выражения от соответствующих формул для случая открытых изотропных цилиндрических волноводов [29].

Из условия непрерывности тангенциальных компонент поля при $\rho = a$ можно получить дисперсионное уравнение, позволяющее определить постоянные распространения мод дакта. Дисперсионное уравнение, имеющее весьма сложный вид, приведено в Приложении в форме, удобной для последующего аналитического рассмотрения. Прежде чем перейти к отысканию решений дисперсионного уравнения, введём обозначения

$$\tilde{Q}_k = k_0 \tilde{q}_k a, \quad S_k = k_0 s_k a, \quad k = 1, 2; \quad (11)$$

$$\tilde{P} = \mathcal{P}(\tilde{N}), \quad P = \mathcal{P}(N_a), \quad \tilde{P}_{b,c} = \mathcal{P}_{b,c}(\tilde{N}), \quad P_{b,c} = \mathcal{P}_{b,c}(N_a). \quad (12)$$

Можно показать, что в диапазоне частот (1) при отсутствии столкновений в плазме величины $\mathcal{P}(N)$, $\mathcal{P}_b(N)$, $\mathcal{P}_c(N)$ являются действительными и удовлетворяют следующим неравенствам:

$$\mathcal{P}_b(N) < \mathcal{P}_c(N) \leq \mathcal{P}(N). \quad (13)$$

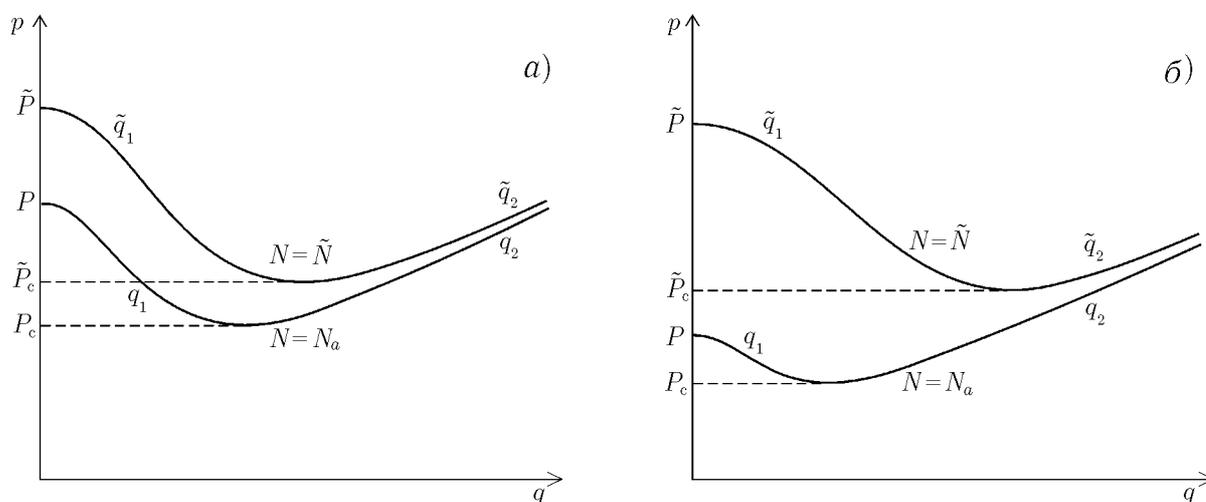


Рис. 1. Поверхности показателя преломления необыкновенной волны в частотном интервале (14) для двух значений плотности плазмы \tilde{N} и N_a ($\tilde{N} > N_a$): а — $\tilde{P}_c < P$; б — $P < \tilde{P}_c$

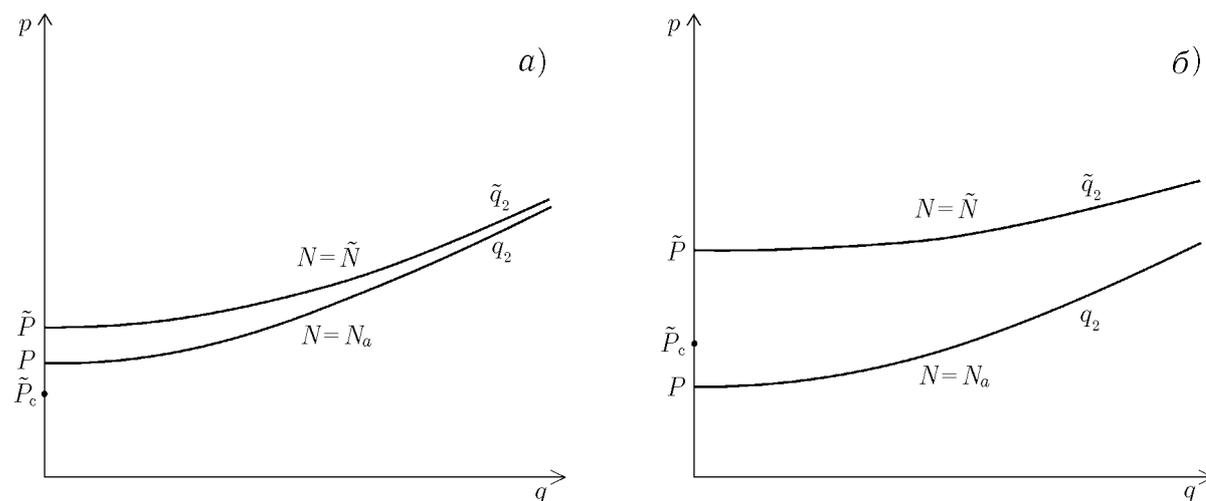


Рис. 2. Поверхности показателя преломления необыкновенной волны в частотном интервале (15) для двух значений плотности плазмы \tilde{N} и N_a ($\tilde{N} > N_a$): а — $\tilde{P}_c < P$; б — $P < \tilde{P}_c$

Напомним, что в рассматриваемом диапазоне распространяющейся является только одна из двух нормальных волн холодной магнитоактивной плазмы. На рис. 1, 2 показаны поверхности показателя преломления этой волны для двух значений плотности плазмы \tilde{N} и N_a ($\tilde{N} > N_a$) применительно к интервалам частот

$$\omega_{LH} < \omega < \omega_H/2 \ll \omega_p, \tag{14}$$

$$\omega_H/2 < \omega < \omega_H \ll \omega_p, \tag{15}$$

принадлежащим диапазону (1). Данные поверхности описываются соответствующими ветвями функций $q_1(p, N)$, $q_2(p, N)$ [1]. Очевидно, что величины \tilde{P} и P являются постоянными распространения вистлера строго вдоль внешнего магнитного поля в однородной плазме при $N = \tilde{N}$ и $N = N_a$ соответственно. В свою очередь, величины \tilde{P}_c и P_c отвечают в частотном интервале (14) (для тех же значений плотности плазмы) так называемым волнам конической рефракции. Как видно из рис. 1, в указанном частотном интервале существует область значений p , где обе ветви поперечного волнового числа — $q_1(p, N)$ и $q_2(p, N)$ — являются действительными. Напротив, в интервале (15) величина $q_1(p, N)$ не может быть чисто действительной ни при каких действительных значениях p (см. рис. 2).

2. СОБСТВЕННЫЕ МОДЫ

Очевидно, что дакт может поддерживать собственные моды с действительными постоянными распространения p , если существуют такие области значений p , в которых величины s_1 и s_2 , относящиеся к фоновой среде, являются одновременно либо действительными ($s_1^2 > 0$ и $s_2^2 > 0$), либо комплексно-сопряжёнными ($s_1 = s_2^*$). Выбор ветвей s_1 и s_2 для собственных мод осуществляется таким образом, чтобы выполнялись условия $\text{Re } s_1 > 0$ и $\text{Re } s_2 > 0$.

Из анализа выражений для поперечных волновых чисел (см. (9)) следует, что в частотном интервале (14) величины s_1 и s_2 являются действительными в области

$$0 < p < P_b \quad (16)$$

и комплексно-сопряжёнными в области

$$P_b < p < P_c. \quad (17)$$

Эти выводы остаются справедливыми и для частотного интервала (15). Кроме того, в частотном интервале (15) величины s_1 и s_2 оказываются действительными также в области

$$P_c < p < P. \quad (18)$$

Оговоримся сразу же, что исследование дисперсионного уравнения не выявило в области (16) решений, отвечающих собственным модам. Как оказалось, такие решения существуют только в областях (17) и (18). Ниже обсуждается поведение решений $p(\omega)$ в этих областях.

2.1. Дисперсионные свойства собственных мод

Как показало численное исследование дисперсионного уравнения (П1), при фиксированном азимутальном индексе m дакт с повышенной плотностью плазмы может поддерживать в диапазоне частот (1) одну собственную моду с действительной постоянной распространения p_m .

На рис. 3 приведены дисперсионные кривые собственных мод с азимутальными индексами $m = 0, \pm 1$ для следующих значений безразмерных параметров: $\tilde{\omega}_p/\omega_p = 1,778$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 1$ ($\tilde{\omega}_p$ и ω_p — плазменная частота электронов внутри и вне дакта соответственно).¹ На рис. 3 дисперсионные кривые $p_m(\omega)$ собственных мод показаны сплошными линиями, зависимость $P(\omega)$ — штриховой линией (кривая 1), зависимости $P_c(\omega)$ и $P_b(\omega)$ — штрихпунктирными линиями (кривые 2 и 3 соответственно). Для удобства соответствующие кривые построены в разных масштабах для частотных областей $\omega > 0,2\omega_H$ (рис. 3а) и $\omega < 0,2\omega_H$ (рис. 3б). В случае $\omega_{LH} \ll \omega < \omega_H$ величины P_b и P_c слабо зависят от ω и подчиняются неравенству $P_b \ll P_c$, которое выполняется с большим запасом (см. рис. 3). При стремлении частоты ω к нижней гибридной частоте ω_{LH} величина P_c уменьшается, а величина P_b резко возрастает, причём на нижней гибридной частоте данные величины оказываются равными: $P_b(\omega_{LH}) = P_c(\omega_{LH})$.

Из рис. 3 видно, что собственные моды имеют нижнюю частоту отсечки, лежащую в области (14), и верхнюю частоту отсечки, лежащую в области (15). На нижней частоте отсечки выполняется либо условие $p_m = P_b$, либо условие $p_m = P_c$, при которых действительные части комплексных величин s_1 и s_2 обращаются в нуль и мода перестаёт быть локализованной. На верхней частоте отсечки выполняется условие $p_m = P$, при котором величина s_2 , являющаяся действительной в окрестности данной

¹ Здесь и далее значения безразмерных параметров, выбранные для численных расчётов, отвечают типичным условиям модельных лабораторных экспериментов по наблюдению каналированного распространения вистлеров в дактах с повышенной плотностью плазмы [2].

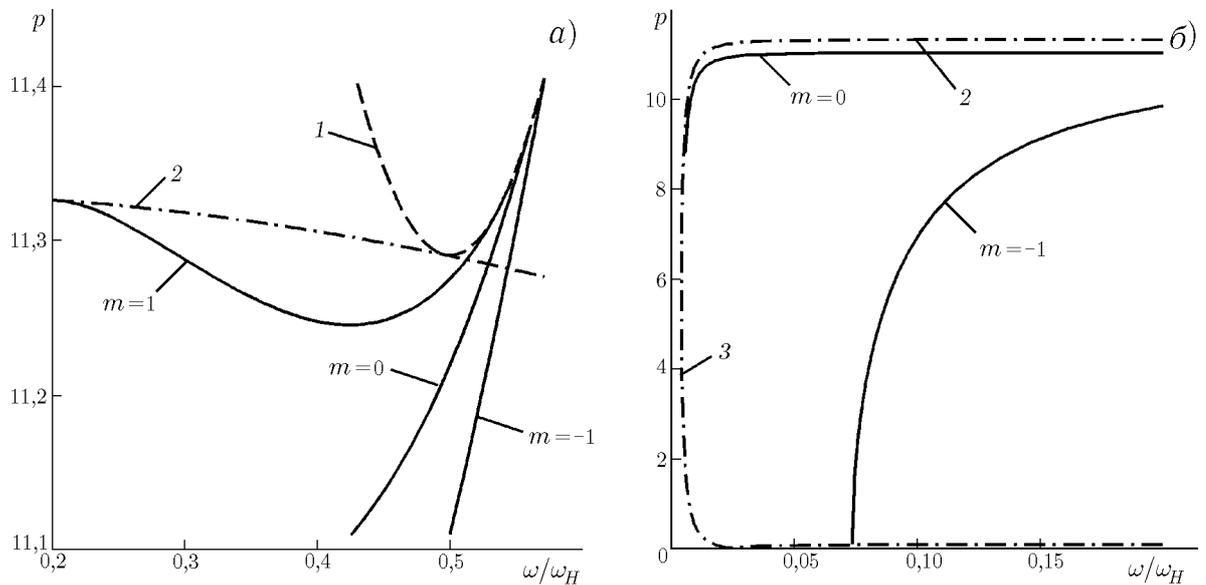


Рис. 3. Дисперсионные кривые собственных мод с азимутальными индексами $m = 0, \pm 1$ и зависимости $P(\omega)$ (кривая 1), $P_c(\omega)$ (кривая 2), $P_b(\omega)$ (кривая 3) при $\tilde{\omega}_p/\omega_p = 1,778$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 1$

частоты (если $P_c < p_m < P$), обращается в нуль, в то время как величина s_1 остаётся действительной и не равной нулю. Следовательно, и в этом случае мода перестаёт быть локализованной. Очевидно, что с удалением от частот отсечки степень локализации поля моды в окружающем пространстве ($\rho > a$) возрастает.

При увеличении плотности плазмы внутри дакта дисперсионные кривые собственных мод сдвигаются в сторону меньших значений p и в предельном случае $\tilde{N}/N_a \rightarrow \infty$ переходят в дисперсионные кривые мод, направляемых идеально проводящим цилиндром в магнитоактивной плазме [30]. Напротив, при уменьшении \tilde{N} постоянные распространения p_m сдвигаются в сторону больших значений. Сдвиг дисперсионных кривых сопровождается изменением частот отсечки и, следовательно, ширины частотного интервала существования мод. В частности, для моды с азимутальным индексом $m = 1$ верхняя и нижняя частоты отсечки при $\tilde{N} \rightarrow N_a$ стремятся к $\omega_H/2$, и ширина интервала существования моды обращается в нуль.

Следует заметить, что поперечные волновые числа \tilde{q}_1 и \tilde{q}_2 , относящиеся к внутренней области дакта, являются либо комплексно-сопряжёнными, либо чисто мнимыми (последнее возможно при $\tilde{P}_c < p_m < P$ в частотном интервале (15)). Поэтому поле является «прижатым» к границе дакта также с её внутренней стороны. Степень локализации поля у границы дакта при $\rho \lesssim a$ возрастает с увеличением плотности плазмы внутри дакта, т. к. в этом случае растёт мнимая часть поперечных волновых чисел \tilde{q}_1, \tilde{q}_2 . Фактически моды данного типа поддерживаются перепадом плотности на границе дакта (заметим, что перепад может быть и плавным [5]). Именно поэтому дисперсионное уравнение рассматриваемых мод при $a \rightarrow \infty$ переходит в дисперсионное уравнение мод, направляемых бесконечной плоской границей между двумя полупространствами с различной плотностью плазмы. Моды, поддерживаемые плоской границей, исследовались в работе [31].

Любопытно отметить, что изученные здесь локализованные моды являются «быстрыми» в том смысле, что их фазовые скорости превосходят (в силу условия $p_m < P$)² фазовую скорость однородной плоской волны, распространяющейся строго вдоль внешнего магнитного поля в однородной

² Как ясно из выполненного рассмотрения, в частотном интервале (14) постоянные распространения мод удовлетворяют даже более сильному неравенству $p_m < P_c$.

плазме с фоновой плотностью N_a . Следует иметь в виду, что данное свойство связано с гиротропией магнитоактивной плазменной среды, окружающей дакт, и не может наблюдаться у собственных мод открытых направляющих систем, находящихся в негиротропной среде. Возможность существования собственных мод с указанным свойством была, по-видимому, впервые показана в работе [30].

2.2. Структура полей собственных мод

Перейдём к рассмотрению структуры полей собственных мод свистового диапазона, направляемых дактом с повышенной плотностью.

На рис. 4 изображены распределения компонент поля собственной моды с азимутальным индексом $m = 1$ по радиальной координате, рассчитанные на основании формул (6) и (7) для $\omega/\omega_H = 0,33$ и тех же значений параметров $\tilde{\omega}_p/\omega_p$, ω_p/ω_H , ω_{LH}/ω_H , ω_{Ha}/c , что и для рис. 3. Постоянная распространения данной моды $p_1 = 11,27$. Представленные зависимости позволяют судить как об относительном вкладе отдельных компонент в поле моды, так и о фазовых соотношениях между ними. Нетрудно убедиться, что в этом случае постоянная распространения лежит в области (17), так что величины s_1 и s_2 , определяющие поведение поля вне дакта, являются комплексными ($s_1 = s_2^*$). Поэтому компоненты поля при $\rho > a$ характеризуются наличием осцилляций, амплитуда которых экспоненциально убывает с увеличением расстояния от дакта.

Для сравнения на рис. 5 представлены распределения компонент поля той же моды при $\omega/\omega_H = 0,515$ (постоянная распространения $p_1 = 11,29$, параметры дакта и плазмы прежние). В этом случае постоянная распространения находится в области (18), где величины s_1 и s_2 являются вещественными, и поле вне дакта экспоненциально спадает с увеличением ρ .

Заметим, что, хотя в обоих рассмотренных случаях поперечные волновые числа \tilde{q}_1 и \tilde{q}_2 , относящиеся к области внутри дакта, являются комплексными, осцилляции поля при $\rho < a$ на рис. 4 и 5 практически не видны. Это объясняется тем, что мнимые части поперечных волновых чисел \tilde{q}_1 и \tilde{q}_2 существенно превышают соответствующие действительные части по абсолютному значению. Поэтому поле собственной моды оказывается сильно «прижатым» изнутри к границе дакта, что приводит к весьма низкой эффективности возбуждения данной моды источниками, расположенными внутри дакта. Это обстоятельство служит, по-видимому, основной причиной того, что в литературе, по сведениям авторов, отсутствуют данные об экспериментальном наблюдении возбуждения и распространения таких собственных мод.

3. НЕСОБСТВЕННЫЕ МОДЫ

Ввиду отмеченных выше свойств собственных мод основной интерес для приложений представляют несобственные (вытекающие) моды, направляемые в свистовом диапазоне дактами с повышенной плотностью. Из-за утечки таких мод в окружающую среду их постоянные распространения являются комплексными даже в случае бесстолкновительной плазмы: $p_{m\nu} = p'_{m\nu} - ip''_{m\nu}$. Здесь m — по-прежнему азимутальный индекс моды ($m = 0, \pm 1, \pm 2, \dots$), ν — порядковый номер моды с фиксированным значением m ($\nu = 1, 2, \dots$).

Как показано в работах [21, 22], посвящённых исследованию дисперсионных свойств азимутально-симметричных мод, несобственные моды в диапазоне (14) являются слабовытекающими ($p'' \ll p'$), если действительные части их постоянных распространения удовлетворяют условию

$$\max(P, \tilde{P}_c) < p' < \tilde{P}. \quad (19)$$

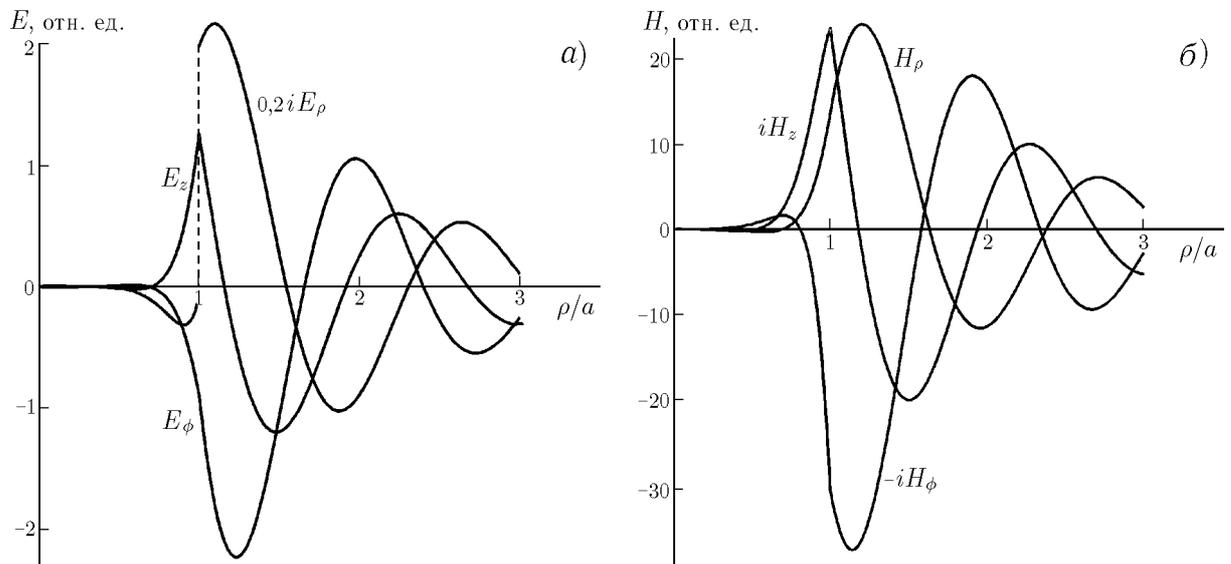


Рис. 4. Распределения компонент поля собственной моды с азимутальным индексом $m = 1$ по поперечной координате при $\omega/\omega_H = 0,33$; параметры $\tilde{\omega}_p/\omega_p, \omega_p/\omega_H, \omega_{LH}/\omega_H, \omega_H a/c$ те же, что и для рис. 3

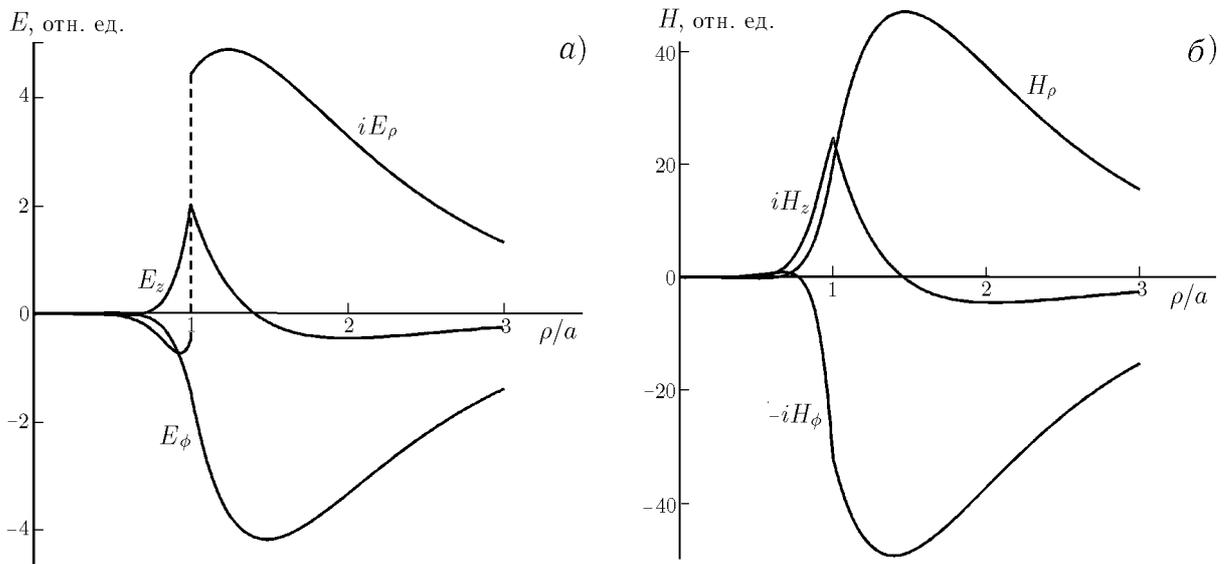


Рис. 5. Распределения компонент поля собственной моды с азимутальным индексом $m = 1$ по поперечной координате при $\omega/\omega_H = 0,515$; параметры $\tilde{\omega}_p/\omega_p, \omega_p/\omega_H, \omega_{LH}/\omega_H, \omega_H a/c$ те же, что и для рис. 3

Заметим, что, хотя несобственные моды могут также существовать в области $\rho' > \tilde{P}$, их утечка в этом случае оказывается весьма значительной, так что несобственные моды в данной области, как правило, не представляют практического интереса. Следует обратить внимание на то, что в случае $\rho' > P$ (ср. с (19)) поле несобственной моды вне дакта содержит как локализованную составляющую, отвечающую ветви s_1 , так и нелокализованную составляющую, отвечающую ветви s_2 . Именно с наличием ветви s_2 , соответствующей вытекающим квазиэлектростатическим волнам, и связана утечка свистовых мод из дактов с повышенной плотностью [21]. Для описания вклада данной ветви в полное поле моды при $\rho > a$ вместо функций Макдональда $K_m(S_2\rho/a)$ удобнее использовать функции Ханкеля первого рода

$H_m^{(1)}(Q_2\rho/a)$, к которым можно перейти с помощью соотношения [32]

$$K_m\left(S_2\frac{\rho}{a}\right) = \frac{\pi}{2}i^{m+1}H_m^{(1)}\left(Q_2\frac{\rho}{a}\right), \quad (20)$$

где $Q_2 = \exp(i\pi/2)S_2$ ($\text{Re } Q_2 > 0$). Выбор ветви функции $Q_2 = k_0aq_2(p)$ согласно указанному правилу связан с тем, что в квазиэлектростатических волнах радиальные компоненты вектора Пойнтинга и волнового вектора имеют противоположные знаки. Поэтому для принятой временной зависимости $\exp(i\omega t)$ использование функций Ханкеля $H_m^{(1)}(Q_2\rho/a)$ удовлетворяет условию излучения, соответствующему утечке энергии из дакта. Отметим, что в рассматриваемом случае выполняется неравенство $\text{Im } Q_2 < 0$, так что поле моды оказывается нелокализованным (как это всегда имеет место для вытекающих волн [29]), несмотря на наличие составляющей, отвечающей ветви s_1 и спадающей до нуля при $\rho \rightarrow \infty$ в силу условия $\text{Re } S_1 > 0$.

Наряду с описанными выше несобственными модами, при $\tilde{P}_c < P$ слабовытекающие моды могут в ряде случаев быть найдены в области $\tilde{P}_c < p' < P$, если плотность плазмы внутри дакта заметно превосходит фоновое значение. При этом обе ветви, s_1 и q_2 , отвечают вытекающим волнам, а функции $K_m(S_1\rho/a)$ удобно заменить на функции Ханкеля второго рода $H_m^{(2)}(Q_1\rho/a)$, где $Q_1 = \exp(-i\pi/2)S_1$ ($\text{Re } Q_1 > 0$).

Как мы увидим далее, постоянные распространения несимметричных вытекающих мод лежат в тех же областях, что и в случае азимутально-симметричных мод. Однако для несимметричных мод существуют некоторые важные частные случаи, требующие отдельного обсуждения.

3.1. Дисперсионные свойства несобственных мод

Для упрощения анализа будем далее считать, что выполняется условие $\omega \gg \omega_{\text{ЛН}}$, и введём удобные обозначения $u = \omega_H^2/\omega^2$, $v = \omega_p^2/\omega^2$ ($v = \tilde{v}$ и $v = v_a$ при $N = \tilde{N}$ и $N = N_a$ соответственно). Очевидно, что при $m \neq 0$, как и для азимутально-симметричных мод ($m = 0$), потери энергии из дакта являются малыми ($p'' \ll p'$), если поперечные волновые числа «свистовой» части поля (ветви \tilde{q}_1 , s_1) и «квазиэлектростатической» его части (ветви \tilde{q}_2 , q_2) сильно различаются по абсолютному значению: $|\tilde{q}_1| \ll |\tilde{q}_2|$, $|s_1| \ll |q_2|$. Данные неравенства выполняются при

$$|p|^2 \gg \tilde{P}_c^2, \quad u \gg 1. \quad (21)$$

Заметим, что $\tilde{P}_c > P_c$ в силу условия $\tilde{N} > N_a$. С учётом первого неравенства в (21) величины \tilde{q}_1 , \tilde{q}_2 , s_1 , q_2 и \tilde{n}_1 , \tilde{n}_2 , n_1 , n_2 описываются сравнительно простыми приближёнными формулами:

$$\begin{aligned} \tilde{q}_1^2 &= \frac{\tilde{v}^2}{p^2u} - p^2, & \tilde{q}_2^2 &= p^2u - 2\tilde{v}; & s_1^2 &= p^2 - \frac{v_a^2}{p^2u}, & q_2^2 &= -s_2^2 = p^2u - 2v_a; \\ \tilde{n}_1 &= \frac{\tilde{v}}{pu^{1/2}} \left(1 + \frac{\tilde{v}}{p^2u}\right), & \tilde{n}_2 &= pu^{1/2} \left(1 - \frac{\tilde{v}}{p^2u}\right), \\ n_1 &= \frac{v_a}{pu^{1/2}} \left(1 + \frac{v_a}{p^2u}\right), & n_2 &= pu^{1/2} \left(1 - \frac{v_a}{p^2u}\right). \end{aligned} \quad (22)$$

Подставляя выражения (22) в строгое дисперсионное уравнение (П1) и опуская в нём некоторые члены, являющиеся малыми при выполнении условий (21), получим после ряда довольно громоздких преобразований, описанных в Приложении, следующее приближённое представление дисперсионного

уравнения:

$$\begin{aligned} & \left[\frac{J_{m+1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} + m \frac{\tilde{\alpha}_1}{\tilde{Q}_1} + \frac{\tilde{Q}_1}{S_1} \left(\frac{K_{m+1}(S_1)}{K_m(S_1)} + m \frac{\alpha_1}{S_1} \right) \right] \times \\ & \times \left[\frac{J_{m+1}(\tilde{Q}_2)}{J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2} - \frac{\eta_a \tilde{Q}_2}{\tilde{\eta} Q_2} \left(\frac{H_{m+1}^{(1)}(Q_2)}{H_m^{(1)}(Q_2)} - m \frac{\beta_2}{Q_2} \right) - \frac{m}{k_0 a p} \left(1 - \frac{\eta_a}{\tilde{\eta}} \right) \right] = \\ & = \kappa(p) \left(\frac{J_{m+1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} - m \frac{\tilde{\beta}_1}{\tilde{Q}_1} - m \frac{\tilde{Q}_1}{(k_0 a \tilde{P})^2} \right) \left(\frac{K_{m+1}(S_1)}{K_m(S_1)} + m \frac{\alpha_1}{S_1} \right) - \\ & - \delta(p) \frac{\tilde{Q}_1}{S_1} \left(\frac{J_{m+1}(\tilde{Q}_2)}{J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2} \right) \left(\frac{H_{m+1}^{(1)}(Q_2)}{H_m^{(1)}(Q_2)} + m \frac{\alpha_2}{Q_2} \right), \quad (23) \end{aligned}$$

где

$$\kappa(p) = \frac{v_a \tilde{Q}_2}{\tilde{v} S_1} \frac{v_a}{p^2 u} \left(\frac{\tilde{v}}{v_a} - 1 \right)^2, \quad \delta(p) = \frac{S_1}{Q_2} \left(\frac{v_a}{p^2 u} \right)^3 \left(\frac{\tilde{v}}{v_a} - 1 \right)^2. \quad (24)$$

При выводе уравнения (23) учтено соотношение (20).

Проанализируем приближённое дисперсионное уравнение (23) для ряда частных случаев. Покажем вначале, что при условии (19) слабывтекающие моды могут поддерживаться дактом, плотность плазмы в котором незначительно превышает фоновое значение, так что

$$(\tilde{v} - v_a)/v_a \ll 1. \quad (25)$$

Предположим также, что наряду с (25) выполняется неравенство $u^{1/2} \gg 4\tilde{v}/v_a$. В этом случае $P \gg \tilde{P}_c$, и для мод с постоянными распространения $p' > \max(P, \tilde{P}_c)$ имеют место условия (21), позволяющие воспользоваться уравнением (23).

Поскольку в случае (25) справедливо неравенство $\tilde{P} - P \ll P$, решение дисперсионного уравнения удобно искать в виде $p = P + \Delta p$ (см. (19)). Заметим, что при достаточно малой разности плотности плазмы внутри и вне дакта уравнение (23) допускает дополнительные упрощения. Во-первых, в правой части (23) можно опустить слагаемое, содержащее параметр δ , который в рассматриваемом случае является малым: $|\delta(p)| \ll 1$. Во-вторых, при $|S_1| \gg 1$, $|S_1| \geq m^2$ (заметим, что здесь $S_1 \approx 2k_0 a (P \Delta p)^{1/2}$) можно также пренебречь первым слагаемым в правой части (23), если только значение $|\Delta p|$ не слишком мало, так что выполняется неравенство

$$|\kappa(p)| \approx \frac{1}{2} \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 \left(\frac{P}{|\Delta p|} \right)^{1/2} \ll 1. \quad (26)$$

В противоположном предельном случае $|S_1| \ll 1$ первым слагаемым в правой части (23) можно пренебречь при соблюдении следующих условий:

$$\begin{aligned} & \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 |m| \ll 4k_0 a |\Delta p|, \quad m < 0; \\ & \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 \ll 4k_0 a |\Delta p \ln(\gamma k_0 a (P \Delta p)^{1/2})|, \quad m = 0; \\ & \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 k_0 a P |\ln(\gamma k_0 a (P \Delta p)^{1/2})| \ll 1, \quad m = 1; \end{aligned}$$

$$\left(\frac{\tilde{v}}{v_a} - 1\right)^2 k_0 a P \ll 2(m-1), \quad m > 1, \quad (27)$$

где $\gamma = \exp C = 1,7811 \dots$, $C = 0,5772 \dots$ — постоянная Эйлера—Маскерони³.

Итак, пренебрегая в нулевом порядке теории возмущений малыми членами в правой части уравнения (23), получаем, что оно распадается на два уравнения:

$$\frac{J_{m+1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} + m \frac{\tilde{\alpha}_1}{\tilde{Q}_1} = -\frac{\tilde{Q}_1}{S_1} \left(\frac{K_{m+1}(S_1)}{K_m(S_1)} + m \frac{\alpha_1}{S_1} \right), \quad (28)$$

$$\frac{J_{m+1}(\tilde{Q}_2)}{J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2} = \frac{\eta_a}{\tilde{\eta}} \frac{\tilde{Q}_2}{Q_2} \left(\frac{H_{m+1}^{(1)}(Q_2)}{H_m^{(1)}(Q_2)} - m \frac{\beta_2}{Q_2} \right) + \frac{m}{k_0 a p} \left(1 - \frac{\eta_a}{\tilde{\eta}} \right), \quad (29)$$

первое из которых описывает слабовытекающие моды «свистового» типа, а второе — сильновытекающие моды «квазиэлектростатического» типа. Следует отметить, что при выполнении условий (21) уравнение (23) распадается на уравнения (28), (29) и за пределами области (19), если условие (25) выполняется с достаточным запасом.

Обсудим вначале уравнение (28). Прежде всего отметим, что решения этого уравнения являются действительными. Это означает, что оно позволяет определить только действительные части $p'_{m\nu}(\omega)$ комплексных постоянных распространения вытекающих мод. Нетрудно убедиться, что при малом перепаде плотности на границе дакта (см. (25)) для мод, удовлетворяющих условию (19), справедливы приближённые соотношения $\tilde{\alpha}_1 \approx \alpha_1 \approx -2$. С учётом этих соотношений, а также известных рекуррентных формул для цилиндрических функций [32] уравнение (28) преобразуется к виду

$$\frac{J_{m-1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} = \frac{\tilde{Q}_1}{S_1} \frac{K_{m-1}(S_1)}{K_m(S_1)}. \quad (30)$$

Легко показать, что уравнение (30) допускает существование дискретных решений $p'_{m\nu} = P + \Delta p_{m\nu}$, где $\nu = 1, 2, \dots$, соответствующих отдельным модам. Например, в случае достаточно широкого дакта, когда $(k_0 a \tilde{P})^2 \gg [\mu_\nu^{(m-1)}]^2 (\tilde{v}/v_a - 1)^{-1}$, где $\mu_\nu^{(m)}$ — ν -й нуль функции Бесселя J_m , из (30) получаем уравнение $J_{m-1}(\tilde{Q}_1) \approx 0$, решения которого записываются следующим образом:

$$p'_{m\nu} = P + \Delta p_{m\nu} = \tilde{P} - \frac{1}{\tilde{P}} \left(\frac{\mu_\nu^{(m-1)}}{2k_0 a} \right)^2. \quad (31)$$

Заметим, что в рамках данного приближения действительные постоянные распространения мод с индексами $m = 1 \pm M$, где $M = 1, 2, \dots$, при одном и том же значении индекса ν оказываются одинаковыми. Для более узкого дакта ($k_0 a \tilde{P} \sim 1$) в случаях, допускающих аналитическое рассмотрение, решения дисперсионного уравнения имеют весьма сложный вид, и мы их здесь не приводим. Обратим лишь внимание на то, что с уменьшением радиуса дакта a дисперсионная кривая $p'_{m\nu}(\omega)$ для фиксированной моды сдвигается к нижней границе $p = P(\omega)$ области (19), а с увеличением a — к верхней границе $p = \tilde{P}(\omega)$. При фиксированном азимутальном индексе m большему значению индекса ν соответствует меньшая постоянная распространения $p'_{m\nu}$. С ростом ν на единицу число крупномасштабных («свистовых») вариаций поля по радиусу (ветви \tilde{q}_1, s_1) также возрастает на единицу. Поэтому такие моды можно условно назвать модами свистового типа, несмотря на то, что в их полях содержатся и

³ Нетрудно убедиться, что при надлежащем выборе радиуса дакта a выполнение условий (27) всегда может быть обеспечено путём уменьшения разности плотности плазмы внутри и вне канала.

квазиэлектростатические составляющие (ветви \tilde{q}_2, q_2) с соответствующим числом мелкомасштабных вариаций по радиусу.

Для отыскания постоянных затухания $p''_{m\nu}(\omega)$ мод свистового типа необходимо учесть малые члены, стоящие в правой части уравнения (23). Хотя соответствующее рассмотрение может быть выполнено аналогично тому, как это делалось в [22] для азимутально-симметричных мод, получающиеся при этом выражения оказываются чрезвычайно громоздкими и поэтому здесь не приводятся. Отметим только, что постоянные затухания $p''_{m\nu}(\omega)$ осциллируют между максимальными и минимальными значениями, отвечающими условиям $J_{m+1}(\tilde{Q}_2) \approx 0$ и $J_m(\tilde{Q}_2) \approx 0$ соответственно, причём $p''_{\min} \ll p''_{\max} \ll p'$. В справедливости сказанного нетрудно убедиться, если учесть, что малые параметры κ, δ удовлетворяют неравенству $|\delta(p)| \ll |\kappa(p)|$.

Помимо слабовытекающих свистовых мод дисперсионное уравнение (23) допускает существование сильновытекающих мод квазиэлектростатического типа, действительные части постоянных распространения которых лежат в области $p' > \tilde{P}_c$, включающей в себя область (19). В нулевом порядке теории возмущений действительные и мнимые части постоянных распространения этих мод могут быть получены из уравнения (29). Заметим, что учёт поправок порядка κ, δ лишь незначительно изменяет соответствующие решения. Сильновытекающие моды с различными значениями индекса ν характеризуются различным числом мелкомасштабных вариаций поля по радиусу. Для того, чтобы моде с бóльшим числом мелкомасштабных вариаций соответствовал больший индекс ν , моды квазиэлектростатического типа следует нумеровать в порядке возрастания действительных частей комплексных постоянных распространения. Поскольку при $(\tilde{\nu} - \nu_a)/\nu_a \rightarrow 0$ утечка этих мод быстро возрастает, практического интереса они не представляют.

Проведённое аналитическое рассмотрение хорошо согласуется с результатами численного решения строгого дисперсионного уравнения, представленными на рис. 6 для следующих значений безразмерных параметров: $\tilde{\omega}_p/\omega_p = 1,05$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_{Ha}/c = 2,5$. На рис. 6а сплошными линиями (с указанными индексами m, ν) показаны зависимости $p''_{m\nu}(\omega)$ для слабовытекающих мод с азимутальными индексами $m = \pm 1$. Кроме того, в качестве примера приведены аналогичные зависимости для двух сильновытекающих мод с $m = 1$ (кривые I и II). Зависимости $\tilde{P}(\omega)$ и $P(\omega)$ показаны штриховыми линиями 1 и 2 соответственно. На рис. 6б изображены кривые $p''_{m\nu}(\omega)$ для слабовытекающих мод с индексами $m = 1, \nu = 1$ и $m = -1, \nu = 1$. Чтобы не усложнять рисунок, мы не приводим имеющую аналогичный вид кривую $p''_{m\nu}$ для слабовытекающей моды с индексами $m = 1, \nu = 2$, а также соответствующие зависимости для сильновытекающих мод.

Отметим, что полученные на основании численных расчётов зависимости $p'_{m\nu}(\omega)$ как для слабовытекающих, так и для сильновытекающих мод практически совпадают с решениями приближённых дисперсионных уравнений (28) и (29) соответственно. Минимумы и максимумы постоянных затухания $p''_{m\nu}(\omega)$ хорошо согласуются с приведёнными выше условиями их существования. С увеличением частоты кривые $p'_{m\nu}(\omega)$ для слабовытекающих мод приближаются к границе $p = \tilde{P}(\omega)$ и при $\omega \lesssim \omega_H/2$ переходят из области (19) в область $p' > \tilde{P}(\omega)$. Следует, однако, иметь в виду, что вне области (19) затухание мод достаточно велико и соответствующие решения дисперсионного уравнения не представляют интереса. С уменьшением частоты дисперсионные кривые $p'_{m\nu}(\omega)$ приближаются к границе $p = P(\omega)$. Точка пересечения дисперсионной кривой с этой границей (рис. 6а) определяет критическую частоту соответствующей моды (заметим, что обрыв кривой $p'_{m\nu}(\omega)$ для моды с индексами $m = -1, \nu = 1$ на рис. 6б отвечает критической частоте данной моды). Исключение составляет лишь мода с индексами $m = 1, \nu = 1$. Эта мода является основной: она не имеет критической частоты и может существовать при сколь угодно малом перепаде плотности на границе дакта. Действительно, в предельном случае $(\tilde{\nu} - \nu_a)/\nu_a \rightarrow 0$ величины $\tilde{Q}_1, S_1, \Delta p$ стремятся к нулю и уравнение (28) при $m = 1$ принимает вид

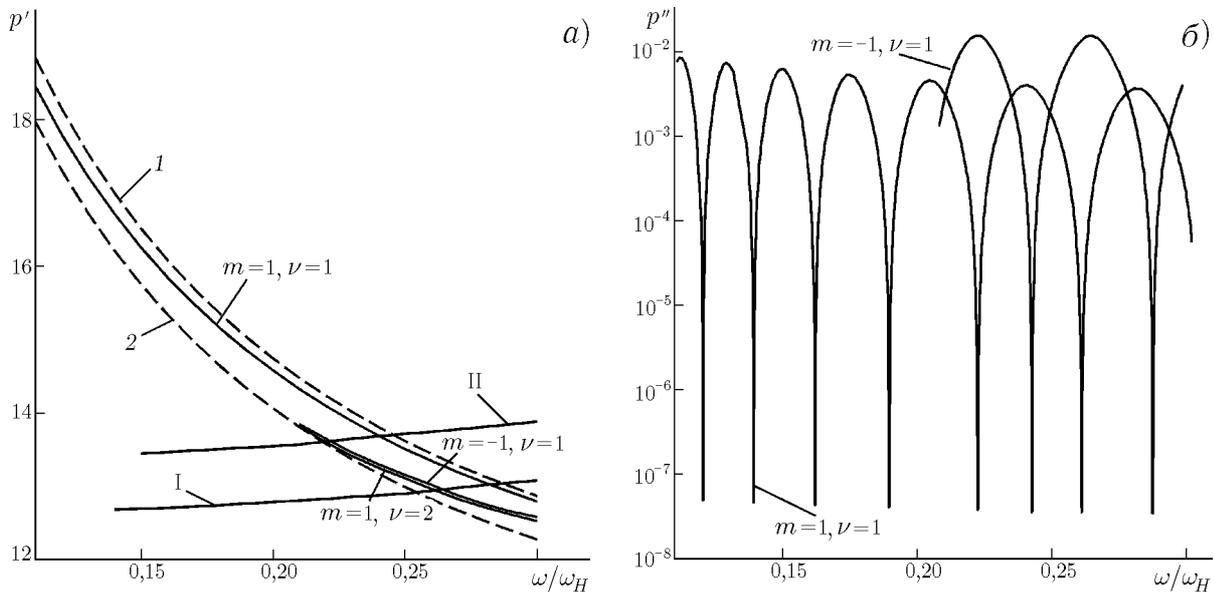


Рис. 6. Дисперсионные кривые вытекающих мод при $\tilde{\omega}_p/\omega_p = 1,05$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 2,5$: а — постоянные распространения $p'_{m\nu}(\omega)$ слабывтекающих мод, постоянные распространения сильновытекающих мод (кривые I и II), зависимости $\tilde{P}(\omega)$ (кривая I) и $P(\omega)$ (кривая 2); б — постоянные затухания $p''_{m\nu}(\omega)$ слабывтекающих мод с индексами $m = \pm 1, \nu = 1$

$$\ln[2/(\gamma S_1)] = 2/\tilde{Q}_1^2. \tag{32}$$

С учётом приближённых соотношений $s_1 \approx 2(P \Delta p)^{1/2}$ и $\tilde{q}_1^2 \approx (\tilde{P}^4 - P^4)/P^2 \approx 2(\tilde{P}^2 - P^2)$, выполняющихся в данном предельном случае, из (32) получаем

$$p'_{11} = P + \frac{1}{(k_0 a \gamma)^2 P} \exp \left[-\frac{2}{(k_0 a)^2 (\tilde{P}^2 - P^2)} \right], \tag{33}$$

откуда следует, что данная мода в самом деле может распространяться при $(\tilde{v} - v_a)/v_a \rightarrow 0$. Нетрудно показать, что постоянная затухания основной моды удовлетворяет неравенству $p''_{11} \ll p'_{11}$. Примечательно, что по своим свойствам эта мода напоминает основную («дипольную») моду круглого диэлектрического волновода [29]. Однако в отличие от случая диэлектрического волновода здесь основная мода является вытекающей.

Перейдём теперь к изучению особенностей волноводного распространения свистовых волн в каналах, плотность плазмы в которых существенно превышает фоновое значение:

$$\tilde{v}/v_a \gg 1. \tag{34}$$

Рассмотрим случай, когда наряду с условием (34) выполняются неравенства (21). Очевидно, что в этом случае для дисперсионного уравнения остаётся справедливым представление (23). Однако теперь дисперсионное уравнение содержит дополнительный малый параметр $\eta_a/\tilde{\eta} \approx v_a/\tilde{v}$ (наряду с параметром δ , для которого по-прежнему имеет место условие $|\delta| \ll 1$). В нулевом порядке теории возмущений по малым параметрам δ и v_a/\tilde{v} уравнение (23) сводится к виду

$$\left[\frac{J_{m+1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} + m \frac{\tilde{\alpha}_1}{\tilde{Q}_1} + \frac{\tilde{Q}_1}{S_1} \left(\frac{K_{m+1}(S_1)}{K_m(S_1)} + m \frac{\alpha_1}{S_1} \right) \right] \left(\frac{J_{m+1}(\tilde{Q}_2)}{J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2} - \frac{m}{k_0 a p} \right) = \kappa(p) \left(\frac{J_{m+1}(\tilde{Q}_1)}{J_m(\tilde{Q}_1)} - m \frac{\tilde{\beta}_1}{\tilde{Q}_1} - m \frac{\tilde{Q}_1}{(k_0 a \tilde{P})^2} \right) \left(\frac{K_{m+1}(S_1)}{K_m(S_1)} + m \frac{\alpha_1}{S_1} \right). \tag{35}$$

Заметим, что параметр $\kappa(p) \approx \tilde{P}^2/(ps_1)$ не является теперь малым. Поэтому в отличие от случая (25) дисперсионное уравнение не распадается на отдельные уравнения (28) и (29) даже при пренебрежении утечкой энергии из канала. Это означает, что вытекающие моды становятся в определённом смысле «гибридными», т. е. одновременно обладают свойствами мод как свистового, так и квазиэлектростатического типов.

Указанное различие между приближёнными дисперсионными уравнениями в случаях сильного и слабого перепада плотности плазмы на границе канала наглядно иллюстрирует рис. 7, на котором изображены зависимости действительных и мнимых частей комплексных постоянных распространения $p_{m\nu} = p'_{m\nu} - ip''_{m\nu}$ от частоты для нескольких вытекающих мод с азимутальным индексом $m = 1$, полученные на основании численного решения строгого дисперсионного уравнения при следующих значениях параметров: $\tilde{\omega}_p/\omega_p = 1,778$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 1$. Кроме того, на рис. 7а штриховыми линиями изображены зависимости $\tilde{P}(\omega)$ и $\tilde{P}_c(\omega)$ (кривые 1 и 2 соответственно). Чтобы не усложнять рис. 7б, мы приводим кривые $p''_{m\nu}$ только для $\nu = 3, \dots, 8$. Заметим, что решения $p_{m\nu}$ дисперсионного уравнения пронумерованы по ν в порядке возрастания значений $p'_{m\nu}$, начиная с нижней границы $p = \tilde{P}_c$ области (19), поскольку такая нумерация оказывается более удобной. Можно убедиться, что приближённое дисперсионное уравнение (35) вполне удовлетворительно описывает поведение кривых $p'_{m\nu}(\omega)$ несмотря на то, что условие (34) выполняется здесь с небольшим запасом ($\tilde{v}/v_a = 3,162$). Мы не приводим результаты сопоставления решений строгого дисперсионного уравнения и приближённого уравнения (35) только из-за недостатка места.

Максимумы и минимумы постоянных затухания на рис. 7б соответствуют условиям

$$\frac{J_{m+1}(\tilde{Q}_2)}{J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2} \approx \frac{m}{k_0 a p} \quad (36)$$

и $J_m(\tilde{Q}_2) \approx 0$. При этом, как следует из рис. 7б, для мод, удовлетворяющих условию (19), выполняются неравенства $p''_{\min} \ll p''_{\max} \ll p'$. Примечательно, что на частотах, отвечающих минимумам постоянных затухания мод, а также в ближайшей окрестности каждой из этих частот поведение действительной части $p'_{m\nu}$ постоянной распространения соответствующей моды приближённо описывается уравнением (28). На данных частотах моды в большей степени обладают свойствами свистовых волн. При удалении от указанных частотных областей моды начинают приобретать свойства квазиэлектростатических волн, а их постоянные затухания увеличиваются. В этом случае поведение $p'_{m\nu}$ приближённо описывается уравнением (29). С ростом частоты дисперсионные кривые $p'_{m\nu}$ в конечном счёте выйдут в область $p' > \tilde{P}$, т. е. в область сильновытекающих мод квазиэлектростатического типа. В данной области решения дисперсионного уравнения при $|p|^2 \gg \tilde{P}^2$ могут быть получены приближённо из уравнения (29). При этом в силу условия (34) уравнение (29) принимает вид (36).

Следует отметить, что с увеличением частоты число мелкомасштабных радиальных вариаций поля моды с фиксированным индексом ν уменьшается на единицу после прохождения участка, на котором постоянная затухания принимает минимальное значение. Поэтому здесь моды различаются как числом мелкомасштабных вариаций поля по радиусу, так и числом крупномасштабных вариаций. В частности, в отличие от случая слабого перепада плотности плазмы на границе дакта две соседние моды на фиксированной частоте различаются не числом крупномасштабных вариаций, а числом мелкомасштабных вариаций: $\tilde{Q}_2(p'_{m\nu+1}) - \tilde{Q}_2(p'_{m\nu}) \approx \pi$; $\tilde{Q}_1(p'_{m\nu}) - \tilde{Q}_1(p'_{m\nu+1}) \ll \pi$. Число крупномасштабных вариаций поля моды с фиксированным индексом ν также не остаётся постоянным при изменении частоты. Например, мода с индексом $\nu = 4$ (см. рис. 7) при $\omega/\omega_H \approx 0,169$, что соответствует минимуму постоянной затухания $p''_{m\nu}$, имеет две крупномасштабные вариации поля по радиусу, а при $\omega/\omega_H \approx 0,296$, что соответствует следующему минимуму $p''_{m\nu}$, — одну крупномасштабную вариацию. Своеобразие свойств вытекающих мод в этом случае делает затруднительным использование для них способа нумерации, применяемого для мод открытых диэлектрических волноводов и пригодного в случае

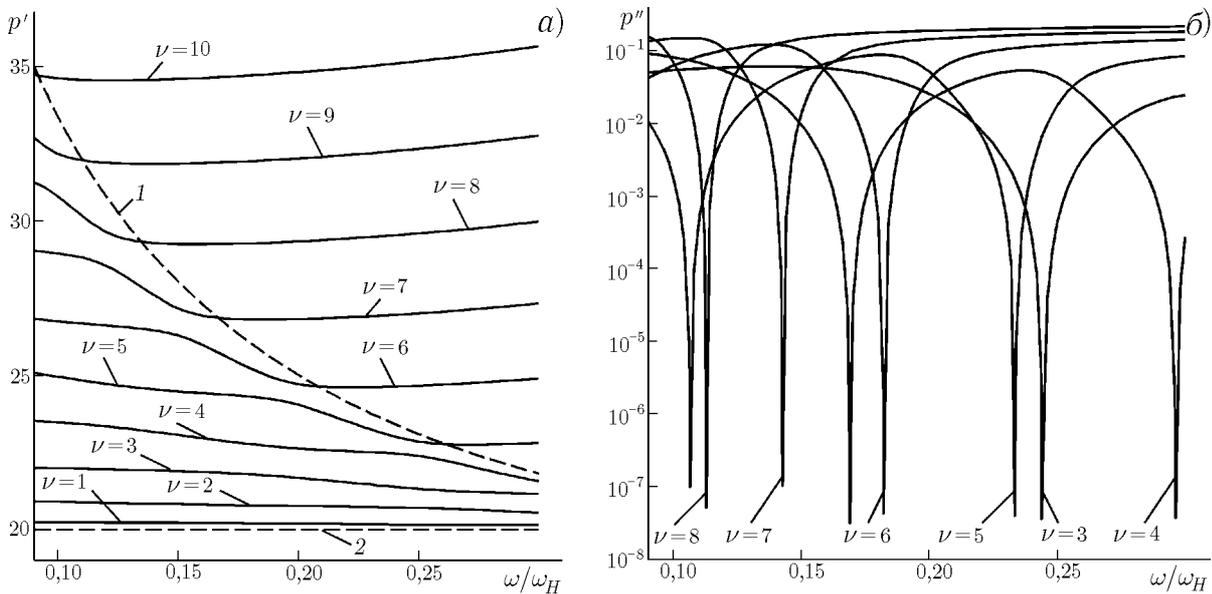


Рис. 7. Дисперсионные кривые вытекающих мод при $\tilde{\omega}_p/\omega_p = 1,778$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 1$: *a* — постоянные распространения $p'_{m\nu}(\omega)$ мод с индексами $m = 1, \nu = 1, 2, \dots, 10$ и зависимости $\tilde{P}(\omega)$ (кривая 1), $\tilde{P}_c(\omega)$ (кривая 2); *б* — постоянные затухания $p''_{m\nu}(\omega)$ мод

слабого перепада плотности плазмы на границе дакта. Именно поэтому здесь целесообразно использовать предложенный выше способ нумерации мод (см. рис. 7).

Проведённое рассмотрение для случаев слабого и сильного перепада плотности плазмы на границе дакта полезно дополнить обсуждением зависимости дисперсионных свойств мод от радиуса дакта. Очевидно, что при достаточно большом радиусе a в уравнении (23) можно опустить члены, содержащие величины $\tilde{\alpha}_k, \alpha_k, \tilde{\beta}_k, \beta_k, (1 - \eta_a/\tilde{\eta})/(k_0 a p), \tilde{Q}_1/(k_0 a \tilde{P})^2$. В этом случае анализ приближённого дисперсионного уравнения практически не отличается от рассмотрения, выполненного в [22] для азимутально-симметричных мод ($m = 0$): следует лишь заменить цилиндрические функции порядка 0 и 1 (см. [22]) на цилиндрические функции порядка m и $m + 1$ соответственно. В противоположном предельном случае, т. е. при стремлении радиуса дакта к нулю, в области (19) остаётся лишь решение, отвечающее основной моде. С уменьшением частоты дисперсионная кривая $p'_{11}(\omega)$ для этой моды переходит из интервала $\tilde{P}_c < p'_{11} < \tilde{P}$ ($P < \tilde{P}_c$) в интервал $P < p'_{11} < \tilde{P}$ ($\tilde{P}_c < P$), сдвигаясь к его нижней границе $p = P$ по мере уменьшения a .

Покажем, что решение $p_{11} = P + \Delta p_{11}$ существует при сколь угодно малых значениях a . Для этого цилиндрические функции, входящие в строгое дисперсионное уравнение, должны быть заменены их асимптотическими представлениями, справедливыми при малых значениях аргументов ($|\tilde{Q}_1| \ll 1, |\tilde{Q}_2| \ll 1, |S_1| \ll 1, |Q_2| \ll 1$) в случае $m = 1$:

$$J_1(\tilde{Q}_{1,2}) \approx \frac{\tilde{Q}_{1,2}}{2}, \quad J_2(\tilde{Q}_{1,2}) \approx \frac{\tilde{Q}_{1,2}^2}{8}, \quad K_1(S_1) \approx \frac{1}{S_1},$$

$$K_2(S_1) \approx \frac{2}{S_1^2} + \ln \frac{2}{\gamma S_1}, \quad H_1^{(1)}(Q_2) \approx -i \frac{2}{\pi Q_2}, \quad H_2^{(1)}(Q_2) \approx -i \frac{4}{\pi Q_2^2}. \quad (37)$$

После подстановки выражений (37) в дисперсионное уравнение (П1) и ряда несложных, но довольно громоздких преобразований члены, содержащие степени S_1 , сокращаются, так что величина S_1 входит в результирующее уравнение только в виде функции $\ln[2/(\gamma S_1)]$. Оказывается, что этот логарифм выражается через следующие величины, которые берутся при $p = P = (\epsilon_a - g_a)^{1/2}$ (т. к. величина Δp_{11}

стремится к нулю при $a \rightarrow 0$):

$$q_2^2(P) = \frac{g_a \eta_a}{\varepsilon_a} \left(1 + \frac{P^2}{\eta_a} \right), \quad n_1(P) = P, \quad n_2(P) = -\eta_a/P,$$

$$\left. \frac{dn_1}{dp} \right|_{p=P} = 2 \frac{\varepsilon_a}{g_a} \left(\frac{2}{1 + \eta_a^{-1} P^2} - 1 \right) - 1. \quad (38)$$

С учётом выражения $s_1 = 2(P \Delta p_{11})^{1/2} (1 + \eta_a^{-1} P^2)^{-1/2}$, справедливого в пределе $\Delta p_{11} \rightarrow 0$, можно показать, что

$$p'_{11} = P + \frac{1 + \eta_a^{-1} P^2}{(k_0 a \gamma)^2 P} \exp \left[-\frac{2}{(k_0 a)^2} \left(\frac{1 + \eta_a^{-1} \tilde{P}^2}{\tilde{P}^2 - P^2} - \frac{P^2}{2\varepsilon_a g_a} \frac{[1 - \eta_a^{-1} (\varepsilon_a + g_a)]^2}{1 + \eta_a^{-1} P^2} \right) \right]. \quad (39)$$

В данном приближении удаётся получить лишь действительную часть комплексной постоянной распространения. Поскольку постоянная затухания является малой ($p''_{11} \ll p'_{11}$), для её отыскания необходим учёт членов следующего порядка малости в представлениях (37), (38). Мы не будем останавливаться здесь на соответствующих расчётах. В предельном случае $(\tilde{v} - v_a)/v_a \rightarrow 0$ выражение (39) даёт полученный ранее результат (33), что и следовало ожидать. Заметим, что при переходе к (33) необходимо учесть неравенства $\varepsilon_a \ll |\eta_a|$, $|g_a| \ll |\eta_a|$, $\tilde{P}^2 \ll |\eta_a|$.⁴ Таким образом, основная вытекающая мода может распространяться с малым затуханием в дакте произвольного радиуса независимо от величины перепада плотности плазмы на границе $\rho = a$.

В заключение данного раздела опишем простые способы «конструирования» приближённых дисперсионных уравнений для рассмотренных выше частных случаев. Оказывается, что эти уравнения можно получить, минуя их громоздкий вывод из строгого дисперсионного уравнения. Так, в случае (25), положив $B_2 = C_2 = 0$ в выражениях для компонент E_ϕ и H_z (см. (6), (7)), из условий непрерывности этих компонент на границе дакта $\rho = a$ сразу же приходим к приближённому дисперсионному уравнению (28) для мод свистового типа. Как показывает анализ поведения различных компонент поля, описанный приём вывода уравнения (28) основан на том обстоятельстве, что вытекающие моды свистового типа при условиях (21) имеют квази-ТЕ-поляризацию. Приближённое дисперсионное уравнение (29) для мод квазиэлектростатического типа можно получить из условий непрерывности радиальной компоненты D_ρ вектора электрической индукции и компоненты E_z на границе дакта, если положить $B_1 = C_1 = 0$.

В случае (34), учитывая, что моды в области (19) являются слабывтекающими ($|C_2| \ll |C_1|$) и имеют квази-ТЕ-поляризацию, положим $C_2 = 0$ в выражениях (7) для компонент поля вне дакта ($\rho > a$). Тогда из условий непрерывности компонент E_ϕ , H_ϕ , H_z при $\rho = a$ получаем уравнение, которое после ряда преобразований сводится к (35).

Подчеркнём, что приближённые представления дисперсионного уравнения, которые выводятся с помощью описанных выше способов, весьма удобны для отыскания действительных частей постоянных распространения мод в качестве первого приближения при численном решении строгого дисперсионного уравнения.

3.2. Структура полей несобственных мод

Рассмотрим структуру полей слабывтекающих мод. Для того, чтобы судить как об абсолютных значениях компонент поля вытекающей моды, так и о фазовых соотношениях между ними, в общем

⁴ Члены порядка ε_a/η_a , g_a/η_a , \tilde{P}^2/η_a приходится, однако, учитывать при выводе формулы (39), т. к. в противном случае не удаётся получить правильное выражение для p'_{11} при $a \rightarrow 0$ и произвольном отношении \tilde{N}/N_a .

случае необходимо проанализировать поведение действительной и мнимой частей каждой компоненты. Оказывается, что для слабоветекающих мод свистового диапазона нет необходимости представлять графики одновременно для действительных и мнимых частей всех компонент поля. Напомним, что в отсутствие источников поле моды определяется с точностью до произвольной комплексной константы. Анализ поведения компонент поля показывает, что указанную константу можно выбрать таким образом, чтобы максимальные значения величин $|\operatorname{Im} E_\rho|$, $|\operatorname{Re} E_\phi|$, $|\operatorname{Re} E_z|$, $|\operatorname{Re} H_\rho|$, $|\operatorname{Im} H_\phi|$, $|\operatorname{Im} H_z|$ существенно превышали максимальные значения величин $|\operatorname{Re} E_\rho|$, $|\operatorname{Im} E_\phi|$, $|\operatorname{Im} E_z|$, $|\operatorname{Im} H_\rho|$, $|\operatorname{Re} H_\phi|$, $|\operatorname{Re} H_z|$ соответственно. Ясно, что в этом случае достаточно привести лишь зависимости величин $\operatorname{Im} E_\rho$, $\operatorname{Re} E_\phi$, $\operatorname{Re} E_z$, $\operatorname{Re} H_\rho$, $\operatorname{Im} H_\phi$, $\operatorname{Im} H_z$ от радиальной координаты. Графики, представленные ниже, построены именно для такой нормировки поля.

На рис. 8 изображены распределения компонент поля основной моды ($m = 1, \nu = 1$) по поперечной координате для случая $\tilde{\omega}_p/\omega_p = 1,02$; $\omega_p/\omega_H = 5,623$; $\omega_{Ha}/c = 2,5$; $\omega/\omega_H = 0,101$. Постоянная распространения данной моды $p_{11} = 18,76 - i3,40 \cdot 10^{-4}$, и, как нетрудно убедиться, её действительная часть лежит в интервале (19). На рис. 8 хорошо видны крупно- и мелкомасштабные составляющие поля, соответствующие захваченным вистлерам и вытекающим квазиэлектростатическим волнам. Наиболее заметны мелкомасштабные осцилляции у радиальной и продольной компонент электрического поля, поскольку именно эти компоненты максимальны в квазиэлектростатических волнах. Значения параметров, для которых построены распределения компонент поля на рис. 8, выбраны таким образом, чтобы постоянная затухания моды была близка к минимальному значению. В этом случае вклад мелкомасштабных вытекающих волн в полное поле моды вне дакта оказывается пренебрежимо малым. При значениях параметров, для которых постоянная затухания принимает максимальное значение, вклад мелкомасштабной составляющей в поле моды возрастает. Заметим, что, поскольку в рассматриваемом случае выполняется условие (25), крупномасштабная составляющая поля моды является слабокаллизированной. Поэтому основную моду в области $\rho > a$ можно рассматривать как плоскую свистовую волну, распространяющуюся вдоль внешнего магнитного поля и обладающую квазилокализованной структурой благодаря замедлению из-за наличия цилиндра с повышенной плотностью плазмы.

На рис. 9 представлены распределения компонент поля моды с индексами ⁵ $m = 1, \nu = 7$ в случае большого перепада плотности на границе дакта при $\tilde{\omega}_p/\omega_p = 1,778$; $\omega_p/\omega_H = 5,623$; $\omega_{Ha}/c = 1$; $\omega/\omega_H = 0,15$. Постоянная распространения моды $p_{17} = 27,21 - i0,019$; действительная часть постоянной распространения по-прежнему принадлежит интервалу (19). При выбранных значениях параметров постоянная затухания заметно превышает минимальное значение, поэтому наличие мелкомасштабных квазиэлектростатических волн весьма существенно сказывается на структуре поля данной моды.

Мы не останавливаемся на анализе структуры полей мод в области $\rho' > \tilde{P}$. В этом случае крупномасштабная составляющая имеет квазиповерхностный характер как внутри, так и вне дакта и сосредоточена вблизи его границы $\rho = a$, а вклад объёмной мелкомасштабной составляющей в полное поле становится преобладающим.

4. ЗАКЛЮЧЕНИЕ

В работе исследованы дисперсионные свойства и структура полей несимметричных мод свистового диапазона, направляемых цилиндрическим дактом с повышенной плотностью плазмы. Как следует из проведённого рассмотрения, в резонансной области (1) свистового диапазона, допускающей существование как крупномасштабных вистлеров, так и мелкомасштабных квазиэлектростатических

⁵ Значения индексов соответствуют обозначениям дисперсионных кривых на рис. 7.

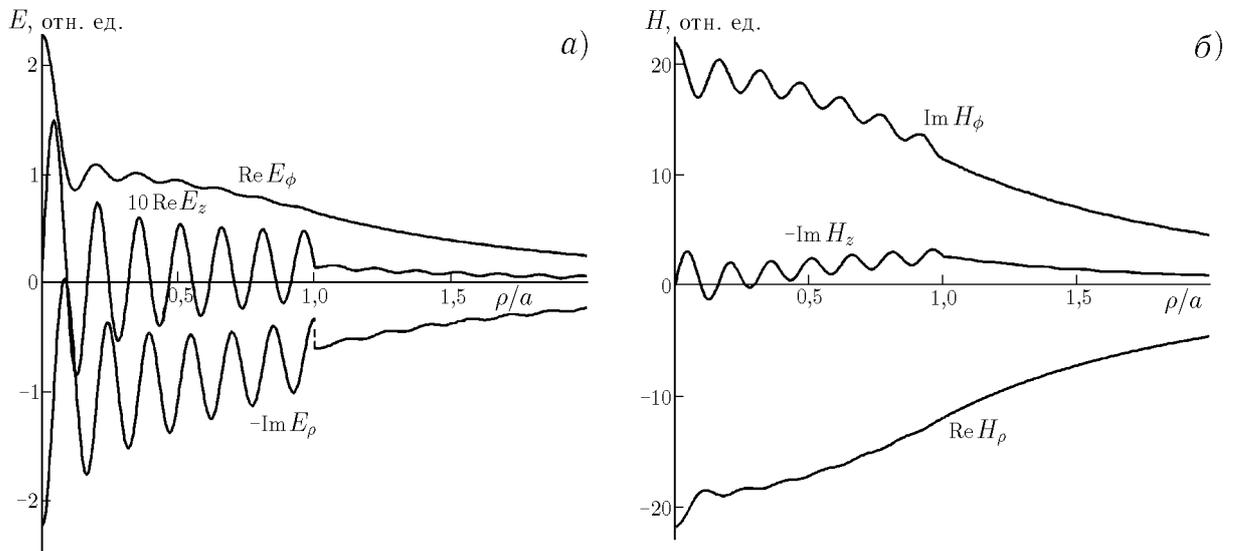


Рис. 8. Распределения по поперечной координате компонент поля основной вытекающей моды ($m = 1$, $\nu = 1$) при $\omega/\omega_H = 0,101$; $\tilde{\omega}_p/\omega_p = 1,02$; $\omega_p/\omega_H = 5,623$; $\omega_{LH}/\omega_H = 4,3 \cdot 10^{-3}$; $\omega_H a/c = 2,5$

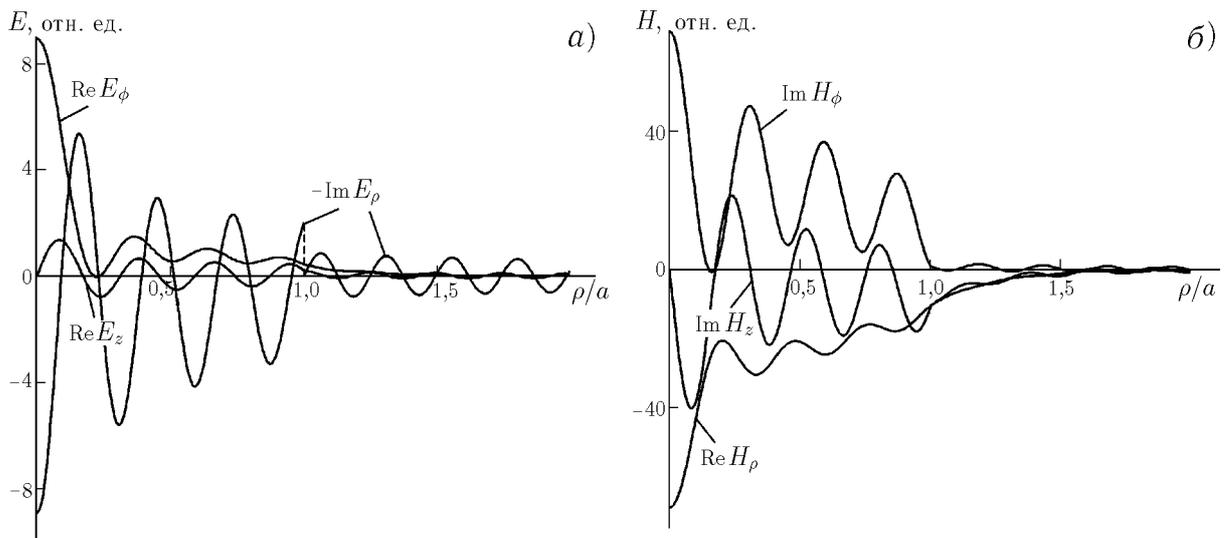


Рис. 9. Распределения по поперечной координате компонент поля вытекающей моды с индексами $m = 1$, $\nu = 7$ при $\omega/\omega_H = 0,15$; значения параметров $\tilde{\omega}_p/\omega_p$, ω_p/ω_H , ω_{LH}/ω_H , $\omega_H a/c$ те же, что и для рис. 7

волн, такие дакты могут поддерживать собственные моды и слабывтекающие несобственные моды. При этом основной модой дакта с повышенной плотностью является вытекающая мода с азимутальным индексом $m = 1$, обладающая достаточно малым радиационным затуханием. Существенно, что методы анализа, использованные в данной работе, могут быть распространены и на другие частотные интервалы, в частности на нерезонансную область $\Omega_H < \omega < \omega_{LH}$ свистового диапазона, где Ω_H — гирочастота ионов. Нетрудно показать, что в этой частотной области дакты с повышенной плотностью могут поддерживать медленные собственные моды.

Заметим, что использованная в настоящей работе модель дакта с резкой границей соответствует в реальных условиях такому профилю $N(\rho)$, когда характерный поперечный масштаб L_N изменения плотности плазмы удовлетворяет условию $L_N \ll 2\pi (k_0 |q_2|)^{-1}$. В случае более плавного профиля

плотности плазмы требуется отдельное рассмотрение, особенно при анализе слабовытекающих мод в области (19), которые представляют значительный интерес для приложений. Можно, однако, показать (так же, как для азимутально-симметричных мод [1, 5]), что основные особенности каналирования несимметричных слабовытекающих мод сохраняются и в случае дакта с плавным профилем плотности плазмы.

Следует подчеркнуть, что выполненное рассмотрение дало возможность установить условия, позволяющие существенно упростить анализ дисперсионных характеристик несимметричных свистовых мод в частных случаях, важных для приложений. Это обстоятельство заметно облегчает интерпретацию результатов лабораторных экспериментов по изучению каналированного распространения вис-тлеров в дактах плотности. Наконец, отметим, что данное исследование является необходимым этапом для последующего рассмотрения особенностей излучения несимметричных источников при наличии дактов плотности в магнитоактивной плазме. Такое исследование представляется весьма актуальным, поскольку в опубликованных к настоящему времени теоретических работах подробно изучено лишь излучение азимутально-симметричных источников, расположенных в дактах плотности.

Работа выполнена при поддержке РФФИ (гранты № 01-02-16949, 01-02-17390, 02-02-06424), программ «Ведущие научные школы» (грант № 00-15-96734) и «Университеты России — фундаментальные исследования», а также Министерства образования РФ (грант Е00-3.5-227) и Минпромна-уки РФ (контракт 40.020.1.1.1171).

ПРИЛОЖЕНИЕ

Дисперсионное уравнение для мод, направляемых цилиндрическим дактом, может быть записано в следующем виде:

$$\Delta_0 - \sum_{i=1}^4 \Delta_i = 0. \quad (\text{П1})$$

Здесь

$$\begin{aligned} \Delta_0 &= M_0 \left[\frac{\tilde{\eta}}{\eta_a} \tilde{n}_2 (n_2 - n_1) J_m^{(1)} J_m^{(2)} - n_2 L_{12} J_m^{(1)} K_m^{(2)} + \tilde{n}_2 L_{21} J_m^{(2)} K_m^{(1)} + n_2 (\tilde{n}_2 - \tilde{n}_1) K_m^{(1)} K_m^{(2)} \right], \\ \Delta_1 &= M_0 \left[\frac{\tilde{\eta}}{\eta_a} \tilde{n}_1 (n_2 - n_1) \hat{J}_m^{(1)} \hat{J}_m^{(2)} - \tilde{n}_1 L_{12} \hat{J}_m^{(1)} \hat{K}_m^{(2)} + n_1 L_{21} \hat{J}_m^{(2)} \hat{K}_m^{(1)} + n_1 (\tilde{n}_2 - \tilde{n}_1) \hat{K}_m^{(1)} \hat{K}_m^{(2)} \right], \\ \Delta_2 &= M_{11} L_{22} \hat{J}_m^{(1)} K_m^{(1)} + M_{22} L_{11} J_m^{(2)} \hat{K}_m^{(2)}, \\ \Delta_3 &= n_1 M_0 L_{22} \left(\hat{J}_m^{(1)} K_m^{(1)} - J_m^{(1)} \hat{K}_m^{(1)} \right), \\ \Delta_4 &= n_2 M_0 L_{11} \left(J_m^{(2)} \hat{K}_m^{(2)} - \hat{J}_m^{(2)} K_m^{(2)} \right), \end{aligned} \quad (\text{П2})$$

где

$$\begin{aligned} J_m^{(1)} &= \frac{J_{m+1}(\tilde{Q}_1)}{\tilde{Q}_1 J_m(\tilde{Q}_1)} + m \frac{\tilde{\alpha}_1}{\tilde{Q}_1^2}, & J_m^{(2)} &= \frac{J_{m+1}(\tilde{Q}_2)}{\tilde{Q}_2 J_m(\tilde{Q}_2)} - m \frac{\tilde{\beta}_2}{\tilde{Q}_2^2}, \\ K_m^{(1)} &= \frac{K_{m+1}(S_1)}{S_1 K_m(S_1)} + m \frac{\alpha_1}{S_1^2}, & K_m^{(2)} &= \frac{K_{m+1}(S_2)}{S_2 K_m(S_2)} - m \frac{\beta_2}{S_2^2}, \\ \hat{J}_m^{(k)} &= J_m^{(k)} + (-1)^k m \frac{\tilde{\alpha}_k + \tilde{\beta}_k}{\tilde{Q}_k^2}, & \hat{K}_m^{(k)} &= K_m^{(k)} + (-1)^k m \frac{\alpha_k + \beta_k}{S_k^2}; \end{aligned}$$

$$M_{ij} = M_0 (\tilde{n}_i - n_j), \quad L_{ij} = \frac{\tilde{\eta}}{\eta_a} n_i - \tilde{n}_j, \quad M_0 = (n_1 - n_2)^{-1} (\tilde{n}_1 - \tilde{n}_2)^{-1},$$

$$k = 1, 2; \quad i = 1, 2; \quad j = 1, 2. \quad (\text{ПЗ})$$

Заметим, что имеют место следующие формулы, вытекающие из выражений для $\tilde{\alpha}_k, \tilde{\beta}_k, \alpha_k, \beta_k$ (см. (9)):

$$\frac{\tilde{\alpha}_{1,2} + \tilde{\beta}_{1,2}}{\tilde{Q}_{1,2}^2} = \frac{1}{(k_0 a)^2 \tilde{g}} \frac{\tilde{\eta} - \tilde{\varepsilon}}{\tilde{\eta}}, \quad \frac{\alpha_{1,2} + \beta_{1,2}}{S_{1,2}^2} = -\frac{1}{(k_0 a)^2 g_a} \frac{\eta_a - \varepsilon_a}{\eta_a}. \quad (\text{П4})$$

При условиях (21) выражение для Δ_0 с учётом формул (22) сводится приближённо к виду

$$\Delta_0 = \left(J_m^{(1)} + K_m^{(1)} \right) \left(\frac{\tilde{\eta}}{\eta_a} J_m^{(2)} + K_m^{(2)} \right). \quad (\text{П5})$$

Член Δ_1 после подстановки (ПЗ), (П4) записывается как

$$\Delta_1 = \frac{P_c^2}{4p^2} \left(\frac{\tilde{\eta}}{\eta_a} J_m^{(1)} + K_m^{(1)} \right) \left(\frac{\tilde{\eta}}{\eta_a} J_m^{(2)} + K_m^{(2)} \right). \quad (\text{П6})$$

При выводе формулы (П6) использованы приближённые соотношения $\tilde{P}^2/P^2 = \tilde{\eta}/\eta_a = \tilde{v}/v_a$, $P_c^2 = 4v_a/u$, выполняющиеся в рассматриваемом случае. Учитывая первое неравенство в (21), нетрудно убедиться, что $|\Delta_1| \ll |\Delta_0|$, т. е. членом Δ_1 в уравнении (П1) можно пренебречь.

Оставшиеся члены $\Delta_2, \Delta_3, \Delta_4$ могут быть упрощены, если использовать приближённые соотношения

$$M_{11}L_{22} \approx \frac{v_a}{p^2 u} \left(\frac{\tilde{v}}{v_a} - 1 \right)^2, \quad M_{22}L_{11} \approx \frac{\tilde{v}}{v_a} \left(\frac{v_a}{p^2 u} \right)^3 \left(\frac{\tilde{v}}{v_a} - 1 \right)^2,$$

$$\frac{\tilde{g}\tilde{\eta}}{\tilde{\varepsilon} - \tilde{\eta}} \approx \tilde{P}^2, \quad \frac{g_a \eta_a}{\varepsilon_a - \eta_a} \approx P^2,$$

а также формулы (22), (ПЗ), (П4). В результате уравнение (П1) принимает вид

$$\left(J_m^{(1)} + K_m^{(1)} \right) \left(\frac{\tilde{\eta}}{\eta_a} J_m^{(2)} + K_m^{(2)} \right) = \frac{v_a}{p^2 u} \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 J_m^{(1)} K_m^{(1)} + \frac{\tilde{v}}{v_a} \left(\frac{v_a}{p^2 u} \right)^3 \left(\frac{\tilde{v}}{v_a} - 1 \right)^2 J_m^{(2)} K_m^{(2)} +$$

$$+ \frac{v_a}{p^2 u} \left(\frac{\tilde{v}}{v_a} - 1 \right) \left[\frac{m}{(k_0 a P)^2} J_m^{(1)} + \frac{m}{(k_0 a \tilde{P})^2} K_m^{(1)} \right] -$$

$$- \frac{\tilde{v}}{v_a} \left(\frac{v_a}{p^2 u} \right)^2 \left(\frac{\tilde{v}}{v_a} - 1 \right) \left[\frac{m}{(k_0 a P)^2} J_m^{(2)} + \frac{m}{(k_0 a \tilde{P})^2} K_m^{(2)} \right]. \quad (\text{П7})$$

Наконец, как нетрудно убедиться, при условиях (21) последнее слагаемое в правой части (П7) даёт малый вклад в уравнение по сравнению с членами, стоящими в левой части, и поэтому может быть отброшено. Окончательно, принимая во внимание формулы (20), (ПЗ), после ряда очевидных преобразований приходим к уравнению (23).

СПИСОК ЛИТЕРАТУРЫ

1. Kondrat'ev I. G., Kudrin A. V., Zaboronkova T. M. Electrodynamics of density ducts in magnetized plasmas. Amsterdam: Gordon and Breach, 1999. 288 p.
2. Вдовиченко И. А., Марков Г. А., Миронов В. А., Сергеев А. М. // Письма в ЖЭТФ. 1986. Т. 44, вып. 5. С. 216.

3. Голубятников Г. Ю., Егоров С. В., Костров А. В., Мареев Е. А., Чугунов Ю. В. // ЖЭТФ. 1989. Т. 96, вып. 6 (12). С. 2009.
4. Заборонкова Т. М., Костров А. В., Кудрин А. В., Тихонов С. В., Тронин А. В., Шайкин А. А. // ЖЭТФ. 1992. Т. 102, вып. 4 (10). С. 1151.
5. Kostrov A. V., Kudrin A. V., Kurina L. E., Luchinin G. A., Shaykin A. A., Zaboronkova T. M. // Physica Scripta. 2000. V. 62, Pt. 1. P. 51.
6. Helliwell R. A. Whistlers and related ionospheric phenomena. Stanford: Stanford University Press, 1965. 365 p.
7. Helliwell R. A. // Modern Radio Science 1993 / Ed. by H. Matsumoto. New York: Oxford University Press, 1993. P. 189.
8. Sazhin S., Hayakawa M., Bullough K. // Ann. Geophys. 1992. V. 10. P. 293.
9. Kondrat'ev I. G., Kudrin A. V., Zaboronkova T. M. // J. Atmos. Sol.-Terr. Phys. 1997. V. 59, No. 18. P. 2475.
10. Kudrin A. V., Lyakh M. Yu., Zaboronkova T. M. // IEEE Trans. Antennas Propagat. 2001. V. 49, No. 12. P. 1645.
11. Агафонов Ю. Н., Бажанов В. С., Исякаев В. Я., Марков Г. А., Похунков А. А., Чугунов Ю. В., Кулишников С. А. // Письма в ЖЭТФ. 1990. Т. 52, вып. 10. С. 1127.
12. Чугунов Ю. В., Марков Г. А. // J. Atmos. Sol.-Terr. Phys. 2001. V. 63, No. 17. P. 1775.
13. Голубятников Г. Ю., Егоров С. В., Ерёмин Б. Г., Литвак А. Г., Стриковский А. В., Толкачёва О. Н., Чугунов Ю. В. // ЖЭТФ. 1995. Т. 107, вып. 2. С. 441.
14. Chen F. F., Boswell R. W. // IEEE Trans. Plasma Sci. 1997. V. 25, No. 6. P. 1245.
15. Эллис В., Буксбаум С., Берс А. Волны в анизотропной плазме. М.: Атомиздат, 1966. 132 с.
16. Кондратенко А. Н. Плазменные волноводы. М.: Атомиздат, 1976. 232 с.
17. Кауфман Р. Н., Рябова Н. А. // Физика плазмы. 1980. Т. 6, вып. 5. С. 1027.
18. Shamrai K. P., Tarapov V. B. // Plasma Sources Sci. Technol. 1996. V. 5, No. 3. P. 474.
19. Arnush D., Chen F. F. // Phys. Plasmas. 1998. V. 5, No. 5. P. 1239.
20. Shivarova A., Tarnev Kh. // Plasma Sources Sci. Technol. 2001. V. 10, No. 2. P. 260.
21. Karpman V. I., Kaufman R. N. // J. Plasma Phys. 1982. V. 27, Pt. 2. P. 225.
22. Заборонкова Т. М., Кудрин А. В., Марков Г. А. // Физика плазмы. 1993. Т. 19, вып. 6. С. 769.
23. Laird M. J., Nunn D. // Planet. Space Sci. 1975. V. 23, No. 12. P. 1649.
24. Кауфман Р. Н. // Изв. вузов. Радиофизика. 1985. Т. 28, № 5. С. 566.
25. Марков Г. А. // Физика плазмы. 1988. Т. 14, вып. 9. С. 1094.
26. Кудрин А. В., Марков Г. А. // Изв. вузов. Радиофизика. 1991. Т. 34, № 2. С. 163.
27. Кудрин А. В., Курина Л. Е., Марков Г. А. // ЖЭТФ. 1997. Т. 112, вып. 4 (10). С. 1285.
28. Гинзбург В. Л. Распространение электромагнитных волн в плазме. М.: Наука, 1967. 684 с.
29. Вайнштейн Л. А. Электромагнитные волны. М.: Радио и связь, 1988. 440 с.
30. Seshadri S. R. // Proc. IEE. 1965. V. 112, No. 8. P. 1497.
31. Adachi S. // Radio Sci. 1966. V. 1, No. 4. P. 671.
32. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука, 1984. 832 с.

¹ Нижегородский государственный технический университет,

² Нижегородский государственный университет
им. Н. И. Лобачевского, г. Нижний Новгород, Россия

Поступила в редакцию
7 мая 2002 г.

NONSYMMETRIC WHISTLER WAVES GUIDED BY CYLINDRICAL DUCTS WITH ENHANCED PLASMA DENSITY

T. M. Zaboronkova, A. V. Kudrin, M. Yu. Lyakh, and L. L. Popova

We study the guided propagation of whistler waves whose fields depend on the azimuthal angle in cylindrical plasma-waveguide channels (density ducts) aligned with an external magnetic field and surrounded by a uniform magnetoplasma. The main attention is paid to ducts with enhanced plasma density. It is shown that under certain conditions, such ducts are capable of supporting proper (eigen) modes and improper leaky modes. We present the results of analysis of the dispersion properties and the field structures of nonsymmetric modes guided by cylindrical ducts in the whistler band.

УДК 537.8.029.6

ЭФФЕКТИВНЫЙ ИМПЕДАНС СЛАБОШЕРОХОВАТОЙ ГРАНИЦЫ ПРОИЗВОЛЬНО АНИЗОТРОПНОЙ СРЕДЫ

С. Н. Шулъга

Рассмотрена краевая задача для электромагнитного поля в произвольно анизотропном полупространстве с неровной границей. Найдено импедансное граничное условие с нелокальным импедансом, которое в рамках теории возмущений учитывает члены четвёртого порядка по малой высоте неровностей. В дополнение к общему случаю рассмотрен случай идеально проводящей поверхности с гауссовскими неровностями.

В предлагаемой статье решена граничная задача для электромагнитного поля в полупространстве с неровной поверхностью. Следуя известной схеме [1, 2], мы преобразуем исходную задачу к эквивалентной в области с ровной границей и несколько иными граничными условиями. Конечная цель — уточнить результаты цитированных работ, которые получены приближённо с учётом квадратичных по малой высоте неровностей членов разложения.

Вначале будем считать неровности детерминированными: для них с помощью точной методики переноса граничных условий получим точное нелокальное граничное условие на ровной плоской поверхности. Предположив затем, что речь идёт об одной из возможных реализаций неровностей, рассмотрим вопрос о нелокальных граничных условиях для статистически средних векторов поля. Итоговое выражение, которое учитывает члены четвёртого порядка малости по высоте неровностей, относится к случаю, когда полупространство над неровной границей заполнено произвольной анизотропной средой. Полученные результаты полезны, к примеру, при вычислении высших поправок теории возмущений в задаче распространения радиоволн в приводном волноводе или в планарном волноводе интегральной оптики [3, 4].

1. Введём декартовую систему координат (x, y, z) так, чтобы ось z для определённости была направлена вверх. Область V^{in} постановки задачи определим условиями $z_S(\mathbf{r}) < z < +\infty$, $-\infty < x < +\infty$, $-\infty < y < +\infty$. Здесь $z_S(\mathbf{r})$ — некоторая функция переменной $\mathbf{r} = (x, y)$, определяющая границу S рассматриваемой области. Среда в области V^{in} считается произвольно-неоднородной, анизотропной и характеризуется материальными параметрами $\hat{\epsilon}(\mathbf{R})$, $\hat{\mu}(\mathbf{R})$. Поверхность S предполагается импедансной и описывается скалярным нелокальным импедансом $\zeta(\mathbf{r})$, зависящим, вообще говоря, от точек \mathbf{r} поверхности S .

Система уравнений для электромагнитного поля включает уравнения Максвелла в области V^{in} :

$$[\nabla, \mathbf{E}(\mathbf{R})] - ik_0 \hat{\mu}(\mathbf{R}) \mathbf{H}(\mathbf{R}) = -(4\pi/c) \mathbf{M}(\mathbf{R}), \quad [\nabla, \mathbf{H}(\mathbf{R})] + ik_0 \hat{\epsilon}(\mathbf{R}) \mathbf{E}(\mathbf{R}) = (4\pi/c) \mathbf{J}(\mathbf{R}), \quad (1)$$

с заданными детерминированными источниками \mathbf{J} , \mathbf{M} и граничное уравнение Леонтовича—Щукина на поверхности S :

$$\mathbf{E}_t(\mathbf{R}_S) - \zeta [\mathbf{n}_S(\mathbf{r}), \mathbf{H}_t(\mathbf{R}_S)] = 0. \quad (2)$$

Здесь $\mathbf{R} = (x, y, z)$, $k_0 = \omega/c$, c — скорость света в вакууме, ω — частота электромагнитного поля, $\mathbf{E}_t(\mathbf{R}_S)$ и $\mathbf{H}_t(\mathbf{R}_S)$ — компоненты электрического и магнитного полей, касательные к поверхности S в точке \mathbf{R}_S , $\mathbf{n}_S(\mathbf{r})$ — единичный вектор нормали к S , направленный внутрь области V^{in} :

$$\mathbf{n}_S(\mathbf{r}) = \frac{\mathbf{z}_0 - \gamma(\mathbf{r})}{\sqrt{1 + \gamma^2}}, \quad (3)$$

$\boldsymbol{\gamma}(\mathbf{r})$ — вектор наклона поверхности S в данной точке: $\boldsymbol{\gamma}(\mathbf{r}) = \nabla z_S(\mathbf{r})$, \mathbf{z}_0 — орт оси z .

Описанная выше задача поставлена в области, граница S которой в общем случае отличается от плоскости.

2. Перейдём согласно намеченной программе к новой задаче, которая эквивалентна в определённом смысле исходной, но задана в новой области V с границей в виде плоскости $z = d$. Здесь d — некоторая константа. Обозначим

$$z_{\max} = \max_{\mathbf{r}} z_S(\mathbf{r}), \quad z_{\min} = \min_{\mathbf{r}} z_S(\mathbf{r}); \quad b = \max(z_{\max}, d), \quad a = \min(z_{\min}, d). \quad (4)$$

Считаем, что величины z_{\max} и z_{\min} — конечные, т. е. неровная поверхность заключена в слое $a < z < b$ конечной толщины $b - a$. Для упрощения выкладок уточним модель среды, положив, что в прилегающем к поверхности S слое $z_S(\mathbf{r}) < z < b$ области V магнитоэлектрическая среда однородная, изотропная и характеризуется материальными параметрами ε_1, μ_1 :

$$\hat{\boldsymbol{\varepsilon}}(\mathbf{R}) = \varepsilon_1 \hat{\mathbf{I}}, \quad \hat{\boldsymbol{\mu}}(\mathbf{R}) = \mu_1 \hat{\mathbf{I}}, \quad (5)$$

где ε_1 и μ_1 — некоторые постоянные, $\hat{\mathbf{I}}$ — единичная диада. Обозначения полей при $z < b$ для ясности снабжены нижним индексом 1.

Запишем для векторов $\mathbf{E}_1(\mathbf{R}) = \mathbf{E}_1(\mathbf{r}, z)$ разложение в ряд Тейлора:

$$\mathbf{E}_1(\mathbf{r}, z + h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} \partial_z^n \mathbf{E}_1(\mathbf{r}, z) = \sum_{p=0}^{\infty} \left[\frac{h^{2p}}{(2p)!} \partial_z^{2p} + \frac{h^{2p+1}}{(2p+1)!} \partial_z^{2p+1} \right] \mathbf{E}_1(\mathbf{r}, z), \quad (6)$$

где $h = \text{const}$. Учтём, что поле $\mathbf{E}_1(\mathbf{r}, z)$ при $z_S(\mathbf{r}) < z < b$ удовлетворяет однородному волновому уравнению

$$(\partial_z^2 + k_1^2 + \nabla_t^2) \mathbf{E}_1(\mathbf{r}, z) = 0. \quad (7)$$

Здесь и далее ∇_t — компонента векторного оператора ∇ , компланарная плоскости $z = 0$; $k_1^2 = k_0^2 \varepsilon_1 \mu_1$. Заменяем ∂_z^{2p} на $(-1)^p v_1^{2p}$, где $v_1 \equiv (k_1^2 + \nabla_t^2)^{1/2}$, в соответствии со следствием из уравнения (7):

$$\partial_z^{2p} = (-1)^p (k_1^2 + \nabla_t^2)^p. \quad (8)$$

Тогда формулы связи для $\mathbf{E}_1(\mathbf{r}, z)$ запишутся следующим образом:

$$\mathbf{E}_1(\mathbf{r}, z + h) = [\cos(hv_1) + \sin(hv_1)v_1^{-1} \partial_z] \mathbf{E}_1(\mathbf{r}, z). \quad (9)$$

Сходная формула связи для $\mathbf{H}_1(\mathbf{r}, z)$ получается совершенно аналогично:

$$\mathbf{H}_1(\mathbf{r}, z + h) = [\cos(hv_1) + \sin(hv_1)v_1^{-1} \partial_z] \mathbf{H}_1(\mathbf{r}, z). \quad (10)$$

В (9), (10) использованы формулы разложения функций $\cos(hv_1)$ и $\sin(hv_1)$ в степенные ряды.

Используя уравнения Максвелла (1), выразим в правой части формул связи векторы поля через горизонтальные (компланарные плоскости $z = 0$) компоненты \mathbf{E}_{1t} , \mathbf{H}_{1t} :

$$\begin{aligned} \mathbf{E}_1(\mathbf{r}, z + h) = & [\cos(hv_1) - \sin(hv_1)v_1^{-1} \mathbf{z}_0 \nabla_t] \mathbf{E}_{1t}(\mathbf{r}, z) + \\ & + \frac{1}{ik_0 \varepsilon_1} [\cos(hv_1) \mathbf{z}_0 \nabla_t + \sin(hv_1)v_1^{-1} (\nabla_t \nabla_t + k_1^2)] [\mathbf{z}_0, \mathbf{H}_{1t}(\mathbf{r}, z)], \end{aligned} \quad (11)$$

$$\begin{aligned} \mathbf{H}_1(\mathbf{r}, z + h) = & [\cos(hv_1) - \sin(hv_1)v_1^{-1} \mathbf{z}_0 \nabla_t] \mathbf{H}_{1t}(\mathbf{r}, z) - \\ & - \frac{1}{ik_0 \mu_1} [\cos(hv_1) \mathbf{z}_0 \nabla_t + \sin(hv_1)v_1^{-1} (\nabla_t \nabla_t + k_1^2)] [\mathbf{z}_0, \mathbf{E}_{1t}(\mathbf{r}, z)]. \end{aligned} \quad (12)$$

Формулы связи по своему построению справедливы только в пределах области $z_S(\mathbf{r}) < z < b$, прилегающей к неровной границе. Легко, однако, проверить, что векторы поля в левой части формул связи подчиняются однородным уравнениям Максвелла с материальными параметрами (5) в любой точке пространства. Тем самым формулы (11) и (12) определяют соответствующий способ аналитического продолжения векторов поля из первоначально избранной области $z_S(\mathbf{r}) < z < b$. Именно такой трактовки соотношений (11), (12) и будем придерживаться в дальнейшем.

Займёмся теперь переносом граничного условия (2) с неровной поверхности S на плоскость $z = d$. Обозначим $\xi(\mathbf{r}) = z_S(\mathbf{r}) - d$ и положим в формулах связи $h = \xi(\mathbf{r})$, $z = d$. В результате получим выражения для векторов $\mathbf{E}_1(\mathbf{r}, z_S)$, $\mathbf{H}_1(\mathbf{r}, z_S)$ через компоненты $\mathbf{E}_{1t}(\mathbf{r}, d)$, $\mathbf{H}_{1t}(\mathbf{r}, d)$. Подставив полученные выражения для $\mathbf{E}_1(\mathbf{r}, z_S)$, $\mathbf{H}_1(\mathbf{r}, z_S)$ в импедансное граничное условие (2), приходим в итоге к соотношению для компонент $\mathbf{E}_{1t}(\mathbf{r}, d)$, $\mathbf{H}_{1t}(\mathbf{r}, d)$:

$$\hat{A}\mathbf{E}_{1t}(\mathbf{r}, d) + \hat{B}[\mathbf{z}_0, \mathbf{H}_{1t}(\mathbf{r}, d)] = 0. \tag{13}$$

Здесь \hat{A} и \hat{B} — операторы, действующие по переменной \mathbf{r} в соответствии с формулами

$$\begin{aligned} \hat{A}(\mathbf{r}) &= \cos(\xi v_1) - \gamma \sin(\xi v_1) v_1^{-1} \nabla_t - \\ &\quad - \frac{\zeta(\mathbf{r})}{\sqrt{1 + \gamma^2}} \frac{1}{ik_0\mu_1} \{ [\mathbf{z}_0, \gamma] \cos(\xi v_1) [\mathbf{z}_0, \nabla_t] + (1 + \gamma\gamma) \sin(\xi v_1) v_1^{-1} ([\mathbf{z}_0, \nabla_t] [\mathbf{z}_0, \nabla_t] + k_1^2) \}, \\ \hat{B}(\mathbf{r}) &= \frac{1}{ik_0\varepsilon_1} [\nabla_t \sin(\xi v_1) v_1^{-1} \nabla_t + k_1^2 \sin(\xi v_1) v_1^{-1}] + \\ &\quad + \frac{\zeta(\mathbf{r})}{\sqrt{1 + \gamma^2}} \{ [\mathbf{z}_0, \gamma] \sin(\xi v_1) v_1^{-1} [\mathbf{z}_0, \nabla_t] - (1 + \gamma\gamma) \cos(\xi v_1) \}. \end{aligned} \tag{14}$$

Напомним, что в этих формулах оператор v_1 не действует по переменной \mathbf{r} , содержащейся в $\xi(\mathbf{r})$ и $\gamma(\mathbf{r})$; напротив, оператор ∇_t действует по переменной \mathbf{r} во всех функциях, стоящих справа от него; два рядом стоящих вектора образуют диаду.

Соотношение (13) имеет смысл нелокального граничного условия для поля на плоскости $z = d$. Выше этой плоскости, в пределах слоя $d < z < b$, поле подчиняется однородным уравнениям Максвелла с материальными параметрами (5), а при $z > b$ оно удовлетворяет неоднородным (со сторонними источниками) уравнениям (1). Таким образом, исходная задача для поля в области V^{in} с неровной границей заменена новой — в области V , граница которой является плоской.

3. Перейдём к ситуации, когда поверхность S испытывает случайные флуктуации как формы, так и физических свойств.

Пусть $z_S(\mathbf{r})$ — функция, определяющая неровную поверхность, и $\zeta(\mathbf{r})$ — импеданс этой поверхности представляют собой случайные величины, зависящие от \mathbf{r} . При этом операторы \hat{A} и \hat{B} также будут случайными. В силу того, что электромагнитное поле удовлетворяет на поверхности $z = d$ граничным условиям со случайными операторами, оно также является случайным.

Пусть $\langle \mathbf{E}(\mathbf{R}) \rangle$ и $\langle \mathbf{H}(\mathbf{R}) \rangle$ — статистически средние компоненты случайных полей $\mathbf{E}(\mathbf{R})$ и $\mathbf{H}(\mathbf{R})$. Построим эквивалентные граничные операторы $\hat{\mathbf{A}}_e(\mathbf{r})$ и $\hat{\mathbf{B}}_e(\mathbf{r})$, такие, чтобы для среднего поля при $z = d + 0$ выполнялось эквивалентное граничное условие вида

$$\hat{\mathbf{A}}_e(\mathbf{r}) \langle \mathbf{E}_t(\mathbf{r}, d) \rangle + \hat{\mathbf{B}}_e(\mathbf{r}) [\mathbf{z}_0, \langle \mathbf{H}_t(\mathbf{r}, d) \rangle] = 0. \tag{15}$$

Для построения операторов $\hat{\mathbf{A}}_e$, $\hat{\mathbf{B}}_e$ воспользуемся алгебраической техникой, изложенной, например, в [5]. Выделим в ζ статистически среднюю и флуктуационную компоненты:

$$\zeta = \bar{\zeta} + \zeta_f, \quad \bar{\zeta} = \langle \zeta \rangle, \quad \zeta_f = \zeta - \langle \zeta \rangle \tag{16}$$

(здесь угловые скобки обозначают статистическое усреднение), и представим операторы $\hat{\mathbf{A}}_e, \hat{\mathbf{B}}_e$ в виде

$$\hat{\mathbf{A}}_e = \hat{\mathbf{I}} + \hat{\mathbf{A}}, \quad \hat{\mathbf{B}}_e = -\bar{\zeta}\hat{\mathbf{I}} + \hat{\mathbf{B}}, \quad (17)$$

где $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ — некоторые случайные операторы, действующие по \mathbf{r} , $\hat{\mathbf{I}}$ — единичная диада, $\hat{\mathbf{I}} = \mathbf{x}_0\mathbf{x}_0 + \mathbf{y}_0\mathbf{y}_0 + \mathbf{z}_0\mathbf{z}_0$. Явный вид операторов $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ легко получается из определения (16) и в целях экономии места здесь не приводится.

Перепишем эквивалентное граничное условие (15) в следующем виде:

$$\mathbf{E}_t(\mathbf{r}, d) - \bar{\zeta}(\mathbf{r})[\mathbf{z}_0, \mathbf{H}_t(\mathbf{r}, d)] = \frac{4\pi}{c}[\mathbf{z}_0, \mathbf{M}_S(\mathbf{r})]. \quad (18)$$

Величина $\mathbf{M}_S(\mathbf{r})$ имеет смысл поверхностных магнитных источников и определяется выражением

$$\mathbf{M}_S(\mathbf{r}) = \hat{\mathbf{V}}_e(\mathbf{r})\mathbf{E}_t(\mathbf{r}) + \hat{\mathbf{V}}_m(\mathbf{r})\mathbf{H}_t(\mathbf{r}). \quad (19)$$

Случайные операторы возмущения $\hat{\mathbf{V}}_e(\mathbf{r})$ и $\hat{\mathbf{V}}_m(\mathbf{r})$ имеют следующий вид:

$$\hat{\mathbf{V}}_e(\mathbf{r}) = \frac{c}{4\pi}[\mathbf{z}_0, \hat{\mathbf{A}}(\mathbf{r})], \quad \hat{\mathbf{V}}_m(\mathbf{r}) = \frac{c}{4\pi}[\mathbf{z}_0, [\hat{\mathbf{B}}(\mathbf{r}), \mathbf{z}_0]]. \quad (20)$$

Запишем решение краевой задачи (1) с граничным условием (18) в виде суммы невозмущённого поля $\mathbf{E}_0(\mathbf{R}), \mathbf{H}_0(\mathbf{R})$ и поля поверхностных источников $\mathbf{M}_S(\mathbf{r})$ [4]:

$$\mathbf{E}(\mathbf{R}) = \mathbf{E}_0(\mathbf{R}) + \int \hat{\mathbf{G}}_{em}(\mathbf{R}, \mathbf{r}')\mathbf{M}_S(\mathbf{r}') d\mathbf{r}', \quad \mathbf{H}(\mathbf{R}) = \mathbf{H}_0(\mathbf{R}) + \int \hat{\mathbf{G}}_{mm}(\mathbf{R}, \mathbf{r}')\mathbf{M}_S(\mathbf{r}') d\mathbf{r}', \quad (21)$$

которое в операторном виде будем записывать как

$$\mathbf{E} = \mathbf{E}_0 + \hat{\mathbf{G}}_{em}\mathbf{M}_S, \quad \mathbf{H} = \mathbf{H}_0 + \hat{\mathbf{G}}_{mm}\mathbf{M}_S. \quad (22)$$

Величины $\mathbf{E}_0(\mathbf{R}), \mathbf{H}_0(\mathbf{R})$ определены как решение «невозмущённой» краевой задачи

$$\begin{aligned} [\nabla, \mathbf{E}_0(\mathbf{R})] - ik_0\hat{\mu}(\mathbf{R})\mathbf{H}_0(\mathbf{R}) &= -(4\pi/c)\mathbf{M}(\mathbf{R}), \\ [\nabla, \mathbf{H}_0(\mathbf{R})] + ik_0\hat{\varepsilon}(\mathbf{R})\mathbf{E}_0(\mathbf{R}) &= (4\pi/c)\mathbf{J}(\mathbf{R}), \\ \mathbf{E}_{0t}(\mathbf{r}, z) - \bar{\zeta}[\mathbf{z}_0, \mathbf{H}_{0t}(\mathbf{r}, z)] &= 0, \end{aligned} \quad (23)$$

где $z = d + 0$, в отсутствие магнитных источников $\mathbf{M}_S(\mathbf{r})$. В формулах (21) $\hat{\mathbf{G}}_{em}(\mathbf{R}, \mathbf{r}')$, $\hat{\mathbf{G}}_{mm}(\mathbf{R}, \mathbf{r}')$ — функции Грина уравнений Максвелла поверхностных источников магнитного типа $\mathbf{M}_S(\mathbf{r})$. Выразив $\mathbf{M}_S(\mathbf{r})$ через векторы поля с помощью формул (19), получим следующие соотношения:

$$\mathbf{E} = \mathbf{E}_0 + \hat{\mathbf{G}}_{em}(\hat{\mathbf{V}}_e\mathbf{E} + \hat{\mathbf{V}}_m\mathbf{H}), \quad \mathbf{H} = \mathbf{H}_0 + \hat{\mathbf{G}}_{mm}(\hat{\mathbf{V}}_e\mathbf{E} + \hat{\mathbf{V}}_m\mathbf{H}). \quad (24)$$

Введём в рассмотрение «векторы состояния»

$$\Psi = \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix}, \quad \Psi_0 = \begin{pmatrix} \mathbf{E}_0 \\ \mathbf{H}_0 \end{pmatrix} \quad (25)$$

и матричные операторы

$$\mathcal{G} = \begin{pmatrix} \hat{\mathbf{G}}_{em} & 0 \\ \hat{\mathbf{G}}_{mm} & 0 \end{pmatrix}, \quad \mathcal{V} = \begin{pmatrix} \hat{\mathbf{V}}_m & \hat{\mathbf{V}}_e \\ 0 & 0 \end{pmatrix}. \quad (26)$$

Тогда пара уравнений (24) относительно случайных векторов \mathbf{E} , \mathbf{H} запишется в форме уравнений Липмана—Швингера:

$$\Psi = \Psi_0 + \mathcal{G}\mathcal{V}\Psi. \quad (27)$$

Именно для уравнений такого типа соответствующая «алгебраическая» техника усреднения изложена в [5]. Следуя ей, построим эквивалентное граничное условие для средней величины $\langle \Psi \rangle$. Для этого нужно располагать детерминированным оператором \mathbf{M} — массовым оператором, обладающим свойством

$$\langle \mathcal{V}\Psi \rangle \equiv \mathbf{M}\langle \Psi \rangle. \quad (28)$$

Последовательность соответствующих выкладок такова. Из (27) для Ψ получаем

$$\Psi = (\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \Psi_0, \quad \langle \Psi \rangle = \langle (\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \rangle \Psi_0. \quad (29)$$

Отсюда вытекает выражение для Ψ_0 :

$$\Psi_0 = \langle (\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \rangle^{-1} \langle \Psi \rangle. \quad (30)$$

Поддействуем оператором \mathcal{V} на Ψ из первого уравнения (29), а затем усредним полученное соотношение. В итоге получаем

$$\langle \mathcal{V}\Psi \rangle = \langle \mathcal{V}(\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \rangle \Psi_0. \quad (31)$$

После подстановки Ψ_0 из (30) получаем окончательное выражение для $\langle \mathcal{V}\Psi \rangle$ из (28), в котором массовый оператор \mathbf{M} имеет вид

$$\mathbf{M} = \langle \mathcal{V}(\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \rangle \langle (\hat{\mathbf{I}} - \mathcal{G}\mathcal{V})^{-1} \rangle^{-1}. \quad (32)$$

Из этого выражения легко получается разложение массового оператора по степеням \mathcal{V} (корреляционным группам [6]):

$$\mathbf{M} = \sum_{n=1}^{\infty} \mathbf{M}_n, \quad (33)$$

где \mathbf{M}_n — величина n -го порядка малости по \mathcal{V} . Первые четыре члена этого разложения таковы:

$$\begin{aligned} \mathbf{M}_1 &= \langle \mathcal{V} \rangle, & \mathbf{M}_2 &= \langle \mathcal{V}\mathcal{G}\mathcal{V} \rangle - \langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V} \rangle, \\ \mathbf{M}_3 &= \langle (\mathcal{V}\mathcal{G})^2 \mathcal{V} \rangle - \langle (\mathcal{V}\mathcal{G})^2 \rangle \langle \mathcal{V} \rangle + \langle \mathcal{V}\mathcal{G} \rangle [\langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V} \rangle - \langle \mathcal{V}\mathcal{G}\mathcal{V} \rangle], \\ \mathbf{M}_4 &= \langle (\mathcal{V}\mathcal{G})^3 \mathcal{V} \rangle - \langle (\mathcal{V}\mathcal{G})^3 \rangle \langle \mathcal{V} \rangle + \langle (\mathcal{V}\mathcal{G})^2 \rangle [\langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V} \rangle - \langle \mathcal{V}\mathcal{G}\mathcal{V} \rangle] + \\ &+ \langle \mathcal{V}\mathcal{G} \rangle [\langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V}\mathcal{G}\mathcal{V} \rangle + \langle (\mathcal{V}\mathcal{G})^2 \rangle \langle \mathcal{V} \rangle - \langle (\mathcal{V}\mathcal{G})^2 \mathcal{V} \rangle - \langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V}\mathcal{G} \rangle \langle \mathcal{V} \rangle]. \end{aligned} \quad (34)$$

С учётом явного вида матричных операторов, входящих в (33), запишем

$$\mathbf{M} = \begin{pmatrix} \hat{\mathbf{M}}_m & \hat{\mathbf{M}}_e \\ 0 & 0 \end{pmatrix}, \quad (35)$$

где $\hat{\mathbf{M}}_e$, $\hat{\mathbf{M}}_m$ — детерминированные диадные операторы, действующие по переменной \mathbf{r} . Они также представляются в виде, аналогичном (33):

$$\hat{\mathbf{M}}_e = \sum_{n=1}^{\infty} \hat{\mathbf{M}}_e^{(n)}, \quad \hat{\mathbf{M}}_m = \sum_{n=1}^{\infty} \hat{\mathbf{M}}_m^{(n)}, \quad (36)$$

где $\hat{\mathbf{M}}_e^{(n)}$, $\hat{\mathbf{M}}_m^{(n)}$ — операторы, имеющие n -й порядок малости относительно операторов возмущения. Первые четыре члена разложения (36) получаются из (34):

$$\begin{aligned}\hat{\mathbf{M}}_\alpha^{(1)} &= \langle \hat{\mathbf{V}}_\alpha \rangle, & \hat{\mathbf{M}}_\alpha^{(2)} &= \langle \hat{\eta} \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta} \rangle \langle \hat{\mathbf{V}}_\alpha \rangle, \\ \hat{\mathbf{M}}_\alpha^{(3)} &= \langle \hat{\eta}^2 \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta}^2 \rangle \langle \hat{\mathbf{V}}_\alpha \rangle + \langle \hat{\eta} \rangle [\langle \hat{\eta} \rangle \langle \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta} \hat{\mathbf{V}}_\alpha \rangle], \\ \hat{\mathbf{M}}_\alpha^{(4)} &= \langle \hat{\eta}^3 \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta}^3 \rangle \langle \hat{\mathbf{V}}_\alpha \rangle + \langle \hat{\eta}^2 \rangle [\langle \hat{\eta} \rangle \langle \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta} \hat{\mathbf{V}}_\alpha \rangle] + \\ &+ \langle \hat{\eta} \rangle [\langle \hat{\eta}^2 \rangle \langle \hat{\mathbf{V}}_\alpha \rangle + \langle \hat{\eta} \rangle \langle \hat{\eta} \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta}^2 \hat{\mathbf{V}}_\alpha \rangle - \langle \hat{\eta} \rangle \langle \hat{\eta} \rangle \langle \hat{\mathbf{V}}_\alpha \rangle],\end{aligned}\quad (37)$$

где индекс α принимает значения e и m ; $\hat{\eta} = \hat{\mathbf{V}}_m \hat{\mathbf{G}}_{mm} + \hat{\mathbf{V}}_e \hat{\mathbf{G}}_{em}$.

Располагая операторами $\hat{\mathbf{M}}_e$, $\hat{\mathbf{M}}_m$, которые также будем называть массовыми, запишем эквивалентное граничное условие для средних векторов $\langle \mathbf{E} \rangle$, $\langle \mathbf{H} \rangle$ в следующем виде:

$$\langle \mathbf{E}_t(\mathbf{r}, d) \rangle - \bar{\zeta} [\mathbf{z}_0, \langle \mathbf{H}_t(\mathbf{r}, d) \rangle] - \frac{4\pi}{c} [\mathbf{z}_0, \hat{I}] [\hat{\mathbf{M}}_e \langle \mathbf{E}_t(\mathbf{r}, d) \rangle + \hat{\mathbf{M}}_m \langle \mathbf{H}_t(\mathbf{r}, d) \rangle] = 0. \quad (38)$$

Последнее выражение наряду с (13) составляет главный результат данной работы. Значимость его состоит в том, что в нём удалось учесть влияние случайных неровностей и флуктуаций импеданса на статистически среднее поле посредством граничного условия на ровной поверхности — плоскости $z = d$. Для элементарных электродинамических моделей варианты эквивалентных граничных условий типа (38), различающиеся степенью точности, с которой находились операторы $\hat{\mathbf{M}}_e$, $\hat{\mathbf{M}}_m$, приведены в работах [4, 7].

4. В электродинамике фундаментальное значение имеет оператор импеданса [4]. Введём его в данном случае как оператор, обеспечивающий выполнение следующего условия для статистически среднего поля:

$$\langle \mathbf{E}_t(\mathbf{r}, d) \rangle - \hat{\mathcal{L}}_e [\mathbf{z}_0, \langle \mathbf{H}_t(\mathbf{r}, d) \rangle] = 0. \quad (39)$$

Сопоставляя эту запись с (38), приходим к строгому уравнению относительно оператора эквивалентного импеданса $\hat{\mathcal{L}}_e$:

$$\left(\hat{\mathbf{I}}_t - \frac{4\pi}{c} [\mathbf{z}_0, \hat{\mathbf{M}}_e] \right) \hat{\mathcal{L}}_e = \bar{\zeta} \hat{\mathbf{I}} - \frac{4\pi}{c} [\mathbf{z}_0, [\hat{\mathbf{M}}_m, \mathbf{z}_0]], \quad (40)$$

где $\hat{\mathbf{I}}_t = \hat{\mathbf{I}} - \mathbf{z}_0 \mathbf{z}_0$. Формальное его решение даётся формулой

$$\hat{\mathcal{L}}_e = \left(\hat{\mathbf{I}}_t - \frac{4\pi}{c} [\mathbf{z}_0, \hat{\mathbf{M}}_e] \right)^{-1} \left[\bar{\zeta} \hat{\mathbf{I}} - \frac{4\pi}{c} [\mathbf{z}_0, [\hat{\mathbf{M}}_m, \mathbf{z}_0]] \right]. \quad (41)$$

Будем пользоваться представлением оператора $\hat{\mathcal{L}}_e$ в виде разложения по степеням операторов возмущения. Оно получается из (41), если разложить оператор $\left(\hat{\mathbf{I}}_t - (4\pi/c) [\mathbf{z}_0, \hat{\mathbf{M}}_e] \right)^{-1}$ в ряд по степеням $\hat{\mathbf{M}}_e$, а затем воспользоваться формулами (36) для массовых операторов.

Не задаваясь целью исследовать в полном объёме свойства $\hat{\mathcal{L}}_e$, отметим следующее. Пусть случайные функции ζ , z_S — статистически однородные и однородно связанные (в узком смысле [2]), а среда в области $V(z > d)$ плоскостойкая:

$$\hat{\epsilon}(\mathbf{R}) = \hat{\epsilon}(z), \quad \hat{\mu}(\mathbf{R}) = \hat{\mu}(z), \quad (42)$$

где $\hat{\epsilon}(z)$, $\hat{\mu}(z)$ — произвольные диадные функции переменной z . Тогда, очевидно, функции $\hat{\mathbf{G}}_{\alpha\beta}(\mathbf{R}, \mathbf{r}')$ будут инварианты относительно сдвига $\mathbf{r} - \mathbf{r}'$. Как следует из (41) или из упомянутого выше разложения оператора $\hat{\mathcal{L}}_e$ по степеням операторов возмущений, действие $\hat{\mathcal{L}}_e$ на любую пространственную

гармонику вида $\exp(i\mathbf{\kappa}\mathbf{r})$ сводится к умножению её на некоторую диадную функцию $\hat{\mathcal{L}}_e$ спектрального параметра $\mathbf{\kappa} = (\kappa_x, \kappa_y, 0)$:

$$\hat{\mathcal{L}}_e(\mathbf{\kappa}) = \exp(-i\mathbf{\kappa}\mathbf{r})\hat{\mathcal{L}}_e \exp(i\mathbf{\kappa}\mathbf{r}). \tag{43}$$

Величина $\hat{\mathcal{L}}_e(\mathbf{\kappa})$ выступает в роли диады эквивалентного импеданса поверхности $z = d$.

Отмеченные выше свойства оператора $\hat{\mathcal{L}}_e$ для статистически однородной поверхности S означают, что с электродинамической точки зрения плоскость $z = d$ является однородной пространственно-диспергирующей границей.

Конкретизируем обсуждаемую ситуацию, положив, что поверхность S является идеально проводящей (и, стало быть, $\bar{\zeta} = \zeta_f \equiv 0$), её положение совпадает с плоскостью $z = d$: $\langle \xi(\mathbf{r}) \rangle \equiv 0$, и неровности являются гауссовскими. Тогда разложение оператора $\hat{\mathcal{L}}_e(\mathbf{\kappa})$ по степеням малой высоты неровностей ξ с учётом членов 4-го порядка малости приводит к следующему результату:

$$\begin{aligned} \hat{\mathcal{L}}_e(\mathbf{\kappa}) = & \frac{c}{4\pi} \frac{1}{(k_0\varepsilon_1)^2} \int d\mathbf{\kappa}' B(\mathbf{\kappa} - \mathbf{\kappa}') (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}'), \mathbf{z}_0]] (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}') - \\ & - \left(\frac{c}{4\pi}\right)^2 \frac{1}{(k_0\varepsilon_1)^2} \int d\mathbf{\kappa}' d\mathbf{\kappa}'' B(\mathbf{\kappa} - \mathbf{\kappa}'') [B(\mathbf{\kappa} - \mathbf{\kappa}') (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}'), \mathbf{z}_0]] \times \\ & \quad [\gamma^2(\mathbf{\kappa}'') - (\mathbf{\kappa}' - \mathbf{\kappa}'')\mathbf{\kappa}''] [\hat{\mathbf{G}}_{em}(\mathbf{\kappa}''), \mathbf{z}_0] (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}'') + \\ & \quad + B(\mathbf{\kappa}'' - \mathbf{\kappa}') \hat{\mathbf{I}}_t [\gamma^2(\mathbf{\kappa}') - (\mathbf{\kappa} - \mathbf{\kappa}')\mathbf{\kappa}'] [\hat{\mathbf{G}}_{em}(\mathbf{\kappa}'), \mathbf{z}_0] (k_1^2 - \mathbf{\kappa}'\mathbf{\kappa}'') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}''), \mathbf{z}_0]] (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}'') - \\ & - \left(\frac{c}{4\pi}\right)^3 \frac{1}{(k_0\varepsilon_1)^4} \int d\mathbf{\kappa}' d\mathbf{\kappa}'' d\mathbf{\kappa}''' B(\mathbf{\kappa} - \mathbf{\kappa}') [\delta(\mathbf{\kappa}' - \mathbf{\kappa}''') + \delta(\mathbf{\kappa}' + \mathbf{\kappa}''' - \mathbf{\kappa}'' - \mathbf{\kappa})] (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}'), \mathbf{z}_0]] \times \\ & \quad \times (k_1^2 - \mathbf{\kappa}'\mathbf{\kappa}'') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}''), \mathbf{z}_0]] (k_1^2 - \mathbf{\kappa}''\mathbf{\kappa}''') [\mathbf{z}_0, [\hat{\mathbf{G}}_{mm}(\mathbf{\kappa}'''), \mathbf{z}_0]] (k_1^2 - \mathbf{\kappa}\mathbf{\kappa}'''). \end{aligned} \tag{44}$$

Здесь $\delta(\mathbf{\kappa})$ — дельта-функция Дирака, $B(\mathbf{\kappa})$ — спектр высот неровностей, $\hat{\mathbf{G}}_{\alpha\beta}(\mathbf{\kappa})$ — фурье-образ функции $\hat{\mathbf{G}}_{\alpha\beta}(\mathbf{R}, \mathbf{r}')$ по разностной переменной $\mathbf{r} - \mathbf{r}'$ на плоскости $z = d$:

$$\begin{aligned} B(\mathbf{\kappa}) = & \frac{1}{2\pi} \int B(\mathbf{r}) \exp(-i\mathbf{\kappa}\mathbf{r}) d\mathbf{r}, \quad B(\mathbf{r} - \mathbf{r}') \equiv \langle \xi(\mathbf{r})\xi(\mathbf{r}') \rangle, \\ \hat{\mathbf{G}}_{\alpha\beta}(\mathbf{\kappa}) = & \lim_{z \rightarrow d} \int \hat{\mathbf{G}}_{\alpha\beta}(\mathbf{R}, \mathbf{r}') \exp[(-i\mathbf{\kappa}(\mathbf{r} - \mathbf{r}'))] d\mathbf{r}. \end{aligned} \tag{45}$$

Первое слагаемое справа в (44), линейное по спектру высот, отвечает приближению Буре для $\hat{\mathcal{L}}_e(\mathbf{\kappa})$; оно было найдено для более частной модели среды в работах [2, 7]. Остальные слагаемые, билинейные по функции B (т. е. четвёртого порядка малости по малым высотам неровностей) получены здесь, видимо, впервые. Как видно из приведённого выражения, если среда в области V симметрична:

$$\hat{\mathbf{e}}(z) = \hat{\mathbf{e}}^T(z), \quad \hat{\boldsymbol{\mu}}(z) = \hat{\boldsymbol{\mu}}^T(z) \tag{46}$$

(T — символ транспонирования диады), то диада эквивалентного импеданса также симметрична в смысле следующего равенства:

$$\hat{\mathcal{L}}_e(-\mathbf{\kappa}) = \hat{\mathcal{L}}_e^T(\mathbf{\kappa}), \tag{47}$$

которое имеет место не только в рассматриваемом приближении, но и в общем случае при выполнении условий (46).

Если же среда в области V такова, что для функции $\hat{\mathbf{G}}_{mm}(\mathbf{\kappa})$ реализуется соотношение

$$\hat{\mathbf{I}}_t \hat{\mathbf{G}}_{mm}(-\mathbf{\kappa}) \hat{\mathbf{I}}_t = \hat{\mathbf{I}}_t \hat{\mathbf{G}}_{mm}(\mathbf{\kappa}) \hat{\mathbf{I}}_t, \tag{48}$$

то, как следует из формулы (44), диада $\hat{\mathcal{L}}_e(\boldsymbol{\kappa})$ является чётной функцией переменной $\boldsymbol{\kappa}$:

$$\hat{\mathcal{L}}_e(-\boldsymbol{\kappa}) = \hat{\mathcal{L}}_e(\boldsymbol{\kappa}). \quad (49)$$

Равенства (47) и (49) выполняются, в частности, когда среда в области V изотропная, т. е. когда

$$\hat{\boldsymbol{\varepsilon}}(z) = \varepsilon(z)\hat{\mathbf{I}}, \quad \hat{\boldsymbol{\mu}}(z) = \mu(z)\hat{\mathbf{I}}. \quad (50)$$

Здесь $\varepsilon(z)$, $\mu(z)$ — некоторые кусочно-гладкие функции переменной z , принимающие на интервале $d < z < b$ постоянные значения ε_1 и μ_1 в соответствии с (5). Подчеркнём, что для статистически шероховатой поверхности под z_{\max} и z_{\min} в определении (4) надлежит понимать максимальное и минимальное значения функции $z_S(\mathbf{r})$ для всех \mathbf{r} во всех возможных её реализациях. Иными словами, предполагаем, что все реализации случайной границы $z = z_S(\mathbf{r})$ локализованы целиком в однородном слое плоскостной среды.

В случае (50) выражение (44) для диады $\hat{\mathcal{L}}_e(\boldsymbol{\kappa})$ можно записать в явном виде следующим образом:

$$\hat{\mathcal{L}}_e(\boldsymbol{\kappa}) = \hat{\mathcal{L}}_e^{(2)}(\boldsymbol{\kappa}) + \hat{\mathcal{L}}_e^{(4)}(\boldsymbol{\kappa}). \quad (51)$$

Здесь $\hat{\mathcal{L}}_e^{(n)}(\boldsymbol{\kappa})$ — величина n -го порядка малости по высоте шероховатостей:

$$\hat{\mathcal{L}}_e^{(n)}(\boldsymbol{\kappa}) = \int \hat{\mathcal{L}}_e^{(n)}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') d\boldsymbol{\kappa}',$$

где $n = 2, 4$;

$$\hat{\mathcal{L}}_e^{(2)}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') = \frac{\mu_1}{\varepsilon_1} B(\boldsymbol{\kappa} - \boldsymbol{\kappa}') (k_1^2 - \boldsymbol{\kappa}\boldsymbol{\kappa}') \hat{\mathbf{H}}(\boldsymbol{\kappa}, \boldsymbol{\kappa}'),$$

$$\begin{aligned} \hat{\mathcal{L}}_e^{(4)}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') &= [\sigma^2(\gamma^2 + \gamma'^2) - 2\chi] \hat{\mathcal{L}}_e^{(2)}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') + \frac{\mu_1}{\varepsilon_1} k_1^2 B(\boldsymbol{\kappa} - \boldsymbol{\kappa}') [\hat{\mathbf{H}}(\boldsymbol{\kappa}, \boldsymbol{\kappa}')\hat{\omega} + \hat{\omega}\hat{\mathbf{H}}(\boldsymbol{\kappa}, \boldsymbol{\kappa}')] - \\ &- \left(\frac{\mu_1}{\varepsilon_1}\right)^2 k_1^2 B(\boldsymbol{\kappa} - \boldsymbol{\kappa}') \int B(\boldsymbol{\kappa}' - \boldsymbol{\kappa}'') [\delta(\boldsymbol{\kappa}' - \boldsymbol{\kappa}''') + \delta(\boldsymbol{\kappa}' + \boldsymbol{\kappa}''' - \boldsymbol{\kappa}'' - \boldsymbol{\kappa})] (k_1^2 - \boldsymbol{\kappa}\boldsymbol{\kappa}') \times \\ &\quad \times \hat{\mathbf{H}}(\boldsymbol{\kappa}', \boldsymbol{\kappa}'') \hat{\mathbf{H}}(\boldsymbol{\kappa}, \boldsymbol{\kappa}''') d\boldsymbol{\kappa}'' d\boldsymbol{\kappa}''', \quad (52) \end{aligned}$$

$$\hat{\mathbf{H}}(\boldsymbol{\kappa}, \boldsymbol{\kappa}') = \mu_1 k_0 [\mathbf{z}_0, \boldsymbol{\kappa}] [\mathbf{z}_0, \boldsymbol{\kappa}'] / [k_1^2 \gamma_\mu(\boldsymbol{\kappa}')] + (k_1^2 \boldsymbol{\kappa}\boldsymbol{\kappa}' / \kappa'^2 - \boldsymbol{\kappa}\boldsymbol{\kappa}') / [\mu_1 k_0 \gamma_\varepsilon(\boldsymbol{\kappa}')],$$

$$i\gamma_\mu(\boldsymbol{\kappa}) = \partial_z \Psi_\mu^+(d+0, \boldsymbol{\kappa}) / \Psi_\mu^+(d+0, \boldsymbol{\kappa}), \quad i\gamma_\varepsilon(\boldsymbol{\kappa}) = \partial_z \Psi_\varepsilon^+(d+0, \boldsymbol{\kappa}) / \Psi_\varepsilon^+(d+0, \boldsymbol{\kappa}),$$

$$\hat{\omega} = \int \boldsymbol{\kappa}' \boldsymbol{\kappa}' B(\boldsymbol{\kappa}') d\boldsymbol{\kappa}', \quad \chi = \int \boldsymbol{\kappa}'^2 B(\boldsymbol{\kappa}') d\boldsymbol{\kappa}',$$

$$\gamma^2 = k_1^2 - \kappa^2, \quad \gamma'^2 = k_1^2 - \kappa'^2, \quad \sigma^2 = B(0). \quad (53)$$

При получении этих выражений использовалось представление $\hat{\mathbf{G}}_{\text{mm}}(\boldsymbol{\kappa})$ из работы [8]; функции $\Psi_\varepsilon^+(z, \boldsymbol{\kappa})$, $\Psi_\mu^+(z, \boldsymbol{\kappa})$ определены ниже.

5. Располагая выражением (51) для диады эквивалентного импеданса идеально проводящей шероховатой плоскости в изотропной среде, рассмотрим вопрос о собственных волнах среднего поля в такой структуре. Легко проверить, что однородные (без источников) уравнения Максвелла (1) допускают решения в виде пространственной гармоника

$$\mathbf{E}(\mathbf{R}) = \mathbf{E}(\boldsymbol{\kappa}, z) \exp(i\boldsymbol{\kappa}\mathbf{r}), \quad \mathbf{H}(\mathbf{R}) = \mathbf{H}(\boldsymbol{\kappa}, z) \exp(i\boldsymbol{\kappa}\mathbf{r}), \quad (54)$$

с волновым вектором $\boldsymbol{\kappa} = (\kappa_x, \kappa_y, 0)$ и векторными амплитудами

$$\begin{aligned}\mathbf{E}(\boldsymbol{\kappa}, z) &= [\mathbf{z}_0, \mathbf{n}] \mathcal{E}(z) - \frac{1}{k_0 \varepsilon(z)} [\mathbf{z}_0 \boldsymbol{\kappa} + i \mathbf{n} \partial_z] \mathcal{H}(z), \\ \mathbf{H}(\boldsymbol{\kappa}, z) &= [\mathbf{z}_0, \mathbf{n}] \mathcal{H}(z) - \frac{1}{k_0 \mu(z)} [\mathbf{z}_0 \boldsymbol{\kappa} + i \mathbf{n} \partial_z] \mathcal{E}(z),\end{aligned}\quad (55)$$

где $\mathbf{n} = \boldsymbol{\kappa}/\kappa$; $\mathcal{E}(z), \mathcal{H}(z)$ — произвольные решения уравнений на интервале $d < z < +\infty$:

$$\left[\varepsilon(z) \partial_z \frac{1}{\varepsilon(z)} \partial_z + k^2(z) - \kappa^2 \right] \mathcal{H}(z) = 0, \quad \left[\mu(z) \partial_z \frac{1}{\mu(z)} \partial_z + k^2(z) - \kappa^2 \right] \mathcal{E}(z) = 0, \quad (56)$$

непрерывные по z вместе с величинами $\partial_z \mathcal{E}(z)/\mu(z), \partial_z \mathcal{H}(z)/\varepsilon(z)$. Из (39) для векторных амплитуд вытекает условие при $z = d + 0$:

$$\mathbf{E}_t(\boldsymbol{\kappa}, z) - \hat{\mathcal{L}}_e[\mathbf{z}_0, \mathbf{H}_t(\boldsymbol{\kappa}, z)] = 0, \quad (57)$$

которое в терминах скалярных потенциалов $\mathcal{E}(z), \mathcal{H}(z)$ принимает вид

$$(\partial_z + k_0 \varepsilon_1 \beta_e) \mathcal{H}(z) = -(\varepsilon_1/\mu_1) l(\boldsymbol{\kappa}) \partial_z \mathcal{E}(z), \quad (\alpha_e \partial_z + i k_0 \mu_1) \mathcal{E}(z) = -i k_0 \mu_1 l(\boldsymbol{\kappa}) \mathcal{H}(z). \quad (58)$$

Здесь $\beta_e(\boldsymbol{\kappa}), \alpha_e(\boldsymbol{\kappa})$ и $l(\boldsymbol{\kappa})$ — компоненты диады эквивалентного импеданса в базисе из векторов $\mathbf{n}, [\mathbf{z}_0, \mathbf{n}]$:

$$\beta_e(\boldsymbol{\kappa}) = \mathbf{n} \hat{\mathcal{L}}_e \mathbf{n}, \quad \alpha_e(\boldsymbol{\kappa}) = [\mathbf{z}_0, \mathbf{n}] \hat{\mathcal{L}}_e [\mathbf{z}_0, \mathbf{n}], \quad l(\boldsymbol{\kappa}) = \mathbf{n} \hat{\mathcal{L}}_e [\mathbf{z}_0, \mathbf{n}] \equiv [\mathbf{z}_0, \mathbf{n}] \hat{\mathcal{L}}_e \mathbf{n}. \quad (59)$$

Ищем решение задачи (56), (58) в виде, который бы обеспечивал выполнение условий излучения при $z \rightarrow +\infty$:

$$\mathcal{H}(z) = A \Psi_\varepsilon^+(z, \boldsymbol{\kappa}), \quad \mathcal{E}(z) = B \Psi_\mu^+(z, \boldsymbol{\kappa}). \quad (60)$$

Функции $\Psi_\varepsilon^+(z, \boldsymbol{\kappa}), \Psi_\mu^+(z, \boldsymbol{\kappa})$ определены как произвольные решения уравнений (56) для \mathcal{H} и \mathcal{E} соответственно, которые удовлетворяют условию излучения при $z \rightarrow +\infty$. Подстановка (60) в граничное условие (58) даёт однородную систему двух линейных алгебраических уравнений относительно коэффициентов A, B , нетривиальное решение которой существует при условии равенства нулю её определителя:

$$[\gamma_\varepsilon(\boldsymbol{\kappa}) + k_0 \varepsilon_1 \beta_e(\boldsymbol{\kappa})] \left[\alpha_e(\boldsymbol{\kappa}) + \frac{k_0 \mu_1}{\gamma_\mu(\boldsymbol{\kappa})} \right] - k_0 l^2(\boldsymbol{\kappa}) \varepsilon_1 \mu_1 = 0. \quad (61)$$

Функции $\gamma_\varepsilon, \gamma_\mu$ переменного $\boldsymbol{\kappa}$ определены выше в формулах (53). Соотношение (61) представляет собой дисперсионное уравнение относительно волновых чисел $\boldsymbol{\kappa}$ дискретного спектра и неспектральных мод в плоскостойкой среде на подложке, характеризуемой некоторой диадой импеданса $\hat{\mathcal{L}}_e$.

Конкретизируем теперь в указанной задаче происхождение этой диады, связав её с гауссовскими шероховатостями идеально проводящей плоскости $z = d$. Величины $\beta_e(\boldsymbol{\kappa}), \alpha_e(\boldsymbol{\kappa})$ и $l(\boldsymbol{\kappa})$ в этом случае допускают следующее разложение по степеням малого параметра, характеризующего высоту неровностей:

$$\beta_e(\boldsymbol{\kappa}) = \beta_e^{(2)}(\boldsymbol{\kappa}) + \beta_e^{(4)}(\boldsymbol{\kappa}), \quad \alpha_e(\boldsymbol{\kappa}) = \alpha_e^{(2)}(\boldsymbol{\kappa}) + \alpha_e^{(4)}(\boldsymbol{\kappa}), \quad l(\boldsymbol{\kappa}) = l^{(2)}(\boldsymbol{\kappa}) + l^{(4)}(\boldsymbol{\kappa}). \quad (62)$$

Здесь верхний индекс указывает на порядок малости соответствующего члена по упомянутому малому параметру. Явный вид этих членов без труда получается из определений (59) после подстановки туда разложения (60); в целях экономии места они здесь не приводятся.

Разложим предполагаемое решение κ дисперсионного уравнения (61) по степеням того же малого параметра:

$$\kappa = \kappa_0 + \delta^{(2)} + \delta^{(4)}, \quad (63)$$

ограничившись учётом членов четвёртого порядка малости. Здесь κ_0 — решение дисперсионного уравнения для среды с гладкой границей:

$$\gamma_\varepsilon(\kappa_0)/\gamma_\mu(\kappa_0) = 0. \quad (64)$$

Тогда решения дисперсионного уравнения (61) для среды с шероховатой границей распадается на две группы. В первой группе параметр κ_0 служит решением уравнения

$$\gamma_\varepsilon(\kappa_0) = 0, \quad (65)$$

описывающего собственные волны и неспектральные моды ТМ-поляризации регулярной среды [9], а поправки $\delta^{(2)}$, $\delta^{(4)}$ даются выражениями

$$\delta^{(2)}(\boldsymbol{\kappa}_0)\gamma'_\varepsilon(\kappa_0) = -\varepsilon_1 k_0 \beta^{(2)}(\boldsymbol{\kappa}_0),$$

$$\delta^{(4)}(\boldsymbol{\kappa}_0)\gamma'_\varepsilon(\kappa_0) = -\varepsilon_1 k_0 \beta^{(4)}(\boldsymbol{\kappa}_0) + \varepsilon_1 \gamma_\mu(\kappa_0) [l^{(2)}(\boldsymbol{\kappa}_0)]^2 - [\delta^{(2)}(\boldsymbol{\kappa}_0)]^2 \gamma''_\varepsilon(\kappa_0)/2 - \varepsilon_1 k_0 \delta^{(2)}(\boldsymbol{\kappa}_0) \beta^{(2)'}(\boldsymbol{\kappa}_0). \quad (66)$$

Здесь и далее направление распространения волны, определяемое единичным вектором $\mathbf{n} = (n_x, n_y, 0)$, считается фиксированным: $\boldsymbol{\kappa} = \kappa \mathbf{n}$, $\boldsymbol{\kappa}_0 = \kappa_0 \mathbf{n}$; штрих означает производную соответствующей величины по переменной κ : $f'(\kappa_0) = \partial f(\kappa)/\partial \kappa|_{\kappa=\kappa_0}$.

Решения второй группы близки к корням уравнения

$$\nu_\mu(\kappa_0) = 1/\gamma_\mu(\kappa_0) = 0, \quad (67)$$

отвечающих собственным волнам и неспектральным модам ТЕ-поляризации регулярной среды [9]. Поправки $\delta^{(2)}$, $\delta^{(4)}$ даются в этом случае следующими формулами:

$$\delta^{(2)}(\boldsymbol{\kappa}_0)\nu'_\mu(\kappa_0) = -\alpha^{(2)}(\boldsymbol{\kappa}_0)/(k_0 \mu_1),$$

$$\delta^{(4)}(\boldsymbol{\kappa}_0)\nu'_\mu(\kappa_0) = \alpha^{(4)}(\boldsymbol{\kappa}_0)/(k_0 \mu_1) + \varepsilon_1 [l^{(2)}(\boldsymbol{\kappa}_0)]^2/(\gamma_\varepsilon(\kappa_0)\mu_1) - [\delta^{(2)}(\boldsymbol{\kappa}_0)]^2 \nu''_\mu(\kappa_0)/2 - \delta^{(2)}(\boldsymbol{\kappa}_0) \alpha^{(2)'}(\boldsymbol{\kappa}_0)/(k_0 \mu_1). \quad (68)$$

Обращаясь вновь к представлению (55) для векторных амплитуд, из (58), (60) заключаем, что при $l(\boldsymbol{\kappa}) \neq 0$ ни один из скалярных потенциалов $\mathcal{H}(z)$, $\mathcal{E}(z)$ не обращается тождественно в нуль. В частности, для волн, которым отвечают выражения (65), (66), функция $\mathcal{H}(z) \sim 1$, а $\mathcal{E}(z) \sim l(\boldsymbol{\kappa})$ — малая величина. Такие волны можно отнести к квази-ТМ-типу, поскольку они приобретают под действием шероховатостей малую продольную (вдоль \mathbf{n}) компоненту магнитного поля. Волны, описываемые соотношениями (67), (68), характеризуются тем, что $\mathcal{H}(z) \sim l(\boldsymbol{\kappa})$ и $\mathcal{E}(z) \sim 1$. Следовательно, электрическое поле волны имеет поперечную (вдоль $[\mathbf{z}_0, \mathbf{n}]$) компоненту и малую продольную компоненту. Поэтому эти волны правомерно отнести к квази-ТЕ-типу.

Легко видеть, что для статистически изотропных шероховатостей, функция корреляции которых $B(\mathbf{r}-\mathbf{r}')$ зависит только от модуля вектора $\mathbf{r}-\mathbf{r}'$, коэффициент $l(\boldsymbol{\kappa}) \equiv 0$. Поэтому в этом и только в этом случае деполяризация собственных волн и неспектральных мод отсутствует: они строго распадаются на волны ТМ-типа, для которых $\mathcal{E} \equiv 0$, и волны ТЕ-типа с $\mathcal{H} \equiv 0$.

Выражения для $\delta^{(2)}(\boldsymbol{\kappa}_0)$ из (66), (68) совпадают с результатами, полученными в работе [10], а поправка $\delta^{(4)}(\boldsymbol{\kappa}_0)$ четвёртого порядка малости найдена впервые в этой работе.

СПИСОК ЛИТЕРАТУРЫ

1. Фейнберг Е. Л. // Исследования по распространению радиоволн / Под ред. Б. А. Введенского. Вып. 2. М.: Изд-во АН СССР, 1948. С. 145.
2. Басс Ф. Г., Фукс И. М. Рассеяние волн на статистически неровной поверхности. М.: Наука, 1972. 424 с.
3. Унгер Х. Г. Планарные и волоконные оптические волноводы. М.: Мир, 1980. С. 656.
4. Фелсен Л., Маркувиц Н. Излучение и рассеяние волн. М.: Мир, 1978. С. 547.
5. Апресян Л. А. // Изв. вузов. Радиофизика. 1974. Т. 17, № 2. С. 165.
6. Финкельберг В. М. // ЖЭТФ. 1967. Т. 53, № 1. С. 401.
7. Брюховецкий А. С., Фукс И. М. // Изв. вузов. Радиофизика. 1985. Т. 28, № 11. С. 1400.
8. Жук Н. П., Третьяков О. А. // Радиотехника и электроника. 1985. Т. 30, № 5. С. 689.
9. Шевченко В. В. Плавные переходы в открытых волноводах. М.: Наука, 1969. С. 191.
10. Жук Н. П., Третьяков О. А., Яровой А. Г. // Изв. вузов. Радиофизика. 1986. Т. 29, № 4. С. 447.

Харьковский национальный университет,
г. Харьков, Украина

Поступила в редакцию
4 октября 2000 г.

EFFECTIVE IMPEDANCE OF A SLIGHTLY ROUGH SURFACE OF AN ARBITRARY ANISOTROPIC MEDIUM

S. N. Shulga

We consider the boundary-value problem for the electromagnetic field in an arbitrary anisotropic half-space with a random surface. In the case of small height of the roughness, we derive the canonical impedance boundary condition with nonlocal impedance in the fourth order of the perturbation theory. In addition to general case, we analyze the case of a perfectly conductive surface with a Gaussian roughness.

УДК 533.9:621.039

О НЕКОТОРЫХ ОСОБЕННОСТЯХ ЭЦР-НАГРЕВА НА ПОЛОВИННОЙ И ФУНДАМЕНТАЛЬНОЙ ГАРМОНИКАХ В МАГНИТНОЙ ЛОВУШКЕ

И. Ю. Костюков, А. Ю. Крячко, М. Д. Токман

В работе сопоставляется эффективность СВЧ нагрева электронов в плотной плазме на фундаментальной ($\omega = eH/(mc)$) и полужелой ($\omega = eH/(2mc)$) гармониках гирочастоты. В частности, проводится анализ недавних экспериментов, продемонстрировавших более высокую эффективность нагрева плазмы на полужелой гармонике в СВЧ резонаторе. Получены уравнения, описывающие нелинейное циклотронное ускорение электронов при резонансе на полужелой гармонике, справедливые для произвольной геометрии СВЧ поля. Для СВЧ поля с «вакуумной» поляризацией мощность энерговыклада в электроны на фундаментальной гармонике на несколько порядков превосходит мощность энерговыклада при нелинейном нагреве на полужелой гармонике. Нами, однако, показано, что в плотной плазме вследствие эффекта депрессии резонансной компоненты высокочастотного поля соответствующее различие составляет не более одного порядка. В этом случае различная эффективность формирования популяции горячих электронов может определяться различием не столько мощности энерговыклада, сколько эффективности удержания. В данной работе, в частности, показано, что двукратное уменьшение величины магнитного поля при переходе от полужелой гармонике к первой может привести к существенному ухудшению устойчивости движения высокоэнергичных электронов и, как следствие, к ухудшению удержания и снижению итоговой эффективности нагрева. Обсуждается возможное влияние на нагрев электронов эффекта квазистатического усиления высокочастотного поля вблизи мишени, помещённой в ловушку для генерации рентгеновского излучения.

ВВЕДЕНИЕ

В различных приложениях (см., например, [1]) требуются компактные и недорогие источники рентгеновского излучения. Одним из перспективных направлений создания подобных источников является использование электронного циклотронного резонанса (ЭЦР) для генерации ускоренных электронов, которые при торможении на специальной мишени генерируют рентгеновское излучение.

Недавно в экспериментах был обнаружен эффективный нагрев электронов в магнитной ловушке на полужелой гармонике гирочастоты [2]. При этом мощность рентгеновского излучения была в 5 раз больше, чем при аналогичном нагреве на фундаментальной гармонике.

Очевидно, что нагрев на полужелой гармонике происходит в результате нелинейного взаимодействия между электромагнитным полем и электронами. Как правило, такой нагрев является менее эффективным, чем обычный ЭЦР-нагрев на фундаментальной гармонике. Основной целью данной работы является анализ возможных причин более эффективного нагрева электронов на полужелой гармонике гирочастоты.

1. ЭЦР-НАГРЕВ ЭЛЕКТРОНОВ НА ПОЛУЦЕЛОЙ ГАРМОНИКЕ ГИРОЧАСТОТЫ

Вначале рассмотрим движение отдельных электронов в условиях циклотронного резонанса на полужелой гармонике. Ранее эта задача обсуждалась в рамках исследования некоторых режимов стимулированного излучения лазеров на свободных электронах [3] и нелинейного нагрева плазмы в тороидальных установках [4]. Здесь мы проведём достаточно общее рассмотрение, справедливое, в частности, и для интересующего нас случая ЭЦР-нагрева в резонаторе магнитной ловушки.

Для большого числа современных и планируемых экспериментов по высокочастотному (ВЧ) нагреву плазмы достаточно ограничиться приближением, в рамках которого амплитуда осцилляций электрона

под действием ВЧ поля и его гирорадиус в магнитном поле много меньше масштаба неоднородности электромагнитного поля. В этом случае нетрудно получить описывающие нагрев уравнения для поля произвольной геометрии. С этой целью вначале получим выражение для второй гармоники силы Лоренца в стационарном ВЧ поле

$$\mathbf{E}(\mathbf{r}, t) = \text{Re}[\mathbf{E}(\mathbf{r}) \exp(-i\omega t)], \quad \mathbf{H}(\mathbf{r}, t) = \text{Re}[\mathbf{H}(\mathbf{r}) \exp(-i\omega t)] \quad (1)$$

при наличии постоянного неоднородного магнитного поля $\mathbf{H}_0(\mathbf{r})$ (отметим, что в зависимость $\mathbf{E}(\mathbf{r})$ мы включаем и гармоническую зависимость фазы электромагнитного поля от координаты, возможную в случае волновой структуры). Чтобы исследовать взаимодействие электрона с электромагнитным полем в рамках указанных выше ограничений, удобно использовать описанный, например, в [5, 6] подход. Для этого выделим в выражении для координаты частицы её осцилляции под действием ВЧ поля:

$$\mathbf{r} = \mathbf{R} + \text{Re} \mathbf{r}_{\sim}, \quad (2)$$

где усреднённая компонента \mathbf{R} обусловлена дрейфом и циклотронным вращением, а осцилляторный радиус-вектор \mathbf{r}_{\sim} связан с комплексной амплитудой $\mathbf{E}(\mathbf{r})$ стандартным соотношением, строго справедливым в случае однородного поля:

$$\mathbf{r}_{\sim} = \hat{\mathbf{T}}\mathbf{E}. \quad (3)$$

Тензор $\hat{\mathbf{T}}$ определяется выражением (см., например, [6])

$$T_{ik} = \frac{e/m}{\omega^2 - \omega_H^2} \left(\delta_{ik} - \frac{\omega_{Hj} \omega_{Hk}}{\omega^2} - i \frac{e_{jkp} \omega_{Hp}}{\omega} \right) \quad (4)$$

и связан с тензором ВЧ поляризуемости холодной плазмы:

$$\hat{\lambda} = -en_e \hat{\mathbf{T}}. \quad (5)$$

Здесь n_e — электронная концентрация, δ_{ik} — тензор Кронекера, e_{ijk} — тензор Леви—Чивита, $\omega_H = eH_0/(mc)$ — гирочастота электрона, $\omega_{Hi} = eH_{0i}/(mc)$, H_{0i} — компоненты магнитного поля в направлении координатных осей, e , m и c — элементарный заряд, масса электрона и скорость света соответственно. Подставив соотношение (3) в выражение для силы Лоренца

$$\mathbf{F}_L = -\frac{e}{c} [\mathbf{V}, \mathbf{H}_0 + \mathbf{H}] - e\mathbf{E}, \quad (6)$$

где \mathbf{V} — скорость частицы, и воспользовавшись уравнением Максвелла

$$\mathbf{H} = \frac{c}{i\omega} [\nabla, \mathbf{E}], \quad (7)$$

а также выражением для осцилляторной скорости $\mathbf{V}_{\sim} = \text{Re}(-i\omega \mathbf{r}_{\sim})$, легко получить квадратичное по полю \mathbf{E} выражение для гармоники силы Лоренца с частотой 2ω :

$$\mathbf{F}_{2\omega} = \text{Re}[\mathbf{F}(\mathbf{R}) \exp(-2i\omega t)] = \text{Re} \left[-\frac{e}{2} T_{jk} E_k \nabla E_j \exp(-2i\omega t) \right] + \mathbf{F}_M. \quad (8)$$

В правой части выражения (8) первое слагаемое обусловлено неоднородностью амплитуды и фазы ВЧ поля, в то время как второе слагаемое связано с неоднородностью постоянного магнитного поля. Это слагаемое можно представить в тензорной форме:

$$F_{Mj} = \text{Re} \left[\exp(-2i\omega t) i\omega \frac{e}{2} e_{jlk} T_{lp} T_{mn} E_p E_n \frac{\partial}{\partial r_m} H_{0k} \right]. \quad (9)$$

Учитывая условие резонанса на полуцелой циклотронной гармонике:

$$\omega = \omega_H/2, \quad (10)$$

можно перейти к соотношениям, описывающим циклотронное ускорение электронов резонансной силой (8). Для этого можно воспользоваться соответствующей процедурой (см., например, [7–12]), развитой для исследования циклотронного ускорения слаборелятивистских частиц резонансным электромагнитным полем. При этом оказывается, что с хорошей точностью итоговые соотношения прямо следуют из соответствующих выражений для задачи об ускорении частиц циркулярно поляризованной составляющей электрического поля волны в случае «стандартного» резонанса $\omega = \omega_H$ (см. [8, 9]). В этом случае надо только вместо фигурирующего в этих работах поля $\mathbf{E}(\mathbf{R}) \exp(-i\omega t)$ подставить эффективное поле $-\mathbf{F}(\mathbf{R}) \exp(-2i\omega t)/e$, где $\mathbf{F}(\mathbf{R})$ — определённая в (8) комплексная амплитуда второй гармоники силы Лоренца. В результате получаем уравнения так называемой второй универсальной модели нелинейного резонанса [13], описываемой следующим гамильтонианом в переменных действие—угол:

$$\mathcal{H} = I\delta - I^2/2 + a\sqrt{I} \cos \theta, \quad (11)$$

где $I = P_{\perp}^2\omega/(m^2c^2\omega_H)$ — нормированный магнитный момент, $a = F_+/(2mc\omega)$ — нормированная действительная амплитуда правополяризованной компоненты ВЧ силы $\mathbf{F}_+ = (\mathbf{F}_x - i\mathbf{F}_y)/2$, θ — фаза циклотронного вращения электрона относительно удвоенной фазы поля $2 \arg[E(\mathbf{r})] - 2\omega t$,

$$\delta = \frac{\omega_H}{2\omega} - 1 - \frac{V_z}{2\omega} \frac{\partial}{\partial z} 2 \arg[E(\mathbf{r})], \quad (12)$$

ось z направлена вдоль магнитного поля $\mathbf{H}_0(\mathbf{r})$.

В уравнениях движения, следующих из гамильтониана (11) и описывающих нелинейное взаимодействие электрона с ВЧ полем на полуцелой гармонике гирочастоты, используется нормированное время $\tau = 2\omega t$. Отметим, что в выражении для параметра δ последнее слагаемое (доплеровский сдвиг) имеет смысл учитывать лишь в случае, когда масштаб неоднородности комплексной фазы поля (обратное волновое число) в направлении магнитной силовой линии существенно превосходит масштаб неоднородности интенсивности поля в том же направлении. Гамильтониан (11) при необходимости можно легко дополнить уравнениями движения частицы вдоль силовой линии (см., например, [7, 8]).

Отметим, что соответствующие частные выражения для плоских волн с конкретными типами поляризации (в случае однородного магнитного поля) были ранее получены в [3, 4, 14, 15]. В этом случае для соответствующих предельных переходов в первое слагаемое в правой части (8) нужно подставить $\nabla = i\mathbf{k}$, где \mathbf{k} — волновой вектор.

Вычислим коэффициент a для интересующих нас случаев. Поскольку в экспериментах мишень и область циклотронного резонанса на фундаментальной и полуцелой гармониках обычно находятся в экваториальной плоскости магнитной ловушки, то можно предположить, что градиент модуля напряжённости постоянного магнитного поля ортогонален направлению поля. Будем также предполагать, что ВЧ поле является резонаторной модой с линейной поляризацией поля $\mathbf{E}(\mathbf{r})$ двух типов:

1) электрическое поле поляризовано вдоль магнитного поля (т. е. вдоль оси z): $\mathbf{E}(\mathbf{r}) = \mathbf{e}_z E(\mathbf{r})$. Тогда параметр a определяется выражением

$$a \approx \frac{e}{8m\omega^2 H_0} \frac{\partial}{\partial x} |E|^2 \approx \frac{e^2 |E(\mathbf{r} = \mathbf{r}_{\text{res}})|^2}{16\omega^3 m^2 c L} = \frac{e^2 |E_0|^2}{16\omega^3 m^2 c L}, \quad (13)$$

где $L = l_E$ — характерный масштаб неоднородности ВЧ поля, \mathbf{r}_{res} — радиус-вектор, определяющий положение точки резонанса;

2) электрическое поле поляризовано поперёк магнитного поля и поперёк градиента его модуля (вдоль оси y). В этом случае предполагаем также однородность поля \mathbf{E} вдоль оси y : $\mathbf{E}(\mathbf{r}) = \mathbf{e}_x E(\mathbf{r})$. Тогда параметр a определяется выражением

$$a \approx \frac{e^2}{48\omega^3 m^2 c} \left| \frac{\partial}{\partial x} |E|^2 + |E|^2 \frac{2}{\omega} \frac{\partial \omega_H}{\partial x} \right| \approx \frac{e^2 |E_0|^2}{16\omega^3 m^2 c L}, \quad (14)$$

где $L = 3 \min(l_E, l_H/4)$, l_H — характерный масштаб неоднородности магнитного поля.

При точном резонансе (10) расстройка δ равна нулю. Предположение, что гирочастота $\omega_H(z)$ слабо зависит от z , приводит к тому, что расстройка тоже плавно зависит от z . Пусть $\omega_H(z) \approx 2\omega(1 + z/l_H)$. Тогда, пренебрегая доплеровским сдвигом в выражении для δ , получаем $\delta \approx z/l_H$. Будем также предполагать, что амплитуда волны почти постоянна вдоль оси z и приближённо равна a . Тогда, используя подход, предложенный в работе [16], нетрудно вычислить среднее приращение энергии электрона в результате взаимодействия с полуволной гармоникой ВЧ силы в приближении постоянной скорости пролёта:

$$W_{1/2} = mc^2 \Delta I \approx \frac{mc^2}{4} \frac{\partial}{\partial I_0} \langle |\Delta I^{(1)}|^2 \rangle \approx mc^2 \frac{\pi a^2 \omega l_H}{V_{\parallel}} = \frac{\pi e^4 |E_0|^4 l_H}{16m^3 \omega^5 L^2 V_{\parallel}}, \quad (15)$$

где приращение энергии вычисляется методом возмущения по a прямым интегрированием гамильтоновых уравнений:

$$\Delta I^{(1)} = \int \frac{\partial \mathcal{H}}{\partial \theta} d\tau \approx \sqrt{I_0} a \int_{-\infty}^{+\infty} \exp\left(i \frac{V_{\parallel} \tau^2}{4\omega l_H}\right) d\tau \approx a(1+i) \sqrt{\frac{2\pi l_H \omega I_0}{V_{\parallel}}}. \quad (16)$$

Здесь угловые скобки означают усреднение по $\theta_0 = \theta(t = -\infty)$, т. е. по фазам влёта электронов в область резонанса, $I_0 = I(t = -\infty)$ — начальное значение переменной действия (нормированной энергии электрона перед взаимодействием), V_{\parallel} — продольная (направленная вдоль ловушки) компонента скорости электрона.

Аналогичным образом может быть найдена работа однородного ВЧ поля над электроном при обычном ЭЦР-нагреве на первой гармонике:

$$\Delta I^{(1)} \approx \frac{eE_0}{mc\omega} (1+i) \sqrt{\frac{2\pi l_H \omega I_0}{V_{\parallel}}}, \quad W_1 \approx \frac{\pi e^2 |E_0|^2 l_H}{m\omega V_{\parallel}}. \quad (17)$$

Сравнивая полученные выражения для ΔW_1 и $\Delta W_{1/2}$, введём параметр эффективности нагрева на половинной гармонике:

$$\eta_{1/2}^2 = \frac{W_{1/2}}{W_1} \approx \left(\frac{e|E_0|}{4m\omega^2 L} \right)^2 \approx \left(\frac{e|E_0|}{mc\omega} \right)^2 \left(\frac{c}{\omega L} \right)^2. \quad (18)$$

Поскольку в современных экспериментах по ЭЦР-нагреву плазмы осцилляторная скорость электрона в ВЧ поле много меньше скорости света, то, естественно, $\eta_{1/2} \ll 1$.

Следует отметить, что, поскольку мишень и зона ЭЦР расположены в экваториальной плоскости ловушки, то, вообще говоря, нужно рассматривать параболический профиль магнитного поля. В этом случае выражения для работы однородного ВЧ поля над электроном при ЭЦР-нагреве на полуволной и первой гармониках изменятся. Однако нетрудно показать, что выражение для параметра эффективности (18), введённого как корень из отношения этих величин, останется неизменным. Таким образом, выражением (18) можно пользоваться и в случае параболического профиля магнитного поля.

Для типичных параметров эксперимента [2, 17, 18] по генерации рентгеновского излучения с помощью ЭЦР-нагрева электронов в магнитной ловушке (мощность электромагнитного излучения $P \approx 1$ кВт, частота ВЧ поля $f \approx 2,45$ ГГц, $l_H \approx 10$ см, $H_0 \approx 850$ Гс для фундаментальной гармоники, $H_0 \approx 1700$ Гс для половинной гармоники) параметр эффективности $\eta_{1/2} \approx 7 \cdot 10^{-3}$. Таким образом, эффективность нагрева электронов на полужелой гармонике гирочастоты намного ниже, чем при нагреве на фундаментальной гармонике.

2. ДЕПРЕССИЯ ПАДАЮЩЕГО ИЗЛУЧЕНИЯ НА ФУНДАМЕНТАЛЬНОЙ ГАРМОНИКЕ

Оценка (13) для случая ЭЦР-нагрева на фундаментальной гармонике может оказаться сильно завышенной вследствие депрессии резонансной компоненты ВЧ поля. Как известно [19, 20], в достаточно плотной плазме резонансная компонента поля (циркулярно поляризованная компонента электрического ВЧ поля, вращающаяся в направлении вращения электрона в магнитном поле) много меньше среднеквадратичной амплитуды поля. Используя известные выражения для диэлектрической проницаемости холодной бесстолкновительной плазмы, при небольшой отстройке от резонанса $\Delta\omega = \omega - \omega_H \ll \omega$ нетрудно получить выражения для поляризации необыкновенной волны в плотной плазме:

$$E_x/E_y = -\varepsilon_{xy}/\varepsilon_{xx} = -i(1+g), \quad (19)$$

где $g = (\Delta\omega/\omega)(1 - 2\omega^2/\omega_p^2) - (\Delta\omega/\omega_p)^2(3 - 4\omega^2/\omega_p^2)$, $\omega_p = \sqrt{4\pi e^2 n_e/m}$ — электронная плазменная частота, $\varepsilon_{ij} = 1 - 4\pi e n_e T_{ij}$ — тензор диэлектрической проницаемости плазмы. Из полученного выражения следует, что в случае точного резонанса ($\Delta\omega = 0$, $\omega = \omega_H$) вектор электрического поля необыкновенной волны вращается в направлении, противоположном направлению вращения электрона в магнитном поле (нерезонансная компонента поля), и амплитуда резонансной компоненты электрического поля равна нулю. При наличии отстройки ($\Delta\omega \neq 0$) необыкновенная волна поляризована эллиптически, и появляется отличная от нуля резонансная компонента поля. В Приложении рассматривается эффект депрессии резонансной компоненты ВЧ поля в заполненном плазмой волноводе.

Рассмотрим случай неоднородного магнитного поля, которое вблизи точки резонанса аппроксимируем линейным профилем: $\omega_H(z) = \omega(1 + z/l_H)$, т. е. $\Delta\omega = \omega z/l_H$. Тогда с помощью выражений (16) и (19), полагая $z(\tau) = V_{\parallel}\tau/(2\omega)$, нетрудно вычислить работу электрического поля над электроном при его пролёте через область циклотронного резонанса с учётом эффекта депрессии:

$$\Delta I^{(1)} \approx \frac{eE_0}{mc\omega} \sqrt{I_0} \int_{-\infty}^{+\infty} d\tau g[z(\tau)] \exp \left[i \int_{-\infty}^{\tau} \frac{\Delta\omega[z(\tau')]}{\omega} d\tau' \right] \approx \frac{eE_0}{mc\omega} \frac{\omega^2}{\omega_p^2} \left(3 - \frac{4\omega^2}{\omega_p^2} \right) (1+i) \sqrt{\frac{\pi V_{\parallel} I_0}{2\omega l_H}},$$

$$W_{\text{depr}} \approx \frac{\pi e^2 |E_0|^2 V_{\parallel}}{4m\omega^3 l_H} \frac{\omega^4}{\omega_p^4} \left(3 - \frac{4\omega^2}{\omega_p^2} \right). \quad (20)$$

Сравнивая полученное выражение с выражением (17) для работы однородного электрического поля над электроном в вакууме (т. е. без учёта эффекта депрессии), введём амплитуду эффективного электрического поля E_{eff} и параметр эффективности нагрева η_{depr} :

$$E_{\text{eff}} = E_0 \eta_{\text{depr}}, \quad W_{\text{depr}} = W_1 \eta_{\text{depr}}^2, \quad \eta_{\text{depr}} \approx \frac{V_{\parallel}}{2l_H \omega} \frac{\omega^2}{\omega_p^2} \left| 3 - 4\omega^2/\omega_p^2 \right|. \quad (21)$$

Аналогично можно вычислить η_{depr} и в случае параболического профиля магнитного поля, когда $\omega_H(z) = \omega(1 + z^2/l_H^2)$:

$$\eta_{\text{depr}} \approx \left(\frac{V_{\parallel}}{l_H \omega} \right)^{4/3} \frac{|(1 - 2\omega^2/\omega_p^2) \Gamma(-1/3)|}{\pi}, \quad (22)$$

где $\Gamma(x)$ — гамма-функция. В обоих случаях эффект депрессии ограничен неоднородностью магнитного поля.

Депрессия резонансного излучения в плазме может быть ограничена и другими процессами: тепловыми и нелинейными эффектами, столкновениями в плазме [8, 21]. Таким образом, в общем случае эффективность нагрева, связанная с депрессией резонансной компоненты фундаментальной гармоники ВЧ поля, определяется оценкой

$$\eta_{\text{depr}} \approx \max \left[\beta_T^2, n_{\parallel} \beta_T, \frac{E_0^2 \omega_H^6}{H_0^2 \omega_p^6}, \frac{\nu \omega_H}{\omega_p^2}, \frac{V_{\parallel}}{2l_H \omega} \frac{\omega^2}{\omega_p^2} |3 - 4\omega^2/\omega_p^2| \right], \quad (23)$$

где $n_{\parallel} = k_{\parallel} c/\omega$, $\beta_T = V_T/c$, $\beta_{\parallel} = V_{\parallel}/c$, V_T — тепловая скорость электронов, ν — транспортная частота соударений электронов. Депрессия излучения на фундаментальной гармонике в η_{depr} раз уменьшает эффективность нагрева на фундаментальной гармонике.

Для типичных параметров экспериментов [2, 17, 18] (энергия теплового движения горячих электронов вдоль магнитного поля и температура электронов фоновой плазмы $T_{\parallel} \approx T \approx 50$ эВ, $n_{\parallel} \sim 1$, $\omega^2/\omega_p^2 \approx 4$) уменьшение эффективности нагрева в результате депрессии излучения связано с тепловыми эффектами и влиянием неоднородности магнитного поля на движение горячих электронов. Оба эти фактора обеспечивают сравнимое уменьшение эффективности нагрева $\eta_{\text{depr}} \approx 0,1$. Отметим, что для резонанса на полуцелой гармонике эффект депрессии резонансной компоненты незначителен, т. к. в данном случае с направлением вращения электронов совпадает вращение не реального поля, а соответствующей компоненты эффективной нелинейной силы.

3. УСТОЙЧИВОСТЬ ДВИЖЕНИЯ ЭЛЕКТРОНОВ В МАГНИТНОЙ ЛОВУШКЕ

Заметим, что учёт депрессии ВЧ поля на фундаментальной гармонике гирочастоты и квазистатического усиления ВЧ поля на полуцелой гармонике недостаточен для объяснения высокой эффективности нагрева на полуцелой гармонике. Согласно нашим расчётам, для эксперимента [2] даже с учётом рассмотренных выше эффектов нагрев на полуцелой гармонике всё равно менее эффективен, чем нагрев на фундаментальной гармонике ($\eta_{1/2}/\eta_{\text{depr}} < 1$), в то время как эксперимент показывает большую эффективность нагрева на полуцелой гармонике. Другим фактором, влияющим на эффективность нагрева, может быть изменение режима удержания горячих электронов в магнитной ловушке, связанное с увеличением магнитного поля при нагреве на полуцелой гармонике гирочастоты. Рассмотрим этот эффект подробнее.

Как следует из экспериментальных данных [17, 18], энергия квантов рентгеновского излучения в случае нагрева на половинной гармонике в несколько раз выше, чем при обычном нагреве. Таким образом, энергия горячих электронов в первом случае также была значительно выше. Одна из возможных причин этого явления может быть связана с устойчивостью захвата электронов в магнитной ловушке. Дело в том, что согласно условиям эксперимента при нагреве на половинной гармонике магнитное поле было в два раза выше, чем при нагреве на фундаментальной гармонике, а распределение магнитного поля существенным образом влияет на движение электронов в ловушке [22].

Исследование устойчивости движения электронов в ловушке в похожем режиме проводилось в работе [17] для объяснения результатов экспериментов [17, 18], при этом полученные оценки хорошо совпадали с экспериментальными результатами. Важно отметить, что обычная теория стохастического ЭЦР-нагрева [22, 23, 24], учитывающая наличие «суперадиабатических» инвариантов движения при больших энергиях электронов, приводит к существенно более низким предельным значениям энергий. На самом деле указанная теория пренебрегает различными факторами (отсутствие строгой аксиальной симметрии магнитной системы, сложная структура ВЧ поля, столкновения и т. д.), возникающими

в реальной ситуации. Как показывают эксперименты (в том числе и обсуждаемые здесь работы [2, 17, 18]), нагрев электронов обычно происходит до более высоких энергий, когда он существенно определяется неадиабатичностью движения (гирорадиус электрона становится соизмеримым с характерным масштабом неоднородности магнитного поля). Таким образом, при достаточно большой энергии поперечного движения электроны переставали удерживаться в поперечном направлении. В этом случае потенциал центробежной силы становился больше эффективного потенциала магнитного поля (см., например, [25]). Следует отметить, что в этих экспериментах, как и в экспериментах [2], использовалась разреженная плазма, и, таким образом, столкновения не играли существенной роли при нагреве (в течение нагрева электрон испытывал лишь несколько столкновений).

Предполагая, что нарушение адиабатичности относительно невелико, мы можем воспользоваться оценкой времени жизни электронов в ловушке, следующей из соответствующей теории возмущений [22]. В этом случае из-за резонанса в неоднородном магнитном поле между гировращением и колебаниями вдоль ловушки электроны стохастически диффундируют в пространстве скоростей и могут попасть в конус потерь. Характерное время жизни электрона в ловушке, введённое как число колебаний электрона вдоль ловушки, в этом случае равно [22]

$$\tau_e \approx \frac{16}{9\pi} \left(\frac{l_H}{R} \right)^{3/2} \frac{\exp(3q/2)}{q} \ln(\lambda\beta_{cr}^2), \quad (24)$$

где $q = 2l_H/(3\rho)$ — параметр неадиабатичности, ρ — гирорадиус электрона, R — расстояние в экваториальной плоскости между осью ловушки и резонансной магнитной силовой линией, $\lambda = H_{\max}/H_{\min}$ — параметр ловушки, $\beta_{cr} = (3\sqrt{\pi}/4)(R/l_H)^{1/4}\sqrt{q}\exp(-q/4)$ — критический питч-угол, соответствующий границе области стохастичности в фазовом пространстве. На рис. 1 показаны зависимости времени жизни τ_e электрона от его энергии T при различной напряжённости H магнитного поля. Поскольку гирорадиус линейно зависит от H , то, как следует из (24), время жизни электрона в ловушке экспоненциально падает с уменьшением магнитного поля, что и демонстрирует рис. 1.

В условиях эксперимента [2] при ЭЦР-нагреве на первой гармонике напряжённость магнитного поля равна 850 Гс. Как следует из рис. 1, при этом время жизни высокоэнергичных электронов ($T > 30$ кэВ) становится порядка наблюдаемого в эксперименте характерного времени τ_{heat} нагрева электронов в магнитной ловушке, составляющего несколько десятков микросекунд (времени, через которое после начала эксперимента наблюдался всплеск рентгеновского излучения). Напротив, при нагреве на половинной гармонике магнитное поле в два раза больше, и время жизни электрона в ловушке на несколько порядков больше, чем время нагрева. При этом электрон может набрать более высокую энергию, чем при нагреве на фундаментальной гармонике. Другими словами, различная эффективность формирования популяции горячих электронов может определяться уже не столько различием мощности энерговыклада, сколько различием эффективности удержания частиц.

Необходимо отметить, что времена нагрева на полуволновой и фундаментальной гармониках гирочастоты, вообще говоря, различаются. Как следует из наших оценок для эффективности нагрева, представленных в разделах 1–3, время нагрева на фундаментальной гармонике гирочастоты с учётом эффекта депрессии примерно на порядок меньше, чем время нагрева на полуволновой гармонике. Поэтому на рис. 1 (где используется логарифмический масштаб) время нагрева для обоих режимов следует описывать горизонтальной полосой с единичной шириной. Однако разница во времени нагрева в обоих

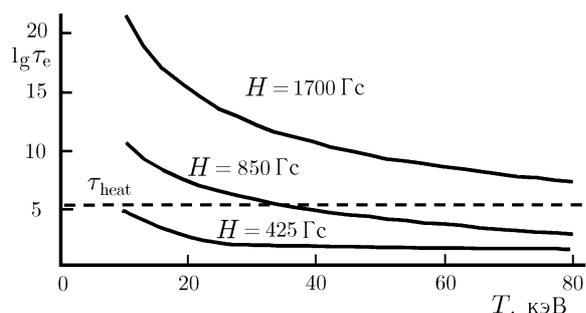


Рис. 1

режимах много меньше расстояния между кривыми, характеризующими время жизни электрона в ловушке для этих режимов (они различаются на несколько порядков). Поэтому шириной полосы можно пренебречь, и сделанные выше качественные выводы об устойчивости электронов для различных режимов нагрева остаются справедливыми.

Другим подтверждением влияния напряжённости магнитного поля на эффективность рентгеновского излучения может служить отсутствие интенсивной эмиссии при нагреве на второй гармонике гирочастоты. Важно отметить, что при нагреве на второй гармонике не происходит депрессии резонансной компоненты ВЧ поля. В случае ЭЦР на второй гармонике напряжённость магнитного поля в центре ловушки в четыре раза меньше, чем в случае ЭЦР на половинной гармонике. Для экспериментальных условий [2] такое магнитное поле соответствует более короткому времени жизни электрона, чем даже в случае нагрева на фундаментальной гармонике. В этом случае электроны не успевают значительно ускориться, что объясняет отсутствие интенсивного рентгеновского излучения.

4. О ВОЗМОЖНОСТИ КВАЗИСТАТИЧЕСКОГО УСИЛЕНИЯ СВЧ ПОЛЯ НА МИШЕНИ

Ещё одним механизмом, который мог бы объяснить высокую эффективность нагрева электронов на полуцелой гармонике гирочастоты, является квазистатическое усиление ВЧ поля вблизи остроконечной мишени. Для генерации рентгеновского излучения в экспериментах [1, 2] использовались мишени, изготовленные из тяжёлых металлов (молибден, вольфрам). Как правило, такие мишени имели иглообразную или плоскую форму и располагались вблизи зоны ЭЦР. Поскольку мишени имели заострённые края, то естественно предположить, что вблизи этих краёв происходит квазистатическое усиление поля.

Рассмотрим условия квазистатического приближения в магнитоактивной плазме. Уравнение, описывающее распределение ВЧ поля вблизи мишени, имеет вид

$$\operatorname{rot} \operatorname{rot} E_i - (\omega/c)^2 \varepsilon_{ij} E_j = 0. \quad (25)$$

Условие квазистатики означает, что в (25) можно пренебречь вторым слагаемым. Полагая, что характерный масштаб квазистатического поля вблизи мишени равен l_{stat} , можно переписать это условие в виде

$$(\omega/c)^2 l_{\text{stat}}^2 |\varepsilon_{ij} E_j| / |E_i| \ll 1. \quad (26)$$

При нагреве на полуцелой гармонике $|\varepsilon_{ij} E_j| / |E_i| \sim 1$, и область квазистатики достаточно велика: $l_{\text{stat}} \approx c/\omega$.

Вблизи зоны резонанса с фундаментальной гармоникой ($\omega \approx \omega_H$) область квазистатики существенно уменьшается в силу расходимости компонент тензора ε_{ij} на циклотронной частоте: $l_{\text{stat}} \approx \approx (c/\omega) \sqrt{\omega \Delta\omega/\omega_p} \rightarrow 0$, и усиление ВЧ поля незначительно. Исключение составляют два случая: 1) ВЧ поле направлено вдоль внешнего магнитного поля (вдоль оси z); 2) поляризация ВЧ поля соответствует поляризации нерезонансной компоненты. При этом параметр $|\varepsilon_{ij} E_j| / |E_i|$ оказывается порядка единицы, и характерный масштаб поля вблизи мишени такой же, как и в случае нагрева на полуцелой гармонике: $l_{\text{stat}} \approx c/\omega$. Однако ни первый, ни второй случаи не представляют интереса. В первом случае проекция электрического поля на плоскость xy равна нулю, следовательно, невозможно возникновение (и последующее усиление) резонансной компоненты поля. Второй случай не реализуется, т. к. циркулярная компонента поля не удовлетворяет граничным условиям на металле.

Однако значительное увеличение эффективности нагрева электронов на полуцелой гармонике в результате квазистатического усиления ВЧ поля вблизи мишени вряд ли возможно. Дело в том, что электроны дрейфуют по резонансной магнитной поверхности и равномерно распределены вдоль неё. Кроме того, необходимо отметить, что в эксперименте [2] расстояние от мишени до оси ловушки составляло

примерно 4,5 см, что значительно больше размера зоны квазистатического усиления ВЧ поля (порядка 1 см). Таким образом, совершая колебания вдоль ловушки, электроны довольно редко пролетают через область квазистатического усиления.

5. ЗАКЛЮЧЕНИЕ

В данной работе нами проанализированы причины увеличения эффективности нелинейного нагрева плазмы на полуцелой гармонике гирочастоты по сравнению с нагревом на первой циклотронной гармонике. В рамках имеющихся экспериментальных данных этот эффект объясняется двумя основными причинами: 1) снижением эффективности нагрева на первой гармонике из-за особенностей поляризации излучения в плотной плазме; 2) ухудшением эффективности удержания горячих электронов в магнитной ловушке при нагреве на первой гармонике гирочастоты вследствие двукратного уменьшения магнитного поля по сравнению с экспериментом по нагреву на полуцелой гармонике.

Кроме того, в данной работе приведён последовательный вывод уравнений движения электронов в условиях резонанса на полуцелой гармонике гирочастоты и проанализирована поляризация СВЧ поля в заполненном плазмой волноводе в условиях ЭЦР.

Отметим, что для полного анализа данного вопроса желательно провести эксперименты не только при различном стационарном магнитном поле, но и при различной частоте СВЧ излучения.

Авторы выражают благодарность Ж.-М. Раксу, М. Бакаль и К. Серебренникову за важные замечания и полезные дискуссии.

ПРИЛОЖЕНИЕ

Найдём собственные моды прямоугольного идеально проводящего волновода, заполненного магнитоактивной плазмой. Будем предполагать, что внешнее магнитное поле однородно и направлено вдоль оси волновода. Введём систему координат, как показано на рис. 2, и будем искать поле внутри волновода в виде

$$\mathbf{E} = \exp(-ihz) [\mathbf{E}_1 \exp(-ik_x x - ik_y y) + \mathbf{E}_2 \exp(ik_x x - ik_y y) + \mathbf{E}_3 \exp(-ik_x x + ik_y y) + \mathbf{E}_4 \exp(ik_x x + ik_y y)]. \quad (\text{П1})$$

Решение (П1) после учёта граничных условий на металле приобретает вид

$$\begin{aligned} E_x &= E_0 \exp(-ihz) \left[\exp(ik_x x) - \frac{\alpha_+}{\alpha_-} \exp(-ik_x x) \right] \sin(k_y y), \\ E_y &= E_0 \alpha_+ \exp(-ihz) \left[\exp(ik_y y) - \frac{\alpha_-}{\alpha_+} \exp(-ik_y y) \right] \sin(k_x x), \\ E_z &= 2iE_0 \beta \exp(-ihz) \sin(k_x x) \sin(k_y y), \end{aligned} \quad (\text{П2})$$

где $k_x a = \pi n_1$; $k_y b = \pi n_2$; $n_1 = 0, \pm 1, \pm 2, \dots$; $n_2 = 0, \pm 1, \pm 2, \dots$, квадрат волнового числа h принимает значения

$$\begin{aligned} h_1^2 &= \frac{\omega^2}{c^2} \varepsilon_{xx} - \frac{k_\perp^2}{2} \left(1 + \frac{\varepsilon_{xx}}{\varepsilon_{zz}} \right) + \frac{1}{2} \sqrt{k_\perp^4 \left(1 - \frac{\varepsilon_{xx}}{\varepsilon_{zz}} \right)^2 + \frac{4\omega^2}{c^2} \frac{\varepsilon_{xy}^2}{\varepsilon_{zz}} k_\perp^2 - \frac{4\omega^4}{c^4} \varepsilon_{xy}^2}, \\ h_2^2 &= \frac{\omega^2}{c^2} \varepsilon_{xx} - \frac{k_\perp^2}{2} \left(1 + \frac{\varepsilon_{xx}}{\varepsilon_{zz}} \right) - \frac{1}{2} \sqrt{k_\perp^4 \left(1 - \frac{\varepsilon_{xx}}{\varepsilon_{zz}} \right)^2 + \frac{4\omega^2}{c^2} \frac{\varepsilon_{xy}^2}{\varepsilon_{zz}} k_\perp^2 - \frac{4\omega^4}{c^4} \varepsilon_{xy}^2}, \\ \alpha_+ &= \frac{(k_0^2 \varepsilon_{xy} - k_x k_y) (k_0^2 \varepsilon_{zz} - k_x^2 - k_y^2) + h^2 k_x k_y}{(k_0^2 \varepsilon_{xx} - h^2 - k_x^2) (k_0^2 \varepsilon_{zz} - k_x^2 - k_y^2) - h^2 k_y^2}, \end{aligned}$$

$$\alpha_- = \frac{(k_0^2 \varepsilon_{xy} + k_x k_y) (k_0^2 \varepsilon_{zz} - k_x^2 - k_y^2) - h^2 k_x k_y}{(k_0^2 \varepsilon_{xx} - h^2 - k_x^2) (k_0^2 \varepsilon_{zz} - k_x^2 - k_y^2) - h^2 k_y^2},$$

$$\beta = \frac{1}{hk_y} [k_0^2 \varepsilon_{xy} - k_x k_y - \alpha_+ (k_0^2 \varepsilon_{xx} - h^2 - k_x^2)].$$

Здесь E_0 — произвольная константа, a и b — геометрические размеры прямоугольного волновода. В случае точного резонанса ($\Delta\omega = 0$) решение (П2) можно записать в виде

$$E_x = E_0 \sin(k_x x) \sin(k_y y), \quad E_y = iE_0 \sin(k_x x) \sin(k_y y), \quad E_z = 0, \quad (\text{П3})$$

откуда следует, что поляризация колебаний соответствует поляризации нерезонансной компоненты поля, а амплитуда резонансной компоненты равна нулю. В случае конечной отстройки $\Delta\omega$ амплитуда резонансной компоненты пропорциональна $|\Delta\omega| \omega / \omega_p^2$.

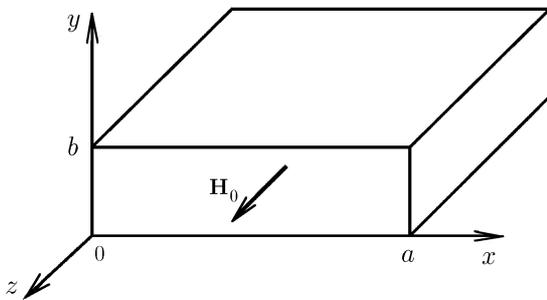


Рис. 2

Таким образом, в холодной магнитоактивной ограниченной плазме, как и в случае безграничной среды, существует депрессия резонансной (циркулярно поляризованной и вращающейся в ту же сторону, что и электрон в магнитном поле) компоненты электрического поля, распространяющейся под конечным углом к магнитному полю. Данный факт, вообще говоря, понятен, т. к. поле, возбуждаемое в ограниченной плазме, представляет собой не что иное, как суперпозицию нормальных волн безграничной плазмы.

Остановимся кратко на возбуждении низших мод (у которых отсутствует зависимость от одной из поперечных координат, т. е. $k_x = 0$ или $k_y = 0$) в области ЭЦР. Как видно из (П3), при этом поле в резонаторе тождественно обращается в нуль, т. е. такие моды вообще не возбуждаются. Очевидно, это связано с тем, что в магнитоактивной плазме собственные волны в общем случае поляризованы эллиптически в плоскости xy , выполнение граничных условий на металле (зануление тангенциальной компоненты поля) автоматически влечёт за собой зануление полного поля в резонаторе. Существование ненулевого поля внутри резонатора возможно лишь для мод с $n_1 \neq 0$ и $n_2 \neq 0$ у которых обращается в нуль полное поле на стенке резонатора.

СПИСОК ЛИТЕРАТУРЫ

1. Bacal M., Gaudin C., Bourdier A., Bruneteau J., Buzzi J. M., Golovanivski K. S., Hay L., Rouille C., Schwartz L. // Nature. 1996. V. 384. P. 421.
2. Golubev S., Zorin V., Vodopyanov A., Gaudin C., Razin S., Rouille C., Bacal M. // J. Plasma Fusion Res. Series. 2001. V. 4. P. 301.
3. Гинзбург Н. С., Токман М. Д. // ЖТФ. 1984. Т. 54. С. 1 062.
4. Tokman M. D. // Proc. Joint Varenna—Lausanna International Workshop on Theory of Fusion Plasmas (Varenna, 1994). Italy. Bologna: Soc. Italiana di Fisica, 1994. P. 437.
5. Литвак А. Г. // Вопросы теории плазмы. Вып. 10 / Под ред. М. А. Леонтовича. М.: Энергоатомиздат, 1980. С. 164.
6. Токман М. Д. // Физика плазмы. 1984. Т. 10. С. 568.
7. Cohen B. I., Cohen R. H., Nevins W. M., Rognlien T. D. // Rev. Mod. Phys. 1991. V. 63. P. 949.
8. Суворов Е. В., Токман М. Д. // Физика плазмы. 1984. Т. 10. С. 586.
9. Нейштадт А. И., Тимофеев А. В. // ЖЭТФ. 1987. Т. 66. С. 1 706.

10. Litvak A. G., Sergeev A. M., Suvorov E. V., Tokman M. D., Khazanov I. V. // Phys. Fluids B. 1993. V. 5. P. 4347.
11. Farina D., Pozzoli R. // Phys. Fluids B. 1991. V. 3. P. 570.
12. Kotel'nikov A., Stupakov G. V. // Phys. Fluids B. 1990. V. 2. P. 882.
13. Арнольд В. И., Козлов В. В., Нейштадт А. Н. // Итоги науки и техники. Серия Современные проблемы математики. Фундаментальные направления. Т. 3. / Под ред. Р. В. Гамкрелидзе. М.: ВИНТИ, 1985. С. 304.
14. Janaki M. S., Dasgupta B. // Phys. Fluids B. 1992. V. 4. P. 2696.
15. Janaki M. S., Dasgupta B. // Phys. Rev. E. 1995. V. 52. P. 6912.
16. Fraiman G. M., Kostyukov I. Yu. // Phys. of Plasmas. 1995. V. 2. P. 926.
17. Serebrennikov K. S., Gaudin C., Buzzi J.-M., Bruneteau J., Rouille C. // Phys. Rev. E. 2001. V. 63. Article No. 046405.
18. Gaudin C., Hay L., Buzzi J.-M., Bacal M. // Rev. Sci. Instrum. 1998. V. 69. P. 890.
19. Железняков В. В. Излучение в астрофизической плазме. М.: Янус-К, 1997.
20. Тимофеев А. В. // Вопросы теории плазмы. Вып. 14 / Под ред. М. А. Леонтовича. М.: Энергоатомиздат, 1989. С. 63.
21. Токман М. Д. // Физика плазмы. 1989. Т. 15. С. 373.
22. Чириков Б. В. // Вопросы теории плазмы. Вып. 13 / Под ред. М. А. Леонтовича. М.: Энергоатомиздат, 1989. С. 1.
23. Jaeger F., Lichtenberg A. J., Lieberman M. // Plasma Phys. 1972. V. 14. P. 1073.
24. Jaeger F., Lichtenberg A. J., Lieberman M. // Plasma Phys. 1973. V. 15. P. 125.
25. Schmidt G. Physics of High Temperature Plasmas: 2nd ed. New York: Academic Press, 1979.

Институт прикладной физики РАН,
г. Нижний Новгород, Россия

Поступила в редакцию
18 июня 2002 г.

SOME FEATURES OF HALF AND FUNDAMENTAL ECR HEATING IN A MAGNETIC TRAP

I. Yu. Kostyukov, A. Yu. Kryachko, and M. D. Tokman

We compare the efficiency of microwave heating of electrons in a dense plasma at the fundamental harmonic ($\omega = eH/(mc)$) and at subharmonics ($\omega = eH/(2mc)$) of the electron gyrofrequency. In particular, recent experimental results showing higher efficiency of the microwave heating at the frequency equal to one half of the electron gyrofrequency are analyzed. Equations describing nonlinear subharmonic electron cyclotron resonance (ECR) heating are derived for an arbitrary geometry of the microwave field. If the microwave field has "vacuum" polarization, then the microwave power absorbed by electrons at the fundamental harmonic of the electron gyrofrequency in rarefied plasmas exceeds by many order of magnitude the corresponding power absorbed by electrons in the case of nonlinear heating at one half of the electron gyrofrequency. However, it is shown that this difference in a dense plasma does not exceed one order of magnitude, which is explained by the effect of depression of the resonance component of the microwave field. In this case, the efficiency of the formation of high-energy electron population can be influenced not only by the energy-deposition rate but mainly by the stability condition of an electron in the magnetic trap. It is shown that a twofold decrease in the magnetic field, necessary to satisfy the ECR condition at half-harmonic of the electron gyrofrequency, leads to a dramatic shortening of hot-electron lifetime in the magnetic trap and, in turn, to a dramatic decrease in the energy-deposition efficiency. We discuss the dependence of electron heating on the effect of quasi-static enhancement of the microwave field near a target located in a magnetic trap for generation of X-ray emission.

УДК 517.9

УНИВЕРСАЛЬНЫЕ ЗАКОНОМЕРНОСТИ ПЕРЕХОДНЫХ ПРОЦЕССОВ*А. А. Короновский, Д. И. Трубецков, А. Е. Храмов, А. Е. Храмова*

В работе проводится анализ переходных процессов на примере логистического отображения. Выявляется зависимость длительности переходных процессов от начальных условий при различных значениях управляющего параметра. Показано, что с увеличением управляющего параметра λ зависимость длительности переходного процесса закономерно усложняется, что обусловлено появлением неустойчивых $2n$ -циклов. В работе также показывается, что для зависимости длительности переходных процессов характерно несколько скейлинговых закономерностей.

Изучение сложного поведения нелинейных динамических систем, как распределённых, так и с сосредоточенными параметрами, привлекает в настоящее время большое внимание исследователей. Во многом это связано с такой фундаментальной проблемой, как исследование закономерностей динамического хаоса. Наиболее простыми и в то же время демонстрирующими многие характерные явления в нелинейных динамических системах являются системы с дискретным временем (отображения).

Однако на фоне активно ведущихся исследований динамического хаоса наблюдается очевидный дефицит результатов, относящихся к анализу переходных процессов, роль которых в динамике нелинейных систем весьма существенна [1–3]. Для ряда динамических систем (системы с импульсным внешним воздействием; системы с очень длительными переходными процессами; системы с мультистабильностью; распределённые автоколебательные системы) переходной процесс в значительной мере определяет их поведение. С проблемой исследования характера переходных процессов связана задача о границах бассейнов притяжения аттракторов динамических систем в случае мультистабильности [4–6]. Детальное понимание закономерностей динамики подобных систем невозможно без анализа систем с малым числом степеней свободы, в том числе и систем с дискретным временем.

Данная работа посвящена исследованию переходных процессов в классической модели нелинейной динамики — логистическом отображении [7]:

$$x_{n+1} = f(x_n) = \lambda x_n (1 - x_n). \quad (1)$$

В работе исследовалась зависимость длительности переходного процесса от начальных условий при различных значениях управляющего параметра λ , лежащего в диапазоне от 1 до 3,57. Иными словами, рассматривалась область субгармонического каскада [7]. Для определения длительности переходного процесса использовался следующий метод. Для фиксированного значения управляющего параметра λ определялся аттрактор, реализующийся в системе, для чего произвольная точка начальных условий x_0 итерировалась $N = 6\,500$ раз; после этого полагалось, что изображающая точка «вышла» на аттрактор. Анализируя полученную последовательность $\{x_n\}_{n=0}^N$ начиная с $n = N - 1, N - 2, \dots$, определялся период режима (неподвижная устойчивая точка, 2-цикл, 4-цикл и т. д.). Затем, последовательно перебирая все возможные начальные условия с шагом разбиения $\Delta x = 10^{-7}$, для каждого начального условия определялся интервал дискретного времени $K(x_0)$, необходимый для того, чтобы изображающая точка «вышла» на найденный ранее аттрактор с абсолютной точностью $\varepsilon = 10^{-7}$. Вышеописанный метод хорошо работает для логистического отображения (1) при всех значениях управляющего параметра λ , за исключением значений λ_i , отвечающих точкам бифуркации удвоения периода, в которых длительность переходного процесса стремится к бесконечности. Следует отметить, что для динамических систем, в которых возможно явление мультистабильности, вышеописанный метод

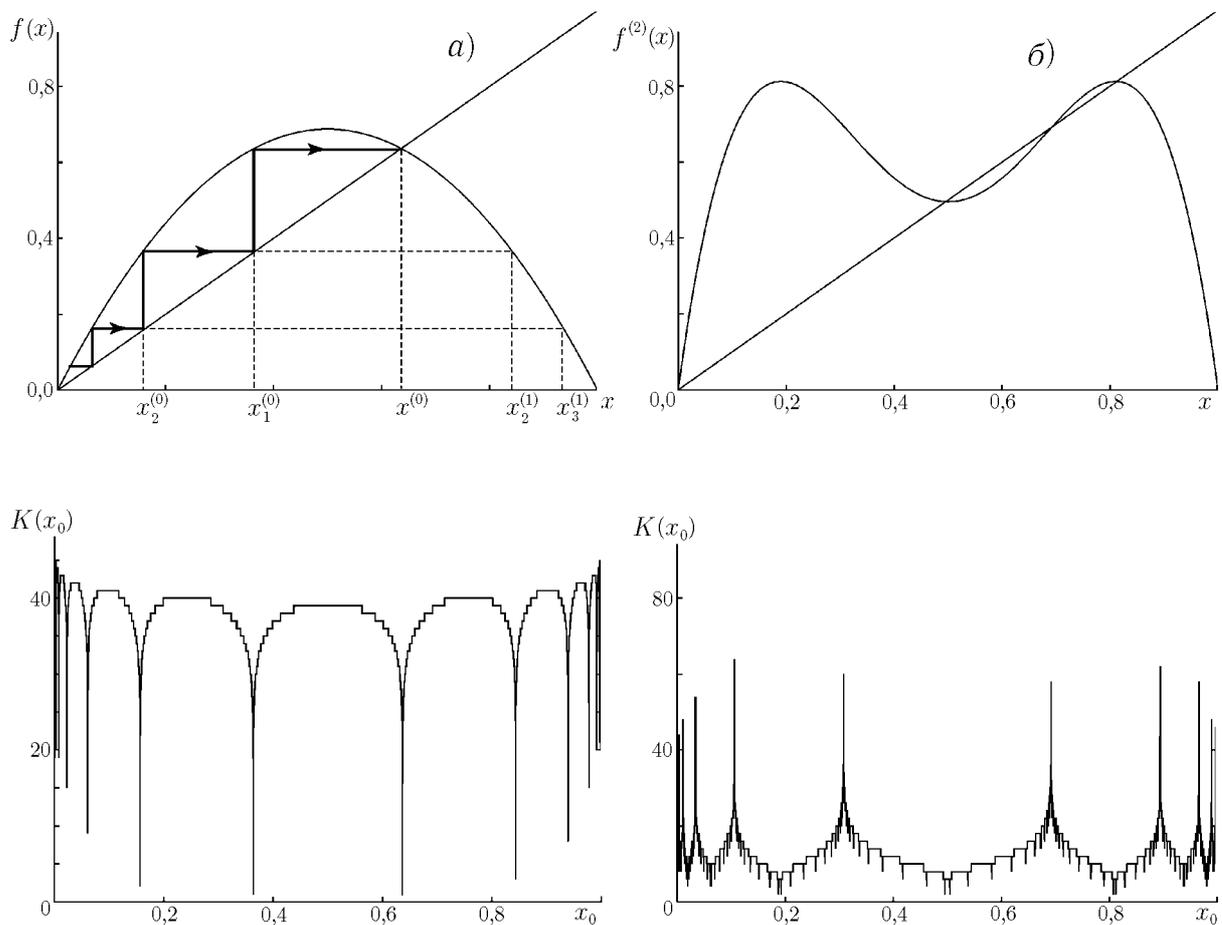


Рис. 1

необходимо модифицировать, определяя аттрактор для каждого начального значения x_0 .

Рис. 1 иллюстрирует зависимость длительности переходных процессов при значениях управляющего параметра λ , соответствующих случаям, когда в системе реализуется неподвижная устойчивая точка (рис. 1а) и устойчивый цикл периода 2 (рис. 1б). На рис. 1а приведена диаграмма Ламерея и зависимость длительности переходного процесса для логистического отображения при управляющем параметре $\lambda = 2,75$, соответствующем случаю, когда в системе реализуется неподвижная устойчивая точка; рис. 1б иллюстрирует двукратно проитерированное логистическое отображение $f^{(2)}(x)$ и зависимость длительности переходного процесса $K(x_0)$ при $\lambda = 3,25$, когда аттрактором является устойчивый цикл периода 2.

Как видно из рис. 1, зависимость длительности K переходного процесса от начальных условий x_0 имеет достаточно сложный, «изрезанный» вид. С ростом управляющего параметра λ зависимость $K(x_0)$ закономерно усложняется.

Глобальные минимумы зависимости длительности переходного процесса приходятся на начальные условия x_0 , совпадающие с точками аттрактора, который существует в фазовом пространстве системы при заданном значении управляющего параметра λ . Иными словами, при значениях параметра $\lambda = 1 \div 3$ глобальный минимум зависимости длительности переходного процесса приходится на устойчивую неподвижную точку $x^0 = (\lambda - 1)/\lambda$, являющуюся аттрактором (см. рис. 1а). С увеличением параметра λ и потерей устойчивости неподвижной точкой x^0 глобальные минимумы длительности переходного процесса сначала приходятся на элементы устойчивого 2-цикла ($\lambda = 3 \div 1 + \sqrt{6}$), затем на

элементы устойчивого цикла периода 4 и т. д.

Поскольку логистическое отображение (1) является необратимым, то для значений управляющего параметра $\lambda = 2 \div 3$ существуют¹ две последовательности точек $\{x_i^{(0)}\}_{i=1}^{\infty}$ и $\{x_i^{(1)}\}_{i=1}^{\infty}$, которые за конечное число итераций отображаются в неподвижную устойчивую точку x^0 :

$$x^0 = f(x_1^{(0)}) = f(f(x_2^{(0)})) = f(f(f(x_2^{(1)}))) = \dots = f^{(n)}(x_n^{(0)}) = f^{(n)}(x_n^{(1)}). \quad (2)$$

Последовательность $\{x_i^{(0)}\}_{i=1}^{\infty}$ сходится к границе $x = 0$ бассейна притяжения аттрактора по закону

$$x_{i+1}^{(0)} = x_i^{(0)}/\lambda, \quad i \rightarrow \infty, \quad (3)$$

а последовательность $\{x_i^{(1)}\}_{i=1}^{\infty}$ — к границе $x = 1$ бассейна притяжения аттрактора по закону

$$x_{i+1}^{(1)} = 1 - x_i^{(1)}/\lambda, \quad i \rightarrow \infty. \quad (4)$$

Так как точки последовательностей $\{x_i^{(0)}\}_{i=1}^{\infty}$ и $\{x_i^{(1)}\}_{i=1}^{\infty}$ отображаются за конечное число итераций в неподвижную устойчивую точку x_0 , то и в этих точках наблюдаются минимумы длительности переходных процессов (см. рис. 1а), причём $K(x_{i+1}^{(k)}) = K(x_i^{(k)}) + 1$, где $k = 0; 1$. Аналогичная ситуация реализуется и для максимумов длительности переходных процессов. Таким образом, можно отметить, что для зависимости длительности переходного процесса K от начальных условий x_0 в рассматриваемой системе имеет место скейлинг с масштабным коэффициентом λ относительно границ бассейна притяжения аттрактора $x = 0$ и $x = 1$.

На рис. 2а показан скейлинг зависимости длительности переходного процесса $K(x_0)$ с масштабным коэффициентом λ относительно границы $x = 0$ бассейна притяжения аттрактора. Область слева от пунктирной линии растягивается по горизонтали в λ раз. После изменения масштаба зависимость длительности переходного процесса повторяет себя, сдвигаясь в область больших значений на 1. Управляющий параметр λ выбран равным 2,75, что соответствует случаю, когда в системе реализуется неподвижная устойчивая точка.

С увеличением управляющего параметра λ в системе происходит каскад бифуркаций удвоения периода [7], и зависимость длительности K переходного процесса от начальных условий x_0 усложняется, что обусловлено несколькими факторами. Во-первых, в связи с тем, что появляются неустойчивые циклы, в точках начальных условий x_0 , совпадающих с элементами указанных циклов, длительность переходного процесса бесконечна.² Вместе с тем длительность переходного процесса оказывается бесконечной и в множестве точек $\{y_i^{(k)}\}_{i=1}^{\infty}$, которые, в силу необратимости отображения (1), за конечное число итераций отображаются в элементы неустойчивых циклов. Во-вторых, число последовательностей $\{x_i^{(k)}\}_{i=1}^{\infty}$ и $\{y_i^{(k)}\}_{i=1}^{\infty}$, элементы которых отображаются за конечное число итераций в элементы либо устойчивых, либо неустойчивых циклов, возрастает.³

Скейлинг зависимости длительности переходного процесса в окрестности границ бассейна притяжения аттрактора имеет место и в том случае, когда в системе реализуются устойчивые $2n$ -циклы, где $n = 1, 2, \dots$ (рис. 2б, 3б). На рис. 2б показан скейлинг зависимости длительности переходного процесса $K(x_0)$ с масштабным коэффициентом λ относительно границы $x = 0$ бассейна притяжения аттрактора. Область слева от пунктирной линии растягивается по горизонтали в λ раз. После изменения масштаба зависимость длительности переходного процесса повторяет себя, сдвигаясь в область

¹ За исключением случая $\lambda = 2$, когда $x^0 = 0,5$ и не существует никаких других точек x , кроме x^0 , таких, что $f(x) = x^0$.

² В силу конечной точности представления начальных условий x_0 максимумы длительности переходных процессов имеют различную величину и не обращаются в бесконечность при численном моделировании.

³ В общем случае при $\lambda > 3$ число таких последовательностей бесконечно.

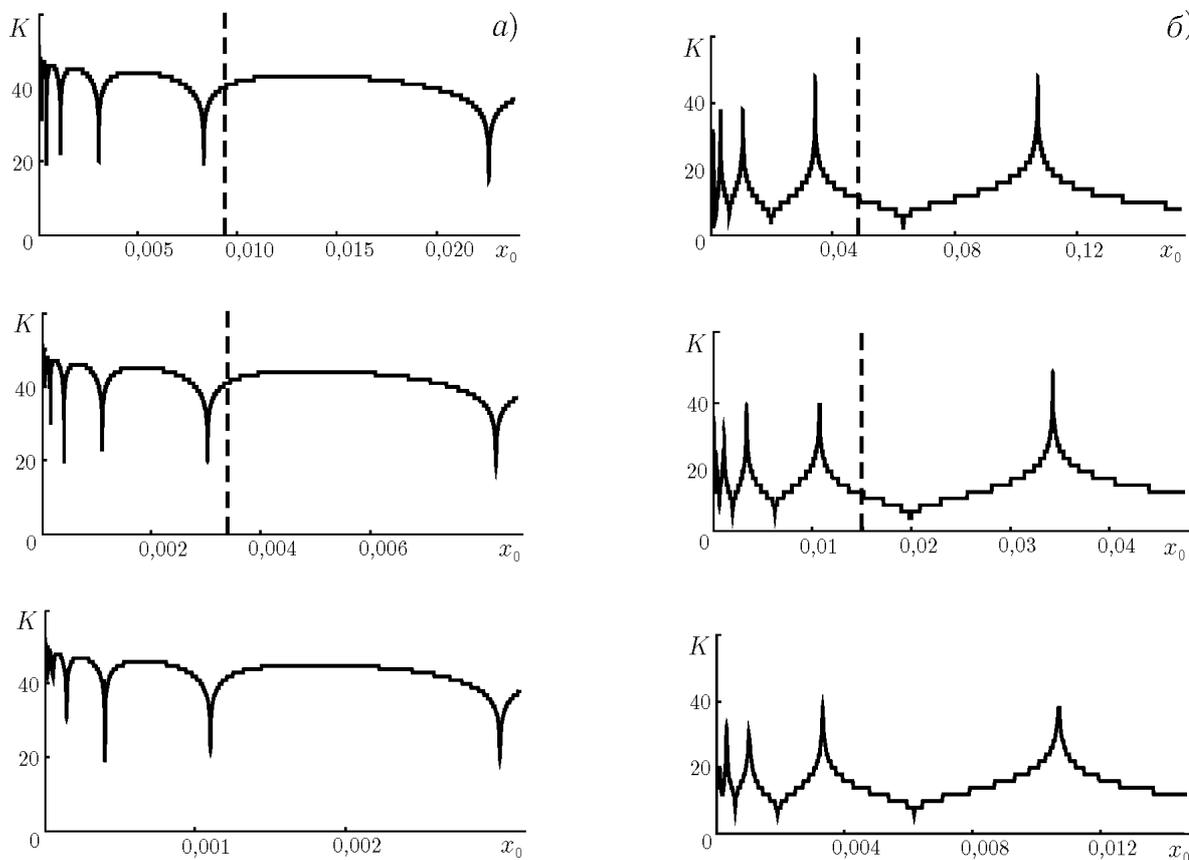


Рис. 2

больших значений на 1. Управляющий параметр $\lambda = 3,236$ соответствует случаю, когда в системе реализуется цикл максимальной устойчивости периода 2.

Аналогичная ситуация показана на рис. 3б для $\lambda = 3,3$. В этом случае в системе реализуется устойчивый цикл периода 2. Возникновение «изрезанности» (по сравнению с рис. 2б) связано с появлением последовательностей $\{z_i^{(k)}\}_{i=0}^{\infty}$, где $k = 0, 1$, элементы которых за конечное число итераций отображаются в элементы устойчивого 2-цикла.

Однако после того, как неподвижная точка x^0 потеряла устойчивость (и появился устойчивый цикл периода 2), в окрестности этой неподвижной неустойчивой точки наблюдается явление скейлинга с масштабным коэффициентом $\mu = df(x^0)/dx$, который является мультипликатором этой неподвижной точки. На рис. 3а показан скейлинг в окрестности неустойчивой неподвижной точки x^0 при $\lambda = 3,25$. Область, заключённая между пунктирными линиями, растягивается по горизонтали с масштабным коэффициентом $\mu = 2,562$. После изменения масштаба зависимость длительности переходного процесса $K(x_0)$ повторяет себя, сдвигаясь в область больших значений на 1. Фактически, появление неустойчивой неподвижной точки порождает две последовательности $\{z_i^{(k)}\}_{i=1}^{\infty}$, где $k = 0, 1$, сходящиеся слева и справа к неустойчивой точке x^0 по закону

$$z_{i+1}^{(k)} = x^0 + \frac{z_i^{(k)} - x^0}{\mu}, \quad i \rightarrow \infty. \quad (5)$$

Каждому элементу последовательностей $\{z_i^{(k)}\}_{i=1}^{\infty}$ соответствует локальный минимум зависимости $K(x_0)$ (см. рис. 3), и аналогично (2) каждый элемент $z_i^{(k)}$ отображается за конечное число итераций

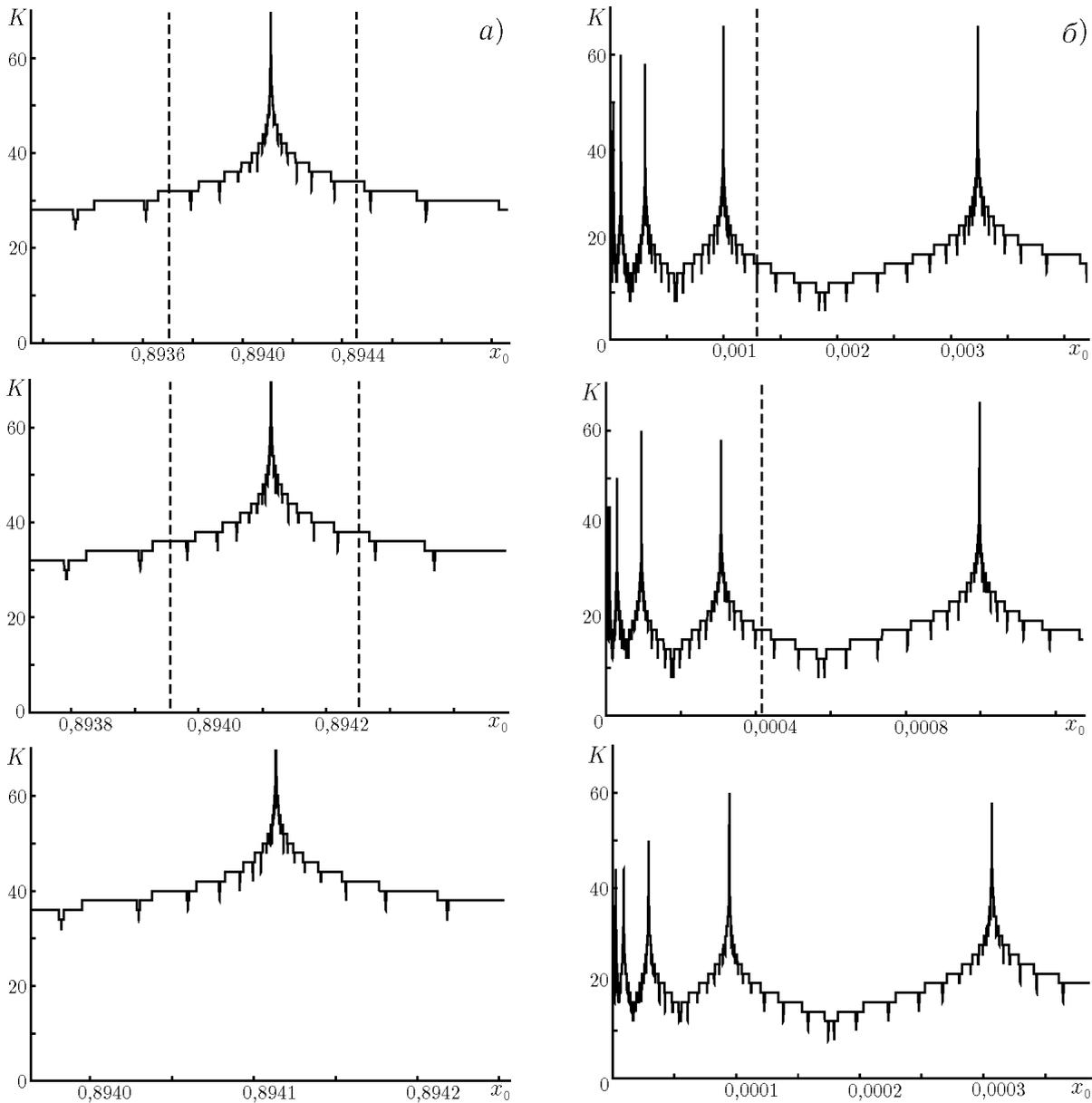


Рис. 3

в элементы устойчивого 2-цикла. Более того, каждый элемент последовательностей $\{z_i^{(k)}\}_{i=1}^{\infty}$ порождает одну или несколько последовательностей $\{x_i^{(k)}\}_{i=1}^{\infty}$, сходящихся к границам $x = 0$ и $x = 1$ бассейна притяжения аттрактора. Иными словами, скейлинг относительно неустойчивой точки x^0 «транслируется» на скейлинг в окрестности границ бассейна притяжения аттрактора (рис. 3б). Исключением является случай, когда в системе реализуется цикл периода 2 максимальной устойчивости (рис. 2б). В этом случае не существует никаких других точек, которые за конечное число итераций отображались бы в элементы устойчивого 2-цикла, кроме самих границ бассейна притяжения аттрактора.

Понятно, что аналогичные явления будут наблюдаться с увеличением управляющего параметра λ и на базе элементов потерявших устойчивость $2n$ -циклов, где $n = 0, 1, \dots$. Таким образом, можно отметить, что с увеличением параметра λ зависимость длительности переходных процессов $K(x_0)$ закономерно усложняется. На рис. 4 приведены зависимости длительности переходного процесса K от

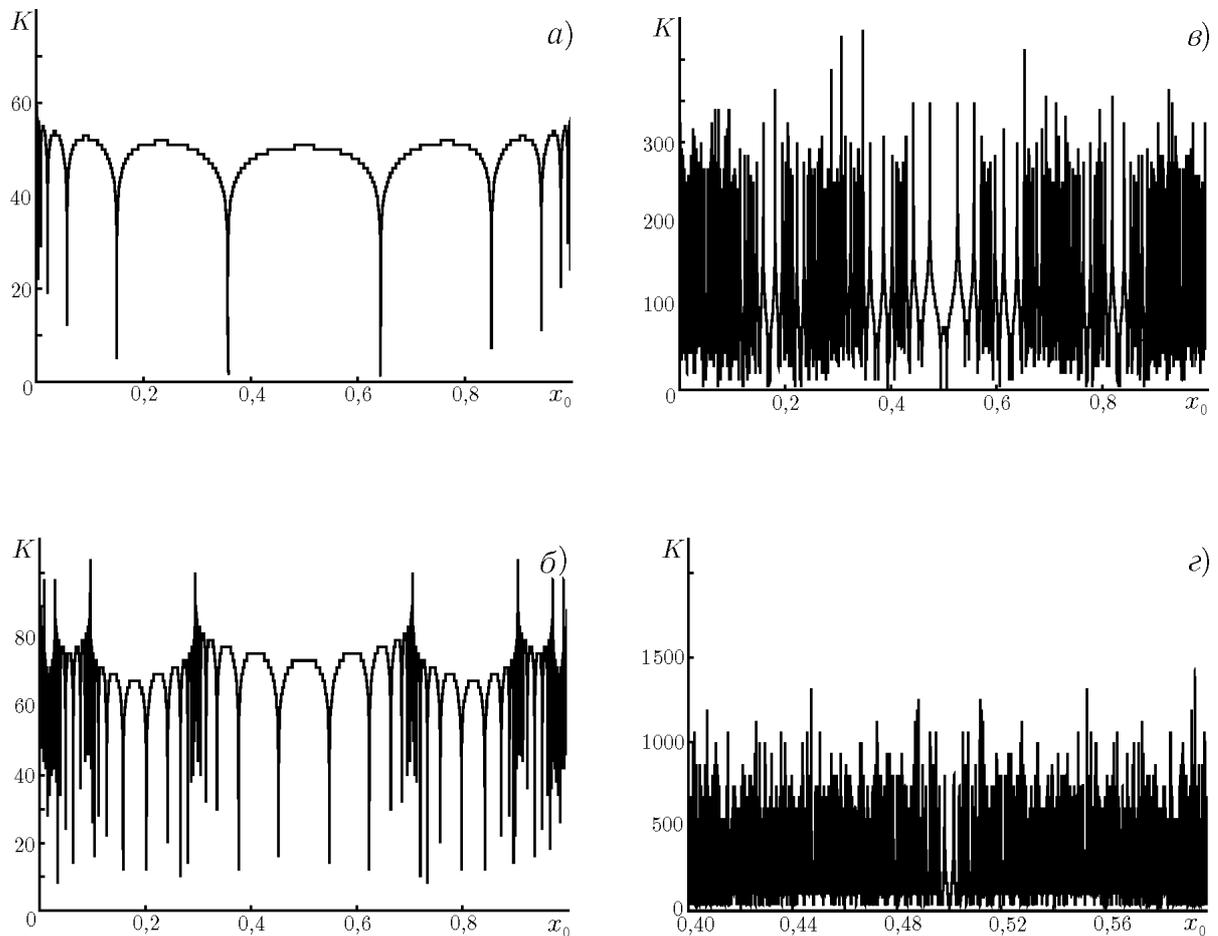


Рис. 4

начальных условий x_0 при разных значениях управляющего параметра: рис. 4а соответствует $\lambda = 2,8$ (неподвижная устойчивая точка), рис. 4б — $\lambda = 3,4$ (цикл периода 2), рис. 4в — $\lambda = 3,55$ (цикл периода 8), рис. 4г — $\lambda = 3,5698$ (цикл периода 64). Усложнение зависимости $K(x_0)$ связано с появлением в результате каскада бифуркаций удвоения периода элементов неустойчивых $2n$ -циклов.

Отметим также, что скейлинг зависимости переходных процессов $K(x_0)$, описанный выше, не является единственным. Существует ещё один вид скейлинга, обусловленный поведением логистического отображения в критической точке [7]. Зависимости длительности переходных процессов для циклов максимальной устойчивости подчиняются скейлинговым закономерностям с константами скейлинга $a = -2,503$ и $b = 2$, что следует из процедуры ренорм-группового анализа [8]. При масштабировании выбранного участка в a раз относительно точки $x = 0,5$ и увеличении длительности переходного процесса в b раз для $2n$ -цикла, где $n = 1, 2, \dots$, вид зависимости $K(x_0)$ соответствует аналогичной зависимости для $(2n - 1)$ -цикла.

Скейлинг с масштабными коэффициентами $a = -2,503$ и $b = 2$ приведён на рис. 5 для циклов максимальной устойчивости периода 2 ($\lambda = 3,236$, рис. 5а), 4 ($\lambda = 3,499$, рис. 5б) и 8 ($\lambda = 3,555$, рис. 5в).

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 01-02-17392), U. S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (проект REC-006) и программы «Университеты России — фундаментальные исследования».

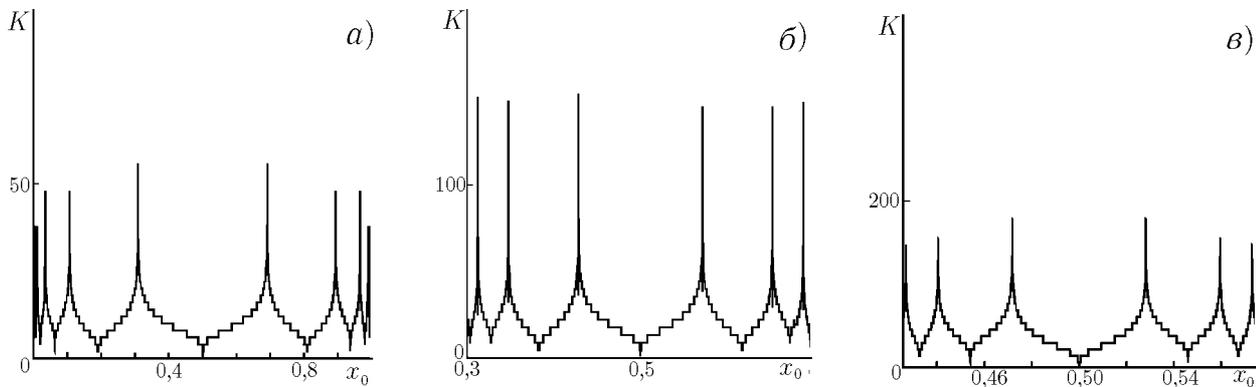


Рис. 5

СПИСОК ЛИТЕРАТУРЫ

1. Астахов А. В. и др. // Радиофизика и электроника. 1993. Т. 38, № 2. С. 291.
2. Кальянов Э. В. // Письма в ЖТФ. 2000. Т. 26, № 15. С. 26.
3. Bezruchko V. P., Dikanov T. V., Smirnov D. A. // Phys. Rev. E. 2001. V. 64. Article no. 036210.
4. Grebogi C., Ott E. // Phys. Rev. Lett. 1983. V. 50, No. 13. P. 935.
5. Nusse H. E., Ott E., Yorke J. A. // Phys. Rev. Lett. 1995. V. 75, No. 13. P. 2482.
6. Grebogi C., Ott E., Yorke J. A. // Phys. Rev. Lett. 1986. V. 56, No. 10. P. 1011.
7. Кузнецов А. П., Кузнецов С. П. // Изв. вузов. Прикладная нелинейная динамика. 1993. Т. 1, № 1–2. С. 15.
8. Feigenbaum M. J. // J. Stat. Phys. 1979. V. 21. P. 669.

Саратовский госуниверситет им. Н. Г. Чернышевского,
ГосУНЦ «Колледж»,
г. Саратов, Россия

Поступила в редакцию
19 ноября 2001 г.

UNIVERSAL RULES OF TRANSIENT PROCESSES

A. A. Koronovsky, D. I. Trubetskov, A. E. Khramov, and A. E. Khramova

We analyze the transients of a logistic map. The dependence of the transient duration on the initial conditions for different values of the control parameter is revealed. It is shown that as the control parameter λ increases, dependence of the transient duration on the initial conditions becomes more complicated, which can be explained by the appearance of unstable $2n$ -cycles. It is also shown that several definite scaling laws exist for the discussed function.

УДК 538.3;539.2

ПОВЕРХНОСТНЫЕ ПЛАЗМЕННЫЕ ВОЛНЫ НА НЕРОВНОЙ ГРАНИЦЕ ТВЁРДОГО ТЕЛА

С. И. Ханкина¹, В. М. Яковенко¹, И. В. Яковенко²

В работе получен и исследован закон дисперсии поверхностных поляритонов в неоднородной плазме, образованной поверхностными электронными состояниями на границе твёрдого тела в условиях, когда масштабы неоднородностей велики по сравнению с длиной волны. Показано, что неоднородность плазмы приводит к возникновению пространственной дисперсии поверхностных электростатических колебаний.

1. Как известно, характеристики распространения поверхностных волн — области существования, спектр, распределение поля и т. п. — зависят от свойств поверхности раздела сред. В частности, они существенно видоизменяются, если граница не является гладкой, а имеет регулярные или случайные неровности, на которых происходит многократное рассеяние поверхностной волны. Особенности распространения волн различной природы (акустических, электромагнитных, волн пространственного заряда и др.) над такими поверхностями изучались в многочисленных работах (см., например, [1–5]). При этом обычно используется математический аппарат, основанный на представлении о неровностях как о малых возмущениях, влияние которых учитывается или в уравнениях, или в граничных условиях, или в тех и других одновременно.

В работах [6, 7] было показано, что на границе твёрдого тела с малыми неровностями могут существовать поверхностные «электронные волны» — поверхностные электронные состояния. В настоящей работе получены и исследованы дисперсионные характеристики поверхностных поляритонов (поверхностных плазменных волн) в неоднородной плазме, возникновение которых обусловлено электронными состояниями на неровной поверхности. Эти волны обладают рядом особенностей. Например, их фазовая скорость меньше, чем у поляритонов, распространяющихся вдоль гладкой поверхности однородной плазмы, концентрация которой совпадает с концентрацией электронов на поверхности неоднородной плазмы.

2. Предположим, что на границе двух сред с различными электромагнитными свойствами (диэлектрик—плазменная среда) создан потенциальный барьер

$$U(x, y, z) = \begin{cases} 0, & y > y_0(x, z); \\ \infty, & y \leq y_0(x, z), \end{cases} \quad (1)$$

форма поверхности которого задаётся функцией $y_0(x, z)$.

Следуя работе [6], предположим сначала, что поверхность является периодически неровной:

$$y_0(x, z) = \zeta_0 \cos(G_x x) \cos(G_z z), \quad (2)$$

где ζ_0 — высота неровности, $G_x = 2\pi/a$, $G_z = 2\pi/b$; a и b — пространственные периоды неровностей границы вдоль осей x и z соответственно.

В этом случае решение стационарного уравнения Шрёдингера, описывающего поведение частицы в области $y > y_0(x, z)$, можно представить в виде

$$\psi_{\mathbf{k}} = \sum_{l_x l_z} a_{l_x l_z} \exp[i(k_x + l_x G_x)x + i(k_z + l_z G_z)z + i k_y l_x l_z y], \quad (3)$$

где $a_{l_x l_z}$ — амплитуды соответствующих гармоник, k_x и k_z — компоненты волнового вектора, l_x и l_z — целые числа,

$$k_{y l_x l_z}^2 = \frac{2mE_{\mathbf{k}}}{\hbar^2} - (k_x + l_x G_x)^2 - (k_z + l_z G_z)^2, \quad (4)$$

$E_{\mathbf{k}}$ и m — энергия и масса частицы (электрона) соответственно.

На границе $y = y_0(x, z)$ выполняется условие

$$\mathbf{n} \nabla \psi_{\mathbf{k}}|_{y=y_0(x, z)} = 0, \quad (5)$$

где \mathbf{n} — нормаль к поверхности $y = y_0(x, z)$.

Если неровности поверхности пологие, т. е. $\zeta_0 G_x \ll 1$, $\zeta_0 G_z \ll 1$, то граничное условие (5) можно перенести с неровной поверхности на гладкую. Тогда на плоскости $y = 0$ выполняется соотношение

$$\frac{\partial \psi_{\mathbf{k}}}{\partial x} \frac{\partial y_0}{\partial x} + \frac{\partial \psi_{\mathbf{k}}}{\partial z} \frac{\partial y_0}{\partial z} - \frac{\partial \psi_{\mathbf{k}}}{\partial y} - y_0 \frac{\partial^2 \psi_{\mathbf{k}}}{\partial y^2} = 0. \quad (6)$$

Спектр поверхностных электронных состояний $E_{\mathbf{k}} = E_{\mathbf{k}}(k_x, k_z)$ в этом приближении можно найти, ограничиваясь рассмотрением взаимодействия нескольких гармоник: a_{00} ; a_{01} ; a_{0-1} ; a_{10} ; a_{-10} ; a_{11} ; a_{-1-1} , т. к. величина связи остальных гармоник с нулевой пропорциональна ζ_0^n , где $n > 2$.

В этом нетрудно убедиться, если поверхность имеет периодические неровности вдоль одного направления, например вдоль оси z . Дисперсионное соотношение $E_{\mathbf{k}} = E_{\mathbf{k}}(k_x, k_z)$, или $k_y = k_y(k_x, k_z)$, определяется из уравнения

$$-i k_{y0} a_0 + \frac{\zeta_0}{2} (k_{y0}^2 - k_z G_z) a_1 + \frac{\zeta_0}{2} (k_{y0}^2 + k_z G_z) a_{-1} = 0,$$

в котором

$$a_1 = -i \frac{\zeta_0}{2k_{y1}} (k_{y0}^2 - k_z G_z) a_0, \quad a_{-1} = -i \frac{\zeta_0}{2k_{y-1}} (k_{y0}^2 + k_z G_z) a_0,$$

$$k_{y0}^2 = \frac{2mE}{\hbar^2} - k_x^2 - k_z^2; \quad k_{y1}^2 = k_{y0}^2 - G_z (G_z + 2k_z); \quad k_{y-1}^2 = k_{y0}^2 - G_z (G_z - 2k_z). \quad (7)$$

Решение (7) находится методом последовательных приближений по малому параметру $\zeta_0 G_z$:

$$k_{y0} = k_{y0}^{(0)} + \delta k_y. \quad (8)$$

При $\zeta_0 = 0$

$$k_{y0}^{(0)} = 0 \quad (9)$$

и

$$k^2 = \frac{2mE_{\mathbf{k}}^{(0)}}{\hbar^2}; \quad k^2 = k_x^2 + k_z^2.$$

В следующем приближении получим величину δk_y , которая имеет наиболее простой вид в предельных случаях: $k_z \ll G_z$ — длинноволновый предел, $k_z \sim G_z$ — резонансный случай, $k_z \gg G_z$ — коротковолновый предел.

Например, в длинноволновом приближении для одномерных периодических неровностей имеем

$$\delta k_y = i \frac{k_z^2 \zeta_0^2 G_z}{2}; \quad \text{Im } k_{y0} > 0; \quad (10)$$

$$E_{\mathbf{k}} = E_{\mathbf{k}}^{(0)} + \delta E_{\mathbf{k}}, \quad \delta E_{\mathbf{k}} = \frac{\hbar^2 (\delta k_y)^2}{2m}.$$

Аналогично находится спектр поверхностных состояний для границы, периодически неровной в двух направлениях [7].

Таким образом, неровности границы раздела двух сред приводят к возникновению поверхностных электронных состояний, волновая функция которых имеет вид

$$\psi_{\mathbf{k}} = a_0 \exp[-|\delta k_y| y + i(k_x x + k_z z - \omega_{\mathbf{k}} t)], \quad (11)$$

$$\hbar \omega_{\mathbf{k}} = E_{\mathbf{k}}^0 + \delta E_{\mathbf{k}}.$$

Из условия нормировки

$$\iiint \psi_{\mathbf{k}} \psi_{\mathbf{k}}^* dx dy dz = 1 \quad (12)$$

определяется амплитуда $a_0 = \sqrt{2|\delta k_y|/S}$, где $S = L_x L_z$, L_x и L_z — размеры образца в направлениях x и z соответственно.

Предположим, что в области $y > y_0(x, z)$ создана неоднородная плазма (такую плазму можно создать, например, путём ионизации примесных атомов, удерживаемых неровной поверхностью диэлектрика, или на поверхности полупроводника). Выразим основные параметры такой плазмы через характеристики неровной поверхности. Концентрацию электронов $n_0(y)$ определим следующим образом:

$$n_0(y) = \sum_{k_x, k_z} \psi_{\mathbf{k}} \psi_{\mathbf{k}}^* n_{\mathbf{k}} = \frac{2}{S} \sum_{k_x, k_z} n_{\mathbf{k}} |\delta k_y| \exp[-2|\delta k_y| y], \quad (13)$$

где $n_{\mathbf{k}}$ — число электронов в состоянии с волновым вектором \mathbf{k} ; суммирование ведётся по всем значениям волновых чисел. При этом минимальные значения \mathbf{k} определяются размерами образца в направлениях x и z , т. е. величинами L_x и L_z , а максимальные — ферми-импульсом $\hbar k_F$ и параметрами решётки G_x , G_z . Полное число частиц в области $y > 0$ равно $\sum_{\mathbf{k}} n_{\mathbf{k}}$, а поверхностная плотность $n_s = \sum_{\mathbf{k}} n_{\mathbf{k}}/S = \int_0^\infty n_0(y) dy$.

Приведём значения $n_0(0)$ и n_s для вырожденного электронного газа при $n_{\mathbf{k}} = 0; 1$.

Если поверхность раздела сред является периодически неровной только вдоль оси z ($k_{z \max} = G_z/2$, $k_{z \min} = 2\pi/L_z$), то при $k_F \ll G_z/2$

$$n_s = k_F^2/(4\pi), \quad n_0(0) = \zeta_0^2 k_F^2 G_z n_s/4, \quad (14)$$

при $k_F \gg G_z/2$

$$n_s = k_F G_z/(2\pi^2), \quad n_0(0) = \zeta_0^2 G_z^3 n_s/12. \quad (15)$$

Например, если $b = 2 \cdot 10^{-5}$ см, $k_F = 6 \cdot 10^6$ см $^{-1}$, то $n_s \approx 10^{11}$ см $^{-2}$, $n_0(0) \approx 10^{15}$ см $^{-3}$.

Для двумерной периодической неровности при $G_z \sim G$, $G_x \sim G$ и $k_F \ll G$

$$n_s = k_F^2/(4\pi), \quad n_0(0) = \zeta_0^2 k_F^2 G n_s/2. \quad (16)$$

Заметим, что плазма заряженных частиц является нейтральной, т. к. мы предполагаем, что заряд электронов с концентрацией $n_0(y)$ полностью компенсируется положительно заряженными частицами.

3. Определим спектр поверхностных электромагнитных колебаний на границе такой плазмы ($y > 0$) и диэлектрика ($y < 0$). Система уравнений, описывающая электромагнитные колебания в плазме, имеет вид

$$\text{rot } \mathbf{E}_1 = -\frac{1}{c} \frac{\partial \mathbf{H}_1}{\partial t}, \quad \text{rot } \mathbf{H}_1 = \frac{1}{c} \frac{\partial \mathbf{D}_1}{\partial t} + \frac{4\pi}{c} \mathbf{j}, \quad \text{div } \mathbf{D}_1 = 4\pi e n(\mathbf{r}, t), \quad e \frac{\partial n}{\partial t} + \text{div } \mathbf{j} = 0; \quad (17)$$

$$\mathbf{D}_1(\mathbf{r}, t) = \int_{-\infty}^t \hat{\varepsilon}_0(t-t') \mathbf{E}_1(\mathbf{r}, t') dt'. \quad (18)$$

Уравнение (18) описывает электромагнитные свойства среды в отсутствие неровностей; фурье-образ диэлектрической проницаемости можно представить в виде

$$\varepsilon_0(\omega) = \int_0^{\infty} \hat{\varepsilon}_0(\tau) \exp(i\omega\tau) d\tau.$$

В уравнениях (17) $n(\mathbf{r}, t)$ — отклонение концентрации электронов от равновесного значения $n_0(y)$, \mathbf{j} — ток, который в отсутствие пространственной дисперсии среды выражается через векторный потенциал \mathbf{A} следующим образом:

$$\mathbf{j} = -\frac{e^2 n_0(y)}{mc} \mathbf{A}(\mathbf{r}, t) \quad (19)$$

(калибровка выбрана таким образом, что потенциал $\phi = 0$ и $\mathbf{E} = -c^{-1} \partial \mathbf{A} / \partial t$), e — заряд электрона, c — скорость света в вакууме.

В диэлектрике

$$\text{rot } \mathbf{E}_2 = -\frac{1}{c} \frac{\partial \mathbf{H}_2}{\partial t}, \quad \text{rot } \mathbf{H}_2 = \frac{\varepsilon_2}{c} \frac{\partial \mathbf{E}_2}{\partial t}. \quad (20)$$

Предположим, что длина электромагнитной волны превосходит высоту и период неровностей поверхности. В этом случае граничные условия для электромагнитных полей такие же, как и на гладкой поверхности, т. е. на плоскости $y = 0$ непрерывны тангенциальные составляющие электрического и магнитного полей.

Зависимость векторного потенциала, как и всех переменных величин в уравнениях (17)–(20), от координат и времени зададим в виде $\mathbf{A}(\mathbf{r}, t) = \mathbf{A}(y) \exp[i(q_x x - \omega t)]$, где q_x и ω — волновое число и частота электромагнитного поля соответственно. Вектор \mathbf{A} имеет компоненты A_x и A_y , а у магнитного поля есть только компонента H_z . Относительно компоненты H_{z1} система уравнений (17) сводится к уравнению

$$\left[\frac{\omega^2}{c^2} \varepsilon_1(y) - q_x^2 \right] H_{z1} + \varepsilon_1(y) \frac{\partial}{\partial y} \left[\frac{1}{\varepsilon_1(y)} \frac{\partial H_{z1}}{\partial y} \right] = 0, \quad (21)$$

$$\varepsilon_1(y) = \varepsilon_0(\omega) - \omega_0^2(y)/\omega^2, \quad \omega_0^2(y) = 4\pi e^2 n_0(y)/m,$$

решение которого ищем в виде

$$H_{z1}(y) = C_1 \exp \left[- \int_0^y q(y') dy' \right]. \quad (22)$$

Для определения $q(y)$ получим из (21) уравнение

$$q^2 + \frac{1}{\varepsilon_1(y)} \frac{\partial \varepsilon_1(y)}{\partial y} q + \frac{\omega^2}{c^2} \varepsilon_1(y) - q_x^2 = 0, \quad (23)$$

которое справедливо при выполнении условия $q^{-1} dq/dy \ll q$. Это условие означает, что относительное изменение глубины проникновения волны в среду $y > 0$ является малым на расстояниях порядка $\lambda = 1/q$. Из двух решений уравнения (23) выбираем то, которое удовлетворяет условию убывания электромагнитного поля при $y \rightarrow \infty$:

$$q_1 = -\frac{1}{2\varepsilon_1(y)} \frac{\partial \varepsilon_1(y)}{\partial y} + \sqrt{q_x^2 - \frac{\omega^2}{c^2} \varepsilon_1(y)}. \quad (24)$$

Здесь $q_x^2 > \omega^2 \varepsilon_1(y)/c^2$ и $q_x^2 - \omega^2 \varepsilon_1(y)/c^2 > [\varepsilon_1^{-1}(y)/\partial \varepsilon_1(y)/\partial y]^2$. Второе неравенство означает, что глубина проникновения электромагнитного поля в плазму меньше масштаба области локализации волновой функции электрона, т. е. $q_1 \sim q_x \gg |\delta k_y|_{\max}$. Например, в случае одномерной периодической решётки $|\delta k_y|_{\max} = \zeta_0^2 G^3/8$, и это условие легко выполняется.

Компоненты электрического поля в плазме выражаются через компоненту H_{z1} следующим образом:

$$E_{x1} = -i \frac{q_1 c}{\omega \varepsilon_1(y)} H_{z1}, \quad E_{y1} = \frac{q_x c}{\omega \varepsilon_1} H_{z1}. \quad (25)$$

В диэлектрике

$$H_{z2} = C_2 \exp(q_2 y); \quad E_{x2} = i \frac{q_2 c}{\omega \varepsilon_2} H_{z2}, \quad E_{y2} = \frac{q_x c}{\omega \varepsilon_2} H_{z2}, \quad (26)$$

где $q_2 = \sqrt{q_x^2 - \omega^2 \varepsilon_2/c^2} > 0$.

Из граничных условий на плоскости $y = 0$ следует

$$q_1(0) = -q_2 \frac{\varepsilon_1(0)}{\varepsilon_2}. \quad (27)$$

Это равенство выполняется, если $\varepsilon_1(0) < 0$. Подставляя в (27) значения q_1 и q_2 , получим дисперсионное уравнение для поперечных поверхностных электромагнитных волн (поверхностных поляритонов):

$$q_x^2 = \frac{\omega^2}{c^2} \frac{\varepsilon_1(0)\varepsilon_2}{\varepsilon_1(0) + \varepsilon_2} - \frac{\varepsilon_1'(0)\varepsilon_2^2 [q_x^2 - \omega^2 \varepsilon_1(0)/c^2]^{1/2}}{\varepsilon_1(0) [\varepsilon_1^2(0) - \varepsilon_2^2]}, \quad \varepsilon_1'(0) = -\frac{4\pi e^2}{m\omega^2} \frac{\partial n_0(y)}{\partial y} \Big|_{y=0}. \quad (28)$$

Уравнение (28) решаем относительно q_x методом последовательных приближений по малому параметру $\varepsilon_1'(0)/q_x \varepsilon_1(0) \ll 1$. В результате получим

$$q_x = q_{x0} + \delta q_x, \quad q_{x0} = \frac{\omega^2}{c^2} \frac{\varepsilon_1(0)\varepsilon_2}{\varepsilon_1(0) + \varepsilon_2}, \quad \delta q_x = \frac{\varepsilon_1'(0)\varepsilon_2^{3/2}}{2|\varepsilon_1(0)|^{1/2}(\varepsilon_1^2(0) - \varepsilon_2^2)}. \quad (29)$$

Величина δq_x положительна, т. к. значение $\varepsilon_1'(0)$ всегда положительно. Действительно, $\varepsilon_1'(0) = 16\pi e^2/(m\omega^2 S) \sum_{k_x, k_z} n_{\mathbf{k}} |\delta k_y|^2 > 0$. Например, в случае границы, периодически неровной в одном направлении, имеем $\varepsilon_1'(0) = [e^2/(m\omega^2)] \zeta_0^4 G_z^4 k_F^2/16$, где $k_F \ll G_z/2$.

Таким образом, наличие поверхностных электронных состояний приводит к возникновению поверхностных поляритонов, закон дисперсии которых отличается от закона дисперсии поляритонов, распространяющихся вдоль гладкой поверхности однородной плазмы. В частности, фазовая скорость полученных волн меньше в результате зависимости концентрации электронов от нормальной координаты.

Для электростатических колебаний ($c \rightarrow \infty$) из (28) следует соотношение

$$\varepsilon_1(0) + \varepsilon_2 = -\frac{\varepsilon_1'(0)\varepsilon_2^2}{\varepsilon_1(0)(\varepsilon_1(0) - \varepsilon_2)q_x}. \quad (30)$$

В этом случае

$$\omega = \omega' + \delta\omega.$$

Если в плазменной среде отсутствует частотная дисперсия диэлектрической проницаемости ($\varepsilon_0 = \text{const}$), то

$$\omega'^2 = \frac{\omega_0^2}{\varepsilon_0 + \varepsilon_2}, \quad \delta\omega = -\frac{\omega' \varepsilon_1'(0)}{4q_x}, \quad (31)$$

т. е. возникает дисперсия частоты, а групповая скорость волны положительна.

Заметим, что в случае $\varepsilon_0(\omega) = \varepsilon_0 - \Omega^2/\omega^2$ частота ω' определяется выражением

$$\omega'^2 = \frac{\Omega^2 + \omega_0^2}{\varepsilon_0 + \varepsilon_2}, \tag{32}$$

где Ω — ленгмюровская частота однородной плазмы.

Если глубина проникновения электромагнитной волны больше области локализации волновой функции электронов ($|q_x| \ll |\delta k_y|$), то на поверхности раздела сред образуется двумерный электронный газ. Можно показать, что дисперсионное уравнение поверхностных электростатических колебаний в этом случае имеет вид

$$\varepsilon_0 + \varepsilon_2 = \frac{\omega_{0s}^2}{\omega^2} |q_x|; \quad \omega_{0s}^2 = \frac{4\pi e^2 n_s}{m}. \tag{33}$$

4. Аналогичные результаты можно получить, если граница раздела сред имеет случайные неровности. В этом случае форма границы описывается случайной функцией координат $y_0 = \zeta(x, z) = \zeta(\mathbf{r})$, относительно свойств которой предполагаем следующее: $\zeta(\mathbf{r})$ — однозначная и достаточно гладкая функция, т. е. дифференцируемая необходимое число раз по всем аргументам, при этом $|\partial\zeta/\partial x| \ll 1$ и $|\partial\zeta/\partial z| \ll 1$ (неровности пологие). Кроме того,

$$\bar{\zeta}(\mathbf{r}) = 0, \quad \overline{\zeta(\mathbf{r})\zeta(\mathbf{r}')} = \bar{\zeta}_0^2 W(|\mathbf{r} - \mathbf{r}'|), \tag{34}$$

где черта означает усреднение по ансамблю реализаций неровной поверхности, $\bar{\zeta}_0$ — среднее отклонение поверхности от плоскости $y = 0$, $W(|\mathbf{r} - \mathbf{r}'|)$ — корреляционная функция шероховатостей поверхности.

При решении уравнения Шредингера с граничным условием (5) используется стандартная процедура определения поля над статистически неровной поверхностью [8]. Соотношение, определяющее спектр поверхностных электронных состояний частиц, имеет вид [7]

$$k_y = -\bar{\zeta}_0^2 \int d\boldsymbol{\kappa} \frac{(\mathbf{k}\boldsymbol{\kappa} - k^2 - k_y^2)^2}{\sqrt{k^2 + k_y^2 - \boldsymbol{\kappa}^2}} W(|\mathbf{k} - \boldsymbol{\kappa}|), \tag{35}$$

где $\mathbf{k} = (k_x, k_z)$, $\boldsymbol{\kappa} = (\kappa_x, \kappa_z)$, $W(|\mathbf{k}|)$ — фурье-преобразование корреляционной функции, вид которого предполагается гауссовым.

Из (35) следует, что при $\bar{\zeta}_0 = 0$ имеем $k_y = 0$. Подставляя далее в (35) $k_y = 0$ и проводя интегрирование, получим выражения для δk_y в различных предельных случаях. Например, в длинноволновом пределе $kl \ll 1$, где l — корреляционная длина,

$$\delta k_y = i \frac{\sqrt{\pi}}{2} \frac{\bar{\zeta}_0^2 k^2}{\sqrt{2} l}, \tag{36}$$

а энергия соответствующего локализованного у поверхности состояния равна

$$E = \frac{\hbar^2 k^2}{2m} \left(1 - \frac{\pi \bar{\zeta}_0^4 k^2}{8l^2} \right).$$

Для определения закона дисперсии волн, распространяющихся вдоль статистически неровной поверхности, при $q_x \ll \delta k_y$ можно воспользоваться формулами (21)–(28). Из них при $k_F l < 1$ следуют выражения для $n_0(0)$ и $\partial n_0/\partial y|_{y=0}$:

$$n_0(0) = \frac{\sqrt{\pi}}{2l\sqrt{2}} \bar{\zeta}_0^2 k_F^2 n_s, \quad \left. \frac{\partial n_0}{\partial y} \right|_{y=0} = \frac{\pi \bar{\zeta}_0^4 k_F^4 n_s}{12l^2}; \quad n_s = \frac{k_F^2}{4\pi}. \tag{37}$$

Для характерных значений (см. [9]) $n_s \approx 10^{12} \text{ см}^{-2}$; $l = 1,5 \cdot 10^{-7} \text{ см}$; $\bar{\zeta}_0 = 5 \cdot 10^{-8} \text{ см}$ имеем $n_0 \approx 10^{17} \text{ см}^{-3}$, $\partial n_0 / \partial y|_{y=0} \approx 10^{22} \text{ см}^{-4}$.

5. Таким образом, на границе твёрдого тела с регулярными или случайными неровностями, представляющими собой малые возмущения поверхности, можно создать неоднородную плазму, параметры которой определяются свойствами поверхности (высотой и периодом неровностей). На границе такой плазмы с диэлектриком в области частот, меньших ленгмюровской, распространяются поверхностные электромагнитные волны, фазовая скорость которых определяется электронной концентрацией на границе раздела сред. Неоднородность плазмы приводит к дисперсии частоты, когда эффектом запаздывания можно пренебречь. Групповая скорость такой волны положительна, а её величина определяется характером убывания электронной концентрации от границы.

Возбуждение такого рода поверхностных поляритонов может дать важную информацию об электронных состояниях на неровной границе.

СПИСОК ЛИТЕРАТУРЫ

1. Rohman T., Maradudin A. // Phys. Rev. B. 1980. V. 21, No. 6. P. 2 137.
2. Equilus A. G., Maradudin A. A. // Phys. Rev. B. 1983. V. 28, No. 2. P. 728.
3. Кондратенко А. Н., Ханкина С. И., Яковенко И. В. // Изв. вузов. Радиофизика. 1983. Т. 26, № 12. С. 1 579.
4. Яковенко В. М. // УФЖ. 1982. Т. 27, № 9. С. 1 424.
5. Гуляев Ю. В., Плесский В. П. // Радиотехника и электроника. 1980. Т. 25, вып. 8. С. 1 569.
6. Pogrebnyak V. A., Yakovenko V. M., Yakovenko I. V. // Phys. Lett. A. 1995. V. 209, No. 3. P. 103.
7. Погребняк В. А., Яковенко В. М., Яковенко И. В. // ФТТ. 1997. Т. 39, № 10. С. 1 874.
8. Басс Ф. Г., Фукс И. М. Рассеяние волн на статистически неровной поверхности. М.: Наука, 1972. 424 с.
9. Андо Т., Фаулер А., Стерн Ф. Электронные свойства двумерных систем. М.: Мир, 1985. 415 с.

¹ Институт радиофизики и электроники им. А. Я. Усикова
НАН Украины,

Поступила в редакцию
6 апреля 2001 г.

² Научно-исследовательский и проектно-конструкторский
институт «Молния» Министерства образования и науки
Украины,
г. Харьков, Украина

SURFACE PLASMA WAVES AT A ROUGH INTERFACE OF A SOLID

S. I. Khankina, V. M. Yakovenko, and I. V. Yakovenko

The dispersion relation for surface polaritons in an inhomogeneous plasma is obtained and studied. The plasma is formed by surface electron states at the interface of a solid under conditions where the scale of inhomogeneity is much greater than the wavelength. It is shown that the nonuniformity of the plasma results in spatial dispersion of the surface electrostatic oscillations.

УДК 519.23

ОЦЕНКИ ПАРАМЕТРОВ СИГНАЛА С МИНИМАЛЬНОЙ КОВАРИАЦИЕЙ

В. И. Турчин

Показано, что существует единственная процедура, обеспечивающая абсолютный минимум среднеарифметического значения ковариации внутри некоторой области изменения параметров и тождественно совпадающая с алгоритмом вычисления оценки, минимизирующей средний риск при квадратичной функции потерь и равномерном априорном распределении параметров. Эта процедура сопоставлена с обычной техникой получения максимально правдоподобных оценок; показано, в частности, что при оценке параметров детерминированного сигнала на фоне гауссовой помехи процедура даёт существенно меньший случайный разброс оценок. В ситуации, когда из физических соображений ковариация оценки не зависит от оцениваемого параметра, с помощью указанной процедуры при произвольном объёме выборки достигается реальная дисперсионная граница.

1. ПОСТАНОВКА ЗАДАЧИ

Методы обработки, сводящиеся к определению нескольких числовых параметров сигнала (частоты, пеленга, времени запаздывания, амплитуды, средней мощности и т. п.), находят самое широкое применение в радио- и гидролокации, навигации, связи и других приложениях (см., например, [1]). Классическая постановка такой задачи обычно формулируется следующим образом. Имеется результат наблюдений — вектор $\tilde{\mathbf{x}}$ длины N , представляющий реализацию случайного вектора \mathbf{x} . Этот случайный вектор характеризуется известной N -мерной плотностью распределения вероятности $w(\mathbf{x}, \boldsymbol{\theta})$, зависящей ещё и от K числовых параметров $\theta_1, \dots, \theta_K$, образующих вектор $\boldsymbol{\theta}$; эти параметры считаются неизвестными. Требуется построить оценку $\hat{\boldsymbol{\theta}}$ вектора $\boldsymbol{\theta}$, которая в самом общем виде может быть представлена как значение некоторой зависящей от N переменных вектор-функции $\mathbf{t}(\mathbf{x})$ — статистики. Это значение находится после подстановки в качестве аргумента результата наблюдений: $\hat{\boldsymbol{\theta}} = \mathbf{t}(\tilde{\mathbf{x}})$. Приемлемая оценка должна быть, очевидно, близка к точному (но неизвестному) вектору $\boldsymbol{\theta}$. Проблема заключается, собственно, в отыскании статистики, с помощью которой может быть получена приемлемая оценка (при этом мы ещё должны ввести некоторый критерий её качества).

Известны в основном две стратегии построения «хороших» оценок [1–4]. Первая из них использует объективный критерий отклонения оценки от точного значения. В качестве меры отклонения естественно выбрать матрицу ковариации оценки $\mathbf{C} = \text{cov}\{\hat{\boldsymbol{\theta}}\}$. Соответственно, приемлемая статистика должна минимизировать, например, диагональные элементы \mathbf{C} . Хорошо известны ограничения, показывающие, что нельзя построить статистику, реализующую произвольно малое отклонение оценки от точного значения. Среди этих ограничений наиболее известно неравенство Крамера—Рао (см., например, [2, 3]):

$$\mathbf{C} \geq \mathbf{F}^{-1}, \quad (1)$$

где \mathbf{F} — матрица Фишера (см., например, [4]), а \mathbf{F}^{-1} представляет так называемую минимальную границу дисперсии (МГД). В большинстве практически интересных случаев оценки, реализующие равенство в (1), не существуют (см., например, [5]), однако известен широкий класс оценок, асимптотически приближающихся к МГД при неограниченном увеличении объёма выборки.¹ Наиболее распростра-

¹ В рассматриваемом случае объёмом выборки можно считать число статистически независимых блоков, на которые может быть разбит вектор \mathbf{x} .

нена оценка по максимуму правдоподобия (МП):

$$\hat{\theta}_{\text{МП}} = \arg \max_{\theta} w(\tilde{\mathbf{x}}, \theta), \quad (2)$$

здесь $w(\tilde{\mathbf{x}}, \theta)$ — функция правдоподобия: плотность распределения вероятности с реализацией случайного вектора \mathbf{x} в качестве аргумента. При большом числе статистически независимых элементов вектора \mathbf{x} МП оценка практически не может быть улучшена, однако при их ограниченном числе вопрос о качестве МП оценки остаётся открытым.

Вторая стратегия использует байесовский подход. Неизвестный параметр θ считается случайной величиной и вводится его субъективная априорная плотность распределения вероятности $w_a(\theta)$, с помощью которой строится апостериорная (условная) плотность распределения вероятности $w_{\text{ар}}(\mathbf{x}, \theta) \sim w(\mathbf{x}, \theta)w_a(\theta)$. Расхождение между некоторой оценкой $\hat{\theta}$ и случайной величиной θ характеризуется функцией потерь $\Pi(\theta, \hat{\theta})$, монотонно возрастающей с увеличением нормы разности между $\hat{\theta}$ и θ . В рамках такого подхода приемлемая оценка должна минимизировать среднее от функции потерь по апостериорной плотности распределения вероятности (так называемый средний риск). Оценка, обеспечивающая минимум среднего риска (МСР), представляется в виде

$$\hat{\theta}_{\text{МСР}} = \arg \min_{\hat{\theta}} \int \Pi(\theta, \hat{\theta}) w(\tilde{\mathbf{x}}, \theta) w_a(\theta) d\theta. \quad (3)$$

МСР-оценки не нашли широкого практического применения, отчасти из-за трудности введения субъективной априорной плотности распределения вероятности и известного произвола при выборе функции потерь [4].

Цель настоящей работы заключается в построении статистик, минимизирующих объективную величину отклонения оценки от точного значения параметра, т. е. ковариацию оценки, для произвольного объёма выборки. Забегая вперёд, отметим, что алгоритм получения таких оценок оказывается схожим с (3), однако введения каких-либо субъективных допущений при этом не требуется.

2. ПРИНЦИП МИНИМУМА СРЕДНЕЙ КОВАРИАЦИИ

Рассмотрим для простоты случай одного вещественного параметра θ ($K = 1$) и некоторую статистику $t(\mathbf{x})$. Ковариация оценки $C(\theta, \{t\})$, зависящая от точного (но неизвестного) значения параметра и являющаяся функционалом относительно $t(\mathbf{x})$, может быть записана в виде

$$C(\theta, \{t\}) = \int [t(\mathbf{x}) - \theta]^2 w(\mathbf{x}, \theta) d\mathbf{x}. \quad (4)$$

Прямая минимизация C по $t(\mathbf{x})$ невозможна из-за логического противоречия, вытекающего из незнания точного значения искомого параметра. Проиллюстрируем это, взяв две различные функции $t_1(\tilde{\mathbf{x}})$ и $t_2(\tilde{\mathbf{x}})$. Каждой из этих функций соответствует некоторая зависимость ковариации оценки от неизвестного параметра θ ; возможный вид таких зависимостей в некоторой интересующей нас области, $\vartheta_1 \leq \theta \leq \vartheta_2$ представлен на рис. 1. Очевидно, что если точное значение параметра находится в области $\vartheta_1 \leq \theta \leq \vartheta'$, использование функции $t_2(\tilde{\mathbf{x}})$ предпочтительнее, а в области $\vartheta' \leq \theta \leq \vartheta_2$ следует воспользоваться функцией $t_1(\tilde{\mathbf{x}})$. Однако значение параметра нам неизвестно, и мы не можем угадать, какому из двух интервалов оно принадлежит, а следовательно, не можем сделать обоснованный выбор между функциями $t_1(\tilde{\mathbf{x}})$ и $t_2(\tilde{\mathbf{x}})$. Это рассуждение иллюстрирует невозможность поиска «наилучшего» алгоритма, дающего оценку с минимальной ошибкой для каждого значения θ . В то же время задача становится осмысленной, если ввести некоторую характеристику ковариации оценки, относящуюся ко всему интервалу изменения параметра.

Такой простейшей характеристикой является среднеарифметическое значение c ковариации $C(\theta)$ в некотором интервале $\theta \in [\vartheta_1, \vartheta_2]$:

$$c = \frac{1}{\vartheta_2 - \vartheta_1} \int_{\vartheta_1}^{\vartheta_2} C(\theta) d\theta = \int [t^2(\mathbf{x})\tau_0(\mathbf{x}) - 2t(\mathbf{x})\tau_1(\mathbf{x}) + \tau_2(\mathbf{x})] d\mathbf{x}, \tag{5}$$

где

$$\tau_m(\mathbf{x}) = \frac{1}{\vartheta_2 - \vartheta_1} \int_{\vartheta_1}^{\vartheta_2} \theta^m w(\mathbf{x}, \theta) d\theta, \quad m = 0, 1, 2. \tag{6}$$

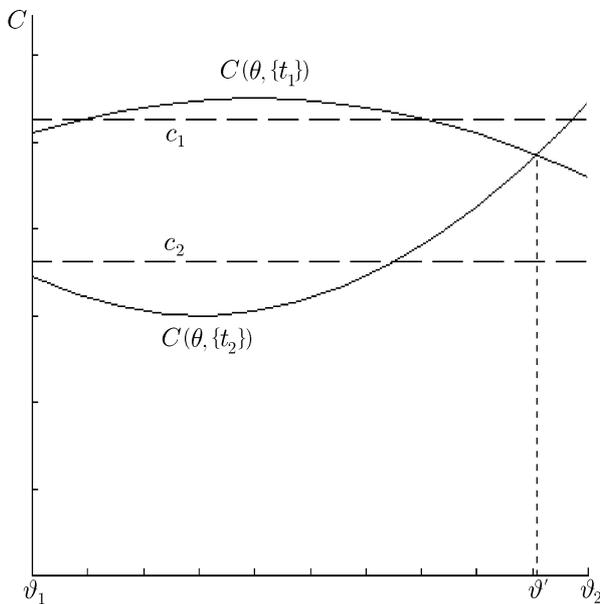


Рис. 1. Возможные зависимости ковариации оценки от оцениваемого параметра θ для двух статистик $t_1(\tilde{\mathbf{x}})$ и $t_2(\tilde{\mathbf{x}})$

Как видно из рис. 1, согласно этому критерию мы должны отдать предпочтение функции t_2 , пожертвовав возможным ухудшением точности оценки в интервале $\vartheta' \leq \theta \leq \vartheta_2$.

Находя первую вариацию (5) по t , получаем алгоритм построения оценки $\hat{\theta}_{\text{МСК}}$, обеспечивающей абсолютный минимум среднеарифметической ковариации:

$$\hat{\theta}_{\text{МСК}} = \tau_1(\tilde{\mathbf{x}}) / \tau_0(\tilde{\mathbf{x}}). \tag{7}$$

Заметим, что границы интервала $[\vartheta_1, \vartheta_2]$ не играют никакой роли при построении МСК-оценки (7) и могут быть расширены до $[-\infty, \infty]$, если это не вступает в противоречие с физическим смыслом оцениваемого параметра.²

Одним из свойств МСК-оценки является её тождественное совпадение с МСР-оценкой в случае, когда априорное распределение w_a выбирается равномерным в интервале $[\vartheta_1, \vartheta_2]$ и используется квадратичная функция потерь $\Pi = (\theta - \hat{\theta})^2$ (явный вид МСР-оценки в этом случае см., например, в [4]).

Таким образом, обе стратегии получения приемлемых оценок приводят к одному и тому же алгоритму обработки. Действительно, равномерная априорная плотность распределения вероятности фактически не меняет апостериорную плотность распределения вероятности, а квадратичная функция потерь ассоциируется со вторым моментом, представляющим определение ковариации (дисперсии для несмещённой оценки). Принципиальная разница, однако, заключается в интерпретации результата.

Обобщение данного подхода на многомерный случай ($K > 1$) не представляет трудностей. Вводя среднеарифметические значения c_k диагональных элементов ковариационной матрицы $C_{kk}(\theta)$ для $\theta \in \Theta$, где Θ — некоторая K -мерная область изменения векторного параметра,

$$c_k = \frac{1}{V_\theta} \int_{\Theta} C_{kk}(\theta) d\theta, \quad V_\theta = \int_{\Theta} d\theta,$$

и беря вариационную производную от c_k по статистике $t_k(\tilde{\mathbf{x}})$, вычисляющей оценку k -го элемента θ , приходим в точности к МСР-оценке при квадратичной функции потерь $\Pi = (\theta - \hat{\theta})^T (\theta - \hat{\theta})$, где

² Например, при оценке такого параметра, как дисперсия, не имеет смысла рассматривать отрицательные значения ϑ_1 .

индекс Т обозначает транспонирование, [4]:

$$\hat{\theta}_{\text{МСК(МСР)}} = \int_{\Theta} \theta w(\tilde{\mathbf{x}}, \theta) d\theta / \int_{\Theta} w(\tilde{\mathbf{x}}, \theta) d\theta. \quad (8)$$

Отметим, что (8), как и (2), представляет собой процедуру определения аргумента, при котором функция правдоподобия достигает глобального максимума по искомому параметру, только с помощью (8) этот аргумент находится не непосредственно, а с использованием техники моментов. Если функция правдоподобия имеет ярко выраженный пик по θ в окрестности точного значения параметра (например, при слабой зашумлённости данных или большом объёме выборки), то МП и МСК(МСР)-оценки дадут близкий результат. При слабо выраженном максимуме процедура определения его положения с помощью моментов будет, по всей видимости, более устойчивой.

В заключение ещё раз подчеркнём, что с точки зрения достижения минимума среднеарифметического значения ковариации оценки (7), (8) не могут быть улучшены в силу единственности решения вариационной задачи. При этом весьма интересной оказывается ситуация, при которой ковариация оценки не зависит от оцениваемого параметра (такой случай будет рассмотрен в следующем разделе). В этом случае среднеарифметическое значение ковариации и собственно величина ковариации совпадают. Тем самым (7), (8) реализуют точный минимум ковариации оценки, равный реально достижимой границе дисперсии. Последняя, естественно, будет превышать МГД, следующую из неравенства Крамера—Рао (1) и недостижимую в большинстве практически интересных случаев. Отметим, что реально достижимая граница дисперсии может быть найдена аналитически, например в случае $K = 1$ после подстановки (7) в (5).

3. ПРАКТИЧЕСКИЙ ПРИМЕР

Проиллюстрируем особенности сформулированного выше алгоритма на простейшем примере: оценке двух параметров детерминированного сигнала на фоне аддитивной гауссовой помехи с нулевым средним и известной матрицей ковариации Φ :

$$x_n = \theta_1 s_n(\theta_2) + \xi_n, \quad (9)$$

где $n = 1, \dots, N$, ξ_n — помеха, θ_1 — неизвестная амплитуда сигнала, s_n характеризует известную форму сигнала, определённую с точностью до числового параметра θ_2 ; в данном примере θ имеет две компоненты, θ_1 и θ_2 . Для простоты положим $\Phi = \sigma^2 \mathbf{I}$, где \mathbf{I} — единичная матрица. Обычно наибольший интерес представляет оценка параметра θ_2 , нелинейно входящего в (9); этой задачей мы и ограничимся. Данные наблюдений x_n будем здесь предполагать комплексными.

Как известно (см., например, [1]), МП алгоритм в данном случае сводится к нахождению глобального максимума функции $\Psi(\theta_2)$:

$$\hat{\theta}_{2\text{МП}} = \arg \max_{\theta} \Psi(\theta), \quad \Psi(\theta) = \frac{|\mathbf{s}^H(\theta) \tilde{\mathbf{x}}|^2}{\mathbf{s}^H(\theta) \mathbf{s}(\theta)}, \quad (10)$$

где $\mathbf{s}(\theta) = (s_1(\theta), \dots, s_N(\theta))^T$, индекс Н обозначает эрмитовое сопряжение.

Чтобы получить явный вид МСК-оценки, мы должны в соответствии с (8) проинтегрировать нормальное распределение вектора \mathbf{x} по параметру θ_1 (амплитуде детерминированного сигнала). Поскольку этот параметр предполагается комплексным, интегрирование должно выполняться по его действительной и мнимой частям; в качестве областей интегрирования выберем всю числовую ось. После

несложных вычислений получаем

$$\hat{\theta}_{2\text{МСК}} = \int_{\vartheta_1}^{\vartheta_2} \theta \exp[\Psi(\theta)/\sigma^2] d\theta / \int_{\vartheta_1}^{\vartheta_2} \exp[\Psi(\theta)/\sigma^2] d\theta, \tag{11}$$

т. е. мы должны подвергнуть статистику Ψ нелинейному преобразованию в соответствии с мощностью помехи σ^2 , а затем найти максимум полученной зависимости по θ , используя метод моментов. Очевидно, что при малом σ^2 или, точнее, при большом значении параметра

$$Q = s^H s \frac{|\theta_1|^2}{\sigma^2}, \tag{12}$$

представляющего отношение сигнал/шум (ОСШ) после обработки, нелинейное искажение резко подчеркнёт глобальный максимум $\Psi(\theta)$ в окрестности точного значения θ_2 , и оценки (10) и (11) окажутся близки. При не слишком большом выходном ОСШ нелинейное искажение будет сглаживать обусловленные шумами локальные экстремумы, что должно повысить точность оценки.

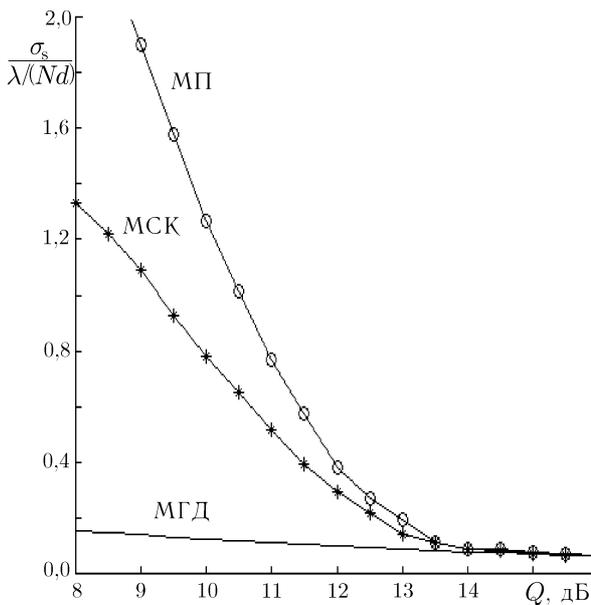


Рис. 2. Нормированные зависимости среднееквдратичного отклонения МП и МСК-оценок синуса пеленга и минимальная граница дисперсии в зависимости от выходного ОСШ Q для 20-элементной антенной решётки

В качестве конкретного примера возьмём один временной отсчёт сигнала на выходе линейной N -элементной антенной решётки. В этом случае $s_n(\theta) = \exp[-i\kappa d(n-1)\theta]$, где κ — волновое число, d — межэлементное расстояние, θ — синус угла прихода плоской волны, падающей на решётку (синус пеленга). В качестве интервала $[\vartheta_1, \vartheta_2]$ естественно принять область изменения синуса вещественного угла $[-1, 1]$. На рис. 2 показаны зависимости среднееквдратичного отклонения (СКО) $\sigma_s = [(\hat{\theta} - \theta)^2]^{1/2}$ МП и МСК-оценок, а также МГД от выходного отношения сигнал/шум Q (12); черта сверху означает усреднение по независимым реализациям (СКО МП и МСК-оценок определялось численным стохастическим моделированием с использованием порядка 10^5 независимых реализаций), среднееквдратичные отклонения нормированы на «естественную» угловую разрешающую способность решётки $\lambda/(Nd)$, где λ — длина волны.

Как следует из рис. 2, СКО МСК(МСР)-оценок синуса пеленга в области небольших выходных ОСШ существенно меньше по сравнению с СКО МП оценки. Например, при допустимом нормированном СКО порядка единицы МСК-оценка допускает снижение ОСШ приблизительно на 1,5 дБ, что для практических приложений является значимой величиной. Аналогичные кривые были получены и для других близких по структуре прикладных задач: оценки частоты тонального сигнала и оценки времени прихода импульса известной формы; при этом вид зависимостей нормированных среднееквдратичных отклонений от выходного ОСШ практически не отличался от показанных на рис. 2. Во всех этих прикладных задачах по физическому смыслу СКО оценки параметра θ_2 , определяющего форму детерминированного сигнала, не должно зависеть от этого параметра, если пренебречь краевыми эффектами возле границ интервала $[\vartheta_1, \vartheta_2]$. Тем самым кривая СКО МСК(МСР)-оценки на рис. 2 фак-

ванном СКО порядка единицы МСК-оценка допускает снижение ОСШ приблизительно на 1,5 дБ, что для практических приложений является значимой величиной. Аналогичные кривые были получены и для других близких по структуре прикладных задач: оценки частоты тонального сигнала и оценки времени прихода импульса известной формы; при этом вид зависимостей нормированных среднееквдратичных отклонений от выходного ОСШ практически не отличался от показанных на рис. 2. Во всех этих прикладных задачах по физическому смыслу СКО оценки параметра θ_2 , определяющего форму детерминированного сигнала, не должно зависеть от этого параметра, если пренебречь краевыми эффектами возле границ интервала $[\vartheta_1, \vartheta_2]$. Тем самым кривая СКО МСК(МСР)-оценки на рис. 2 фак-

тически представляет реальную дисперсионную границу, которая может быть достигнута, но не может быть улучшена.

4. ЗАКЛЮЧЕНИЕ

Данная работа устанавливает тождественность двух стратегий оценки параметров: байесовской (субъективной) и стратегии минимизации объективной погрешности оценки, характеризуемой ковариацией оценки. С этой точки зрения оценки, реализующие минимум среднеарифметической ковариации (минимум среднего риска), представляют большой практический интерес, т. к. объективно являются наиболее точными независимо от объёма выборки (как известно, оценки по минимуму среднего риска не получили широкого практического распространения, в частности, из-за внесения субъективного элемента в оценку). При независимости оценки от оцениваемого параметра строго реализуется достижимая минимальная дисперсионная граница, отличающаяся от широко известной границы Крамера—Рао, недостижимой в большинстве практически интересных приложений. Для оценки параметров детерминированного сигнала на фоне гауссовой помехи получен аналитический вид процедуры и показано, что она даёт практически значимый выигрыш по сравнению с общеизвестной оценкой максимума правдоподобия.

Автор благодарит А. А. Мальцева и В. Т. Ермолаева за плодотворное обсуждение данной работы и А. А. Родионова, выполнившего численное моделирование. Работа поддержана Российским фондом фундаментальных исследований (гранты № 00–02–16438, 00–15–96741).

СПИСОК ЛИТЕРАТУРЫ

1. Тихонов В. И., Харисов В. Н. Статистический анализ и синтез радиотехнических устройств и систем. М.: Радио и связь, 1991. 608 с.
2. Леман Э. Теория точечного оценивания: Пер. с англ. М.: Наука, 1991. 444 с.
3. Закс Ш. Теория статистических выводов: Пер. с англ. М.: Мир, 1975. 776 с.
4. Бард Й. Нелинейное оценивание параметров: Пер. с англ. М.: Статистика, 1979. 349 с.
5. Кендалл М., Стьюарт А. Статистические выводы и связи: Пер. с англ. М.: Наука, 1973. 900 с.

Институт прикладной физики РАН,
г. Нижний Новгород, Россия

Поступила в редакцию
23 июля 2002 г.

MINIMUM-COVARIANCE ESTIMATE OF SIGNAL PARAMETERS

V. I. Turchin

We show the existence of the only procedure ensuring the absolute minimum of arithmetic mean of covariance inside a certain parameter-variation range and coinciding identically with the algorithm for calculating the estimate minimizing the average risk for quadratic loss function and uniform a priori distribution of parameters. This procedure is compared to the standard method for obtaining maximum likely estimates. It is shown, in particular, that the procedure ensures a significantly smaller random scatter of estimates during the estimation of the deterministic-signal parameters against a background of Gaussian interference. In the case where the estimate covariance is independent of the parameter estimated for physical reasons, using the above procedure with an arbitrary sample volume, we reach the actual variance boundary.

УДК 519.216.3

ПРЕДЕЛЬНОЕ ВРЕМЯ ПРЕДСКАЗУЕМОСТИ ПРОЦЕССОВ АВТОРЕГРЕССИИ—СКОЛЬЗЯЩЕГО СРЕДНЕГО

С. Н. Мусеев

Показано, что для оптимального по минимуму среднего квадрата ошибки прогноза процесса авторегрессии—скользящего среднего время предсказуемости может превышать время корреляции в бесконечное число раз уже для процесса авторегрессии первого порядка—скользящего среднего первого порядка.

ВВЕДЕНИЕ

Линейные процессы авторегрессии s -го порядка—скользящего среднего q -го порядка (АРСС(s, q)) широко используются в качестве моделей реальных случайных процессов в самых различных областях исследований, в том числе при обработке сигналов, теоретическом описании отклика радиотехнических систем, анализе временных рядов различной физической природы и т. д. Связано это с возможностью экономичного с точки зрения количества параметров описания процессом АРСС(s, q) практически любого линейного случайного процесса с заданной точностью, развитой техникой подгонки процессов АРСС(s, q) к наблюдаемым данным и простотой моделирования этих процессов [1]. В работах [2, 3] на многочисленных примерах было показано, что время предсказуемости τ_{pred} различных линейных временных рядов при использовании линейного прогноза не может превосходить время корреляции τ_{cor} более чем в $1,1 \div 1,4$ раза. В связи с широкой распространённостью в различных приложениях модели процесса АРСС(s, q) представляет практический и теоретический интерес подробное исследование соотношения между его временем предсказуемости и временем корреляции.

В этой работе ставится цель показать, что для линейных процессов АРСС(s, q) при определённом сочетании коэффициентов модели время предсказуемости может в бесконечное число раз превышать время корреляции уже для процесса АРСС(1, 1).

1. ПОСТАНОВКА ЗАДАЧИ

Пусть наблюдаемый случайный процесс y_t , где $t = 0, \pm 1, \pm 2, \dots$, является процессом АРСС(s, q):

$$y_t - \varphi_1 y_{t-1} - \dots - \varphi_s y_{t-s} = \xi_t - \theta_1 \xi_{t-1} - \dots - \theta_q \xi_{t-q}. \quad (1)$$

Здесь ξ_t — произвольный белый шум с независимыми в моменты времени t и $t - k$, где $k \neq 0$, значениями, имеющий нулевое среднее и дисперсию σ_ξ^2 . Процесс АРСС(s, q) стационарен, если корни уравнения $x^s - \varphi_1 x^{s-1} - \dots - \varphi_s = 0$ лежат внутри единичного круга, и обратим, если корни уравнения $x^q - \theta_1 x^{q-1} - \dots - \theta_q = 0$ лежат внутри единичного круга. Обратимость процесса АРСС(s, q) означает, что его можно представить в виде процесса авторегрессии бесконечного порядка.

Для каждого фиксированного порядка авторегрессии s и скользящего среднего q необходимо подобрать такие коэффициенты φ_i и θ_j , где $i = 1, 2, \dots, s$; $j = 1, 2, \dots, q$, чтобы для процесса АРСС(s, q) наблюдалось бы максимальное отношение $\tau_{\text{pred}}/\tau_{\text{cor}}$ в случае линейного прогноза, т. е. максимальное относительное время предсказуемости. Для этого введём следующую функцию:

$$g(s, q) = \max_{\substack{\varphi_i, \theta_j; \\ i=1,2,\dots,s; j=1,2,\dots,q}} \{\tau_{\text{pred}}/\tau_{\text{cor}}\}. \quad (2)$$

Функция $g(s, q)$ характеризует потенциальные возможности линейного прогноза процесса $APCC(s, q)$ для оптимального сочетания коэффициентов φ_i, θ_j , где $i = 1, 2, \dots, s; j = 1, 2, \dots, q$, с точки зрения отношения $\tau_{\text{pred}}/\tau_{\text{cor}}$. Задача заключается в исследовании поведения функции $g(s, q)$ в зависимости от s и q . В дальнейшем будем рассматривать только невырожденный случай $s > 0, q > 0$, поскольку процессы авторегрессии $AP(s)$ при $q = 0$ и скользящего среднего $CC(q)$ при $s = 0$ требуют отдельного исследования.

2. ПАРНЫЕ И МНОЖЕСТВЕННЫЕ КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

Спектральная плотность процесса $APCC(s, q)$ с точностью до множителя совпадает с произведением спектральных плотностей процессов $AP(s)$ и $CC(q)$. Поскольку в теории интегральных преобразований произведение изображений означает свёртку оригиналов, то нормированную автокорреляционную функцию процесса $APCC(s, q)$ в стационарном режиме можно представить в следующем виде:

$$r_k = \sum_{i=-q}^q r_i^{\text{CC}} r_{k-i}^{\text{AP}} / \sum_{i=-q}^q r_i^{\text{CC}} r_i^{\text{AP}},$$

где $r_k = M(y_t y_{t+k}) / [M(y_t^2) M(y_{t+k}^2)]^{1/2}$, M — оператор усреднения, r_i^{AP} и r_i^{CC} — нормированные автокорреляционные функции процессов $AP(s)$ и $CC(q)$ соответственно; $k \geq 0$. Коэффициенты корреляции r_i^{AP} для стационарного режима находятся из системы уравнений

$$r_0^{\text{AP}} = 1, \quad r_i^{\text{AP}} = \sum_{j=1}^s \varphi_j r_{i-j}^{\text{AP}},$$

где $i = 1, 2, \dots$. Функция r_k^{CC} имеет следующий вид [1]:

$$r_k^{\text{CC}} = \begin{cases} 1, & k = 0; \\ \left(-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{i+k} \right) / \left(1 + \sum_{i=1}^q \theta_i^2 \right), & k = 1, 2, \dots, q; \\ 0, & k > q. \end{cases}$$

Отметим также полезное рекуррентное соотношение для функции r_k при $k > q$:

$$r_k = \sum_{j=1}^s \varphi_j r_{k-j}.$$

Время корреляции τ_{cor} рассчитывается по значениям нормированной автокорреляционной функции r_k , состоящей из коэффициентов парных корреляций. Время предсказуемости τ_{pred} для оптимального линейного прогноза рассчитывается по значениям функции ρ_k , состоящей из последовательности коэффициентов множественной корреляции. Коэффициент множественной корреляции между отсчётом y_{t+k} и p отсчётами $y_t, y_{t-1}, \dots, y_{t-p+1}$ определяется как коэффициент корреляции между отсчётом y_{t+k} и линейной комбинацией p отсчётов $y_{t+k}^* = \sum_{i=0}^{p-1} a_i y_{t-i}$, максимизированный по весам a_i :

$$\rho_{k(0, \dots, p-1)} = \max_{a_i; i=0, 1, \dots, p-1} \left\{ M(y_{t+k} y_{t+k}^*) / [M(y_{t+k}^2) M(y_{t+k}^{*2})]^{1/2} \right\}. \quad (3)$$

Из (3) следует, что при $p > p'$ всегда справедливы соотношения:

$$\rho_{k(0, \dots, p-1)} \geq \rho_{k(0, \dots, p'-1)} \geq \rho_{k(0)} = |r_k|.$$

Процесс АРСС(s, q) теоретически нельзя описать марковским процессом конечного порядка, т. е. процессом АР(p) конечного порядка ($p < \infty$). Поэтому необходимо положить $p = \infty$ в (3), тем самым выбирая максимальный из всех возможных коэффициентов множественной корреляции:

$$\rho_k = \lim_{p \rightarrow \infty} \rho_{k(0, \dots, p-1)}. \tag{4}$$

Отметим, что функция ρ_k — невозрастающая: $\rho_{k+1} \leq \rho_k$, где $k \geq 0$, в то время как для $\rho_{k(0, \dots, p-1)}$ как функции k это в общем случае не верно.

Коэффициент множественной корреляции просто выражается через коэффициенты парной корреляции r_k [4, 5]:

$$\rho_{k(0, \dots, p-1)} = \sqrt{1 - \det \mathbf{R}_k / \det \mathbf{R}}. \tag{5}$$

Здесь $\det \mathbf{R}$ — детерминант матрицы \mathbf{R} , $R_{ij} = r_{i-j}$, где $i = 1, 2, \dots, p$; $j = 1, 2, \dots, p$, — нормированная корреляционная матрица размера $p \times p$ отсчётов $y_t, y_{t-1}, \dots, y_{t-p+1}$, а \mathbf{R}_k — нормированная корреляционная матрица размера $(p+1) \times (p+1)$ отсчётов $y_{t+k}, y_t, y_{t-1}, \dots, y_{t-p+1}$, которая получается из матрицы \mathbf{R} добавлением сверху строки $(r_k, r_{k+1}, \dots, r_{k+p-1})$ и слева столбца $(1, r_k, r_{k+1}, \dots, r_{k+p-1})^T$; индекс T обозначает операцию транспонирования. Из формулы (5) следует, что $\rho_{k(0, \dots, p-1)}$ полностью определяется первыми $p - 1$ коэффициентами парной корреляции r_1, r_2, \dots, r_{p-1} и p коэффициентами $r_k, r_{k+1}, \dots, r_{k+p-1}$. Из (5) также следует, что $\rho_{k(0, \dots, p-1)} = 0$, если $r_m = 0$ для всех $m \geq k$.

Коэффициент множественной корреляции наглядно выражается через корреляцию отсчёта с его оптимальным прогнозом, через дисперсию прогноза или через дисперсию ошибки прогноза. Оптимальный по минимуму среднего квадрата ошибки (СКО) прогноз отсчёта y_{t+k} , который даётся в момент времени t на момент времени $t + k$, при p известных отсчётах $y_t, y_{t-1}, \dots, y_{t-p+1}$ для линейных процессов выглядит следующим образом:

$$\hat{y}_{t+k} = \sum_{i=0}^{p-1} \Psi_{pk}(i) y_{t-i}, \tag{6}$$

где веса $\Psi_{pk}(i)$ предсказывающих переменных $y_t, y_{t-1}, \dots, y_{t-p+1}$ находятся из системы линейных уравнений

$$\sum_{i=0}^{p-1} \Psi_{pk}(i) r_{i-j} = r_{k+j},$$

где $j = 0, 1, \dots, p - 1$, или в явном виде выражаются через коэффициенты корреляции [5]:

$$\Psi_{pk}(i) = (-1)^i \det \mathbf{R}_{1\ i+2} / \det \mathbf{R},$$

где $i = 0, 1, \dots, p - 1$, $\mathbf{R}_{1\ i+2}$ — матрица размера $p \times p$, которая получается из матрицы \mathbf{R}_k , если из неё вычеркнуть первую строку и $(i + 2)$ -й столбец.

В работах [6, 7] показано, что оптимальный линейный прогноз, полученный из условия минимума СКО, по сравнению с другими линейными прогнозами обладает максимальной корреляцией с прогнозируемым отсчётом. Отсюда следует, что для такого оптимального прогноза коэффициент множественной корреляции можно рассчитать как коэффициент корреляции между отсчётом ряда y_{t+k} и оптимальным прогнозом \hat{y}_{t+k} :

$$\rho_{k(0, \dots, p-1)} = M(y_{t+k} \hat{y}_{t+k}) / [M(y_{t+k}^2) M(\hat{y}_{t+k}^2)]^{1/2} = \left[\sum_{i=0}^{p-1} \Psi_{pk}(i) r_{k+i} \right]^{1/2}.$$

Наконец, в силу того, что коэффициент множественной корреляции является частным случаем множественного корреляционного отношения [6, 8], для него справедливы следующие формулы:

$$\rho_{k(0,\dots,p-1)} = \sigma_{\text{pred}}(k)/\sigma_y = \sqrt{1 - \sigma_p^2(k)/\sigma_y^2}, \quad (7)$$

где $\sigma_y^2 = M(y_{t+k}^2) = \sigma_\xi^2 (1 + \sum_{i=1}^q \theta_i^2) \sum_{i=-q}^q r_i^{\text{CC}} r_i^{\text{AP}} / (1 - \sum_{i=1}^s \varphi_i r_i^{\text{AP}})$ — дисперсия процесса y_t , $\sigma_{\text{pred}}^2(k) = M(\hat{y}_{t+k}^2) = M(y_{t+k} \hat{y}_{t+k})$ — дисперсия оптимального по минимуму СКО прогноза на k шагов вперёд, $\sigma_p^2(k) = M[(y_{t+k} - \hat{y}_{t+k})^2] = \sigma_y^2 \det \mathbf{R}_k / \det \mathbf{R}$ — СКО оптимального прогноза (6) на k шагов вперёд. Из последнего равенства в (7) хорошо видно, что коэффициент множественной корреляции $\rho_{k(0,\dots,p-1)}$ является наглядной и понятной мерой точности линейного прогноза на k шагов вперёд, т. к. его квадрат совпадает с относительной разностью между дисперсией процесса и СКО прогноза.

3. ВРЕМЯ КОРРЕЛЯЦИИ И ВРЕМЯ ПРЕДСКАЗУЕМОСТИ

После того, как найдены удобные в вычислительном плане формулы для r_k и ρ_k , необходимо определить τ_{cor} и τ_{pred} . Время корреляции τ_{cor} и время предсказуемости τ_{pred} определяются как характерные времена спадаения до нуля соответственно нормированной автокорреляционной функции и коэффициента множественной корреляции. Наиболее часто на практике как для монотонно спадающих, так и для осциллирующих нормированных автокорреляционных функций $R(\tau)$ используют следующие определения τ_{cor} [7, 9, 10]:

$$\tau_{\text{cor}}^{(1)} = \int_0^\infty |R(\tau)|^\alpha d\tau, \quad 0 < \alpha < \infty; \quad (8)$$

$$\tau_{\text{cor}}^{(2)} = \max\{\tau : |R(\tau)| \geq \varepsilon\}, \quad 0 < \varepsilon < 1. \quad (9)$$

Как правило, в (8) полагают $\alpha = 1$. Геометрически таким образом заданное τ_{cor} равно основанию прямоугольника с высотой 1, имеющего ту же площадь, что и площадь под кривой $|R(\tau)|$.

Для процессов, заданных в дискретные моменты времени $t = 0, \pm 1, \pm 2, \dots$, к которым относятся изучаемые нами процессы АРСС(s, q), аналогичные (8) и (9) определения τ_{cor} и τ_{pred} будут выглядеть следующим образом:

$$\tau_{\text{cor}}^{(1)} = \sum_{k=1}^\infty |r_k|^\alpha, \quad \tau_{\text{pred}}^{(1)} = \sum_{k=1}^\infty \rho_k^\alpha, \quad 0 < \alpha < \infty; \quad (10)$$

$$\tau_{\text{cor}}^{(2)} = \max\{k : |r_k| \geq \varepsilon\}, \quad \tau_{\text{pred}}^{(2)} = \max\{k : \rho_k \geq \varepsilon\}, \quad 0 < \varepsilon < 1. \quad (11a)$$

Определения (11a) могут приводить к вырожденным случаям, когда $\tau_{\text{cor}}^{(2)} = 0$, $\tau_{\text{pred}}^{(2)} \neq 0$. Для этих вырожденных случаев будем использовать несколько видоизменённый вариант определений (11a):

$$\tau_{\text{cor}}^{(2)} = \max\{\tau : |r(\tau)| \geq \varepsilon\}, \quad \tau_{\text{pred}}^{(2)} = \max\{\tau : \rho(\tau) \geq \varepsilon\}, \quad 0 < \varepsilon < 1, \quad (11б)$$

где функции $r(\tau)$ и $\rho(\tau)$ в непрерывном времени получаются соответственно из функций r_k и ρ_k в дискретном времени путём линейной интерполяции между соседними значениями.

4. МАКСИМАЛЬНОЕ ОТНОСИТЕЛЬНОЕ ВРЕМЯ ПРЕДСКАЗУЕМОСТИ

Максимальное относительное время предсказуемости $g(s, q)$ для определений (10), (11) находится аналитически. Для этого достаточно рассмотреть простейший невырожденный случай — процесс АРСС(1, 1). В этом случае имеем

$$r_k = r_1 \varphi_1^{k-1}, \quad \rho_{k(0, \dots, p-1)} = \rho_{1(0, \dots, p-1)} |\varphi_1^{k-1}|,$$

где $k > 0, r_1 = (\varphi_1 - \theta_1)(\varphi_1 \theta_1 - 1) / [\theta_1(\varphi_1 - \theta_1) + (\varphi_1 \theta_1 - 1)]$. Для определений $\tau_{\text{cor}}^{(1)}, \tau_{\text{pred}}^{(1)}$ (10) получаем

$$\tau_{\text{cor}}^{(1)} = |r_1|^\alpha / (1 - |\varphi_1|^\alpha), \quad \tau_{\text{pred}}^{(1)} = \rho_{1(0, \dots, p-1)}^\alpha / (1 - |\varphi_1|^\alpha).$$

Следовательно,

$$\tau_{\text{pred}}^{(1)} / \tau_{\text{cor}}^{(1)} = \rho_{1(0, \dots, p-1)}^\alpha / |r_1|^\alpha. \tag{12}$$

Положим

$$\theta_1 = \text{sgn}(\varphi_1) \left(1 - \sqrt{1 - |\varphi_1|} \right)^{\sqrt{2a/(1-a)}}, \tag{13}$$

где $0 \leq a \leq 1$,

$$\text{sgn}(x) = \begin{cases} 1, & x \geq 0; \\ -1, & x < 0. \end{cases}$$

Нетрудно показать, что в этом случае

$$\lim_{|\varphi_1| \rightarrow 1} r_k = \text{sgn}(\varphi_1^k) a, \quad \lim_{|\varphi_1| \rightarrow 1} \rho_{k(0, \dots, p-1)} = a \sqrt{p/[a(p-1) + 1]}, \tag{14}$$

где $k > 0$. Формулы (14) показывают, что при θ_1 в виде (13) и $|\varphi_1| \rightarrow 1$ функции r_k и $\rho_{k(0, \dots, p-1)}$ будут бесконечно медленно стремиться к нулю с ростом задержки k . Подставляя предельные значения (14) в формулу (12), получаем

$$\lim_{|\varphi_1| \rightarrow 1, a \rightarrow 0} \tau_{\text{pred}}^{(1)} / \tau_{\text{cor}}^{(1)} = p^{\alpha/2}, \quad \lim_{|\varphi_1| \rightarrow 1, p \rightarrow \infty} \tau_{\text{pred}}^{(1)} / \tau_{\text{cor}}^{(1)} = a^{-\alpha/2}.$$

Отсюда сразу же следует, что при выполнении условий

$$|\varphi_1| \rightarrow 1, \quad a \rightarrow 0, \quad p \rightarrow \infty$$

предельное относительное время предсказуемости процесса АРСС(1, 1) для определений $\tau_{\text{cor}}^{(1)}, \tau_{\text{pred}}^{(1)}$ (10) равно бесконечности:

$$g(1, 1) = \infty. \tag{15}$$

Условия $|\varphi_1| \rightarrow 1, a \rightarrow 0$ означают, что функции r_k и $\rho_{k(0, \dots, p-1)}$, где $k > 0$, стремятся к нулю. Однако r_k стремится к нулю быстрее, чем $\rho_{k(0, \dots, p-1)}$, в результате чего реализуется соотношение (15).

Для определений $\tau_{\text{cor}}^{(2)}, \tau_{\text{pred}}^{(2)}$ (11) максимальное отношение $\tau_{\text{pred}}^{(2)} / \tau_{\text{cor}}^{(2)}$ достигается при выполнении условий

$$r_k < \varepsilon, \quad \rho_{k(0, \dots, p-1)} > \varepsilon, \quad k > 0. \tag{16}$$

Эти условия могут быть выполнены, если в качестве r_k и $\rho_{k(0, \dots, p-1)}$ взять рассмотренные выше предельные значения (14). Полагая в (14) $p = 2$, преобразуем условия (16) для этого случая в единственное простое условие для величины a в формулах (13), (14):

$$\varepsilon \left(\varepsilon + \sqrt{\varepsilon^2 + 8} \right) / 4 < a < \varepsilon. \tag{17}$$

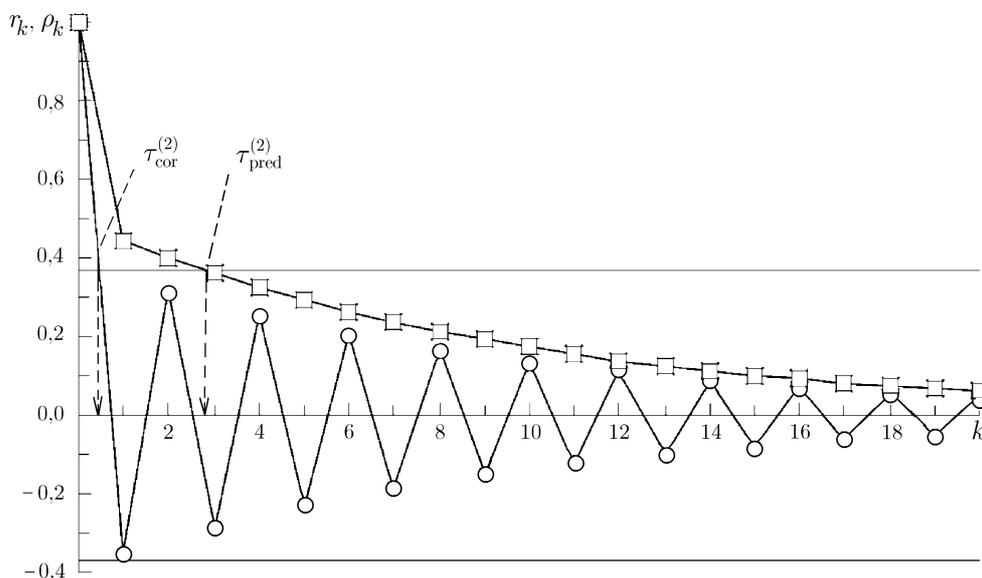


Рис. 1

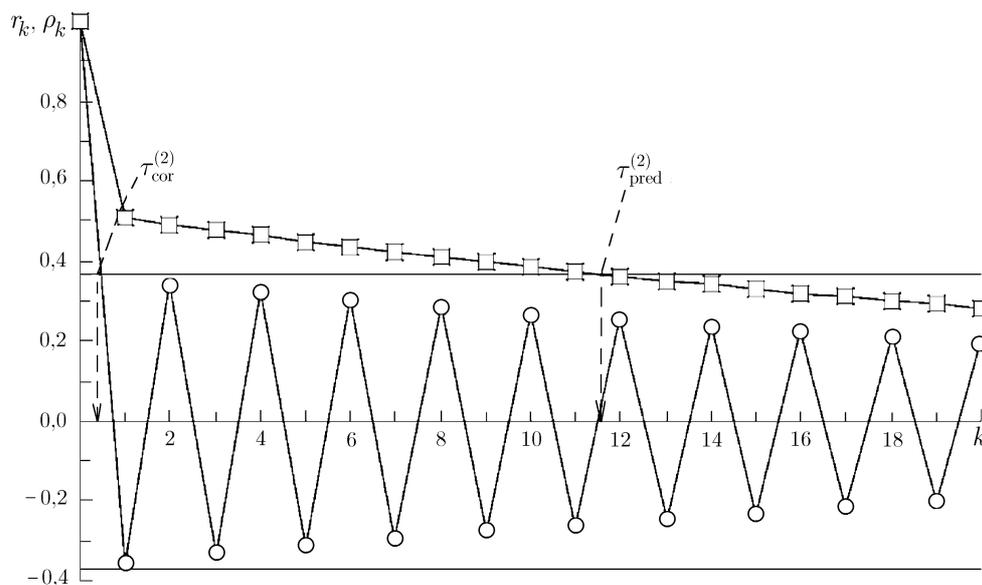


Рис. 2

Следовательно, если для процесса АРСС(1, 1) выбрать θ_1 в виде (13), где a удовлетворяет (17), то при $|\varphi_1| \rightarrow 1$ предельное относительное время предсказуемости процесса АРСС(1, 1) для определений $\tau_{\text{cor}}^{(2)}, \tau_{\text{pred}}^{(2)}$ (11) будет равно бесконечности: $g(1, 1) = \infty$.

Этот вывод иллюстрируют рис. 1 и 2, где кружками показаны значения нормированной автокорреляционной функции r_k в зависимости от задержки k , а квадратиками — значения коэффициента множественной корреляции ρ_k процесса АРСС(1, 1). Кружки, соединённые линиями, дают функцию $r(\tau)$, а соединённые линиями квадратики — функцию $\rho(\tau)$ в формулах (11б). Горизонтальными линиями на рисунках показаны уровни $\pm \varepsilon = \pm \exp(-1) \approx \pm 0,368$. Величина θ_1 рассчитывалась по формуле (13), в которой полагалось $a = 1/3$ (это значение удовлетворяет соотношению (17)). Для рис. 1 $\varphi_1 = -0,9$;

$\theta_1 \approx -0,684; \tau_{\text{cor}}^{(2)} \approx 0,47; \tau_{\text{pred}}^{(2)} \approx 2,8; \tau_{\text{pred}}^{(2)}/\tau_{\text{cor}}^{(2)} \approx 6$. Для рис. 2 $\varphi_1 = -0,97; \theta_1 \approx -0,83; \tau_{\text{cor}}^{(2)} \approx 0,47; \tau_{\text{pred}}^{(2)} \approx 11,6; \tau_{\text{pred}}^{(2)}/\tau_{\text{cor}}^{(2)} \approx 25$. Из рис. 1 и 2 хорошо видно, что по мере приближения $|\varphi_1|$ к единице $\tau_{\text{cor}}^{(2)}$ практически не меняется, в то время как $\tau_{\text{pred}}^{(2)}$ быстро растёт. Очевидно, что при $|\varphi_1| \rightarrow 1$ отношение $\tau_{\text{pred}}^{(2)}/\tau_{\text{cor}}^{(2)} \rightarrow \infty$.

Таким образом, максимальные отношения $\tau_{\text{pred}}^{(1)}/\tau_{\text{cor}}^{(1)}$ и $\tau_{\text{pred}}^{(2)}/\tau_{\text{cor}}^{(2)}$ для процессов АРСС(1, 1) равны бесконечности. Очевидно, что это будет справедливо и для процессов АРСС(s, q), где $s > 0, q > 0$, частным случаем которых является процесс АРСС(1, 1). Другими словами, используя терминологию работы [11], процессы АРСС(s, q) имеют бесконечный горизонт предсказуемости при любых $s > 0, q > 0$.

Отметим также, что при $|\varphi_1| \rightarrow 1$ отношение $\tau_{\text{pred}}^{(2)}/\tau_{\text{cor}}^{(2)}$ значительно быстрее стремится к бесконечности по сравнению с отношением $\tau_{\text{pred}}^{(1)}/\tau_{\text{cor}}^{(1)}$.

5. ОПРЕДЕЛЕНИЕ ЧИСЛА p ПРЕДСКАЗЫВАЮЩИХ ПЕРЕМЕННЫХ

Оптимальный прогноз процесса АРСС(s, q) теоретически опирается на бесконечное число предсказывающих переменных, т. е. в (6) надо положить $p = \infty$. Однако на практике p всегда конечно. При уменьшении p относительное время предсказуемости $g(s, q)$ будет также уменьшаться. Так, при минимальном возможном значении $p = 1$ всегда $g(s, q) = 1$ как для определений $\tau_{\text{cor}}, \tau_{\text{pred}}$ (10), так и для определений (11). Минимальное p , при котором ещё не наступает существенного ухудшения точности линейного прогноза, можно определить с помощью функции $\beta_m(p)$, состоящей из последовательности коэффициентов частной множественной корреляции. Коэффициент частной множественной корреляции определяется как коэффициент обычной корреляции между отсчётом y_{t+1} и линейной комбинацией m отсчётов $y_m^* = \sum_{i=1}^m a_i y_{t-p+2-i}$ при фиксированных $p - 1$ промежуточных отсчётах $\mathbf{y}_p = (y_t, y_{t-1}, \dots, y_{t-p+2})$, максимизированный по весам a_i :

$$\beta_m(p) = \max_{a_i; i=1,2,\dots,m} \{M(y_{t+1}y_m^* | \mathbf{y}_p) / [M(y_{t+1}^2 | \mathbf{y}_p)M(y_m^{*2} | \mathbf{y}_p)]^{1/2}\}.$$

Значение $\beta_m(p)$ легко рассчитать по следующим формулам:

$$\beta_m^2(p) = 1 - \sigma_{p-1+m}^2(1) / \sigma_{p-1}^2(1) = (\rho_{1(0,\dots,p-2+m)}^2 - \rho_{1(0,\dots,p-2)}^2) / (1 - \rho_{1(0,\dots,p-2)}^2), \tag{18}$$

где $\sigma_0^2(1) = \sigma_y^2, \rho_{1(-1)}^2$. Из (18) видно, что величина $\beta_m^2(p)$ совпадает с относительным уменьшением СКО прогноза на шаг вперёд $\sigma_{p-1}^2(1)$ в результате увеличения числа предсказывающих переменных в прогнозе (6) с $p - 1$ до $p - 1 + m$. Поэтому минимальное значение p для прогноза (6) совпадает с максимальным значением p , при котором ещё можно считать, что величина $\beta_m^2(p)$ много больше нуля. Наибольший практический интерес для определения подходящего значения p представляют случаи $m = 1$ и $m = \infty$. При $m = 1$ вместо $\beta_1(p)$ обычно используют частную автокорреляционную функцию $\beta(p)$, состоящую из коэффициентов корреляции между y_{t+1} и y_{t-p+1} при фиксированных $p - 1$ промежуточных отсчётах $y_t, y_{t-1}, \dots, y_{t-p+2}$, которая рассчитывается по формуле

$$\beta(p) = (-1)^{p+1} \det \mathbf{R}_{1p+1} / \sqrt{\det \mathbf{R}_{11} \det \mathbf{R}_{p+1p+1}},$$

где \mathbf{R}_{ij} — матрица размера $p \times p$, которая получается из матрицы \mathbf{R}_1 , если из неё вычеркнуть i -тую строку и j -й столбец. Отметим, что

$$\beta_1(p) = |\beta(p)|, \quad \beta_m(1) = \rho_{1(0,\dots,m-1)}, \quad \beta_{m+1}(p) \geq \beta_m(p).$$

Для примера рассмотрим процедуру определения p для процесса АРСС(1, 1) при θ_1 в виде (13) и $|\varphi_1| \rightarrow 1$, т. е. для функций r_k и $\rho_{k(0, \dots, p-1)}$, заданных в виде (14). В этом предельном случае

$$\beta_m^2(p) = ma^2 / [(ap - a + 1)(ap + am - 2a + 1)]. \quad (19)$$

Поведение функций (19) иллюстрирует рис. 3, где кружками, соединёнными сплошными линиями 1, показаны значения $\beta_1^2(p)$, а квадратиками, соединёнными сплошными линиями 2, — значения $\beta_\infty^2(p)$. Для всех кривых на рис. 3 полагалось $a = 1/3$. Предположим, что приемлемые с практической точки зрения потери в точности прогнозирования при уменьшении числа предсказываемых переменных составляют 10 %. Тогда минимальное значение p , соответствующее этим потерям, определяется по точкам пересечения функциями (19) уровня 0,1, который показан на рис. 3 горизонтальной линией. Из рис. 3 видно, что $\beta_1^2(p)$ пересекает уровень 0,1 примерно при $p = 1$. Следовательно, если выбрать в прогнозе (6) $p = 1$, то увеличение числа предсказываемых переменных до $p = 2$ приведёт к относительному уменьшению СКО прогноза на шаг вперёд $\sigma_1^2(1)$ на величину $\beta_1^2(2) = 6,25\% < 10\%$. Однако увеличение числа предсказываемых переменных с 1 до бесконечности приведёт к уменьшению $\sigma_1^2(1)$ на $\beta_\infty^2(2) = 25\%$, что составляет неприемлемо большую величину. Поэтому p правильнее выбирать по точке пересечения уровня 0,1 функцией $\beta_\infty^2(p)$. Из рис. 3 видно, что в этом случае $p = 8$, и увеличение числа предсказываемых переменных с 8 до бесконечности приведёт к уменьшению СКО прогноза $\sigma_8^2(1)$ всего лишь на $\beta_\infty^2(9) \approx 9,1\%$, что соответствует заданному уровню приемлемых потерь.

На практике часто удобнее рассчитывать выигрыш (потери) в прогнозировании по среднеквадратическим отклонениям ошибок прогнозов $\sigma_p(1)$, а не по СКО $\sigma_p^2(1)$. Для этих целей введём функцию, состоящую из приведённых коэффициентов частной множественной корреляции:

$$b_m(p) = 1 - \sigma_{p-1+m}(1) / \sigma_{p-1} = 1 - \sqrt{1 - \beta_m^2(p)}. \quad (20)$$

Величина $b_m(p)$ совпадает с относительным уменьшением среднеквадратической ошибки прогноза на шаг вперёд $\sigma_{p-1}(1)$ в результате увеличения числа предсказываемых переменных в прогнозе (6) с $p-1$ до $p-1+m$.

На рис. 3 кружками, соединёнными штриховыми линиями 3, отложены значения $b_1(p)$, а квадратиками, соединёнными штриховыми линиями 4, — значения $b_\infty(p)$, которые рассчитывались по формулам (19), (20). Определение подходящей величины p в прогнозе (6) по поведению функций (20) аналогично определению p по поведению функций (19). Так, если выбрать $p = 3$, то увеличение числа предсказываемых переменных с 3 до бесконечности приведёт к уменьшению среднеквадратической ошибки прогноза $\sigma_3(1)$ на $b_\infty(4) \approx 8,7\%$.

Работа выполнена при поддержке CRDF и Минобразования РФ (проект № vz-010-0).

СПИСОК ЛИТЕРАТУРЫ

1. Бокс Дж., Дженкинс Г. Анализ временных рядов, прогноз и управление. Вып. 1. М.: Мир, 1974. 406 с.

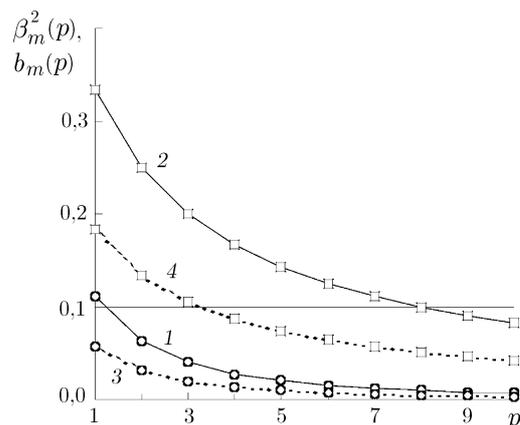


Рис. 3

2. Аносов О. Л., Бутковский О. Я., Кравцов Ю. А. // Радиотехника и электроника. 1995. Т. 40, № 12. С. 1 886.
3. Anosov O. L., Butkovskii O. Yu., Kravtsov Yu. A., Protopopescu V. A. // Physics of Vibrations. 1999. V. 7, No. 2. P. 61.
4. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика. Исследование зависимостей. М.: Финансы и статистика, 1985. 488 с.
5. Коваленко И. Н., Филиппова А. А. Теория вероятностей и математическая статистика. М.: Высшая школа, 1973. 368 с.
6. Ивченко Г. И., Медведев Ю. И. Математическая статистика. М.: Высшая школа, 1984. 248 с.
7. Тихонов В. И. Статистическая радиотехника. М.: Радио и связь, 1982. 624 с.
8. Моисеев С. Н. // Радиотехника и электроника. 1999. Т. 44, № 6. С. 715.
9. Романенко А. Ф., Сергеев Г. А. Вопросы прикладного анализа случайных процессов. М.: Сов. радио, 1968. 256 с.
10. Мирский Г. Я. Характеристики стохастической взаимосвязи и их измерения. М.: Энергоиздат, 1982. 320 с.
11. Кравцов Ю. И. // Успехи физических наук. 1989. Т. 158, вып. 1. С. 93.

Воронежский госуниверситет,
г. Воронеж, Россия

Поступила в редакцию
27 августа 2002 г.

LIMITING PREDICTABILITY TIME OF AUTOREGRESSION – MOVING AVERAGE PROCESSES

S. N. Moiseev

It is shown that the predictability time for optimal minimum rms error of the prediction of autoregression – moving average process can exceed the correlation time infinitely often already for the first-order autoregression – first-order moving average process.