МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Известия высших учебных заведений

РАДИОФИЗИКА

ежемесячный научно-технический журнал

Издаётся с апреля 1958 г.

Том XLVIII № 5 **Нижний Новгород** **2005**

Содержание

УДК 52-77+621.391:53.08+520.86+004.93'1

ТОМОГРАФИЯ ПРИ ОГРАНИЧЕННОМ ЧИСЛЕ ПРОЕКЦИЙ. II. РАДИОАСТРОНОМИЧЕСКИЙ МЕТОД CLEAN В ПРИЛОЖЕНИИ К ТРЁХМЕРНЫМ ЗАДАЧАМ

М. И. Агафонов, О. И. Шарова

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

Радиоастрономический метод CLEAN распространён на задачи трёхмерной (3D) томографической реконструкции. Рассмотрены два варианта: $3D_{1D}$ — реконструкция по одномерным проекциям — и $3D_{2D}$ — реконструкция на основе двумерных проекций. Деконволюция с введением трёхмерного синтезированного луча (эквивалентной суммарной передаточной функции) позволяет сократить в несколько раз число ракурсов по сравнению с традиционным подходом. Установлена связь максимального уровня боковых лепестков синтезированного луча, построенного на основе передаточных функций гауссового вида, с числом используемых проекций. Достижение в $3D_{2D}$-случае качества реконструкции, аналогичного $2D_{1D}$-случаю, требует использования в обоих случаях равного числа проекций при условии их равномерного распределения в пространстве. Использование одномерных проекций при $3D_{1D}$-реконструкции требует удвоения их числа. Процесс реконструкции иллюстрирован примером трёхмерной модели оптически тонкого излучающего объекта. Рассмотрены возможности применения разработанного подхода в астротомографии и дистанционном зондировании с введением передаточных функций, которые определяют разрешение приёмных диаграмм и спектрографов, а также временно́е разрешение локационных профилей.

ВВЕДЕНИЕ

Задача реконструкции трёхмерной внутренней структуры объектов возникает во многих областях современных научных исследований. Методы её решения представляют интерес для широкого круга развивающихся технологий. К ним следует отнести исследования лабораторной плазмы, энергетику, медицину, астротомографию, неразрушающий контроль и бесконтактный мониторинг широкого круга объектов [1–4]. По ряду причин во многих случаях количество проекций ограничено, наблюдения доступны лишь для малого числа направлений. Для целой серии задач может быть также ограничен и сектор расположения углов. Наличие указанных обстоятельств стимулировало поиск путей решения проблемы, отличающихся от традиционного томографического подхода, количество проекций для которого согласно [5] должно быть не менее

$$N_{\mathrm{BR}} \geq \pi \omega_{\mathrm{b}} D, \tag{1}$$

где ω_{b} — граничная пространственная частота, или частота среза, а D — диаметр объекта. Число проекций N_{BR} было названо в [6] числом Брейсуэлла—Риддла.

Решение задачи методом деконволюции в пространстве изображения положено в основу радиоастрономического подхода [6], применение которого позволяет существенно уменьшить количество необходимых проекций по сравнению с числом N_{BR}. Его основными отличительными признаками являются также введение синтезированного луча — эквивалентной суммарной передаточной функции (synthesized beam, SB), процедура чистки (CLEAN) для исключения откликов от боковых лепестков (sidelobes, SL) этого виртуального луча, а также метод двух чисток для определения области допустимых решений в сложных случаях 2-CLEAN DSA (determination of permissible solutions area). Алгоритмы чистки [7] — это радиоастрономические реализации итерационного алгоритма с нелинейными ограничениями [8, 9]. Ранее в работе [6] была детально рассмотрена реконструкция распределения яркости двумерного (2D) изображения при ограниченном

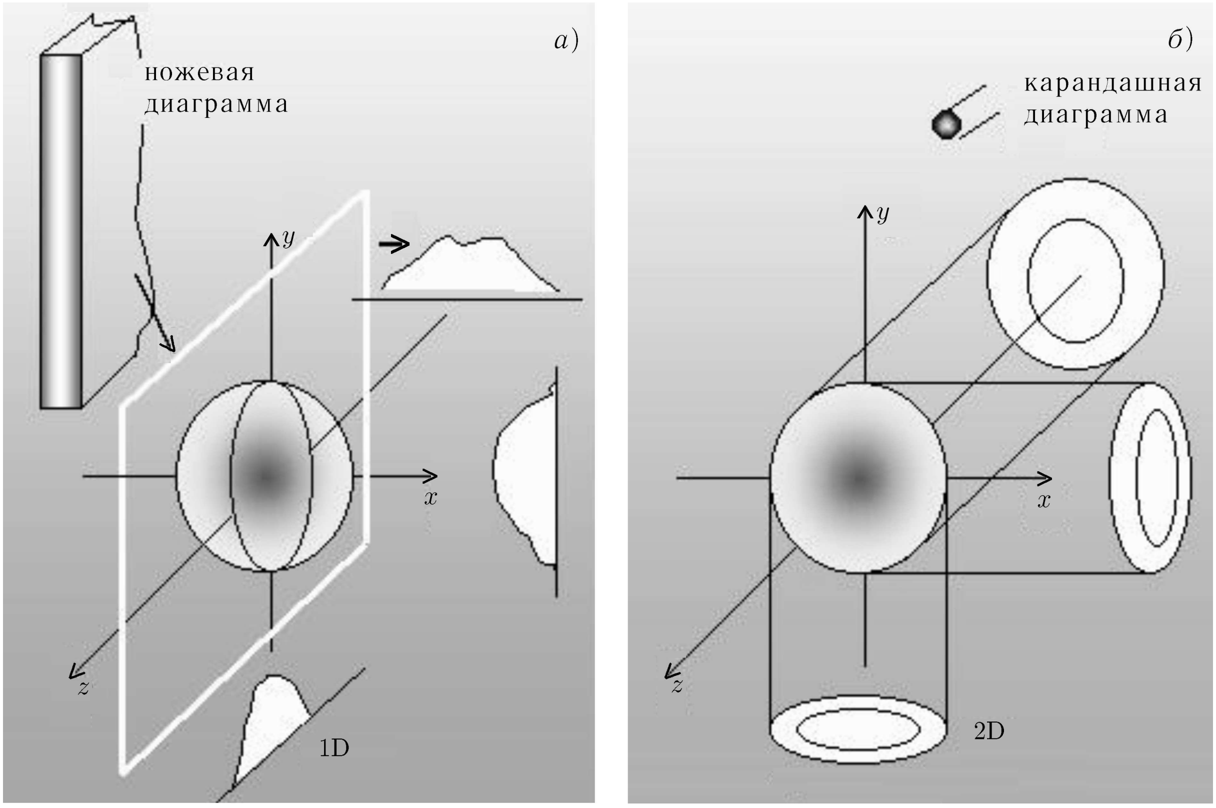

Рис. 1. Варианты трёхмерной реконструкции: по одномерным проекциям (*а*) и на основе двумерных проекций (*б*)

числе одномерных (1D) проекций. Было обосновано и название подхода — радиоастрономический. Разработанный способ позволяет весьма привлекательным образом провести реконструкцию при наличии малого числа проекций.

Целью настоящей работы является разработка принципиальных вопросов переноса разработанного ранее в [6] подхода на случай трёхмерных (3D) задач. В качестве примера рассмотрена реконструкция трёхмерного распределения яркости оптически тонкого излучающего объекта конечных размеров. Графическими примерами иллюстрирована процедура построения синтезированных лучей. Рассмотрен вопрос взаимосвязи уровней их боковых лепестков и количества проекций для различных вариантов задачи. Сделана оценка числа проекций, необходимого для восстановления трёхмерной структуры с такой же точностью, какая была достигнута при реконструкции двумерного распределения яркости на основе одномерных проекций.

1. ПОДХОД К РЕКОНСТРУКЦИИ

Рассматриваемые в настоящей работе варианты трёхмерной реконструкции показаны на рис. 1. Вариант на рис. 1*а* соответствует трёхмерной реконструкции ($3D_{1D}$) на основе одномерных проекций на разных ракурсах. Одномерные профили соответствуют сканированию ножевыми приёмными диаграммами вдоль заданных направлений. Каждая точка проекции излучающего оптически тонкого объекта является интегральной интенсивностью слоя. В случае на рис. 1*б* трёхмерная реконструкция ($3D_{2D}$) проводится на основании полученных с различных направлений двумерных изображений объекта. Элементы изображения соответствуют интегральной интенсивности на луче зрения. Разрешение таких двумерных проекций определяется полушириной передаточной функции приёмной системы. Регистрация изображений эквивалентна сканированию карандашными диаграммами.

Решение задачи трёхмерной реконструкции при ограниченном числе проекций можно провести аналогично рассмотренному ранее в [6, 10] случаю реконструкции двумерного изображения. Требуется восстановить трёхмерное распределение яркости (плотности) $f(x,y,z)$ по ограниченному числу двумерных или одномерных проекций объекта u_i, полученных с помощью приёмной диаграммы h_i. Одномерные проекции, полученные ножевым лучом, плоскость которого перпендикулярна направлению $l_i(\varphi_i,\theta_i)$, вдоль этого направления имеют вид, показанный на рис. 1*а*. В системе координат x,y,z они принимают вид

$$u_i(x,y,z)=\iiint f(x',y',z')h_i(x\cos\varphi_i\sin\theta_i+y\sin\varphi_i\sin\theta_i+z\cos\theta_i-z')\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z'.$$

Двумерные проекции u_i, полученные карандашным лучом с осью вдоль направления $l_i(\varphi_i,\theta_i)$, на плоскости, перпендикулярной этому направлению, показаны на рис. 1*б*. В системе координат x,y,z проекции принимают вид

$$u_i(x,y,z)=\\=\iiint f(x',y',z')h_i(x\cos\varphi_i\cos\theta_i+y\sin\varphi_i\cos\theta_i-z\sin\theta_i-x',y\cos\varphi_i-x\sin\varphi_i-y')\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z'.$$

На основе набора проекций по схеме обратного проецирования строится так называемое суммарное, или «грязное», изображение $g(x,y,z)$:

$$g(x,y,z)=\sum_{i=1}^{N}u_i(x,y,z). \tag{2}$$

Наблюдение объекта с нескольких направлений с диаграммами направленности $h_i(x,y,z)$ эквивалентно наблюдению с использованием синтезированного луча P:

$$P=\sum_{i=1}^{N}h_i(x,y,z).$$

Суммарное изображение, синтезированный луч (эквивалентная суммарная передаточная функция) и распределение яркости связаны соотношением

$$g(x,y,z)=\iiint f(x',y',z')P(x-x',y-y',z-z')\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z'+n(x,y,z), \tag{3}$$

где слагаемое $n(x,y,z)$ описывает шум. Для решения уравнения (3) используются итерационные методы с нелинейными ограничениями: стандартная чистка ST-CLEAN [11] или чистка по контуру TC-CLEAN [12]. Алгоритм поиска решения сводится к проведению серии расчётов распределения яркости с различными коэффициентами усиления итерации или уровня контура, вычислению контрольных проекций, определению невязки с исходными проекциями. В качестве решения выбирается вариант, соответствующий минимуму невязки. В сложных случаях может использоваться метод двух чисток (2-CLEAN DSA), т. е. оба алгоритма. При отсутствии априорной информации это даёт возможность определить область допустимых решений. Блок-схема метода реконструкции аналогична изображённой ранее на рис. 3*в* в работе [6].

Проиллюстрируем процесс восстановления трёхмерной структуры на простом примере. Распределение яркости примем в виде суммы двух гауссовых функций:

$$f(x,y,z) = A\exp\left\{-\frac{[(x-x_\mathrm{a})^2+(y-y_\mathrm{a})^2+(z-z_\mathrm{a})^2]\,d}{S_\mathrm{a}^2}\right\} + \\ + B\exp\left\{-\frac{[(x-x_\mathrm{b})^2+(y-y_\mathrm{b})^2+(z-z_\mathrm{b})^2]\,d}{S_\mathrm{b}^2}\right\}, \quad (4)$$

с параметрами $A=10$; $B=20$; $d=4\ln 2$; $S_\mathrm{a}=0{,}7$; $S_\mathrm{b}=0{,}5$; $x_\mathrm{a}=0{,}3$; $x_\mathrm{b}=-0{,}8$; $y_\mathrm{a}=0{,}2$; $y_\mathrm{b}=0{,}7$; $z_\mathrm{a}=-0{,}5$; $z_\mathrm{b}=0{,}1$. Рассмотрим сканирование объекта приёмными диаграммами простой формы с трёх направлений: по осям x,y и z. В этом случае суммарное изображение и синтезированный луч можно представить в аналитическом виде.

1.1. Синтезированный луч (эквивалентная суммарная передаточная функция)

Диаграммы направленности для ножевого луча h_k и для карандашного луча h_p примем в виде

$$h_\mathrm{k}(\alpha) = \exp\left(-\frac{\alpha^2 d}{S^2}\right), \qquad h_\mathrm{p}(\alpha,\beta) = \exp\left(-\frac{\alpha^2+\beta^2}{S^2}\,d\right), \quad (5)$$

где α принимает значения x,y и z, а β — значения y,z и x соответственно. Ширина диаграммы направленности на уровне 0,5 равна S и характеризует разрешение одномерных проекций, разрешение двумерных соответствует $S\times S$. Синтезированные лучи P_k и P_p определяются формулами

$$P_\mathrm{k}(x,y,z) = \exp\left(-\frac{x^2 d}{S^2}\right) + \exp\left(-\frac{y^2 d}{S^2}\right) + \exp\left(-\frac{z^2 d}{S^2}\right),$$

$$P_\mathrm{p}(x,y,z) = \exp\left(-\frac{x^2+y^2}{S^2}\,d\right) + \exp\left(-\frac{y^2+z^2}{S^2}\,d\right) + \exp\left(-\frac{z^2+x^2}{S^2}\,d\right)$$

и представляют собой характеристики виртуальных датчиков, измеряющих распределение яркости (плотности) внутри объекта. Синтезированные лучи показаны на рис. 2 в виде последовательности поверхностей одинакового уровня интенсивности (в виде изоповерхностей). Область вычисления функций ограничивается кубом с заранее заданным размером ребра.

Синтезированный луч на основе трёх ножевых диаграмм показан на рис. 2*а*. Горизонтальная плоскость в случаях *1* и *2* является реальным элементом поверхности, в остальных случаях она используется как опорный элемент. Синтезированный луч на основе трёх карандашных диаграмм показан на рис. 2*б*. Здесь горизонтальная поверхность используется только как опорный элемент. Уровни интенсивности для поверхностей приведены в табл. 1, где также даны размеры рёбер кубов, ограничивающих область представления каждой фигуры (на рис. 2 они соответствуют размерам квадрата в горизонтальной плоскости). Максимумы интенсивности синтезированных лучей расположены в центре системы координат и нормированы на единицу. С понижением уровня появляются так называемые боковые лепестки. Максимальный уровень боковых лепестков синтезированных лучей равен 0,33 в случае трёх карандашных диаграмм и 0,66 в случае трёх ножевых диаграмм. Соответствующие фигуры выделены рамкой. При дальнейшем повышении уровня изоповерхности принимают вид замкнутых контуров. Их контуры становятся близкими по форме.

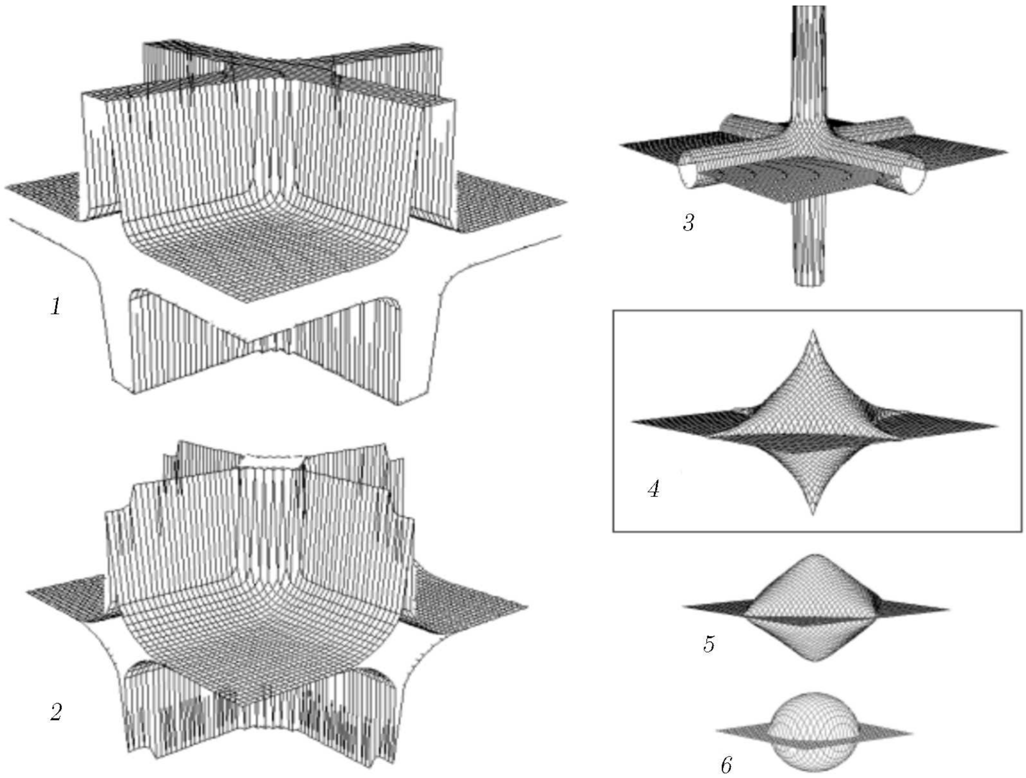

Рис. 2*а*. Синтезированный луч из трёх ножевых диаграмм (поверхности *1*–*6* соответствуют уровням интенсивности, указанным в табл. 1)

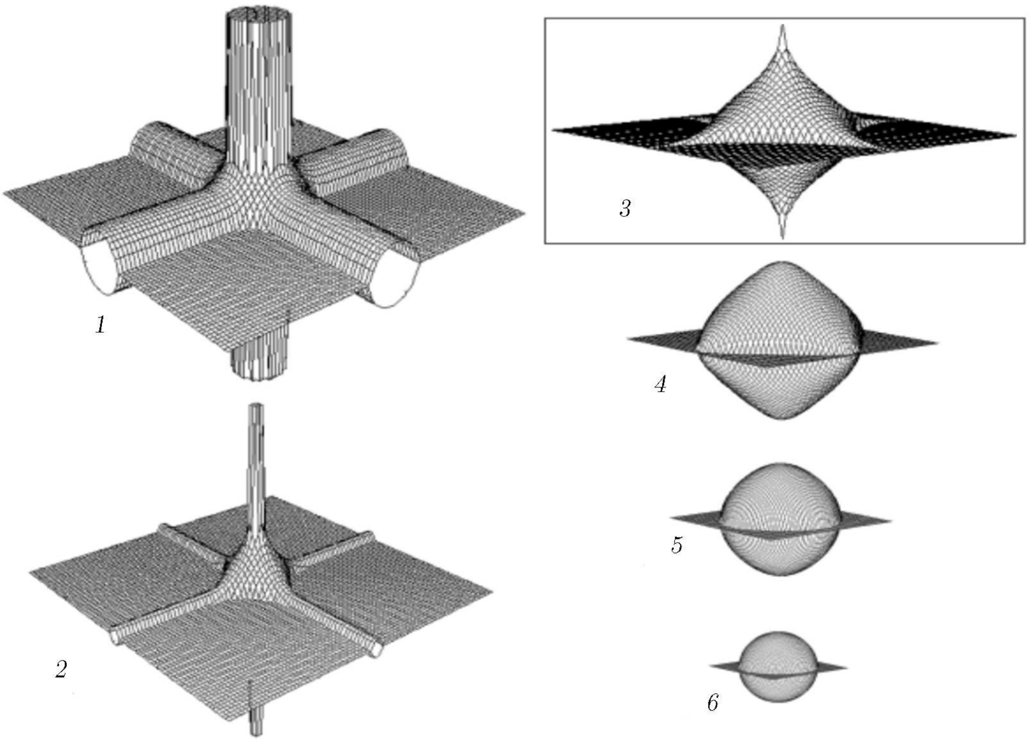

Рис. 2*б*. Синтезированный луч из трёх карандашных диаграмм (поверхности *1*–*6* соответствуют уровням интенсивности, указанным в табл. 1)

Таблица 1

Номер поверхности (контура) на рис. 2	Уровень контура (интенсивность)		Размер ребра куба, ограничивающего область представления поверхности	
	рис. 2*а*	рис. 2*б*	рис. 2*а*	рис. 2*б*
1	0,17	0,10	8,0 S	8,0 S
2	0,33	0,30	8,0 S	8,0 S
3	0,50	0,33	8,0 S	3,6 S
4	0,66	0,40	3,6 S	2,0 S
5	0,70	0,60	2,0 S	1,6 S
6	0,97	0,97	0,6 S	0,6 S

Это хорошо видно на рис. 2*а* и *б*. Максимальные уровни боковых лепестков для рассматриваемых случаев отличаются в два раза.

При отсутствии боковых лепестков или при большом числе проекций, когда боковыми лепестками можно пренебречь, сканирование трёхмерного объекта эквивалентно свёртке с функцией, отличной от нуля только в центральной области. В этом случае изоповерхности синтезированных лучей имеют форму сфер, интенсивность уменьшается с увеличением радиуса. Результатом свёртки распределения яркости трёхмерного объекта с такими лучами будет изображение, обладающее минимальными искажениями. Диаметр сферы на уровне половины интенсивности отвечает за угловое разрешение.

Обсудим вопрос о зависимости максимального уровня боковых лепестков от количества проекций и размерностей объекта и приёмной диаграммы. Рассмотрим случай приёмной диаграммы с одним максимумом. Синтезированный луч нормирован на единицу, уровень каждого из боковых лепестков не выше $1/N$, что следует из алгоритма построения синтезированного луча. В случае $2\mathrm{D}_{1\mathrm{D}}$ и $3\mathrm{D}_{2\mathrm{D}}$ это и есть максимальный уровень боковых радиальных лепестков. Выше этого уровня изоповерхность синтезированного луча замкнута. В случае ножевых лучей уровень $1/N$ соответствует плоскостям ножевого луча. Будем считать, что направления сканирования распределены в пространстве равномерно, т. е. по одной линии пересекаются только два ножевых луча. Тогда боковым лепесткам, образующимся при пересечении двух ножевых диаграмм, соответствует уровень $2/N$. Это и есть максимальный уровень боковых лепестков. Таким образом, максимальный уровень боковых лепестков $\mathrm{SL}_{\max}$ можно представить формулой

$$\mathrm{SL}_{\max} = (M - m)/N, \tag{6}$$

где M и m — размерности объекта и проекций соответственно, N — число проекций. Формула (6) справедлива для вариантов $2\mathrm{D}_{1\mathrm{D}}$ и $3\mathrm{D}_{2\mathrm{D}}$. В случае $3\mathrm{D}_{1\mathrm{D}}$ она справедлива при условии равномерного распределения направлений сканирования в пространстве. Если по одной прямой пересекаются K ножевых лучей, то максимальный уровень боковых лепестков составляет K/N. В случае, когда все ножевые лучи пересекаются по одной прямой, изоповерхность синтезированного луча при любом уровне интенсивности будет незамкнутой. Максимальный уровень боковых лепестков одинаков в случаях $2\mathrm{D}_{1\mathrm{D}}$ и $3\mathrm{D}_{2\mathrm{D}}$ при одном и том же числе проекций. Число одномерных проекций для трёхмерной реконструкции должно быть удвоено по сравнению с числом двумерных проекций для решения той же задачи.

1.2. Суммарные изображения

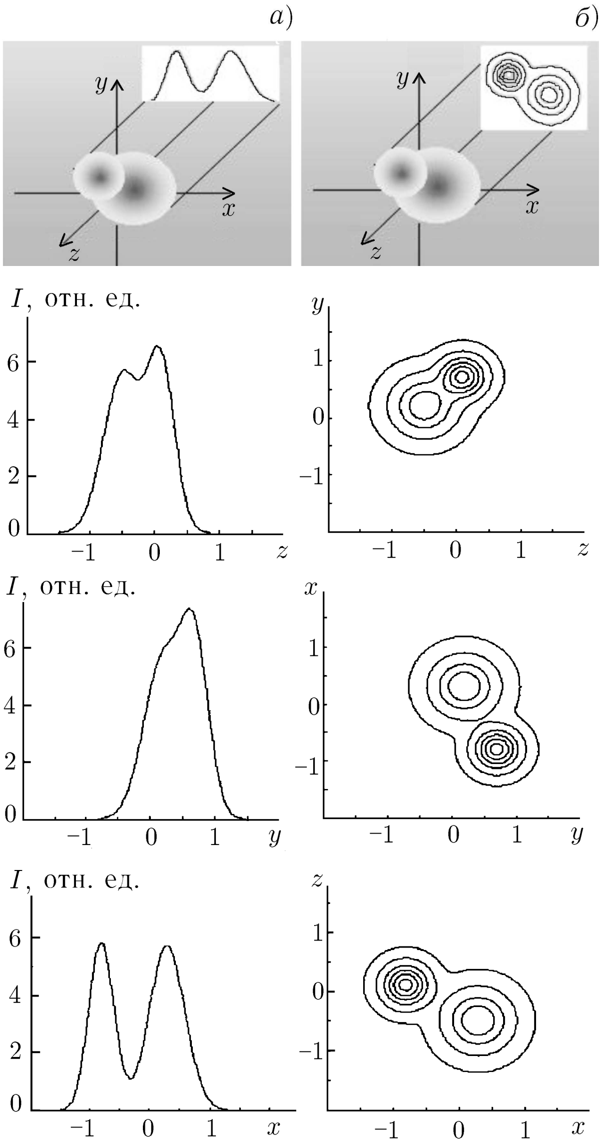

Рис. 3. Расчётные проекции (распределения интенсивности I) в трёх направлениях для принятой модели: одномерные при ножевых лучах (*а*) и двумерные при карандашных лучах (*б*)

Проекции являются свёрткой распределения яркости с диаграммой направленности. В случае сканирования вдоль осей координат выражение для проекций имеет наиболее простой вид. Проекции, полученные ножевым лучом, перемещающимся вдоль оси координат α, можно представить в виде $u_i(x,y,z) = V(\alpha)$, где

$$V(\alpha) = \iiint f(\mu_1,\mu_2,\mu_3) \times \times \mathrm{d}\mu_j\,\mathrm{d}\mu_m\, h_{\mathrm{k}}(\alpha-\mu_i)\,\mathrm{d}\mu_i, \quad (7)$$

где $i = 1,2,3$; $j = 2,3,1$; $m = 3,1,2$ соответственно для $\alpha \equiv x,y,z$. В случае карандашного луча, перемещающегося параллельно оси γ, двумерные проекции в плоскости $\alpha\beta$ имеют вид $u_i(x,y,z) = W(\alpha,\beta)$, где

$$W(\alpha,\beta) = \iiint f(\mu_1,\mu_2,\mu_3)\,\mathrm{d}\mu_m \times \times h_{\mathrm{p}}(\alpha-\mu_i,\beta-\mu_i)\,\mathrm{d}\mu_i\,\mathrm{d}\mu_j, \quad (8)$$

где $i = 1,2,3$; $j = 2,3,1$; $m = 3,1,2$ соответственно для $\alpha \equiv x,y,z$, $\beta \equiv y,z,x$, $\gamma = z,x,y$.

На рис. 3 показаны полученные результаты расчёта проекций для принятой модели при ножевых и карандашных лучах. Иллюстрация к проецированию для каждого из вариантов приводится в верхней части рисунка.

Построение трёхмерных массивов данных для суммарных изображений, или, в астрономической терминологии, грязных карт, выполнялось по схеме обратного проецирования, аналогичной случаю реконструкции двумерного изображения. Для визуализации модели, суммарных изображений и результатов реконструкции использовалось представление набором сечений. Полученные сечения, соответствующие трёхмерным массивам, изображены на рис. 4 для пяти различных положений на оси z. На рис. 4*а* показаны сечения модели, на рис. 4*б* и *в* — сечения двух суммарных изображений, построенных по трём одномерным проекциям и трём двумерным проекциям соответственно. Цифрами *1*–*6* на рис. 4 обозначены номера изофот, интенсивность которых соответствует 1, 10, 30, 50, 70 и 90 % от максимума трёхмерной функции.

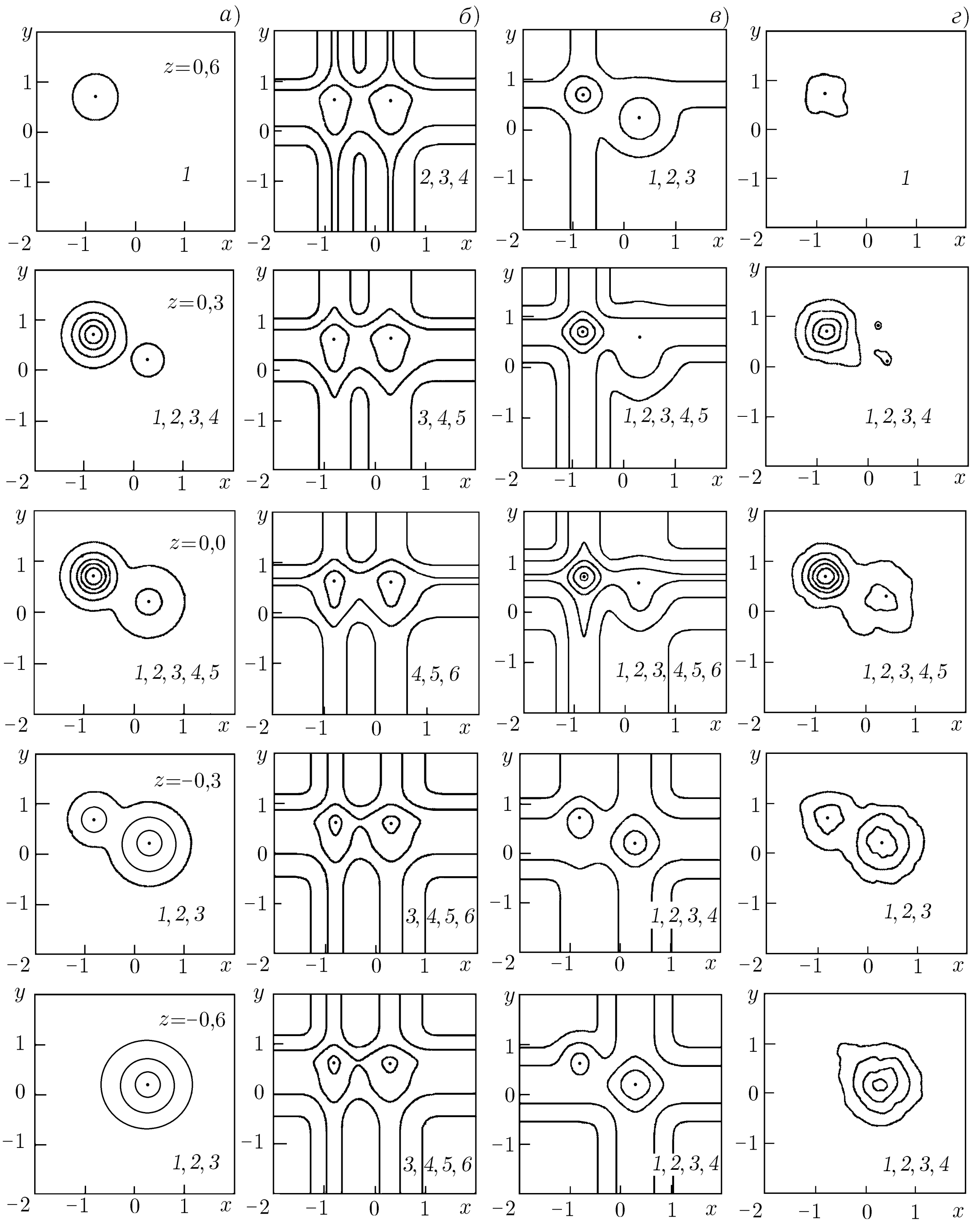

Рис. 4. Изофоты в сечениях трёхмерной модели (*а*) и суммарного трёхмерного изображения для проецирования при трёх ножевых (*б*) и трёх карандашных (*в*) диаграммах соответственно; результат реконструкции при чистке суммарного изображения (*в*) с использованием метода ST-CLEAN показан на панели *г*)

1.3. Чистка

На рис. 4*г* представлен результат трёхмерной реконструкции модели, полученный на основании трёх её двумерных проекций. Чистка суммарного изображения (см. рис. 4*в*) проводилась методом ST-CLEAN. В результате реконструкции мы имеем некоторое распределение плотности $\hat{f}$ излучающих частиц с разрешением, эквивалентным тому, которое можно получить от датчика с радиусом S, помещённого внутрь объекта. Введённый датчик должен регистрировать усреднённую по объёму S^3 плотность частиц. Поэтому сравнение результатов реконструкции следует проводить не с моделью, а с несколько сглаженным изображением f, полученным путём её свёртки с диаграммой направленности $h_0 = \exp[-(x^3 + y^3 + z^3)/S^3]$. Наилучшее совпадение результата с моделью достигнуто при коэффициенте усиления в итерациях $\lambda = 0{,}15$. Минимальная среднеквадратичная ошибка трёхмерной реконструкции $3\mathrm{D}_{2\mathrm{D}}$, вычисляемая по формуле

$$\sigma = \sqrt{\sum_i [f(T_i) - \hat{f}(T_i)]^2 \Big/ \sum_i [f(T_i)]^2}\,,$$

при этом составила 13 %. Отметим, что трёх одномерных проекций оказалось явно недостаточно для достижения удовлетворительного результата в случае чистки суммарного изображения, приведённого на рис. 4*б*. Вариант реконструкции $3\mathrm{D}_{1\mathrm{D}}$ требует большего числа ракурсов.

Использование второго метода чистки TC-CLEAN для трёхмерной реконструкции принципиально не отличается от его приложения в случае двумерной реконструкции, рассмотренном ранее [6]. Алгоритм адаптируется к трёхмерному восстановлению: приспособленный контур становится поверхностью в трёхмерном пространстве, а в итерациях выполняются трёхмерные быстрые преобразования Фурье. В итоге метод двух чисток также адаптируется к трёхмерной реконструкции. Это согласуется с приведённым в [13] утверждением о том, что методы обработки одномерных ($m = 1$) и двумерных ($m = 2$) сигналов или данных имеют существенные отличия, однако дальнейшее повышение размерности ($m \geq 2$) не приводит к заметным отличиям от двумерного случая, кроме усложнения вычислений.

Рассмотренная несложная модель позволила наглядно иллюстрировать общий подход, эффективность чистки и возможность реконструкции трёхмерной структуры двухкомпонентного объекта при наличии лишь трёх двумерных проекций.

2. ОБСУЖДЕНИЕ

2.1. Особенности реконструкции

Проанализируем рассмотренные варианты трёхмерной реконструкции, принимая во внимание максимальные уровни боковых лепестков синтезированных лучей при ножевых и карандашных приёмных диаграммах (см. табл. 1 и рис. 2), а также учитывая структуру синтезированного луча для двумерной реконструкции [6]. Формулой (6) была показана связь уровня боковых лепестков $\mathrm{SL}_{\max}$ синтезированного луча с количеством проекций N для реконструкции в пространстве размерности M при учёте размерности проекций m. Величина искажений на суммарных изображениях возрастает вместе с уровнем боковых лепестков соответствующих синтезированных лучей и уменьшается с увеличением числа используемых проекций. Построение эквивалентных суммарных передаточных функций по гауссовским профилям для трёх вариантов реконструкции показывает, что при одинаковом числе проекций N максимальный уровень боковых лепестков синтезированного луча при трёхмерной реконструкции на основе двумерных проекций

$$\mathrm{SL}_{\max}(3\mathrm{D}_{2\mathrm{D}}) = \mathrm{SL}_{\max}(2\mathrm{D}_{1\mathrm{D}}) \tag{9}$$

равен максимальному уровню боковых лепестков синтезированного луча при двумерной реконструкции на основе одномерных проекций. Можно утверждать, что и степень искажений в случаях $3D_{2D}$ и $2D_{1D}$ одинакова при условии равномерного распределения проекций. Вместе с тем достижение аналогичного качества при трёхмерной реконструкции при использовании ножевых лучей ($3D_{1D}$) требует удвоенного числа профилей $2N$. Всё вышеизложенное позволяет сделать вывод о возможности проведения предварительного менее сложного моделирования ряда задач в двумерном варианте с последующим переносом методических результатов на случай трёхмерной реконструкции.

Заметим, что неравномерное распределение проекций в пространстве будет приводить к различиям разрешения по разным направлениям. «Экваториальный» вариант расположения того же числа двумерных проекций приведёт к ухудшению разрешения в горизонтальной плоскости. Это будет отражать форма изоповерхностей трёхмерного синтезированного луча. Поверхности с уровнем выше значений отсечения боковых лепестков сожмутся по вертикальной оси. При этом разрешение трёхмерного изображения будет различно в горизонтальном и вертикальном направлениях.

2.2. Сопоставление

Разработанный ранее подход для двумерной реконструкции [6, 14] с использованием метода двух чисток мы распространили на два варианта трёхмерной задачи. Большинство известных методов реконструкции (см., например, [1, 2, 4, 15]) также могут быть адаптированы к трёхмерным задачам. Методы, основанные на фильтрации, наиболее распространены, однако при малом числе проекций они неэффективны и не позволяют бороться с искажениями. Стандартные алгоритмы обратных фильтрованных проекций основаны на классической работе [5]. Модификации метода различаются лишь формой фильтрующей функции (линейной в варианте Рема—Лака, косинусной в варианте Шеппа—Логана и т. д.) в пределах установленной частоты усечения. Для проведения корректной реконструкции такими методами требуется снижение пространственной частоты среза $\omega_{\rm b}$ до величины, соответствующей полному заполнению области пространственных частот. При использовании трёх двумерных проекций такая частота настолько мала, что не позволяет обеспечить необходимое разрешение двух компонент у выбранной модели. Наличие боковых лепестков синтезированного луча свидетельствует о малом числе проекций и приводит к провалам в заполнении области пространственных частот суммарной передаточной функцией. Методы, основанные на фильтрации, не устраняют откликов от боковых лепестков. Интерполирующая процедура оказывает положительный эффект, однако он невелик. Методы чистки, являясь итерационными алгоритмами с нелинейными ограничениями, имеют принципиальное отличие, поскольку обладают нелинейными свойствами. Разработанный подход позволяет в несколько раз (при высоком отношении сигнала к шуму — почти десятикратно) сократить необходимое число ракурсов по сравнению с методом обратных фильтрованных проекций. Расплатой за столь сильную экономию в сложных случаях может служить возникновение неоднозначности, некоторой области допустимых решений. Однако метод двух чисток позволяет определить её границы на основании двух экстремальных вариантов решений [6]. В итоге может быть выполнен необходимый анализ и распознавание структуры.

Реконструкция с использованием метода максимальной энтропии эффективна при обработке больших массивов данных. Это связано с введением статистического подхода. Метод позволяет свести к минимуму влияние шумов, но при наличии большого числа проекций. В [1, 4] описано применение для реконструкции итерационного алгоритма Гершберга—Папулиса с введением спектральных ограничений. Использование априорной информации для выделения некоторой области спектра пространственных частот является сильным средством. Итерационная процедура заключается в последовательной подгонке данных. Чистка имеет существенные отличия от этого метода, описанного в [1], на этапе выбора компонентов решения и при дальнейшем их последовательном исключении при итерациях непосредственно в пространстве суммарного изображения. Эффективность метода связана с введением синтезированного луча. Кроме того, способность к реконструкции в широкой полосе пространственных частот без априорной информации о спектре реализуется благодаря использованию метода двух чисток и является важной особенностью метода. Введение ограничений, например информации о спектре, обеспечивает дополнительную возможность сокращения числа проекций. Достоинствами разработанного подхода является его относительная простота и тот немаловажный факт, что чистка уже хорошо известна астрономам, т. к. применяется в системах апертурного синтеза. Заметим, что в современных промышленных томографах используют полное заполнение области пространственных частот, а методы и программы реконструкции, как правило, основаны на фильтрации.

2.3. Перспективы использования

Разработанный подход к реконструкции томограмм может быть адаптирован к различным задачам. Применение метода не ограничивается излучающими объектами. Перспективно его приложение к трансмиссионным (с использованием профилей поглощения) и локационным томографическим задачам. Вычисление синтезированных лучей на основе суммирования приёмных диаграмм при разных ракурсах отражает лишь один из возможных вариантов их построения. В зависимости от решаемой задачи возможно введение синтезированных виртуальных лучей на основе различных передаточных функций, соответствующих, например, временно́му или спектральному разрешению проекций. В этой связи следует кратко остановиться на двух важных задачах астрофизики и дистанционного зондирования, для которых приложение метода, на наш взгляд, представляет несомненный интерес.

Дистанционное зондирование (варианты $\mathbf{2D_{1D}}$ и $\mathbf{3D_{1D}}$). Реконструкция с введением синтезированного луча из передаточных функций, соответствующих временно́му разрешению, перспективна и может быть адаптирована для применения в радиолокационных задачах. Ширина отдельной передаточной функции, форма которой аналогична ножевой диаграмме, определяет в этом случае разрешение по дальности. Значения одномерной проекции отражённого сигнала, соответствующие временно́й оси, привязываются к пространственным координатам с учётом направления зондирования. В [16] отмечалось, что томографические методы реконструкции могут использоваться в качестве основы формирования радиолокационных изображений. Актуальность развития таких методов для решения задач построения и распознавания локационных изображений при землеобзоре, в метеорологии, высокоинформативных изображений летательных аппаратов подчёркивалась неоднократно [16, 17]. Использование метода синтезированной апертуры для повышения углового разрешения имеет предел из-за нестабильности, возникающей при увеличении времени синтеза. Некогерентная томографическая обработка может быть более перспективной для повышения информативности. Отказ от учёта фазы и обработка только амплитуд предусматривают переход от высокоточного аппаратного оснащения к потребности в вычислительных ресурсах. Как отмечалось в [18], практическое использование такой обработки при ограничении числа проекций, сектора расположения углов, а также их неравномерного распределения требует разработки новых алгоритмов цифровой некогерентной обработки ракурсной информации, обеспечивающих получение на изображениях тангенциального разрешения, равного радиальному, независимо от дальности. Использование разработанного подхода более перспективно по сравнению с методами реконструкции, основанными на преобразовании

Радона и фильтрации. После привязки локационных профилей к координатной сетке эквивалентная суммарная передаточная функция будет определять как радиальное, так и тангенциальное разрешение исследуемой области. Принцип введения суммарной передаточной функции ($2D_{1D}$) для двумерной томографической реконструкции изображений в задаче бокового землеобзора в прожекторном режиме показан нами в [19]. Многопозиционная локация позволяет проводить реконструкцию трёхмерных изображений объектов с использованием синтезированного луча на основе передаточных функций, ответственных за временно́е разрешение для направлений регистрации проекций. Поворот объекта в пространстве также даёт набор ракурсов. Трёхмерная суммарная передаточная функция вводится аналогично построению синтезированного луча для варианта ($3D_{1D}$), изображённого на рис. 2*а*. Разработанный подход адаптируется к упомянутому выше варианту двумерного землеобзора и может служить основой развития томографической реконструкции трёхмерных высокоинформативных локационных изображений.

Астрофизика (варианты $\mathbf{2D_{1D}}$ и $\mathbf{3D_{1D}}$). В последние годы томографические методы обработки активно используются для получения данных о структуре космических объектов [20]. Такие исследования объединены общим названием астротомография. В [21] показаны перспективы использования разработанного нами двумерного варианта метода [6] в подобных астрономических задачах. Применение его целесообразно для реконструкции доплеровских томограмм [22] при исследованиях двойных систем. Основой служат одномерные профили спектров эмиссионных линий. При знании эфемерид, преобразованных к значениям орбитальных фаз $\phi(t)$, они пересчитываются в кривые лучевых скоростей $V_{\mathrm{R}}(\phi)$. Для реконструкции доплеровских томограмм, как правило, используется метод обратных фильтрованных проекций. В доплеровской томографии он получил распространение с 1991 года [23] и широко применяется до сих пор [24, 25]. Несмотря на вращение объектов, равномерное распределение спектрограмм удаётся получить далеко не всегда. Недостаток наблюдательного времени и отсутствие благоприятных условий являются причинами возникновения скважности. Полученный в итоге ряд может быть неполным, число орбитальных фаз, или ракурсов, ограничено, что делает применение разработанного метода весьма перспективным. Для построения синтезированных лучей в этом случае необходимо использовать передаточные функции, ответственные за спектральное разрешение профилей линий, пересчитанное в разрешение лучевых скоростей. Метод доплеровской томографии, предложенный в [22], является косвенным способом получения информации о космических объектах. Компоненты двойных систем, которые не разрешаются прямыми измерениями непосредственно из наблюдений, разрешаются в пространстве скоростей. Томограммы соответствуют распределению интенсивности излучения $I(V_x, V_y)$ в пространстве скоростей на частотах некоторой эмиссионной линии. Доплеровская карта легче поддаётся интерпретации, чем исходные спектрограммы. Она может указывать на характерные особенности течения вещества. При некоторых предположениях возможно решение второй задачи — преобразование распределения интенсивности из пространства скоростей в пространство координат: $I(V_x, V_y) \to I(x, y)$. О росте интереса к теме свидетельствуют также отечественные работы [26, 27], в первой из которых представлены синтетические доплеровские томограммы, построенные по результатам трёхмерных расчётов, а во второй — результаты трёхмерного численного моделирования течения вещества в двойных системах.

Широкое распространение получила реконструкция двумерных доплеровских томограмм. Согласно [22] было введено упрощение, связанное с предположением о равенстве нулю третьей компоненты скорости: $V_z = 0$. Это приближение оправдано при исследованиях многих двойных систем. Однако, как отмечено недавно в [28], введение третьей компоненты скорости было бы весьма полезно для систем, в которых скорости вещества значительны вне орбитальной плоскости. Трёхмерная реконструкция может быть полезна также при прецессии оси системы. Реконструкция трёхмерной доплеровской томограммы должна выполняться на основании одномерных спектральных профилей, что соответствует рассмотренному варианту $3D_{1D}$. Применение в задаче разработанного подхода может быть наиболее эффективно.

Вариант $\mathbf{3D_{2D}}$. Наличие двумерных проекций — изображений излучающих оптически тонких объектов для различных ракурсов — позволяет провести реконструкцию их внутренней структуры (вариант $3D_{2D}$). Реализовать подобные исследования в астрофизике возможно лишь для вращающихся объектов при достаточном разрешении и стационарности объекта за период наблюдений. При исследованиях Солнца такая задача могла бы решаться при одновременном получении изображений с использованием одного или нескольких космических аппаратов.

В задачах с использованием большого числа равномерно распределённых проекций замена метода обратного фильтрованного проецирования на разработанный радиоастрономический метод также может быть перспективна, т. к. позволяет существенно расширить при реконструкции динамический диапазон.

3. ЗАКЛЮЧЕНИЕ

Рассмотрена возможность приложения радиоастрономического метода CLEAN к двум вариантам трёхмерной томографической реконструкции. Решение задачи методом деконволюции в пространстве изображений с введением синтезированного луча и процедуры чистки позволяет сократить в несколько раз число проекций по сравнению с традиционными способами реконструкции. Подход к реконструкции, разработанный ранее для двумерной задачи и включающий использование метода двух чисток, распространён на трёхмерные задачи. Его использование иллюстрировано примером реконструкции двухкомпонентной модели.

Основные выводы проведённого исследования:

1) Предложенный метод реконструкции трёхмерной структуры в целом подобен рассмотренному ранее случаю двумерной реконструкции. Он позволяет использовать меньшее по сравнению с традиционным подходом число проекций для восстановления в широкой полосе пространственных частот.

2) Анализ формы синтезированных лучей (эквивалентных суммарных передаточных функций), построенных с использованием функций с гауссовским профилем, позволяет утверждать, что достижение качества реконструкции, аналогичного двумерному варианту ($2D_{1D}$), требует в случае трёхмерной реконструкции на основе двумерных проекций ($3D_{2D}$) использования равного числа проекций при условии их равномерного распределения в пространстве. Достижение аналогичного качества трёхмерной реконструкции при использовании одномерных проекций ($3D_{1D}$) требует удвоенного числа проекций в сравнении с количеством двумерных проекций при трёхмерной реконструкции ($3D_{2D}$).

3) Предварительная отработка ряда методических вопросов трёхмерного моделирования возможна на основе менее сложного в вычислительном отношении двумерного варианта.

Разработанный подход к реконструкции трёхмерных изображений перспективен для применения в различных задачах при ограниченном числе ракурсов. Он позволяет проводить реконструкцию томограмм в координатном пространстве и в пространстве скоростей. Допускается введение различных вариантов эквивалентных суммарных передаточных функций. Такие виртуальные синтезированные лучи могут быть построены не только на основе приёмных диаграмм, но и с использованием набора других передаточных функций. В задачах радиолокации эти функции соответствуют временно́му разрешению исходных одномерных профилей, в доплеровской астротомографии — разрешению радиальной скорости. Метод достаточно прост в вычислительном отношении, легко адаптируется к введению ограничений, основанных на априорной информации.

Авторы выражают благодарность Д. В. Бисикало и Е. А. Карицкой за внимание к работе. Работа выполнена при частичной поддержке Минобрнауки России в рамках научно-технической программы «Научные исследования высшей школы по приоритетным направлениям науки и техники» (проект 209.01.01.003), а также РФФИ (грант No. 04–02–16924а) и программы по поддержке ведущих научных школ Российской Федерации (проект НШ–1483.2003.2).

СПИСОК ЛИТЕРАТУРЫ

1. Левин Г. Г., Вишняков Г. Н. Оптическая томография. М.: Радио и связь, 1989.
2. Физика визуализации изображений в медицине: в 2-х томах. Пер. с англ. / Под ред. С. Уэбба. М.: Мир, 1991.
3. Бейтс Р. Х. Т., Гарден К. Г., Петерс Т. М. // ТИИЭР. 1983. Т. 71, No. 3. С. 84
4. Низкотемпературная плазма. Т. 13. Томография плазмы / В. В. Пикалов, Т. С. Мельникова. Новосибирск: Наука, 1995.
5. Bracewell R. N., Riddle A. C. // Astrophys. J. 1967. V. 150. P. 427.
6. Агафонов М. И. // Изв. вузов. Радиофизика. 2004. Т. 47, No. 2. С. 94.
7. Томпсон А. Р., Моран Д. М., Свенсон Д. У. Интерферометрия и синтез в радиоастрономии. М.: Физматлит, 2003.
8. Шафер Р., Мерсеро Р., Ричардс М. // ТИИЭР. 1981. Т. 69, No. 4. С. 432.
9. Василенко Г. И., Тараторин А. М. Восстановление изображений. М.: Радио и связь, 1986.
10. Агафонов М. И., Подвойская О. А. // Изв. вузов. Радиофизика. 1989. Т. 32, No. 6. С. 742.
11. Hogbom J. A. // Astron. Astrophys. Suppl. Ser. 1974. V. 15, No. 3. P. 417.
12. Steer D. G., Dewdney P. E., Ito M. R. // Astron. Astrophys. 1984. V. 137, No. 2. P. 159.
13. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов: Пер. с англ. М.: Мир, 1988.
14. Агафонов М. И., Подвойская О. А. // Изв. вузов. Радиофизика. 1990. Т. 33, No. 10. С. 1 185.
15. Наттерер Ф. Математические аспекты компьютерной томографии: Пер. с англ. М.: Мир, 1990.
16. Кононов А. Ф. // Зарубежная радиоэлектроника. 1991. No. 1. С. 35.
17. Вопросы перспективной радиолокации / Под ред. А. В. Соколова. М.: Радиотехника, 2003.
18. Опаленов Ю. В., Потапов А. А. // Радиотехника и электроника. 2000. Т. 45, No. 12. С. 1 447.
19. Агафонов М. И., Шарова О. И. // Докл. XI Междунар. научно-техн. конф. «Радиолокация, навигация, связь». Т. 3. Воронеж, 2005. С. 1 647.
20. Cameron A. C., Schwope A., Vrielman S. // Astron. Nachrichten. 2004. V. 325, No. 3. P. 179.
21. Agafonov M. I. // Abstr. IAU XXV General Assembly (JD9-Astrotomography), Sydney, 2003. P. 193.
22. Marsh T. R., Horn K. // Mon. Not. R. Astron. Soc. 1988. V. 235. P. 269.
23. Horn K. // Fundamental Properties of Cataclysmic Variable Stars / Ed. by A. W. Shafter. San Diego State University, 1991. P. 53.
24. Richards M. // Astron. Nachrichten. 2004. V. 325, No. 3. P. 229.
25. Morales-Rueda L. // Astron. Nachrichten. 2004. V. 325, No. 3. P. 193.
26. Кузнецов О. А., Бисикало Д. В., Боярчук А. А. и др. // Астрон. ж. 2001. Т. 78, No. 11. С. 997.
27. Бисикало Д. В., Боярчук А. А., Кузнецов О. А., Чечёткин В. М. // Астрон. ж. 2000. Т. 77, No. 1. С. 31.
28. Steeghs D. // Astron. Nachrichten. 2004. V. 325, No. 3. P. 185.

Поступила в редакцию 1 июля 2003 г.; принята в печать 23 мая 2005 г.

FEW PROJECTIONS TOMOGRAPHY. II. RADIOASTRONOMICAL METHOD CLEAN IN THE 3-DIMENTIONAL APPLICATIONS

M. I. Agafonov and O. I. Sharova

The radioastronomical method CLEAN is expanded into the problems of 3D tomographical reconstruction. The two variants are analyzed: $3D_{1D}$-reconstruction with 1D-projections and $3D_{2D}$-reconstruction on the basis of 2D-projections. Deconvolution with the introduction of the synthesized beam (the equivalent summary transfer function) allows to several times reduce the number of the required projections in comparison with the traditional approach. The connection between the level of the sidelobes of the synthesized beam constructed on the basis of Gaussian form of transfer functions and the number of the used projections is established. To achieve the quality of reconstruction in a $3D_{2D}$-case on the analogy of $2D_{1D}$ it is necessary to use an equal number of projections in the condition of their even distribution in the space. To achieve the same quality of $3D_{1D}$-reconstruction it is necessary to use a doubled number of 1D-projections. The approach to the problem is illustrated with the example of a 3D-model of an optically thin emitting object. The perspectives of the application of the developed approach in astrotomography and remote sensing with the introduction of various transfer functions, which can be responsible for the resolution of reception diagrams, resolutions of spectrographs, and time-resolution of radar profiles, are discussed.

УДК 621.371

АНИЗОТРОПНАЯ СТРУКТУРА МЕЛКОМАСШТАБНОЙ ИОНОСФЕРНОЙ ТУРБУЛЕНТНОСТИ

В. А. Алимов, Ф. И. Выборнов, А. В. Рахлин

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

Рассмотрены различные теоретические модели спектра мелкомасштабной ионосферной турбулентности. Отмечается особая роль обобщённой модели спектра ионосферной турбулентности, учитывающей зависимость анизотропии (вытянутости) мелкомасштабных неоднородностей верхней ионосферы вдоль направления магнитного поля Земли от поперечных масштабов этих неоднородностей. Приведены результаты первых специализированных экспериментов по радиозондированию среднеширотной ионосферы сигналами орбитальных ИСЗ на частотах 150 и 400 МГц в условиях повышенной солнечной активности. Эксперименты проводились на радиофизическом полигоне в Нижегородской области в 2003 году. Исследовались статистические характеристики амплитудных флуктуаций принимаемых сигналов при различном угле ϑ между лучом зрения с ИСЗ на наземный пункт приёма и направлением магнитного поля Земли. В ходе эксперимента была обнаружена зависимость наклона спектра амплитудных флуктуаций принимаемого излучения от угла ϑ. Полученный результат согласуется с обобщённой моделью спектра ионосферной турбулентности и может свидетельствовать в пользу резко выраженной анизотропной структуры мелкомасштабных неоднородностей электронной концентрации среднеширотной ионосферы в возмущённых геофизических условиях.

1. ТЕОРЕТИЧЕСКИЕ МОДЕЛИ СПЕКТРОВ МЕЛКОМАСШТАБНОЙ ТУРБУЛЕНТНОСТИ СРЕДНЕШИРОТНОЙ ИОНОСФЕРЫ

В исследованиях неоднородной структуры околоземной плазмы широко распространена степенная форма спектра флуктуаций электронной концентрации неоднородностей среднеширотной ионосферы [1, 2]:

$$\Phi_N(\boldsymbol{\kappa}) \propto [1 + \kappa_{0\perp}^{-2}\,(\kappa_\perp^2 + a^2\kappa_z^2)]^{-p/2}. \tag{1}$$

Здесь $\kappa_\perp^2 = \kappa_x^2 + \kappa_y^2$, κ_x, κ_y и κ_z — координаты пространственных волновых чисел неоднородностей в декартовой системе координат, связанной с магнитным полем Земли $\mathbf{H}_0$ (орт $\mathbf{z}_0 \uparrow\uparrow \mathbf{H}_0$), $\kappa_{0\perp} = 2\pi/l_{0\perp}$ — волновое число, соответствующее внешнему масштабу ионосферной турбулентности $l_{0\perp}$ в плоскости, перпендикулярной $\mathbf{H}_0$, коэффициент a характеризует анизотропию (вытянутость) ионосферных неоднородностей вдоль магнитного поля $\mathbf{H}_0$, p — показатель спектра неоднородностей.

Наряду с моделью (1) в работах по искусственной ионосферной турбулентности (ИИТ) была предложена модель анизотропного спектра (см. [3, 4])

$$\Phi_N(\boldsymbol{\kappa}) \propto \kappa_\perp^{-p}\Phi_N(\kappa_z) \approx \kappa_\perp^{-p}\exp\,[-\kappa_z^2 l_{0\|}^2/4], \tag{2}$$

которая подчёркивает резкую анизотропию ИИТ относительно направления магнитного поля Земли ($l_{0\|}$ — характерный масштаб ИИТ вдоль направления $\mathbf{H}_0$, $l_{0\|} \gg l_{0\perp}$).

В конце 90-х годов в [5] нами была предложена обобщённая модель спектра ионосферной турбулентности, учитывающая как анизотропные свойства крупномасштабной фракции турбулентности, так и зависимость анизотропии мелкомасштабных неоднородностей верхней ионосферы вдоль направления магнитного поля Земли от поперечных масштабов этих неоднородностей. Для среднеширотной ионосферы эта модель может быть описана следующим соотношением [5]:

$$\Phi_N(\boldsymbol{\kappa}) \propto \begin{cases} (\kappa_\perp^2 + \kappa_z^2)^{-p/2}, & l_\perp > l_\chi; \\ [\kappa_\perp^2 + (\kappa_\perp/\kappa_\chi)^{2\alpha}\,\kappa_z^2]^{-p/2}, & l_\perp \le l_\chi. \end{cases} \tag{3}$$

Здесь $\kappa_\chi = 2\pi/l_\chi$, l_χ — характерный масштаб неоднородностей в спектре ионосферной турбулентности, разделяющий крупномасштабную ($l_\perp > l_\chi$) и мелкомасштабную ($l_\perp < l_\chi$) фракции в модели анизотропной турбулентности верхней ионосферы, α — показатель анизотропии мелкомасштабных неоднородностей.

Измеряемый в экспериментах по просвечиванию ионосферы сигналами орбитальных ИСЗ одномерный частотный спектр амплитудных флуктуаций принимаемого излучения $\Phi_A(\nu)$ имеет существенно различные угловые зависимости для указанных моделей мелкомасштабной ионосферной турбулентности (здесь ν — спектральная частота флуктуаций) [1, 5].

Так, для модели (1) [4]

$$\Phi_A(\nu) \propto \nu^{-(p-1)} \tag{4}$$

при любой ориентации ϑ луча зрения (с единичным вектором $\mathbf{n}$) с ИСЗ на наземный пункт приёма относительно направления магнитного поля Земли $\mathbf{H}_0$ ($\vartheta = \widehat{\mathbf{n}\mathbf{H}_0}$).

Для модели (3) [5]

$$\Phi_A(\nu) \propto \begin{cases} \nu^{-(p-1)}, & \vartheta \le \vartheta_\chi; \\ \nu^{-[(\alpha+1)\,p-1]}, & \vartheta > \vartheta_\chi, \end{cases} \tag{5}$$

где ϑ_χ — малый (несколько градусов) угол [4, 5].

Для модели (2) справедливо соотношение (5) при $\alpha p \equiv 1$:

$$\Phi_A(\nu) \propto \begin{cases} \nu^{-(p-1)}, & \vartheta \le \vartheta_\chi; \\ \nu^{-p}, & \vartheta > \vartheta_\chi. \end{cases} \tag{6}$$

В рамках этой модели изменение наклона (показателя) спектра амплитудных флуктуаций в области углов $\vartheta \approx 0$ (при луче зрения вдоль направления $\mathbf{H}_0$) всегда составляет единицу.

В отличие от анизотропной модели спектра (2) обобщённая модель мелкомасштабной ионосферной турбулентности (3) предсказывает возможность вариаций наклона спектра амплитудных флуктуаций принимаемого излучения от ориентации луча зрения с ИСЗ относительно направления магнитного поля Земли для различных геофизических условий (параметры α и p для модели (3) независимы).

Модель (2) была диагностирована в экспериментах по просвечиванию высокоширотной ионосферы и среднеширотной ионосферы (при воздействии на неё мощным коротковолновым радиоизлучением) сигналами орбитальных ИСЗ в метровом и дециметровом диапазонах длин волн [4, 6]. В то же время для среднеширотной ионосферы в естественных условиях исследования анизотропной структуры мелкомасштабной ионосферной турбулентности не проводились, прежде всего, из-за низкого уровня мелкомасштабной ионосферной турбулентности на средних широтах в спокойных геофизических условиях. В настоящее время с учётом технических возможностей современных радиоизмерительных комплексов сигналов орбитальных ИСЗ в метровом и дециметровом диапазонах длин волн и, особенно, с учётом современных возможностей получения необходимой оперативной информации о гео- и гелиообстановке в околоземной плазме такие исследования становятся реальными. Наиболее интересны такие наблюдения при повышенной геофизической возмущённости в ионосфере средних широт, когда и должны проявиться особенности анизотропии мелкомасштабной турбулентности среднеширотной ионосферы.

Результаты первого подобного эксперимента, ориентированного на прямую проверку предложенной в [5] обобщённой модели спектра мелкомасштабной ионосферной турбулентности в среднеширотной ионосфере, представлены ниже.

2. СХЕМА ЭКСПЕРИМЕНТА

Эксперимент по приёму сигналов ИСЗ в метровом и дециметровом диапазонах длин волн проводился в Нижегородской области в 2003 году. Измерения амплитудных флуктуаций сигналов орбитальных ИСЗ на частотах 150 и 400 МГц при просвечивании ими среднеширотной ионосферы осуществлялись с помощью наземного радиоизмерительного комплекса НИРФИ. Ранее этот комплекс использовался в экспериментах по изучению неоднородной структуры среднеширотной ионосферы в условиях искусственной модификации ионосферы мощным коротковолновым радиоизлучением [6].

Для проведения измерений в 2003 году приёмный комплекс НИРФИ был модернизирован. Сигналы ИСЗ записывались с помощью измерительной платы в цифровом виде на персональный компьютер. В качестве измерительной платы в эксперименте использовался аудиоконтроллер SB Live 5.1. Для записи и последующей обработки сигналов использовалась программа SpectLab (версия 4.32.07). Для компенсации частотной характеристики платы SB Live 5.1 проводилась калибровка шумовым сигналом. Во время записи сигнал наблюдается на экране дисплея и прослушивается через наушники.

Измерительный комплекс позволяет проводить одновременную регистрацию двух сигналов на частотах 150 и 400 МГц. При обработке вычислялись текущие частотные спектры $\Phi_A(\nu)$ амплитуды сигналов в диапазоне до $\nu \sim 10$ Гц на отдельных временны́х интервалах длительностью $8\div12$ с.

Проведению экспериментов по исследованию статистических характеристик сигналов ИСЗ предшествовал предварительный расчёт моментов времени нахождения спутника в зоне радиовидимости с наземного пункта. Вычислялись азимут $\varphi_{\text{м}}$ и угол места $\vartheta_{\text{м}}$ спутника в каждую минуту его полёта. Для этой цели использовалась программа «Прогноз», разработанная в НИРФИ и ранее использовавшаяся в эксперименте [6]. Во время эксперимента 2003 года указанная программа была модернизирована с учётом соответствующей программы TRAKSTAR26 (версия 2.65, см. сайт [7]), а также программа Micro Orbites 3.0 (версия 4.2, см. сайт [8]). Программа TRAKSTAR26 позволяла рассчитать угол места, азимут и моменты времени для пролётного ИСЗ, а программа Micro Orbites 3.0 позволяла на экране дисплея наблюдать положение ИСЗ на карте Земли в реальном времени.

Основной задачей проводимых экспериментальных исследований является тестирование обобщённой модели ионосферной турбулентности (3) с помощью наземного радиоизмерительного комплекса сигналов орбитальных ИСЗ метрового и дециметрового диапазонов длин волн с различной ориентацией луча зрения с ИСЗ на наземный пункт приёма относительно направления магнитного поля Земли — выделенного направления для анизотропных мелкомасштабных неоднородностей электронной концентрации ионосферной плазмы (метод Ерухимова) [4]. Именно метод Ерухимова лежит в основе нашего эксперимента. Фактически, речь идёт об измерении угловой зависимости частотных спектров амплитудных флуктуаций $\Phi_A(\nu,\vartheta)$ принимаемых сигналов от бортовых передатчиков орбитальных ИСЗ после дифракции их в случайно-неоднородной ионосфере. Согласно развитым выше теоретическим представлениям функции $\Phi_A(\nu,\vartheta)$ носят специфический характер для каждой из трёх указанных теоретических моделей спектров мелкомасштабной ионосферной турбулентности (см. соотношения (1), (4); (2), (6) и (3), (5) соответственно). Поэтому анализ получаемых в эксперименте спектров амплитудных флуктуаций сигналов орбитальных ИСЗ $\Phi_A(\nu,\vartheta)$ должен указать предпочтительную модель спектра мелкомасштабных неоднородностей электронной концентрации для среднеширотной ионосферы в условиях естественных геофизических возмущений, обусловленных явлениями повышенной солнечной активности.

3. РЕЗУЛЬТАТЫ ИЗМЕРЕНИЙ

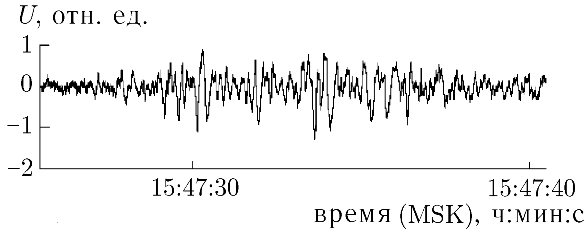

Рис. 1

В начале эксперимента для отработки методики измерений проводилась запись сигналов пролётных ИСЗ в спокойных геофизических условиях. В этих сеансах, как и ожидалось, флуктуации амплитуды сигналов ИСЗ на частотах 150 и 400 МГц практически отсутствовали из-за низкого уровня плазменной турбулентности среднеширотной ионосферы.

В дальнейшем наблюдения за сигналами орбитальных ИСЗ были переведены в режим сопровождения явлений повышенной солнечной активности. Наиболее информативные записи сигналов были получены в период с 31 октября по 4 ноября 2003 года, когда наблюдалась резко выраженная солнечная активность (мощные солнечные всплески), а индекс $K_{\text{р}}$ магнитного поля Земли изменялся от 3 до 8 в отдельные периоды этого временнóго интервала (информация получена с сайта [9]).

В это время регистрировались довольно интенсивные амплитудные флуктуации принимаемых сигналов от передатчиков орбитальных ИСЗ на частотах 150 и 400 МГц. Амплитуда флуктуаций варьировалась для отдельных временны́х интервалов наблюдения [1], но наклоны нормированных спектров амплитудных флуктуаций принимаемого излучения на отдельных временны́х интервалах практически совпадали. Следует заметить, что такая ситуация была характерна для случаев, когда находившиеся в зоне радиовидимости орбитальные ИСЗ были заметно удалены от плоскости магнитного меридиана.

Наибольший интерес в наших экспериментальных исследованиях представляют записи сигналов орбитальных ИСЗ, когда они находятся вблизи плоскости магнитного меридиана при больших углах места. В этих условиях (для геометрии нашего эксперимента при азимуте ИСЗ $\varphi_{\text{м}} \simeq 188^\circ$ и угле места $\vartheta_{\text{м}} \simeq 71^\circ$) луч зрения с орбитального ИСЗ на наземный пункт приёма практически совпадает с направлением магнитного поля Земли $\mathbf{H}_0$ (угол $\vartheta \approx 0$) и, исходя из обобщённой модели анизотропной мелкомасштабной турбулентности, в это время можно ожидать всплеска амплитудных флуктуаций и повышения их частоты (ср. [4]).

В наших наблюдениях за пролётными ИСЗ в условиях повышенной геофизической возмущённости в период с 31 октября по 4 ноября 2003 года такая ситуация создалась лишь в одном сеансе 31 октября в 15:40 MSK, когда на частоте 150 МГц принималось излучение от пролётного ИСЗ, находившегося вблизи $\varphi_{\text{м}} \approx 188^\circ$ и $\vartheta_{\text{м}} \approx 71^\circ$ (магнитный индекс $K_{\text{р}}$ в это время был равен 6).

На рис. 1 приведён образец записи сигнала ИСЗ в этом сеансе на частоте 150 МГц. Отчётливо виден всплеск амплитудных флуктуаций сигнала на временнóм интервале длительностью около 10 с. Спектры амплитудных флуктуаций принимаемого излучения для этого временнóго интервала (кривая *1*) и второго десятисекундного интервала наблюдений, спустя 1 минуту после первого (кривая *2*), приведены на рис. 2. Из рис. 2 видно, что наклон амплитудного спектра заметно изменяется в зависимости от ориентации ϑ луча зрения с ИСЗ на наземный пункт приёма относительно направления магнитного поля Земли. Для первого временнóго интервала (15:47:27÷15:47:37 MSK) расчётный угол между радиолучом и геомагнитным полем с учётом погрешности орбитальных параметров составлял $\vartheta \approx 0\div1^\circ$, а наклон спектра составлял $p_1 \approx 3$ в

[1] Это явление, по-видимому, было связано с повышенной интенсивностью мелкомасштабных неоднородностей электронной концентрации верхней ионосферы, которая в обычных (невозмущённых) условиях носит резко выраженный облачный характер: мелкомасштабные неоднородности верхней ионосферы с повышенным уровнем относительных флуктуаций электронной концентрации группируются в виде отдельных изолированных облаков протяжённостью в несколько десятков километров [10].

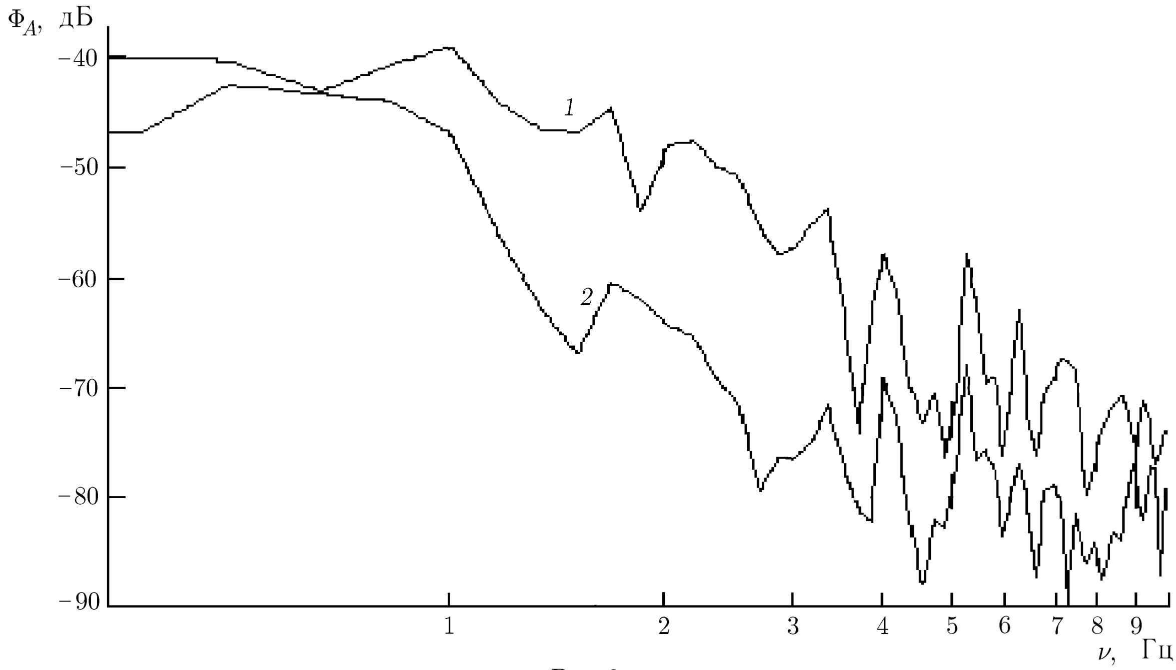

Рис. 2

области спектральных частот сигнала $\nu \approx 1{\div}4$ Гц. Для второго интервала $\vartheta \simeq 30^\circ$ и наклон спектра был заметно больше: $p_2 \approx 5{,}6$ (в области $\nu \approx 1{\div}3$ Гц) [2].

В рамках обобщённой модели анизотропной мелкомасштабной турбулентности (см. (5)) $p_1 \approx \approx p-1$, а $p_2 \approx \alpha p + p_1$. Соответственно, в нашем случае ($p_1 \approx 3$, $p_2 \approx 5{,}6$) находим, что показатель спектра ионосферной турбулентности был равен $p \approx 4$, а показатель анизотропии спектра составлял $\alpha \approx 0{,}65$.

Отличие в показателях спектров для разных временны́х интервалов (вблизи и вдали от луча зрения с ИСЗ относительно направления магнитного поля Земли) составляло $p_2 - p_1 \approx 2{,}6$.

Этот результат не противоречит предложенной концепции обобщённой анизотропной модели мелкомасштабной турбулентности среднеширотной ионосферы (см. (3) и (5)), но не соответствует модели резко анизотропной ИИТ (см. (2) и (6)), для которой разность измеренных показателей спектров должна быть равна 1.

По результатам нашего эксперимента можно оценить нижнюю границу характерного масштаба l_χ неоднородностей в спектре ионосферной турбулентности, разделяющей мелкомасштабную и крупномасштабную фракции в модели анизотропной турбулентности верхней ионосферы (см. (3)). Эта величина представляет собой френелевский масштаб $l_{\text{ф}}$, который в наших измерениях был равен $l_{\text{ф}} \approx v_l/\nu_{\text{ф}} \approx 2{,}5$ км, где $v_l \approx 2{,}5$ км/с — скорость луча зрения на орбитальный ИСЗ на высоте слоя F_2 ионосферы [1], $\nu_{\text{ф}}$ — френелевская спектральная частота осцилляций сигнала, $\nu_{\text{ф}} \approx 1$ Гц, см. рис. 2). Величина $l_\chi \approx l_{\text{ф}} \approx 2{,}5$ км находится в хорошем соответствии с ожидаемой величиной $l_\chi \approx 1{\div}2$ км для среднеширотной ионосферы [5].

Таким образом, обнаруженная в нашем эксперименте зависимость наклона спектра амплитудных флуктуаций принимаемого излучения от ориентации луча зрения с ИСЗ на наземный пункт приёма относительно направления магнитного поля Земли может свидетельствовать в пользу обобщённой модели спектра мелкомасштабной ионосферной турбулентности. Эта модель учитывает зависимость анизотропии (вытянутости) мелкомасштабных неоднородностей верхней

[2] Изменения в наклоне спектров на частоте $\nu \geq 4$ Гц (для кривой *1* на рис. 2) и $\nu \geq 3$ Гц (для кривой *2* на рис. 2) обусловлены влиянием шумов.

ионосферы вдоль направления магнитного поля Земли от поперечных размеров этих неоднородностей.

Для получения достоверной информации о параметрах предполагаемой (теоретической) обобщённой модели спектра мелкомасштабной турбулентности среднеширотной ионосферы необходимы, конечно, дальнейшие эксперименты по наземному приёму сигналов орбитальных ИСЗ метрового и дециметрового диапазонов длин волн в возмущённых геофизических условиях среднеширотной ионосферы с использованием метода Ерухимова.

Работа выполнена в рамках проекта РФФИ (грант № 03–02–17303).

СПИСОК ЛИТЕРАТУРЫ

1. Гершман Б. Н., Ерухимов Л. М., Яшин Ю. Я. Волновые явления в ионосфере и космической плазме. М.: Наука, 1984.
2. Rino C. L., Fremow E. J. // J. Atmos. Terr. Phys. 1977. V. 39, No. 8. P. 859.
3. Ерухимов Л. М., Метелёв С. А., Митякова Э. Е. и др. // Тепловые нелинейные явления в плазме. Горький, 1979. С. 7.
4. Erukhimov L. M., Kosolapenko V. I., Lerner A. M., et al. // Planet. Space Sci. 1981. V. 29, No. 9. P. 93.
5. Алимов В. А., Ерухимов Л. М., Мясников Е. Н. и др. // Изв. вузов. Радиофизика. 1997. Т. 40, № 4. С. 446.
6. Выборнов Ф. И., Ерухимов Л. М., Комраков Г. П. и др. // Изв. вузов. Радиофизика. 1986. Т. 29, № 4. С. 491.
7. http://www.celestrak.com.
8. http://www.sat-net.com/winorbit.
9. http://www.sel.noaa.gov.
10. Алимов В. А., Ерухимов Л. М., Рахлин А. В. и др. // Ионосферные исследования. М: Сов. радио, 1980. № 30. С. 102.

Поступила в редакцию 16 мая 2004 г.; принята в печать 28 октября 2004 г.

ANISOTROPIC STRUCTURE OF SMALL-SCALE IONOSPHERIC TURBULENCE

V. A. Alimov, F. I. Vybornov, and A. V. Rakhlin

We consider various theoretical models of the small-scale ionospheric turbulence spectrum. The particular role of the generalized model of the ionospheric-turbulence spectrum, which takes into account the dependence of the anisotropy (extension) of small-scale irregularities of the upper ionosphere along the Earth's magnetic field lines on the transverse scale of those irregularities, is pointed out. The results of the first specialized experiments on radio-wave sounding of the midlatitude ionosphere by signals from orbital satellites at frequencies 150 and 400 MHz under conditions of increased solar activity are presented. The experiments were performed at the radiophysical facility in the Nizhny Novgorod region in 2003. We studied statistical characteristics of the amplitude fluctuations of the received signals for different angles ϑ between the geomagnetic-field direction and the line of sight from a satellite to a ground-based reception point. The experiments revealed that the slope of the spectrum of amplitude fluctuations of the received radiation is a function of the angle ϑ. The obtained result agrees with the generalized model of ionospheric turbulence spectrum and can be an argument in favor of the pronounced anisotropic structure of small-scale electron-density irregularities of the midlatitude ionosphere under disturbed geophysical conditions.

УДК 621.372

УСЛОВИЯ ОРТОГОНАЛЬНОСТИ ВЫТЕКАЮЩИХ МОД

А. Б. Маненков

Институт физических проблем им. П. Л. Капицы РАН, г. Москва, Россия

Рассматриваются различные подходы к нормировке вытекающих мод открытых волноводов. Исследованы общие свойства таких мод и их связь с модами непрерывного спектра. Обсуждаются вопросы применимости теории возмущений для расчёта смещения констант распространения мод при малых изменениях структуры волновода. Общие соотношения проиллюстрированы на примере волновода в виде канала в диэлектрике.

ВВЕДЕНИЕ

Среди большого разнообразия открытых волноводов существует обширный класс систем, которые могут эффективно направлять вытекающие (квазисобственные) моды [1–8]. В качестве одного из простых примеров укажем волновод, представляющий собой вакуумный (или воздушный) канал в диэлектрике [4, 8]. Слабовытекающими являются также моды в оптических волокнах, у которых показатель преломления внешней среды (покрытия) больше или равен показателю преломления сердцевины [8].

Хотя эти моды известны достаточно давно [1, 2, 8], их анализу уделяли недостаточное внимание в основном из-за того, что основную роль на практике обычно отводили направляемым (поверхностным) модам. В последние годы интерес к структурам, в которых могут эффективно распространяться вытекающие моды, резко повысился в связи с исследованиями новых типов волноводов, таких, как многослойные (брэгговские) и микроструктурированные волокна, а также фотонные кристаллы [5–7]. Оказалось, что волноводы такого типа весьма перспективны для различных приложений, поскольку они обладают достаточно уникальными дисперсионными характеристиками и, кроме того, могут работать при больших уровнях мощности и в широком диапазоне частот. Следует также отметить, что вытекающие моды (и подобные им) достаточно давно используют в различных электронных приборах, однако при их анализе, как правило, применяют упрощённые подходы. В частности, расчёты таких систем часто основывают на теории закрытых волноводов, а влияние радиационных полей учитывают, вводя, например, эффективный коэффициент затухания. Такое приближение не является достаточно точным, особенно для приборов, работающих в режимах больших мощностей.

Вытекающие моды (ВМ) не являются собственными модами открытых волноводов [1, 2, 9, 10], тем не менее при определённых условиях в достаточно большой области пространства их поля могут доминировать. Эти моды можно выделить из разложений по непрерывному спектру (по радиационным модам), используя аналитические свойства спектральных представлений. Свойства ВМ существенно отличаются от свойств направляемых мод (НМ); в частности, поля ВМ экспоненциально возрастают при удалении от оси волновода (в любой поперечной плоскости). Это свойство не позволяет применить многие «стандартные» формулы теории НМ для расчёта характеристик ВМ; например, нельзя использовать обычное выражение для нормы. В работах [9, 10] был предложен способ нормировки полей ВМ, который основан на переходе к комплексным координатам. Следует, однако, заметить, что при решении целого ряда вопросов такой подход не всегда удобен. В частности, при использовании указанного способа возникает ряд трудностей при выводе различных формул методами теории возмущений, например, при расчёте сдвига фазовых скоростей при малом изменении характеристик волновода.

В настоящей работе рассмотрен альтернативный способ нормировки, который применúм к модам различного типа; он позволяет существенно упростить вывод многих соотношений теории открытых волноводов, а также ввести единообразные формулы для определения условий ортогональности мод различного вида. На основе полученных формул выведено несколько частных выражений для нормы вытекающих мод, которые могут быть использованы в сочетании с различными численными методами.

1. ВЫДЕЛЕНИЕ ВЫТЕКАЮЩИХ МОД

Свойства ВМ тесно связаны со свойствами радиационных мод (РМ), поэтому, прежде чем рассматривать вопросы нормировки, остановимся очень коротко на свойствах последних. В первых разделах будем рассматривать двумерную структуру — плоский диэлектрический волновод, который представляет собой диэлектрическую пластину (слой) в общем случае с переменным профилем показателя преломления $n(y)$ при $|y| < d$, где $2d$ — толщина пластины. Волновод лежит на подложке с показателем преломления n_1, показатель преломления верхнего покрытия обозначим через n_2. Для простоты предполагаем, что диэлектрические потери отсутствуют, т. е. $\mathrm{Im}\, n = 0$. В основном будем рассматривать симметричную систему, для которой $n(y)$ — чётная функция y и $n_1 = n_2$. Возможные обобщения на несимметричный случай и на более сложные системы проводятся по сходной схеме и будут описаны ниже без детального обоснования. Будем анализировать TE-задачу, когда электрическое поле поляризовано вдоль оси x (рис. 1). Как обычно, временной множитель $\exp(-i\omega t)$ опущен во всех формулах; здесь $\omega = kc$ — частота, k — волновое число, c — скорость света в пустоте. Согласно принципу предельного поглощения будем считать число k комплексным с малой положительной мнимой частью, которую в окончательных формулах будем полагать равной нулю, т. е. считаем $\mathrm{Im}\, k \to +0$.

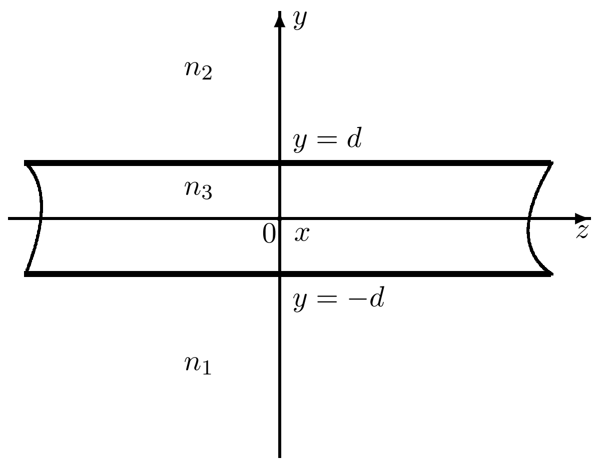

Рис. 1. Планарный волновод

Рассмотрим структуру радиационных мод, из множества которых затем выделим поля ВМ. Волновое уравнение имеет частные решения в виде волн [9], поле которых $E_x = \Psi(y) \exp(i\beta_\kappa z)$, где функция $\Psi(y)$ при $|y| > d$ определяется соотношениями

$$\Psi(y) = \begin{cases} u_\kappa^{(1)} \exp[i\kappa_1 (y+d)] + v_\kappa^{(1)} \exp[-i\kappa_1 (y+d)], & y < -d; \\ u_\kappa^{(2)} \exp[-i\kappa_2 (y-d)] + v_\kappa^{(2)} \exp[i\kappa_2 (y-d)], & y > d. \end{cases} \tag{1}$$

Здесь $\beta_\kappa = \sqrt{k^2 n_1^2 - \kappa_1^2}$ — коэффициенты распространения мод, κ_1 и κ_2 — поперечные волновые числа в подложке и покрытии ($\kappa_2 = \sqrt{k^2\,(n_2^2 - n_1^2) + \kappa_1^2}$). Поскольку κ_2 выражается через κ_1, то в дальнейшем будем считать, что все величины являются функциями $\kappa = \kappa_1$. Заметим, что для рассматриваемой системы при условии $n_1 \geq n_2$ поперечные волновые числа собственных мод непрерывного спектра κ_1 в подложке вещественны и неотрицательны ($\mathrm{Im}\, \kappa = 0$). В формуле (1) амплитудные функции $u_\kappa^{(1)}$, $v_\kappa^{(1)}$, $u_\kappa^{(2)}$, $v_\kappa^{(2)}$ не зависят от координат. При $|y| > d$ функция $\Psi(y)$ описывает четыре плоские волны. Амплитуды падающих на диэлектрический слой двух плоских волн $u_\kappa^{(1)}$, $u_\kappa^{(2)}$ можно взять произвольными, тогда амплитуды отражённых и прошедших волн $v_\kappa^{(1)}$, $v_\kappa^{(2)}$ определяются из условий непрерывности внешних и внутренних полей на границах $y = \pm d$. Связь указанных функций формально можно записать в виде операторного равенства [9, 10]

$$\hat{\mathbf{S}}\,(u_\kappa^{(1)}, u_\kappa^{(2)})^{\mathrm{T}} = (v_\kappa^{(1)}, v_\kappa^{(2)})^{\mathrm{T}}, \tag{2}$$

где T — символ операции транспонирования, $\hat{\mathbf{S}}$ — линейный оператор рассеяния плоских волн (S-матрица), который переводит амплитуды падающих волн в амплитуды прошедших и отражённых волн. Для РМ эти амплитуды должны быть собственными векторами (матричными столбцами) S-матрицы, т. е. должно быть выполнено соотношение

$$\hat{\mathbf{S}}\,(u_{m\kappa}^{(1)}, u_{m\kappa}^{(2)})^{\mathrm{T}} = s_m(\kappa)\,(u_{m\kappa}^{(1)}, u_{m\kappa}^{(2)})^{\mathrm{T}}, \tag{3}$$

где $s_m(\kappa)$ — собственное значение, $m = 1, 2$ — номер ветви РМ. С учётом приведённых выше формул поля́ собственных РМ можно записать в виде $E_x = U_{m\kappa}(y)\exp(i\beta_\kappa z)$, где при $|y| > d$ функция $U_{m\kappa}(y)$ определяется выражением

$$U_{m\kappa}(y) = \begin{cases} u_{m\kappa}^{(1)}\,\{\exp[-i\kappa_1\,(y-d)] + \exp[i\kappa_1\,(y-d)]/s_m(\kappa)\}, & y < -d; \\ u_{m\kappa}^{(2)}\,\{\exp[i\kappa_2\,(y+d)] + \exp[-i\kappa_2\,(y+d)]/s_m(\kappa)\}, & y > d. \end{cases} \tag{4}$$

Для рассматриваемой задачи элементы S-матрицы $[S_{ij}]$, где индексы i и j принимают значения 1 либо 2, можно определить из решения двух вспомогательных задач. В первой задаче рассматривается отражение и прохождение плоской волны с осевой компонентой волнового числа β_κ, падающей на средний слой снизу. В этом случае определяются элементы S_{11} и S_{21}. Два других элемента вычисляются из второй вспомогательной задачи, когда такая же волна падает на слой сверху. Применяя стандартную методику, из (3) получаем квадратное уравнение для собственных значений $s_m(\kappa)$, корни которого для двух ветвей определяются соотношением

$$2s_m(\kappa) = S_{11} + S_{22} \mp \sqrt{(S_{11} - S_{22})^2 + S_{12}S_{21}}\,, \tag{5}$$

где $m = 1, 2$. Таким образом, S-матрица разделила вырожденное множество РМ на две ветви, причём моды этих ветвей, как следует из свойств S-матрицы, являются ортогональными [9, 10].

Для волновода с постоянным профилем показателя преломления все выражения могут быть получены в аналитической форме. В качестве примера рассмотрим двумерный канал в диэлектрике, когда $n_3 < n_1$, где n_3 — постоянный показатель преломления внутренней среды (среднего слоя) при $|y| < d$. Для чётной ветви имеем

$$U_{2\kappa}(x) = \begin{cases} A_{2\kappa}\cos(g_\kappa y), & |y| < d; \\ u_{2\kappa}^{(1)}\,\{\exp[i\kappa\,(|y|-d)] + \exp[-i\kappa\,(|y|-d)]/s_2(\kappa)\}, & |y| > d, \end{cases} \tag{6}$$

$$s_2(\kappa) = \frac{\kappa\cos(g_\kappa d) + ig_\kappa\sin(g_\kappa d)}{\kappa\cos(g_\kappa d) - ig_\kappa\sin(g_\kappa d)}\,, \qquad A_{2\kappa} = \frac{[1 + s_2(\kappa)]u_{2\kappa}^{(1)}}{s_2(\kappa)\cos(g_\kappa d)}\,, \tag{7}$$

где $g_\kappa = \sqrt{k^2\,(n_3^2 - n_1^2) + \kappa^2}$. Функция g_κ имеет точку ветвления κ_{c}, где $\kappa_{\mathrm{c}} = k\sqrt{n_1^2 - n_3^2}$. При $\kappa_1 < \kappa_{\mathrm{c}}$ в формуле (6) удобно заменить тригонометрические функции на гиперболические. Как видно из этих уравнений, собственное значение $s_m(\kappa)$ входит в выражения для амплитуд РМ и, как будет показано ниже, в условие ортогональности мод.

Подобным образом рассматривается случай, когда в среднем слое $n_3 > n_1$. При этом основное отличие от примера, исследованного выше, заключается в том, что для такой системы в спектре мод появляются направляемые моды (моды дискретного спектра). Полюсы $s_m(\kappa)$ являются поперечными волновыми числами НМ открытого волновода, т. е. дисперсионное уравнение может быть записано в виде $1/s_m(\kappa) = 0$. При отсутствии диэлектрических потерь эти волновые числа κ_i (здесь i — номер НМ) лежат на мнимой оси комплексной плоскости κ. Кроме того, изменяется положение точки ветвления функции g_κ; при отсутствии потерь она расположена на мнимой оси волнового числа κ. Примерный вид комплексной плоскости κ показан на рис. 2, где кружками отмечены значения поперечных волновых чисел κ_i направляемых мод, а пунктирными линиями показаны разрезы функций β_κ и g_κ. В трёхмерном случае (см. ниже) к этим разрезам добавляется разрез, выходящий из точки $\kappa = 0$. Спектр мод становится более сложным для несимметричного волновода [11], когда, например, $n_1 > n_2$. В этом случае множество РМ разбивается на три ветви. Для такой структуры волновые числа κ_1 остаются вещественными, тогда как κ_2 могут быть комплексными. Для первой ветви, у которой $0 < \kappa < \kappa_0$, где $\kappa_0 = k\sqrt{n_1^2 - n_2^2}$, поля РМ экспоненциально убывают при $y > d$. При $\kappa > \kappa_0$ поля́ двух других ветвей РМ имеют ту же структуру, что и рассмотренные выше.

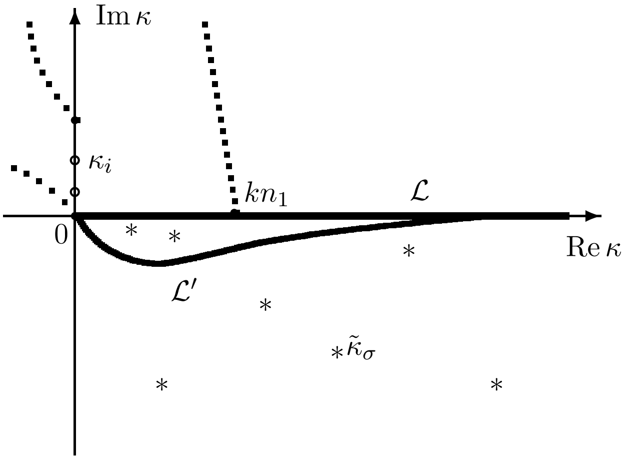

Рис. 2. Комплексная плоскость поперечного волнового числа κ. Пунктирные линии — разрезы функций, сплошные кривые — контуры интегрирования

При возбуждении волновода сторонним током $\mathbf{j} = j_x\mathbf{e}_x$, где $\mathbf{e}_x$ — орт оси x, поле E_x можно разложить по собственным модам волновода в виде суммы полей НМ и интегралов по модам непрерывного спектра. Правее объёма, занятого источниками, для несимметричного волновода имеем [11]

$$E_x = \sum_i C_i U_i(y)\exp(i\beta_i z) + \int\limits_0^{\kappa_0} C_{0\kappa}U_{0\kappa}(y)\exp(i\beta_\kappa z)\,\mathrm{d}\kappa + \sum_{m=1}^{2}\int\limits_{\kappa_0}^{\infty} C_{m\kappa}U_{m\kappa}(y)\exp(i\beta_\kappa z)\,\mathrm{d}\kappa, \qquad (8)$$

где U_i и $U_{m\kappa}$ — распределения полей НМ и трёх ветвей РМ ($m = 0, 1, 2$), β_i и β_κ — коэффициенты распространения, C_i и $C_{m\kappa}$ — амплитуды. Пределы интегрирования в (8) зависят от показателей преломления волновода в первом и третьем слоях. Заметим, что для симметричной структуры, у которой $n_1 = n_2$, имеем $\kappa_0 = 0$, и в формуле (8) второе слагаемое отсутствует. С учётом условий излучения [1, 10] в выражении для β_κ выбираем ту ветвь корня, у которой $\mathrm{Im}\,\beta_\kappa > 0$ при $\mathrm{Im}\,k > 0$. Амплитудные коэффициенты равны [9–11]

$$C_i = \frac{k}{\zeta_\mathrm{v}\beta_i N_i}\int j_x U_i \exp(-i\beta_i z)\,\mathrm{d}y\,\mathrm{d}z, \qquad C_{m\kappa} = \frac{k}{\zeta_\mathrm{v}\beta_\kappa D_m(\kappa)}\int j_x U_{m\kappa}\exp(-i\beta_\kappa z)\,\mathrm{d}y\,\mathrm{d}z, \qquad (9)$$

где $m = 0, 1, 2$; ζ_v — импеданс вакуума, N_i и $D_m(\kappa)$ — норма направляемой моды и нормирующий множитель радиационной моды соответственно. В этих формулах интегрирование проводится по области, занятой током. Амплитуды мод определяются с помощью условий ортогональности, которые имеют вид

$$\langle U_i, U_{i'} \rangle = N_i \delta_{ii'}, \qquad \langle U_i,\, U_{m\kappa} \rangle = 0, \tag{10}$$

$$\langle U_{m\kappa}, U_{m'\kappa'} \rangle = D_m(\kappa)\delta_{mm'}\delta(\kappa - \kappa'), \tag{11}$$

где индексы m и m' принимают значения 0, 1 и 2, $\delta_{mm'}$ — символ Кронекера. Здесь и ниже используем обозначение для интеграла от произведения двух произвольных функций $F_1(y)$ и $F_2(y)$:

$$\langle F_1, F_2 \rangle = \int\limits_{-\infty}^{+\infty} F_1(y) F_2(y)\, \mathrm{d}y. \tag{12}$$

В формулах (9) множители $k/(\zeta_{\mathrm{v}}\beta_i)$ и $k/(\zeta_{\mathrm{v}}\beta_\kappa)$ возникли из-за того, что при их выводе была использована скалярная нормировка полей. Для РМ нормирующий множитель можно выразить через собственные значения и собственные векторы S-матрицы:

$$D_m(\kappa) = 2\pi \left\{\kappa_1\, [u_{m\kappa}^{(1)}]^2 + \kappa_2\, [u_{m\kappa}^{(2)}]^2\right\} / [\kappa_1 s_m(\kappa)], \tag{13}$$

где по определению $\kappa = \kappa_1$. Из (9) и (13) следует, что в комплексной плоскости κ амплитуды $C_{m\kappa}$ могут иметь полюсы в нулях функций $1/s_m(\kappa)$.

Чтобы выделить поля нескольких вытекающих мод из разложения (8), продолжим аналитически все подынтегральные функции в комплексную плоскость параметра κ. Для простоты будем опять рассматривать симметричную геометрию. В этом случае $\kappa_0 = 0$, и интегрирование по поперечному волновому числу κ проводится вдоль луча $(0, +\infty)$, который идёт вдоль действительной оси (линия $\mathcal{L}$ на рис. 2). Как следует из условий излучения [9, 10, 12], точка ветвления $\kappa = kn_1$ должна обходиться снизу.

Сместим исходный контур $\mathcal{L}$ в комплексной плоскости κ вниз [9, 10]. На рис. 2 новый контур, который, как и исходный, проходится слева направо, обозначен символом $\mathcal{L}'$. Обозначим также через $\mathcal{M}$ область, ограниченную линиями $\mathcal{L}$ и $\mathcal{L}'$. Будем предполагать, что при $\mathrm{Re}\,\kappa > |kn_1|$ контур $\mathcal{L}'$ идёт вдоль вещественной оси (т. е. смещена только конечная, левая часть $\mathcal{L}$), поэтому область $\mathcal{M}$ конечна. Амплитуды $C_{m\kappa}$ аналитичны в области $\mathrm{Re}\,\kappa > 0$ и $\mathrm{Im}\,\kappa < 0$ за исключением в общем случае счётного числа полюсов [12, 13], расположенных в точках $\tilde{\kappa}_\sigma$, которые соответствуют поперечным волновым числам ВМ (здесь σ — номер моды). В дальнейшем все величины, относящиеся к ВМ, будем отмечать тильдой. На рис. 2 полюсы амплитуд $C_{m\kappa}$ отмечены звёздочками. Контур $\mathcal{L}$ смещается так, чтобы «нужные» полюсы были расположены внутри $\mathcal{M}$ (см. ниже). Для того, чтобы связать значения интегралов по непрерывному спектру вдоль линий $\mathcal{L}$ и $\mathcal{L}'$, используем теорему о вычетах [14]. Нетрудно показать, что разность интегралов по новому и старому контурам пропорциональна сумме вычетов подынтегральных функций в полюсах $\tilde{\kappa}_\sigma$, которые были пересечены при деформации $\mathcal{L}$ и оказались внутри области $\mathcal{M}$. Знаки у вычетов находятся стандартным способом (см., например, [14], стр. 80) с учётом ориентации контуров $\mathcal{L}$ и $\mathcal{L}'$. В результате этих преобразований получим новое представление для поля, которое включает поля нескольких вытекающих мод [9, 10]:

$$E_x = \sum_i C_i U_i(y) \exp(i\beta_i z) + \sum_\sigma \tilde{C}_\sigma \tilde{U}_\sigma(y) \exp(i\tilde{\beta}_\sigma z) + \sum_{m=1}^{2} \int\limits_{\mathcal{L}'} C_{m\kappa} U_{m\kappa}(y) \exp(i\beta_\kappa z)\, \mathrm{d}\kappa. \tag{14}$$

При выводе (14) для простоты считалось, что выделенные моды не вырождены и полюсы первого порядка.

В приведённой выше формуле амплитуды ВМ пропорциональны вычетам подынтегральных функций:

$$\tilde{C}_\sigma = \frac{k}{\zeta_{\rm v}\tilde{\beta}_\kappa \tilde{N}_\sigma} \int j_x \tilde{U}_\sigma \exp(-i\tilde{\beta}_\sigma z)\,{\rm d}y\,{\rm d}z, \tag{15}$$

где величину

$$\tilde{N}_\sigma = \frac{i}{2\pi}\,\frac{{\rm d}D_m(\kappa)}{{\rm d}\kappa}\bigg|_{\kappa=\tilde{\kappa}_\sigma} \tag{16}$$

называют нормой ВМ. Поля симметричных (чётных) ВМ для системы, описанной выше, имеют вид

$$\tilde{U}_\sigma(y) = \begin{cases} \tilde{A}_\sigma \cos(\tilde{g}_\sigma y), & |y| < d; \\ \tilde{u}_\sigma^{(1)} \exp[i\tilde{\kappa}_\sigma\,(|y| - d)], & |y| > d. \end{cases} \tag{17}$$

Заметим, что вне волновода поля́ ВМ представляют собой две уходящие от оси z плоские волны. Волновые числа $\tilde{\kappa}_\sigma$ удовлетворяют тому же дисперсионному уравнению $1/s_m(\kappa) = 0$, что и числа НМ, но лежат в другой области. При отсутствии диэлектрических потерь вблизи значений корней слабовытекающих мод имеем $s_m(\kappa) \propto (\kappa - \kappa_\sigma^*)/(\kappa - \kappa_\sigma)$ [12, 13], где звёздочка обозначает операцию комплексного сопряжения. Поля ВМ растут при $y \to \pm\infty$ [10, 15], поскольку $\operatorname{Im}\tilde{\kappa}_\sigma < 0$. Для несимметричной геометрии поля́ растут, по крайней мере, в одном из направлений оси y. Отметим, что, как правило, целесообразно выделять из интегралов в (8) только ВМ с малым затуханием [9, 15], у которых полюсы лежат вблизи отрезка $[0, kn_1]$, т. е. при условии $\operatorname{Im}\tilde{\kappa}_\sigma \ll kn_1$ и $\operatorname{Im}\tilde{\beta}_\sigma \ll \operatorname{Re}\tilde{\beta}_\sigma$. Такое выделение имеет смысл лишь в ограниченной области пространства. В частности, вдали от источников амплитуды полей ВМ будут существенно меньше амплитуд полей пространственной волны [15]. Отметим также, что в интегралах по РМ подынтегральные функции при $\operatorname{Im}\kappa \to -\infty$ обычно не убывают, так что свести интегралы (8) только к сумме вычетов в общем случае не удаётся.

Выражение (16) для $\tilde{N}_\sigma$ трудно использовать для вычисления нормы ВМ, особенно в случае, когда расчёты проводятся численно. В то же время выражение, подобное первой формуле в (10), к ним неприменимо из-за роста полей в поперечной плоскости. Можно модифицировать (10), вводя комплексную координату y вдали от оси волновода. В этом случае формула (16) может быть преобразована к виду [10]

$$\int\limits_{-\infty\exp(i\gamma)}^{+\infty\exp(i\gamma)} \tilde{U}_\sigma \tilde{U}_{\sigma'}\,{\rm d}y = \tilde{N}_\sigma \delta_{\sigma\sigma'}. \tag{18}$$

Символ $y = +\infty\exp(i\gamma)$ в верхнем пределе интеграла означает, что сначала интегрирование по y ведётся вдоль вещественной оси, а затем, когда достигается большое вещественное значение координаты $y_1 \gg d$, интегрирование проводится по лучу $y = y_1 + t\exp(i\gamma)$, где параметр $t \geq 0$. Угол γ выбирается так, что $\operatorname{Im}[i\kappa_\sigma \exp(i\gamma)] < 0$ и $\operatorname{Im}[i\kappa_{\sigma'} \exp(i\gamma)] < 0$. Аналогичный смысл имеет символ $y = -\infty\exp(i\gamma)$ в нижнем пределе. Переход к комплексным координатам является регуляризацией расходящихся интегралов, содержащих поля ВМ; он подобен регуляризации расходящихся интегралов Фурье.

Покажем тождественность исходной формулы (16) и выражения (18). Для простоты рассмотрим случай чётных ВМ симметричного диэлектрического волновода. Применим формулу Грина к двум собственным функциям $U_{m\kappa}(y)$ и $U_{m\kappa'}(y)$ (см. выше) на отрезке $[-y_{\rm b}, y_{\rm b}]$, у которых параметры κ и κ' стремятся к κ_σ. После несложных преобразований [1, 10] получим следующее

соотношение:
$$\int\limits_{-y_{\mathrm{b}}}^{y_{\mathrm{b}}} \tilde{U}_\sigma^2 \,\mathrm{d}y = 2i\tilde{u}_\sigma^2 \left[\frac{\partial}{\partial\kappa} \left(\frac{1}{s_m(\kappa)} \right) \bigg|_{\kappa_\sigma} - \frac{\exp(2i\kappa_\sigma y_{\mathrm{b}})}{2\kappa_\sigma} \right]. \tag{19}$$

При выводе этой формулы опущены слагаемые, которые при $\kappa = \kappa' = \kappa_\sigma$ обращаются в нуль. Формула (19) верна не только для вещественных, но и для комплексных y_{b} ($|y_{\mathrm{b}}| \gg d$), т. е. в случае, когда интегрирование по y проводится по контуру, лежащему в комплексной плоскости координаты y. В качестве такого контура возьмём ломаную линию, состоящую из трёх отрезков: из вещественного отрезка $[-y_{\mathrm{l}}, y_{\mathrm{l}}]$ и двух дополнительных отрезков $[-y_{\mathrm{b}}, -y_{\mathrm{l}}]$ и $[y_{\mathrm{l}}, y_{\mathrm{b}}]$, где y_{b} — комплексное число, у которого $\mathrm{Re}\, y_{\mathrm{b}} > 0$, $\mathrm{Im}\, y_{\mathrm{b}} > 0$. Дополнительные отрезки расположены на лучах, выходящих из точек $\pm y_{\mathrm{l}}$ под углами γ и $\gamma - \pi$ к действительной оси комплексной координаты y. Предполагаем, что $|y_{\mathrm{b}}| \gg y_{\mathrm{l}} \gg d$. Первое слагаемое в правой части формулы (19) стремится к нулю, если $y_{\mathrm{b}} \to +\infty \exp(i\gamma)$, когда угол γ подчиняется условиям, приведённым выше. Далее, учитывая связь между производными функций $D_m(\kappa)$ и $s_m(\kappa)$, которая следует из (13), получим требуемое тождество. Заметим, что общий случай рассматривается по такой же схеме [1, 10].

2. ОРТОГОНАЛЬНОСТЬ С ВЕСОМ

Выражение (18) для нормы ВМ существенно удобнее исходного соотношения (16). Например, для рассматриваемой двумерной задачи с постоянными диэлектрическими проницаемостями подложки и покрытия интегралы для полей вне среднего слоя легко вычислить, в результате получим
$$\tilde{N}_\sigma = \int\limits_{-d}^{d} \tilde{U}_\sigma \tilde{U}_{\sigma'} \,\mathrm{d}y + i\big\{[\tilde{u}_\sigma^{(1)}]^2/\tilde{\kappa}_{1\sigma} + [\tilde{u}_\sigma^{(2)}]^2/\tilde{\kappa}_{2\sigma}\big\}/2. \tag{20}$$

С помощью этой формулы легко рассчитать норму в тех случаях, когда поле ВМ находится численно, причём для её применения не надо знать поле вне диэлектрического волновода. Тем не менее и при использовании соотношения (18) возникает ряд проблем. Эти проблемы связаны, в первую очередь, с тем, что, применяя описанный подход, нельзя установить ортогональность ВМ и мод непрерывного спектра с волновыми числами κ, которые лежат на новом контуре $\mathcal{L}'$ (рис. 2), т. е. в разложении (14) в обычном смысле функции не ортогональны. Из-за этого, например, сложно применять к ВМ стандартный вариант метода возмущений (см. ниже). Недостатком теории следует признать и тот факт, что для разных типов мод условия ортогональности определяются разными формулами. В частности, формулы, подобные (18), неприменимы к РМ с волновыми числами κ, лежащими на смещённом контуре $\mathcal{L}'$, поскольку в представлении (4) при комплексных κ и y одно из слагаемых будет всегда экспоненциально большим.

Более общий способ регуляризации расходящихся интегралов основан на введении во все выражения быстро спадающей на бесконечности весовой функции [13, 16], которая «гасит» экспоненциальный рост подынтегральных функций. Переопределим интеграл (12) в соответствии с формулой
$$\langle F_1,\, F_2 \rangle_\alpha = \lim_{\alpha\to+0} \int\limits_{-\infty}^{+\infty} F_1(y) F_2(y) w_\alpha(y) \,\mathrm{d}y, \tag{21}$$

где вес $w_\alpha(y) = \exp(-\alpha y^2)$, $\alpha > 0$. Введённая здесь весовая функция быстро убывает при $|y| \gg d$ (быстрее, чем растут поля ВМ или РМ) и тем самым обеспечивает сходимость интегралов.

Метод ортогонализации, основанный на приведённом выше соотношении (21), предъявляет весьма слабые требования к поведению полей на бесконечности, поэтому он более универсален, чем методы, описанные выше. Для примера выведем условия ортогональности радиационных и вытекающих мод, причём будем исследовать общий случай, когда поперечные волновые числа РМ комплексны. Будем рассматривать представление для поля (14), в котором выделено несколько слагаемых, соответствующих полям ВМ, а интегралы по непрерывному спектру мод берутся вдоль смещённого в комплексную плоскость контура $\mathcal{L}'$. Пусть l — вещественный параметр ($l \geq 0$), а $\kappa = \eta(l)$ — комплексная функция, которая задаёт координаты точек на кривой $\mathcal{L}'$. С учётом этих определений условия ортогональности рассматриваемых мод могут быть записаны в виде

$$\langle \tilde{U}_\sigma, \tilde{U}_{\sigma'} \rangle_\alpha = \tilde{N}_\sigma \delta_{\sigma\sigma'}, \qquad \langle \tilde{U}_\sigma, U_{m\kappa} \rangle_\alpha = 0, \tag{22}$$

$$\langle U_{m\kappa}, U_{m'\kappa'} \rangle_\alpha = D_m(\kappa) \delta_{mm'} \delta(\kappa - \kappa'). \tag{23}$$

В формуле (23) комплексные числа κ и κ' лежат на кривой $\mathcal{L}'$. Входящую в формулу (23) дельта-функцию с комплексными значениями аргумента определяем следующим соотношением [16]:

$$\delta(\kappa - \kappa') = \frac{\mathrm{d}l}{\mathrm{d}\kappa}\, \delta(l - l'), \tag{24}$$

где $\kappa = \eta(l)$ и $\kappa' = \eta(l')$, а $\delta(l - l')$ — обычная дельта-функция вещественных аргументов. При выводе формулы (23) учтены соотношения

$$\int\limits_{-\infty}^{+\infty} \exp(-\alpha y^2) \cos[(\kappa - \kappa')\, y]\, \mathrm{d}y = \sqrt{\pi/\alpha}\, \exp[-(\kappa - \kappa')^2/(4\alpha)], \tag{25}$$

$$\lim_{\alpha \to +0} \sqrt{\pi/\alpha}\, \exp[-(\kappa - \kappa')^2/(4\alpha)] = 2\pi \delta(\kappa - \kappa'), \tag{26}$$

где κ и κ' — произвольные комплексные числа на контуре $\mathcal{L}'$. Формула (26) верна, если угол наклона касательной к кривой $\mathcal{L}'$ меньше $\pi/4$. Используя приведённые формулы, нетрудно доказать тождественность всех трёх приведённых выше выражений (16), (18) и (22) для нормы $\tilde{N}_\sigma$. Подчеркнём, что основное достоинство условий ортогональности (22) и (23) заключается в том, что они применимы ко всем типам мод, включая направляемые, вытекающие и радиационные, в том числе и в случае, когда их поперечные волновые числа комплексны, т. е. при использовании модифицированного разложения (14) для поля в волноводе. Это свойство позволяет упростить вывод многих соотношений теории диэлектрических волноводов [16]. Например, тождественность формул (16) и (22) для нормы ВМ следует из тождественности выражений для амплитуд, которые можно получить двумя способами: на основе теоремы о вычетах и из условий ортогональности мод в разложении (14).

3. ПРИМЕНЕНИЯ ФОРМУЛ

Используя выведенные выше формулы, рассмотрим несколько вопросов теории вытекающих мод. Прежде всего, выведем условия, при выполнении которых можно приближённо вычислить норму ВМ, ограничиваясь интегралами по конечной области сечения волновода. Подобное приближение часто используется для оценки нормы. Рассмотрим вначале конкретный пример — волновод в виде канала в диэлектрике при $n_1 = n_2 > n_3$. При условии $\tilde{V} = kd\sqrt{n_1^2 - n_3^2} \gg 1$ для первой вытекающей моды получим

$$\tilde{g}_\sigma d \approx \frac{\pi}{2} \left(1 - \frac{i}{\tilde{V}}\right), \qquad \tilde{\kappa}_\sigma d \approx \tilde{V} + \frac{\pi^2}{8\tilde{V}} \left(1 - \frac{2i}{\tilde{V}}\right). \tag{27}$$

В выражении для нормы (20) второе слагаемое соответствует интегралу по внешней области поперечного сечения волновода; этот интеграл обозначим через $I_{\rm r}$, а интеграл по внутренней области обозначим через $I_{\rm d}$. Из (20) и (27) получаем

$$|I_{\rm r}/I_{\rm d}| \approx \pi^2/(4\tilde{V}^3) \approx |\operatorname{Im}\tilde{\kappa}_\sigma/\operatorname{Re}\tilde{\kappa}_\sigma|. \qquad (28)$$

Таким образом, при расчёте нормы ВМ внешним интегралом можно пренебречь только в том случае, когда мнимая часть поперечного волнового числа мала, т. е., что достаточно очевидно, когда мода слабовытекающая. Формула (28) даёт количественную оценку условия слабого вытекания. Можно показать, что эта оценка верна для широкого класса волноводов с различными полуотражающими стенками, например для диэлектрических трубок [3, 4]. Заметим, что в формулу (28) входят значения поперечного, а не продольного волнового числа. Этот факт достаточно важен, т. к. отношение $|\operatorname{Im}\tilde{\beta}_\sigma/\operatorname{Re}\tilde{\beta}_\sigma|$ может быть мало даже в случае, когда $|\operatorname{Im}\tilde{\kappa}_\sigma/\operatorname{Re}\tilde{\kappa}_\sigma| \sim 1$. Например, для рассмотренного выше волноводного канала подобная ситуация реализуется при условии $kd \gg 1$; в этом случае имеем $|\operatorname{Im}\tilde{\beta}_\sigma/\operatorname{Re}\tilde{\beta}_\sigma| \sim |\operatorname{Im}\tilde{\kappa}_\sigma/[(kd)^2\operatorname{Im}\tilde{\kappa}_\sigma]| \ll 1$. Аналогичная ситуация возможна также в периодических волноводах с выпуклыми зеркалами.

Рассмотрим теперь кратко задачу о сдвиге постоянных распространения ВМ при небольшом изменении профиля показателя преломления волновода. Пусть $n^{(0)}(y)$ — невозмущённый профиль показателя преломления, а $n^{(1)}(y)$ — возмущённый, причём, как обычно, возмущение мало. Для НМ сдвиг постоянной распространения в первом приближении равен

$$[\beta_i^{(1)}]^2 - [\beta_i^{(0)}]^2 = k^2\,\langle\delta\varepsilon\,U_i^{(0)}, U_i^{(0)}\rangle/N_i^{(0)}, \qquad (29)$$

где $U_i^{(0)}$ и $N_i^{(0)}$ — поле и норма рассматриваемой НМ в невозмущённой системе, $\delta\varepsilon = [n^{(1)}]^2 - [n^{(0)}]^2$, а $\beta_i^{(0)}$ и $\beta_i^{(1)}$ — продольные волновые числа в невозмущённом и возмущённом волноводах соответственно, i — номер моды. Аналогичным образом, используя нормировку с весовой функцией $w_\alpha(y)$, для сдвига постоянной распространения ВМ в первом приближении теории возмущений получим

$$[\tilde{\beta}_\sigma^{(1)}]^2 - [\tilde{\beta}_\sigma^{(0)}]^2 = k^2\,\langle\delta\varepsilon\,\tilde{U}_\sigma^{(0)}, \tilde{U}_\sigma^{(0)}\rangle_\alpha/\tilde{N}_\sigma^{(0)}, \qquad (30)$$

где $\tilde{U}_\sigma^{(0)}$ и $\tilde{N}_\sigma^{(0)}$ — поле и норма ВМ в невозмущённом волноводе. Подобная формула была ранее выведена для квантовомеханических задач в [13]. В следующем приближении метода теории возмущений [12] в выражении для сдвига $\tilde{\beta}_\sigma$ дадут вклад поля́ РМ:

$$[\tilde{\beta}_\sigma^{(1)}]^2 - [\tilde{\beta}_\sigma^{(0)}]^2 = k^2\,\langle\delta\varepsilon\,\tilde{U}_\sigma^{(0)}, \tilde{U}_\sigma^{(0)}\rangle_\alpha/\tilde{N}_\sigma^{(0)} + \sum_{m=1}^{2}\int\limits_{\mathcal{L}'} \frac{Q_{\sigma m\kappa}^2\,\mathrm{d}\kappa}{[\tilde{\beta}_\sigma^{(0)}]^2 - \beta_\kappa^2}\,, \qquad (31)$$

$$Q_{\sigma m\kappa} = k^2\,\langle\delta\varepsilon\,\tilde{U}_\sigma^{(0)}, U_{m\kappa}^{(0)}\rangle_\alpha/[D_{m\kappa}^{(0)}\tilde{N}_\sigma^{(0)}]^{1/2}. \qquad (32)$$

При выводе соотношения (31) существенную роль играл тот факт, что в соответствии с формулой (22) поля ВМ ортогональны полям РМ, у которых волновые числа κ лежат на смещённом контуре $\mathcal{L}'$. Чтобы не усложнять формулы, предполагалось, что волновод симметричный ($\kappa_0 = 0$) и в нём не распространяются НМ. В общем случае вклад этих мод легко учесть стандартным способом. Для НМ также нетрудно учесть влияние РМ, вводя дополнительные слагаемые, содержащие интегралы по непрерывному спектру с естественной заменой $\tilde{U}_\sigma$ на U_i и т. д.

Интересно отметить, что область применимости формулы (30) для смещения волнового числа $\tilde{\beta}_\sigma$ оказывается иной, чем область применимости формулы (29) для сдвига постоянной распространения НМ. Формулы (29) и (30), а также следующие приближения, дают правильную оценку смещения значений β_i и $\tilde\beta_\sigma$, если это смещение мало, например при условии

$$k^2\,\langle|\delta\varepsilon\, U|, |U|\rangle \ll |N|, \tag{33}$$

где $U = U_i^{(0)}$ для формулы (29) и $U = \tilde U_\sigma^{(0)}$ для формулы (30), а N — норма направляемой или вытекающей моды в невозмущённом диэлектрическом волноводе. Для НМ это условие выполнено, если $|\delta\varepsilon| \ll |n|^2$ и площадь области возмущения показателя преломления невелика, причём точность формул растёт, когда возмущённая область удаляется от оси z. Для ВМ эти условия недостаточны для применимости формул (30) и (31). В качестве примера опять рассмотрим волновод в виде канала в диэлектрике. Предположим, что возмущение функции $n(y)$ мало и расположено на небольшом отрезке оси y с центром в точке y_{p}. Если $n(y)$ возмущается внутри канала или вблизи границ $y = \pm d$, то второй член в правой части соотношения (31) мал, т. к. он пропорционален $|\delta\varepsilon|^2$. При этих условиях сдвиг $\tilde\beta_\sigma$ также мал и правильно описывается указанной формулой. Будем теперь сдвигать интервал вдоль оси y в сторону бо́льших положительных y. Когда отрезок попадёт в область экспоненциального роста поля $\tilde U_\sigma$, формулы (30) и (31) перестанут работать, хотя на самом деле смещение величины $\tilde\beta_\sigma$ может оставаться небольшим. В этом заключается недостаток использованной выше версии теории возмущений для анализа ВМ. Для оценки смещения в этом случае можно использовать точную формулу

$$[\tilde\beta_\sigma^{(1)}]^2 - [\tilde\beta_\sigma^{(0)}]^2 = k^2\,\langle\delta\varepsilon\,\tilde U_\sigma^{(0)}, \tilde U_\sigma^{(1)}\rangle_\alpha / \langle\tilde U_\sigma^{(0)}, \tilde U_\sigma^{(1)}\rangle_\alpha. \tag{34}$$

Это соотношение выводится с помощью формулы Грина, в которую подставляются поля в невозмущённом и возмущённом волноводах $\tilde U_\sigma^{(0)}$, $\tilde U_\sigma^{(1)}$ и весовая функция w_α. Нетрудно показать, что, как правило, для двумерных систем при увеличении y_{p} величина $\tilde\beta_\sigma$ колеблется с малой амплитудой. Этот эффект легко объяснить с помощью соображений геометрической оптики, рассматривая задачу о прохождении лучей через несколько диэлектрических слоёв. Для трёхмерных задач величина $|\tilde\beta_\sigma^{(1)} - \tilde\beta_\sigma^{(0)}|$ обычно убывает, если площадь области возмущения показателя преломления постоянна и эта область удаляется от оси волновода. Ошибочность результатов, которые получаются из формул (29), связана с тем, что в общем случае ряд теории возмущений является асимптотическим, слагаемые в нём могут расти, но при этом иметь разные знаки. Подобный эффект существует в теории возбуждения ВМ: выделять эти моды целесообразно только вблизи оси волновода. Вдали от оси поля́ ВМ и интегралы по РМ растут, в то время как полное поле, которое описывает цилиндрическую или сферическую волны, убывает.

Приведённые выше формулы можно использовать не только для расчёта смещения постоянных распространения мод. С помощью (30) можно вывести простое соотношение, которое удобно применять для расчёта нормы $\tilde N_\sigma$. Для вывода формулы рассмотрим плоский симметричный волновод, у которого при $|y| < d$ профиль показателя преломления имеет вид

$$n^2(y) = n_1^2 + \tau\Delta_\varepsilon(y), \tag{35}$$

где, как и выше, n_1 — показатель преломления окружающей волновод среды. В выражение для $n(y)$ введён вспомогательный параметр τ, который будем варьировать вблизи значения $\tau = 1$. Применяя метод возмущений, нетрудно получить соотношение

$$\left.\frac{\partial\tilde\beta_\sigma^2}{\partial\tau}\right|_{\tau=1} = k^2\,\langle\Delta_\varepsilon\tilde U_\sigma, \tilde U_\sigma\rangle_\alpha / \tilde N_\sigma, \tag{36}$$

откуда следует выражение для нормы ВМ:

$$\tilde N_\sigma = k^2\,\langle\Delta_\varepsilon\tilde U_\sigma, \tilde U_\sigma\rangle_\alpha / (\partial\tilde\beta_\sigma^2/\partial\tau), \tag{37}$$

где все величины рассчитываются при $\tau = 1$. Приведённое соотношение удобно тем, что при его вычислении не надо знать поле $\tilde{U}_\sigma$ в окружающей волновод среде, т. к. там $\Delta_\varepsilon = 0$. Подобная формула особенно полезна при численном анализе векторных трёхмерных задач (см. ниже). Из приведённого выражения можно получить ряд приближённых соотношений, например, в области высоких частот.

4. ВЕКТОРНЫЕ ЗАДАЧИ

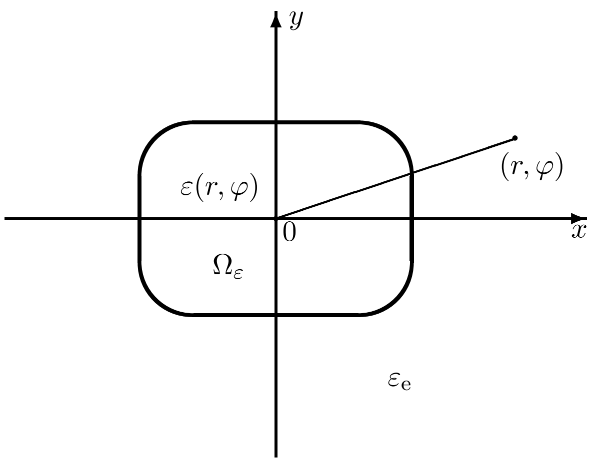

Рис. 3. Поперечное сечение трёхмерного волновода и системы координат

Описанный выше аппарат может быть обобщён на более сложные системы и, в первую очередь, на трёхмерные волноводы с векторными модами. Рассмотрим достаточно простой случай — диэлектрический волновод с однородным изотропным покрытием. Поперечное сечение такого волновода и системы координат показаны на рис. 3. Относительная диэлектрическая проницаемость ε внутри сердцевины волокна (в области Ω_ε) является скалярной гладкой функцией поперечных цилиндрических координат r и φ (рис. 3). Переход к разрывным распределениям осуществляется стандартным способом. Вне сердцевины (в оболочке) диэлектрическая проницаемость постоянна и равна ε_{e}. Система собственных мод такого диэлектрического волновода[1] включает в себя направляемые моды, поля которых обозначим через $\mathbf{E}_i$, $\mathbf{H}_i$, а также радиационные моды, поля которых обозначим через $\mathbf{E}_{m\kappa}$, $\mathbf{H}_{m\kappa}$. Индекс i, нумерующий НМ, пробегает конечное множество целых чисел. Ветви РМ нумеруются дискретным индексом m; непрерывный параметр κ (поперечное волновое число в оболочке) пробегает все действительные значения на луче $0 < \kappa < +\infty$. Заметим, что в трёхмерных задачах число ветвей РМ в общем случае счётно. Собственные моды вводятся с помощью оператора рассеяния $\hat{\mathbf{S}}$. Для их построения внешние поля представляют в виде рядов цилиндрических гармоник, амплитуды которых связаны S-оператором. Схема построения таких мод была описана ранее [9, 10, 16] и во многом повторяет схему, рассмотренную в разделе 1. Как и для планарных структур, ВМ можно выделить из интегралов по РМ (по непрерывному спектру), смещая вниз контур интегрирования в комплексной плоскости κ.

Рассмотрим несколько подробнее вопросы нормировки собственных мод. Введём обозначение для интеграла от векторных функций $\mathbf{E}$, $\mathbf{H}$ по произвольному сечению $z = z_0$:

$$\langle \mathbf{E}, \mathbf{H} \rangle_\alpha = \lim_{\alpha \to +0} \int\limits_{z=z_0} w_\alpha \, [\mathbf{E} \times \mathbf{H}] \, \mathbf{e}_z \, \mathrm{d}x \, \mathrm{d}y, \tag{38}$$

где $\mathbf{e}_z$ — единичный вектор, направленный вдоль оси z волновода, косым крестом обозначена операция векторного произведения. В подынтегральную функцию введён быстро убывающий вес $w_\alpha = \exp(-\alpha r^2)$. Условия ортогональности направляемых и радиационных мод для векторных задач имеют вид

$$2 \langle \mathbf{E}_i, \mathbf{H}_{i'} \rangle_\alpha = N_i \delta_{ii'}, \qquad \langle \mathbf{E}_i, \mathbf{H}_{m\kappa} \rangle_\alpha = 0, \tag{39}$$

[1] Для простоты рассматриваем случай, когда в волноводе нет собственных мод других типов, таких как присоединённые и комплексные.

$$2\,\langle \mathbf{E}_{m\kappa}, \mathbf{H}_{m'\kappa'}\rangle_\alpha = D_m(\kappa)\delta_{mm'}\delta(\kappa-\kappa'). \tag{40}$$

Как и выше, нормирующий множитель $D_m(\kappa)$ выражается через собственные векторы и собственное значение S-оператора [10]. При выводе всех формул прямые и встречные моды выбирались так, чтобы поперечные компоненты их электрических полей были равны, а поперечные компоненты магнитных полей отличались знаком. Вывод формул проводился с помощью леммы Лоренца, которая применялась к полям собственных мод, заключённых в объёме между двумя бесконечно близкими поперечными плоскостями [9, 16].

Исходное выражение для нормы ВМ имеет точно такой же вид (16), как для двумерных скалярных задач (производная от нормирующего множителя $D_m(\kappa)$). Используя лемму Лоренца [9, 10], из этой формулы можно вывести выражение для нормы, аналогичное (22). Условие ортогональности ВМ имеет вид

$$2\,\langle \tilde{\mathbf{E}}_\sigma, \tilde{\mathbf{H}}_{\sigma'}\rangle_\alpha = \tilde{N}_\sigma\delta_{\sigma\sigma'}. \tag{41}$$

Другой способ регуляризации расходящихся интегралов в формуле для $\tilde{N}_\sigma$ основывается на интегрировании в сечении $z = z_0$ вдоль комплексной радиальной координаты $r = |r|\exp(i\gamma)$ [9, 10]; получаемое соотношение аналогично формуле (18). Как и для двумерных задач, этот способ нормировки неприменим к РМ, которые определены на смещённом контуре интегрирования (рис. 2).

Для трёхмерных диэлектрических волноводов нетрудно обобщить почти все формулы, которые были выведены для планарных систем. В качестве примера оценим смещение волновых чисел ВМ при изменении диэлектрической проницаемости волновода. Пусть $\varepsilon^{(0)}$ — невозмущённый профиль проницаемости, а $\varepsilon^{(1)}$ — возмущённый, причём, как обычно, возмущение мало. Для ВМ сдвиг постоянной распространения равен

$$\tilde{\beta}_\sigma^{(1)} - \tilde{\beta}_\sigma^{(0)} = \frac{k\zeta_{\rm v}^{-1}\,\langle \delta\varepsilon\,\tilde{\mathbf{E}}_\sigma^{(0)}, \tilde{\mathbf{E}}_\sigma^{(1)}\rangle_\alpha}{\langle \tilde{\mathbf{E}}_\sigma^{(1)}, \tilde{\mathbf{H}}_\sigma^{(0)}\rangle_\alpha + \langle \tilde{\mathbf{E}}_\sigma^{(0)}, \tilde{\mathbf{H}}_\sigma^{(1)}\rangle_\alpha}\,, \tag{42}$$

где $\delta\varepsilon = \varepsilon^{(1)} - \varepsilon^{(0)}$, а $\tilde{\mathbf{E}}_\sigma^{(0)}$ и $\tilde{\mathbf{E}}_\sigma^{(1)}$ — поля ВМ в невозмущённом и возмущённом диэлектрическом волноводе соответственно. Дальнейшее упрощение формулы легко проводится в двух случаях. Если $\delta\varepsilon$ мало (см. выше), то можно положить $\tilde{\mathbf{E}}_\sigma^{(1)} \approx \tilde{\mathbf{E}}_\sigma^{(0)}$ и получить явную формулу для смещения $\beta_i^{(0)}$. В другом случае, когда площадь области возмущения мала, можно использовать квазистатическое приближение, рассчитывая поле $\tilde{\mathbf{E}}_\sigma^{(1)}$ в области возмущения из вспомогательной статической задачи [17].

Из формулы (42) можно получить другое выражение для нормы ВМ. Для его вывода рассмотрим волновод, у которого внутри области Ω_ε профиль проницаемости имеет вид $\varepsilon(x,y) = \varepsilon_{\rm e} + \tau\Delta_\varepsilon(x,y)$, где $\varepsilon_{\rm e}$ — проницаемость окружающей волновод среды. В выражение для ε введён вспомогательный параметр τ, который, как и выше, будем варьировать вблизи значения $\tau = 1$. Применяя метод возмущений, нетрудно получить следующее выражение для нормы ВМ:

$$\tilde{N}_\sigma = \frac{k\zeta_{\rm v}^{-1}\,\langle \Delta_\varepsilon\tilde{\mathbf{E}}_\sigma, \tilde{\mathbf{E}}_\sigma\rangle_\alpha}{\partial\tilde{\beta}_\sigma/\partial\tau}\,, \tag{43}$$

где все величины рассчитываются для $\tau = 1$.

Заметим, что во многих численных методах [18] в результате расчётов определяются не все поля, а, например, только электрическое поле, причём лишь внутри волновода или в непосредственной близости от сердцевины. Процедура продолжения поля наружу, а также расчёт магнитного поля обычно требует дополнительных вычислений, зачастую они являются достаточно трудоёмкими и могут сопровождаться потерей точности. Применяя выведенное выше соотношение, можно обойти указанные трудности и достаточно быстро вычислить величину $\tilde{N}_\sigma$. Поля ВМ растут в окружающей волновод среде, поэтому, используя формулу (43), избегаем дополнительных проблем, связанных с этим эффектом.

Для трёхмерных задач существует аналог формулы (20), которая ранее была выведена для планарного волновода. Для простоты рассмотрим случай слабонаправляющих [8] трёхмерных структур, когда задача сводится к скалярной. Построим в поперечной плоскости окружность с центром на оси z и радиусом $r = r_{\rm c}$, расположенную вне области Ω_ε. При $r > r_{\rm c}$ внешнее поле ВМ можно представить в виде ряда цилиндрических гармоник:

$$\tilde{E}(r,\varphi) = \sum_{n=-\infty}^{\infty} \tilde{u}_{n\sigma} H^{(1)}_{|n|}(\tilde{\kappa}_\sigma r)\exp(in\varphi), \tag{44}$$

где $H^{(1)}_n$ — функции Ханкеля, $\tilde{u}_{n\sigma}$ — амплитуды гармоник. При использовании этого разложения интеграл $I_{\rm r}$ по внешности окружности $r = r_{\rm c}$ вычисляется в явной форме:

$$I_{\rm r} = \pi r_{\rm c}^2 \sum_{n=-\infty}^{\infty} \tilde{u}_{n\sigma}\tilde{u}_{-n\sigma} \left\{ H^{(1)}_{|n|-1}(\tilde{\kappa}_\sigma r_{\rm c}) H^{(1)}_{|n|+1}(\tilde{\kappa}_\sigma r_{\rm c}) - [H^{(1)}_{|n|}(\tilde{\kappa}_\sigma r_{\rm c})]^2 \right\}. \tag{45}$$

С помощью (45) можно оценить отношение интегралов по сердцевине волокна и его оболочке. Для слабовытекающих мод это отношение обычно мало (см. выше).

Для простоты выше рассматривались системы, у которых диэлектрические потери отсутствуют. Однако все выведенные формулы были записаны в таком виде, что они применимы для расчёта диэлектрических волноводов, у которых среды имеют такие потери [9, 19]. В частности, в выражениях для норм не использовались комплексно-сопряжённые функции [10]. Формулы с комплексно-сопряжёнными функциями справедливы только при отсутствии диэлектрических потерь [20]. В нескольких недавно опубликованных работах [21, 22] этот факт не был учтён, поэтому часть приведённых в них соотношений (для систем с потерями) не является корректной. Отметим также, что вопрос о влиянии потерь в материале диэлектрических волноводов достаточно важен и возникает на практике при анализе оптических устройств, использующих электроабсорбционный эффект [23]. Этот вопрос, как оказалось [22, 24], весьма важен и при анализе процессов распространения мод при частотах, близких к частотам отсечки.

5. ЗАКЛЮЧЕНИЕ

В работе рассмотрены различные подходы, которые можно использовать для нормировки полей вытекающих мод открытых диэлектрических волноводов. Выведены формулы для норм ВМ, которые могут применяться в сочетании с современными численными методами, такими, как методы интегральных уравнений и конечных элементов. Эти подходы могут быть использованы для расчёта норм мод других типов, например антиповерхностных и комплексных. С небольшими изменениями методики применимы для вычисления норм мод в анизотропных (гиротропных) и периодических волноводах, а также норм квазисобственных колебаний в открытых резонаторах.

Автор признателен рецензенту за полезные замечания. Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты № 02–02–17317 и 03–02–16161).

СПИСОК ЛИТЕРАТУРЫ

1. Вайнштейн Л. А. Открытые резонаторы и открытые волноводы. М.: Сов. радио, 1966.

2. Tamir T., Oliner A. A. // Proc. IEEE. 1963. V. 51, No. 2. P. 317.
3. Маненков А. Б. // Радиотехника и электроника. 1977. Т. 22, № 10. С. 2 043.
4. Маненков А. Б., Мелёхин В. Н. // Радиотехника и электроника. 1979. Т. 24, № 7. С. 1 282.
5. Steel M. J., Osgood R. M. J. // J. Lightwave Technol. 2001. V. 19, No. 4. P. 495.
6. Xu Y., Ouyang G., Lee R., Yariv A. // J. Lightwave Technol. 2002. V. 20, No. 3. P. 428.
7. Issa N. A., Poladian L. // J. Lightwave Technol. 2003. V. 21, No. 4. P. 1 005.
8. Адамс М. Введение в теорию оптических волноводов. М.: Мир, 1984.
9. Маненков А. Б. // Изв. вузов. Радиофизика. 1970. Т. 13, № 5. С. 739.
10. Вайнштейн Л. А., Маненков А. Б. // Лекции по электронике СВЧ и радиофизике. Саратов: СГУ, 1986. Кн. 1. С. 141.
11. Marcuse D. Theory of Dielectric Optical Waveguides. New York: Academic Press, 1974.
12. Ландау Л. Д., Лифшиц Е. М. Квантовая механика (нерелятивистская теория). М.: Наука, 1989.
13. Базь А. И., Зельдович Б. Я., Переломов А. М. Рассеяние, реакции и распады в нерелятивистской квантовой механике. М.: Наука, 1966.
14. Лаврентьев М. А., Шабат Б. В. Методы теории функций комплексного переменного. М.: Гос. изд. физ.-мат. лит., 1958.
15. Маненков А. Б. // Изв. вузов. Радиофизика. 1975. Т. 18, № 7. С. 1 025.
16. Manenkov A. B. // IEE Proc. Part J. 1993. V. 140, No. 3. P. 206.
17. Никольский В. В. Вариационные методы для внутренних задач электродинамики. М.: Наука, 1967.
18. Клеев А. И., Маненков А. Б., Рожнев А. Г. // Радиотехника и электроника. 1993. Т. 38, № 11. С. 1 938.
19. Маненков А. Б. // Изв. вузов. Радиофизика. 1981. Т. 24, № 1. С. 84.
20. Dasyras N. F., Tigelis I. G., Tsigopoulos A. D., Manenkov A. B. // J. Opt. Soc. Am. A. 2004. V. 21, No. 9. P. 1 740.
21. Benech P., Khalil D. A. M., Andre F. S. // Opt. Commun. 1992. V. 88, No. 2–3. P. 96.
22. Romanova E. A. // Opt. Commun. 2002. V. 208, No. 1–3. P. 91.
23. Хансперджер Р. Интегральная оптика. М.: Мир, 1987.
24. Arnbak J. // Electron. Lett. 1969. V. 5, No. 3. P. 41.

Поступила в редакцию 12 июля 2004 г.; принята в печать 7 февраля 2005 г.

ORTHOGONALITY CONDITIONS FOR LEAKY MODES

A. B. Manenkov

Different approaches for the normalization of the open-waveguide leaky modes are considered. General characteristics of such modes and their relation with the modes of the continuous spectrum are investigated. The problem of the perturbation-theory applicability to the calculation of their propagation-constant shifts under small changes of the waveguide structure is discussed. The general equations are illustrated by the example of the channel waveguide.

УДК 621.372

РЕЗОНАНСЫ В МИКРОПОЛОСКОВОЙ СИСТЕМЕ ДИАФРАГМА—КОРОТКОЗАМЫКАЮЩИЙ ПОРШЕНЬ

Д. А. Усанов, С. С. Горбатов

Саратовский госуниверситет, г. Саратов, Россия

Обнаружено резонансное отражение в микрополосковой системе диафрагма—короткозамыкающий поршень при расстояниях между диафрагмой и поршнем, значительно меньших длины волны в линии передачи. Показано, что исследованная система может быть использована при конструировании новых типов СВЧ устройств с регулируемыми характеристиками, в том числе полупроводниковых, характеризующихся уменьшенными габаритами.

При конструировании волноводных систем обычно исходят из предположения, что в них обеспечены условия распространения только для волны основного типа. Из этого же предположения обычно исходят при конструировании устройств с резонансными характеристиками (индуктивный штырь с ёмкостным зазором, резонансная диафрагма и т. д.). Устройства такого типа широко распространены на практике, например при создании полупроводниковых устройств СВЧ, предназначенных для работы при высоких уровнях мощности. Анализируя физику работы таких устройств, обычно отмечают, что в окрестности неоднородности (штыря, диафрагмы) поле волны основного типа может быть существенным образом искажено, что связано с возбуждением быстрозатухающих волн высших типов. В то же время, если вблизи неоднородности находится, например, отражающая поверхность, то роль ближнего поля вблизи неоднородности может быть существенной. В частности, в [1, 2] описаны резонансные явления в волноводной системе, включающей ёмкостную диафрагму и короткозамыкающий поршень. Было обнаружено возникновение резонансов при расстоянии между диафрагмой и короткозамыкателем, равном приблизительно $\lambda_{\text{В}}/100$, где $\lambda_{\text{В}}$ — длина волны в волноводе [1, 2].

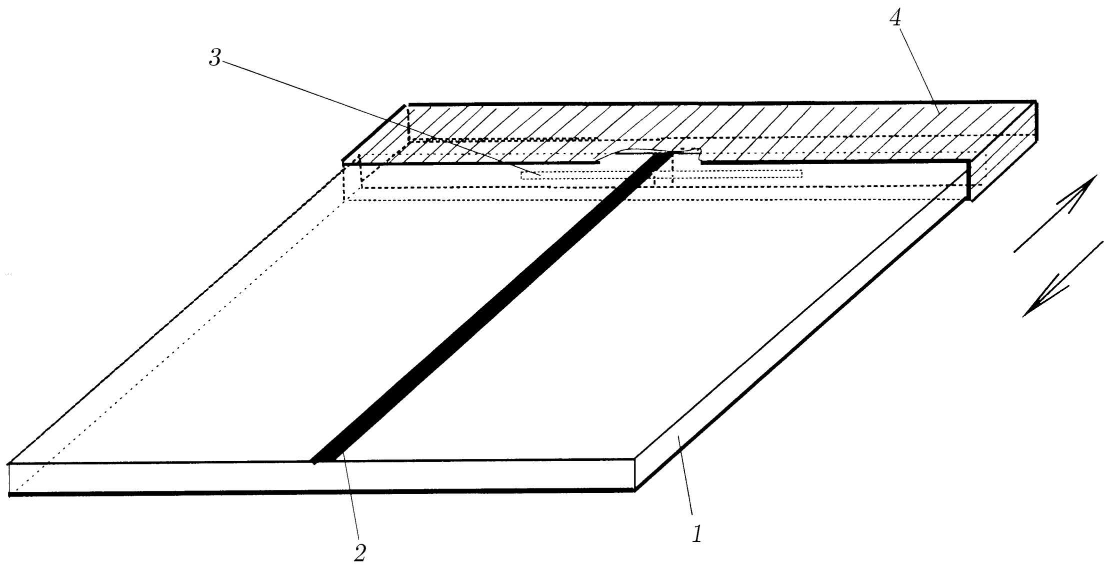

Рис. 1

Целью настоящей работы является исследование возможности возникновения такого рода резонансных характеристик отражения в микрополосковой системе. Схематическое изображение предлагаемого микрополоскового аналога описанных в [1, 2] волноводных систем приведено на рис. 1. Размеры диэлектрической платы *1* микрополосковой линии составляли 60 × 42 × 1,5 мм, ширина микрополоска *2* — 1,8 мм, толщина — 18 мкм. Торец диэлектрической платы металлизировался, таким образом, чтобы в металлизации оставалось прямоугольное отверстие *3* с размерами 0,2 × 15 мм, расположенное симметрично по центру торца относительно микрополоска. Металлизированный торец вставлялся в металлический короткозамыкатель *4* прямоугольной формы с размерами, соответствующими сечению микрополосковой линии (60 × 10 × 1,5 мм), имеющий возможность перемещения с трением скольжения вдоль её ребра в пределах 10 мм так, что во всех положениях осуществляется гальванический контакт короткозамыкателя с микрополоском. Толщина стенок материала короткозамыкателя составляла 200 мкм. На рис. 1 поверхность короткозамыкателя заштрихована. Для наглядности связи короткозамыкателя с диафрагмой на нём на рис. 1 сделан вырез. Стрелками обозначено направление перемещения короткозамыкателя. Отверстие выполняло функции ёмкостной диафрагмы. Диафрагма служила элементом связи с резонатором, образованным металлизированным торцом микрополосковой платы и короткозамыкателем.

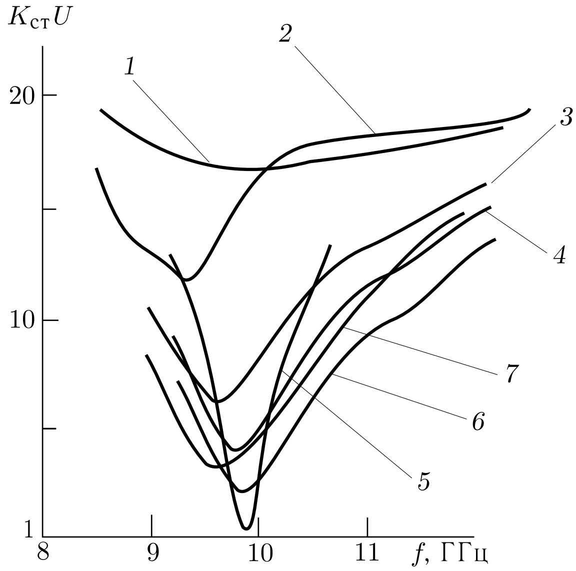

Рис. 2

На рис. 2 приведены результаты измерений коэффициента стоячей волны $K_{ст}U$ исследуемой системы в полосе частот 8,5÷12 ГГц для различных расстояний L между короткозамыкателем и диафрагмой. Кривая *1* соответствует $L = 0$, кривая *2* — $L = 200$ мкм, кривая *3* — $L = 400$ мкм, кривая *4* – $L = 800$ мкм, кривая *5* — $L = 1{,}5$ мм, кривая *6* — $L = 2{,}0$ мм, кривая *7* — 2,5 мм.

Результаты измерений свидетельствуют о том, что $K_{ст}U$ изменяется от величин, бо́льших 10, при нулевом расстоянии между короткозамыкателем и диафрагмой до 1,25 в точках резонанса

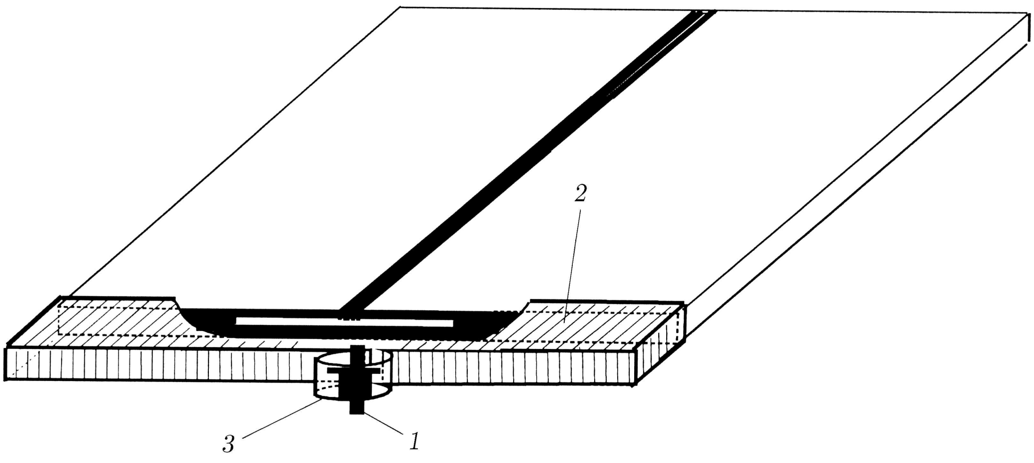

Рис. 3

при $L = 1{,}5$ мм. Причём при $L = 0$ зависимость $K_{\text{ст}}U$ от частоты слабая и нерезонансная, а резонансный характер зависимости проявляется при расстояниях L, равных 200 мкм и больше. Наиболее сильно резонансный характер зависимости $K_{\text{ст}}U$ от частоты проявляется для $L = 1{,}5$ мм (кривая *5*). При расстояниях L, равных 2,0 мм и больше, резонанс сдвигается в сторону низких частот, причём $K_{\text{ст}}U$ при резонансе увеличивается.

Отметим, что при увеличении L от 800 мкм до 1,5 мм резонанс сдвигается в сторону высоких частот аналогично тому, как это наблюдается в волноводной системе диафрагма—близко расположенный поршень [1], где увеличение частоты резонанса с увеличением расстояния между поршнем и диафрагмой в определённом интервале расстояний объяснялось возбуждением распространяющихся волн высших типов. Вклад в резонанс системы возбуждающихся волн высших типов носит индуктивный характер. Этот вклад при небольших расстояниях между диафрагмой и поршнем может быть существенным. Увеличение расстояния между поршнем и диафрагмой при малых расстояниях приводит к увеличению частоты наблюдаемых резонансов, что связано с уменьшением индуктивности короткозамкнутого отрезка линии передачи.

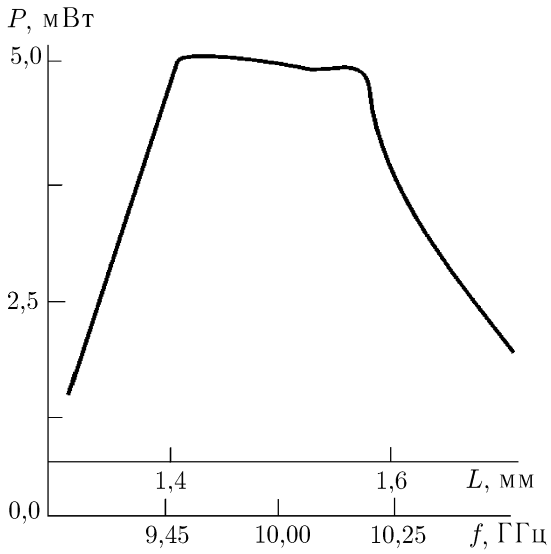

Рис. 4

Исследованная система была использована при конструировании микрополоскового генератора на диоде Ганна с механической перестройкой частоты (рис. 3). Диод *1* типа 3А 703 размещался в короткозамыкателе *2* в специальной технологической выемке *3*, сделанной на передней поверхности короткозамыкателя так, как это показано на рис. 3. Центр выемки был смещён по горизонтали относительно центра передней поверхности короткозамыкателя на 5 мм. Диаметр выемки составлял 6 мм. На диод подавалось постоянное напряжение смещения, равное 8,5 В. Результаты измерения зависимости мощности и частоты генератора от расстояния между короткозамыкателем и диафрагмой приведены на рис. 4. Как следует из этих результатов, при изменении расстояния L от 1,4 до 1,6 мм частота перестраивается от 9,4 ГГц до 10,2 ГГц при изменении мощности генерируемого сигнала относительно среднего значения 5 мВт в пределах $\pm 0{,}2$ мВт.

Таким образом, показано, что исследованная система может быть успешно использована при конструировании новых типов микрополосковых СВЧ устройств с регулируемыми характеристиками, в том числе полупроводниковых.

СПИСОК ЛИТЕРАТУРЫ

1. Усанов Д. А., Вениг С. Б., Орлов В. Е. // Радиотехника и электроника. 2000. Т. 45, № 6. С. 761.
2. Усанов Д. А., Горбатов С. С., Вениг С. Б., Орлов В. Е. // Письма в ЖТФ. 2000. Т. 26, вып. 18. С. 47.

Поступила в редакцию 1 марта 2004 г.; принята в печать 10 ноября 2004 г.

THE RESONANCE IN THE MICROSTRIP WITH DIAPHRAGM—SHORT CIRCUITED PLUNGER SYSTEM

D. A. Usanov and S. S. Gorbatov

The resonance reflection in the microstrip with diaphragm—short circuited plunger system for the distance between diaphragm and plunger significantly smaller wavelength in the microstrip to discovery. It is show, the researching system in so doing design the new type microwave devices with small dimensions can to use with controlling characteristics.

УДК 533.9

ОБ ИСПОЛЬЗОВАНИИ МЕТОДА ВОЗМУЩЕНИЙ ПРИ КИНЕТИЧЕСКОМ АНАЛИЗЕ НЕЛИНЕЙНЫХ ПРОБЛЕМ ЭЛЕКТРОДИНАМИКИ ПЛАЗМЫ

М. Д. Токман, А. Ю. Крячко

Институт прикладной физики РАН, г. Нижний Новгород, Россия

Обсуждается корректность метода возмущений (разложения по степеням амплитуды высокочастотного поля) при решении кинетического уравнения. Рассматривается случай формального нарушения условий применимости данного метода для «сильно локализованной» функции распределения. Показано, что при этом асимптотические выражения для моментов функции распределения остаются корректными не только в линейном, но и в квадратичном по полю приближении.

ВВЕДЕНИЕ

Широко распространённым методом исследования различных проблем электродинамики плазмы является асимптотическое решение кинетического уравнения методом разложения функции распределения по степеням амплитуд высокочастотных полей (см., например, [1]). Формальным параметром малости, определяющим условие корректности такого разложения, является отношение изменения импульса частиц под действием высокочастотного поля к характерному масштабу изменения исходной (невозмущённой) функции распределения частиц. Как следствие, условия применимости подобного метода возмущений при любой конечной амплитуде поля нарушаются в случае «сильно локализованных» в фазовом пространстве функций распределения. Для корректного анализа динамики ансамбля частиц при таких начальных функциях распределения приходится переходить к непосредственному суммированию траекторий различных частиц (в качестве примера см. [2]) или пользоваться гидродинамическим приближением. При этом переход к стандартному гидродинамическому описанию невозможен, например, в случае моноэнергетических пучков в магнитном поле с разбросом по питч-углам и/или фазам циклотронного вращения.

Нарушение условий применимости решения кинетического уравнения методом разложения возмущения функции распределения по степеням поля для «сильно локализованных» функций распределения является, в известном смысле, парадоксальным фактом, т. к. размер области, занятой частицами в фазовом пространстве, конечно, никак не влияет на корректность применения метода возмущений по степеням амплитуд полей непосредственно к уравнениям движения, являющимся характеристиками кинетического уравнения. В рамках линейной теории волн в плазме этот парадокс снимается. Дело в том, что при переходе в соответствующих «кинетических» выражениях к невозмущённой функции распределения в виде δ-функции, т. е. к моноэнергетическим пучкам, или к «холодной» плазме, выражения для линейных колебаний зарядов и токов (моментов функции распределения) оказываются корректными несмотря на нарушение условий, обеспечивающих применимость соответствующих приближений при определении функции распределения как таковой [1, 3]. Основная цель данной работы состоит в доказательстве корректности предельного перехода к «сильно локализованной» функции распределения при кинетическом расчёте не только линейных, но и квадратичных по полю электродинамических откликов электронного ансамбля (необходимых для анализа, например, параметрических волновых процессов [4]). Важность этого исследования определяется высокой эффективностью кинетического подхода при исследовании принципиально новых режимов электродинамики плазмы [1].

Далее, в разделе 1, для простой, но достаточно общей модели демонстрируется нарушение формальных условий применимости метода разложения функции распределения по степеням амплитуды поля при переходе к «сильно локализованному» распределению. В разделе 2 для этой же модели доказано, что использование асимптотических «кинетических» выражений для определения вариаций моментов функции распределения (например, токов) остаётся корректным не только в линейном приближении, но и при учёте квадратичных по полю величин. В разделе 3 в качестве ещё одного примера проведено сопоставление результатов кинетического и гидродинамического подходов для более сложной системы.

1. УСЛОВИЯ ПРИМЕНИМОСТИ МЕТОДА ВОЗМУЩЕНИЙ

Рассмотрим уравнение Лиувилля для бесстолкновительного ансамбля частиц (электронов), динамика которых описывается уравнениями движения достаточно общего вида, пригодными для анализа широкого класса задач о взаимодействии частиц с электромагнитными полями [2] [7]:

$$\left\{\frac{\partial}{\partial t}+[\Omega(\mu,t)+\varepsilon\omega(\mu,\theta,t)]\,\frac{\partial}{\partial\theta}+\varepsilon\xi(\mu,\theta,t)\,\frac{\partial}{\partial\mu}\right\}f=0, \tag{1}$$

где μ и θ — обобщённый импульс и обобщённая координата соответственно. Слагаемое $\Omega(\mu,t)\times\times\,\partial f/\partial\theta$ описывает «свободное» движение частиц (в т. ч., возможно, циклотронное вращение), пропорциональные ε члены описывают воздействие внешних высокочастотных полей, параметр ε определяется амплитудой поля. Малость ε позволяет представить решение уравнения (1) в виде стандартного разложения по степеням амплитуды поля:

$$f(\mu,\theta,t)=f_0(\mu)+\varepsilon f_1(\mu,\theta,t)+\varepsilon^2f_2(\mu,\theta,t)+\,\ldots\,+\varepsilon^nf_n(\mu,\theta,t)+\,\ldots\,, \tag{2}$$

где $f_0(\mu)$ — невозмущённая функция распределения, слагаемое εf_1 представляет собой линейное по полю возмущение, ε^2f_2 — квадратичное и т. д. Построение асимптотического разложения (2) основано на очевидном предположении

$$\{f_n,\partial f_n/\partial\mu,\,\ldots\}\gg\{\varepsilon f_{n+1},\varepsilon\partial f_{n+1}/\partial\mu,\,\ldots\}, \tag{3}$$

где $n=0,1,2,\,\ldots$, которое налагает определённые формальные ограничения на величину производных невозмущённой функции распределения $f_0(\mu)$. Действительно, стандартный вид соотношений линейного приближения [3]:

$$\frac{\partial f_1}{\partial t}+\Omega(\mu,t)\,\frac{\partial f_1}{\partial\theta}=-\xi(\mu,\theta,t)\,\frac{\partial f_0}{\partial\mu}\ , \tag{4а}$$

$$f_1=-\mu_1(\mu,\theta,t)\,\partial f_0/\partial\mu, \tag{4б}$$

$$\mu_1=\int\limits_0^t\xi\left(\mu,\theta-\int_\tau^t\Omega(\mu,\tilde\tau)\,\mathrm{d}\tilde\tau,\tau\right)\,\mathrm{d}\tau, \tag{4в}$$

[1] В качестве примеров приведём исследование мазерной неустойчивости релятивистского пучка электронов [3] или построение классического аналога квантового лазера без инверсии [5, 6].

[2] Хотя мы не предполагаем явно гамильтоновой формы соответствующих уравнений движения, всё рассмотрение проводится для систем, сохраняющих в процессе эволюции фазовый объём.

[3] Для определённости рассмотрим начальное условие $f_1(t=0)=0$.

и структура уравнений для возмущений следующих порядков, например для квадратичного по полю приближения:

$$\frac{\partial f_2}{\partial t} + \Omega\,\frac{\partial f_2}{\partial \theta} = -\xi\,\frac{\partial f_1}{\partial \mu} - \omega\,\frac{\partial f_1}{\partial \theta}\ , \tag{5}$$

требуют для выполнения (3) удовлетворения следующего неравенства:

$$\partial^n f_0/\partial\mu^n \gg (\varepsilon\mu_1)\,\partial^{n+1} f_0/\partial\mu^{n+1}, \tag{6}$$

где величина $\varepsilon\mu_1$ введена в (4в) и имеет смысл изменения обобщённого импульса в линейном приближении по внешнему полю (см. [1]). Поскольку $\partial^n f_0/\partial\mu^n \sim f_0/(\mu_T)^n$, где μ_T — характерный размер области фазового пространства, занятой частицами, то при попытке использования, например, δ-образной функции распределения, когда $\mu_T \to 0$, при любой конечной величине ε будут нарушены условия (6).

2. УЧЁТ КВАДРАТИЧНЫХ ПОПРАВОК ДЛЯ ТОКОВ

Найдём выражение для входящей в разложение (2) функции $f_2(\mu,\theta,t)$. Рассматривая начальное условие $f(t=0) = f_0(\mu)$ из (5) и (4в) получим

$$f_2 = \frac{1}{2}\,\mu_1^2\,\frac{\partial^2 f_0}{\partial\mu^2} - \mu_2\,\frac{\partial f_0}{\partial\mu}\ , \tag{7}$$

где

$$\mu_2 = -\int\limits_0^t \left(\xi\,\frac{\partial\mu_1}{\partial\mu} + \omega\,\frac{\partial\mu_1}{\partial\theta}\right)\Bigg|_{\theta\to\theta-\int\limits_\tau^t \Omega(\mu,\tilde\tau)\,\mathrm{d}\tilde\tau}\ \mathrm{d}\tau. \tag{8}$$

Найдём момент функции распределения:

$$I = \int J(\mu,\theta) f(\mu,\theta,t)\,\mathrm{d}\mu\,\mathrm{d}\theta, \tag{9}$$

такой, что $\int J(\mu,\theta) f_0(\mu)\,\mathrm{d}\mu\,\mathrm{d}\theta = 0$ и $J(\mu=0,\theta) = 0$. Подставив (2), (4в), (7) и (8) в (9), после интегрирования по частям получим

$$I = \int \left\{\varepsilon\,\partial(J\mu_1)/\partial\mu + \varepsilon^2\left[\frac{1}{2}\,\partial^2(J\mu_1^2)/\partial\mu^2 + \partial(J\mu_2)/\partial\mu\right]\right\} f_0\,\mathrm{d}\mu\,\mathrm{d}\theta. \tag{10}$$

Теперь построим решение уравнения Лиувилля (1), не требуя выполнения неравенств (6). Для этого сперва построим формальное решение (1) методом характеристик, не предполагая изначально малости ε. Затем в итоговом выражении для момента функции распределения корректно перейдём к малым ε.

Итак, представим начальную функцию распределения $f_0(\mu)$ как $F[(\mu - \langle\mu\rangle)/\mu_T]$, где $\langle\mu\rangle$ — соответствующее среднее значение обобщённого импульса. Решение уравнения (1) с начальным условием $f(t=0) = f_0(\mu)$ представим как функцию интеграла движения:

$$f = f_0[\mu - \Delta\mu(\mu,\theta,t,\varepsilon)], \tag{11а}$$

где интеграл движения представлен в виде

$$\mu - \Delta\mu(\mu,\theta,t,\varepsilon) = \mu_0 = \mathrm{const}, \tag{11б}$$

μ_0 — начальное значение обобщённого импульса, $\Delta\mu$ — возмущение, равное 0 при $t = 0$ и стремящееся к 0 при $\varepsilon \to 0$. Выражение (11а) пригодно для любых величин μ_T, поэтому возможен корректный переход к случаю $\mu_T \to 0$, в том числе и в случае, когда начальная функция распределения сводится к δ-функции Дирака:

$$f = \frac{N}{2\pi}\,\delta[\mu - \Delta\mu(\mu,\theta,t,\varepsilon) - \langle\mu\rangle], \tag{12}$$

где N — концентрация частиц. Заметим, что следующая далее формальная процедура справедлива как для функции распределения в виде δ-функции, так и для функции распределения общего вида. Подставив (11а) в (9) в соответствии с правилами замены аргумента функции распределения перейдём к интегрированию по ансамблю начальных значений μ_0:

$$I = \int J(\mu^*,\theta)\,[1 - \partial(\Delta\mu)/\partial\mu]^{-1}_{\mu=\mu^*}\, f_0(\mu_0)\,\mathrm{d}\theta\,\mathrm{d}\mu_0, \tag{13а}$$

где

$$\mu^* - \Delta\mu(\mu^*,\theta,t,\varepsilon) = \mu_0. \tag{13б}$$

Теперь сопоставим два асимптотических разложения величины I по степеням ε: следующего из (13) и определяемого из (10). Сначала найдём разложение входящей в (11)–(13) величины $\Delta\mu$ по степеням ε с учётом квадратичных слагаемых:

$$\mu - \mu_0 = \Delta\mu = \varepsilon M_1(t,\mu,\theta) + \varepsilon^2 M_2(t,\mu,\theta). \tag{14}$$

Для этого представим решение соответствующей (1) системы уравнений движения

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \varepsilon\xi(\mu,\theta,t), \qquad \frac{\mathrm{d}\theta}{\mathrm{d}t} = \Omega(\mu,t) + \varepsilon\omega(\mu,\theta,t) \tag{15}$$

в виде ряда [4] по степеням ε:

$$\begin{gathered}\mu - \mu_0 = \Delta\mu = \varepsilon\eta_1(t,\mu_0,\theta_0) + \varepsilon^2\eta_2(t,\mu_0,\theta_0) + \mathrm{O}(\varepsilon^3),\\ \theta - \theta_0 = \Theta(t,\mu_0) + \varepsilon\theta_1(t,\mu_0,\theta_0) + \mathrm{O}(\varepsilon^2),\end{gathered} \tag{16}$$

где μ_0 и θ_0 — соответствующие начальные условия. Подставив (16) в (15), получаем

$$\Theta = \int\limits_0^t \Omega(\mu_0,\tau)\,\mathrm{d}\tau, \tag{17}$$

$$\eta_1 = \int\limits_0^t \xi(\mu_0,\theta_0+\Theta,\tau)\,\mathrm{d}\tau, \tag{18а}$$

$$\theta_1 = \int\limits_0^t \left[\omega(\mu_0,\theta_0+\Theta,\tau) + \left.\frac{\partial\Omega}{\partial\mu}\right|_{\mu=\mu_0}\eta_1\right]\mathrm{d}\tau, \tag{18б}$$

[4] Заметим, что для резонансных частиц такое разложение может быть справедливым лишь на ограниченном временно́м интервале.

$$\eta_2 = \int_0^t \left[\frac{\partial \xi}{\partial \mu} \bigg|_{\substack{\mu=\mu_0 \\ \theta=\theta_0+\Theta}} \eta_1 + \frac{\partial \xi}{\partial \theta} \bigg|_{\substack{\mu=\mu_0 \\ \theta=\theta_0+\Theta}} \theta_1 \right] \mathrm{d}\tau. \tag{18в}$$

Чтобы получить разложение (14), необходимо перейти от представления возмущения обобщённого импульса $\Delta\mu$ в виде функции начальных условий (16) к представлению $\Delta\mu$ как функции «текущих» координат μ и θ (как в (11б) и (14)). Для линейного возмущения получаем очевидное тождество

$$M_1(t, \mu, \theta) = \eta_1(\theta_0 = \theta - \Theta(t, \mu_0), \mu_0 = \mu) \equiv \mu_1(\mu, \theta, t). \tag{19}$$

Для квадратичного члена получаем

$$M_2(t, \mu, \theta) = \eta_2[\theta_0 = \theta - \Theta(t, \mu_0), \mu_0 = \mu] - \mu_1 \frac{\partial \mu_1}{\partial \mu} - \theta_1[\theta_0 = \theta - \Theta(t, \mu_0), \mu_0 = \mu] \frac{\partial \mu_1}{\partial \theta}, \tag{20}$$

где величина $\mu_1(\mu, \theta, t)$ определена в (4в). Дифференцируя (20) с учётом (15) и приводя подобные члены, получаем

$$\frac{\mathrm{d}M_2}{\mathrm{d}t} = -\xi(\mu, \theta, t) \frac{\partial \mu_1}{\partial \mu} - \omega(\mu, \theta, t) \frac{\partial \mu_1}{\partial \theta}. \tag{21}$$

Сопоставляя последнее соотношение с (8), получаем тождество, в известном смысле аналогичное [5] (19):

$$M_2(t, \mu, \theta) \equiv \mu_2(\mu, \theta, t). \tag{22}$$

Из (13б), (14), (19) и (22) можно получить соотношение

$$\mu^* = \mu_0 + \varepsilon \mu_1(\mu_0, \theta, t) + \varepsilon^2 [\mu_2 + \mu_1 (\partial \mu_1 / \partial \mu)]|_{(\mu_0, \theta, t)} + \mathrm{O}(\varepsilon^3), \tag{23}$$

подставив которое в (13а) и обозначив переменную интегрирования μ_0 как μ, получаем

$$I = \int f_0(\mu) \left\{ \left[J + \varepsilon \mu_1 \frac{\partial J}{\partial \mu} + \frac{\varepsilon^2}{2} \mu_1^2 \frac{\partial^2 J}{\partial \mu^2} + \varepsilon^2 \mu_2 \frac{\partial J}{\partial \mu} + \varepsilon^2 \mu_1 \frac{\partial \mu_1}{\partial \mu} \frac{\partial J}{\partial \mu} + \mathrm{O}(\varepsilon^3) \right] \times \right.$$
$$\left. \times \left[1 + \varepsilon \frac{\partial \mu_1}{\partial \mu} + \varepsilon^2 \mu_1 \frac{\partial^2 \mu_1}{\partial \mu^2} + \varepsilon^2 \left(\frac{\partial \mu_1}{\partial \mu} \right)^2 + \varepsilon^2 \frac{\partial \mu_2}{\partial \mu} + \mathrm{O}(\varepsilon^3) \right] \right\} \mathrm{d}\mu \, \mathrm{d}\theta. \tag{24}$$

Из (24) после объединения членов при одинаковых степенях ε следует (10)

3. ПРИМЕР БОЛЕЕ СЛОЖНОЙ СИСТЕМЫ: ЭЛЕКТРОМАГНИТНО ИНДУЦИРОВАННАЯ ПРОЗРАЧНОСТЬ

В качестве ещё одного примера сравним результаты кинетического и гидродинамического анализа параметрического эффекта электромагнитно индуцированной прозрачности (ЭИП) в плазме. Первоначально эффект ЭИП обсуждался для квантовых многоуровневых сред [8]; это явление

[5] Эквивалентность функций $M_1(t, \mu, \theta)$, $M_2(t, \mu, \theta)$ и $\mu_1(\mu, \theta, t)$, $\mu_2(\mu, \theta, t)$ в известном смысле очевидна, т. к. следует из сопоставления (2), (4) и (7) с разложением выражения (11а) по степеням ε при выполнении условий (6), однако соответствующий формальный вывод является довольно громоздким. В частности, необходимо учитывать соотношения, следующие из особенностей интегрирования по невозмущённым траекториям: $\mathrm{d}(\partial\mu_1/\partial\theta)/\mathrm{d}t = \partial\xi/\partial\theta$, $\mathrm{d}(\partial\mu_1/\partial\mu)/\mathrm{d}t = \partial\xi/\partial\mu - (\partial\Omega/\partial\mu)(\partial\mu_1/\partial\theta)$.

состоит в «просветлении» области резонансного поглощения сигнальной волны при наличии мощной накачки, сопровождающемся также существенным уменьшением групповой скорости. С точки зрения общей теории колебаний и волн ЭИП является «параметрическим» вариантом известного в теории колебаний эффекта динамического демпфирования [9, 10]; прямым классическим аналогом ЭИП квантовых систем является ЭИП для волны в плазме на частоте циклотронного резонанса [9–12]. В последнем случае рассматривается распространение необыкновенной волны с частотой ω_1 в окрестности электронно-циклотронного резонанса вдоль магнитного поля; волна накачки (распространяющаяся в ту же сторону) имеет частоту ω_2, связанную с ω_1 условием параметрического резонанса $\omega_1 - \omega_2 = \omega_{\rm p}$, где $\omega_{\rm p}$ — плазменная частота. Все особенности распространения электронно-циклотронной волны в этих условиях описываются её «эффективным» показателем преломления, зависящим от интенсивности накачки как от параметра. Гидродинамическая теория этого параметрического эффекта предсказывает следующее выражение для показателя преломления [12]:

$$N_{\rm hyd}^2 = 1 - \frac{\omega_{\rm p}^2}{\omega_1}\;\frac{Z_{\rm p} + \xi_{\rm EC}\omega_{\rm L}\omega_2\,[2k_1/k_2 - 1 - \omega_1/(2\omega_2)]}{(\omega_1 - \omega_H + i\gamma)\,Z_{\rm p} - \omega_{\rm L}^2\omega_1\xi_{\rm EC}}\;, \tag{25}$$

где k_1 и k_2 — волновое число сигнальной волны и волны накачки соответственно, $\omega_{\rm L} = \omega_1 - \omega_2$, ω_H — электронная гирочастота, γ — эффективная частота соударений, $Z_{\rm p} = \omega_{\rm L}^2 - \omega_{\rm p}^2 - 3k_{\rm L}^2V_T^2 + i\gamma\omega_{\rm L}$, $\xi_{\rm EC} = |eE_2/[2m\,(\omega_2 - \omega_H + i\gamma)]|^2\,(k_2/\omega_2)^2$, E_2 — амплитуда электрического поля волны накачки, V_T^2 — квадрат тепловой скорости электронов, e и m — элементарный заряд и масса электрона соответственно ($e > 0$).

Стандартный подход в рамках кинетической теории, основанный на аналогичном описанному в разделе 1 асимптотическом решении кинетического уравнения с простейшим интегралом соударений (так называемое τ-приближение [13]), с точностью до квадратичных по полю накачки членов приводит к существенно более сложному выражению [12]:

$$N_{\rm kin}^2 = N_{0\rm kin}^2 - \xi_{\rm EC}\,\frac{A(\omega_1, k_1)}{D(\omega_{\rm L}, k_{\rm L}) - \xi_{\rm EC}B(\omega_1, k_1)}\;. \tag{26}$$

Здесь $N_{0\rm kin}^2$ — выражение для «линейного» показателя преломления необыкновенной волны, следующее из кинетической теории [1]:

$$N_{0\rm kin}^2 = 1 - \frac{\omega_{\rm p}^2}{\omega_1^2}\int\limits_{-\infty}^{+\infty}\frac{(\omega_1 - k_1V_\parallel)\,F_0 - k_1\,\partial(\langle V_\perp^2\rangle\,F_0)/\partial V_\parallel}{\omega_1 - \omega_H - k_1V_\parallel + i\gamma}\,{\rm d}V_\parallel, \tag{27}$$

$D(\omega_{\rm L}, k_{\rm L})$ — «продольная» диэлектрическая проницаемость магнитоактивной плазмы [1]:

$$D(\omega_{\rm L}, k_{\rm L}) = 1 + \frac{\omega_{\rm p}^2}{\omega_{\rm L}}\int\limits_{-\infty}^{+\infty}\frac{(\partial F_0/\partial V_\parallel)\,V_\parallel\,{\rm d}V_\parallel}{\omega_{\rm L} - k_{\rm L}V_\parallel + i\gamma}\;, \tag{28}$$

$k_{\rm L} = k_1 - k_2$, A и B — функции, не зависящие от $\xi_{\rm EC}$:

$$A = \frac{\omega_{\rm p}^2}{\omega_1^2}\left[\int\limits_{-\infty}^{+\infty}\frac{\omega_{\rm p}^2}{\omega_{\rm L}Z_{\rm L}}\left(\frac{k_1}{k_2}\;\frac{\omega_2}{Z_2^*} - \frac{\omega_1}{Z_1}\right)F_0\,{\rm d}V_\parallel\right]\left[\int\limits_{-\infty}^{+\infty}\frac{|Z_2^{(0)}|^2}{Z_1Z_{\rm L}}\left(\frac{\omega_2}{Z_2} - 1 + \frac{k_{\rm L}}{k_2}\;\frac{\omega_2Z_1}{Z_2Z_{\rm L}}\right)F_0\,{\rm d}V_\parallel\right],$$

$$B = \int\limits_{-\infty}^{+\infty}\frac{\omega_{\rm p}^2}{\omega_{\rm L}Z_{\rm L}}\;\frac{|Z_2^{(0)}|^2}{Z_1Z_2}\left[\frac{\omega_2}{Z_2} - 1 + \frac{Z_1\omega_1}{2Z_{\rm L}^2}\right]F_0\,{\rm d}V_\parallel, \tag{29}$$

$F_0(V_\parallel) = N_0^{-1} \int f_0 \,\mathrm{d}^2 V_\perp$ — невозмущённая функция распределения по продольным скоростям, нормированная на единицу, $Z_1 = \omega_1 - \omega_H - k_1 V_\parallel + i\gamma$, $Z_2 = \omega_2 - \omega_H - k_2 V_\parallel + i\gamma$, $Z_\mathrm{L} = \omega_\mathrm{L} - k_\mathrm{L} V_\parallel + i\gamma$, $\langle V_\perp^2 \rangle = \int f_0 V_\perp^2 \,\mathrm{d}^2 V_\perp \left(\int f_0 \,\mathrm{d}^2 V_\perp\right)^{-1}$, $Z_2^{(0)} = Z_2(V_T = 0)$. В пределе холодной столкновительной плазмы оба выражения сводятся к единой формуле:

$$N^2 = 1 - \frac{\omega_\mathrm{p}^2}{\omega_1} \, \frac{(\omega_\mathrm{L}^2 - \omega_\mathrm{p}^2 + i\gamma\omega_\mathrm{L}) + \xi_\mathrm{EC}\omega_\mathrm{L}\omega_2 \left(\dfrac{2k_1}{k_2} - 1 - \dfrac{\omega_1}{2\omega_2}\right)}{(\omega_1 - \omega_H + i\gamma)\,(\omega_\mathrm{L}^2 - \omega_\mathrm{p}^2 + i\gamma\omega_\mathrm{L}) - \omega_\mathrm{L}^2 \omega_1 \xi_\mathrm{EC}} \,, \tag{30}$$

хотя формально расчёт определяющих (26) квадратичных по полю вариаций функции распределения становится некорректным при стремлении тепловой скорости частиц V_T к нулю, а гидродинамика, естественно, такой предельный переход вполне допускает.

Итак, в этом разделе приведён пример корректности кинетического расчёта тока в плазме в квадратичном по полю приближении для $V_T \to 0$, причём этот пример относится к заметно более сложной системе, учитывающей трёхмерное движение частиц и (хотя бы модельно) соударения между ними.

4. ЗАКЛЮЧЕНИЕ

В качестве заключения отметим следующее. Стандартное решение кинетического уравнения методом возмущений по амплитудам высокочастотных полей формально справедливо при условии малости характерных вариаций обобщённого импульса $\Delta\mu$ по сравнению с размером области локализации ансамбля частиц в импульсном пространстве μ_T. Тем не менее при анализе проблем собственно электродинамики плазмы в подобных случаях отнюдь не всегда необходимо усложнять процедуру расчёта. Как показано в данной работе, корректность использования асимптотических «кинетических» выражений с целью определения не только линейных, но и квадратичных по полю вариаций моментов функции распределения (например, токов), по крайней мере, для весьма широкого круга задач определяется не соотношением между величинами $\Delta\mu$ и μ_T. Достаточно лишь малости $\Delta\mu$, обеспечивающей применимость метода возмущений по степеням амплитуд полей при решении уравнений движения [6].

Работа поддержана грантами РФФИ (№ 03–02–17234) и NWO-РФФИ (№ 047.016.016).

СПИСОК ЛИТЕРАТУРЫ

1. Ахиезер А. И., Ахиезер И. А., Половин Р. В. и др. Электродинамика плазмы. М.: Наука, 1974.
2. Гинзбург Н. С. // ЖТФ. 1986. Т. 56. С. 1 433.
3. Железняков В. В. // Изв. вузов. Радиофизика. 1960. Т. 3. С. 57.
4. Силин В. П. Параметрическое воздействие излучения большой мощности на плазму. М.: Наука, 1973.
5. Гапонов-Грехов А. В., Токман М. Д. // ЖЭТФ. 1997. Т. 112. С. 1 176.
6. Ерухимова М. А., Токман М. Д. // ЖЭТФ. 2000. Т. 118. С. 291.
7. Тимофеев А. В. // Вопросы теории плазмы / Под ред. М. А. Леонтовича. М.: Энергоатомиздат, 1985. Т. 14. С. 56.
8. Harris S. E. // Phys. Today. 1997. V. 50. P. 36.

[6] В зависимости от конкретной ситуации это условие может определяться, например, релятивистскими эффектами, соотношением между масштабом неоднородности полей и вызываемыми ими вариациями координат частиц, соотношением между изменением импульса $\Delta\mu$ и «средним» импульсом $\langle\mu\rangle$.

9. Литвак А. Г., Токман М. Д. // Изв. вузов. Радиофизика. 2001. Т. 44. С. 407.
10. Litvak A. G., Tokman M. D. // Phys. Rev. Lett. 2002. V. 88. Article no. 095 003.
11. Shvets G., Wurtele J. S. // Phys. Rev. Lett. 2002. V. 89. Article no. 115 003.
12. Крячко А. Ю., Литвак А. Г., Токман М. Д. // ЖЭТФ. 2002. Т. 122. С. 805.
13. Ландау Л. Д., Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. М.: Наука, 1979.

Поступила в редакцию 12 апреля 2004 г.; принята в печать 5 ноября 2004 г.

ON APPLICATION OF PERTURBATION THEORY FOR KINETIC ANALYSIS OF NONLINEAR PROBLEMS OF PLASMA ELECTRODYNAMICS

M. D. Tokman and A. Yu. Kryachko

The correctness of perturbation theory (decomposition by the degrees of HF field) for the investigation of kinetic equation is discussed. We consider the conditions, when this approach can not be applied formally for the strongly localized distribution function. It is shown, that not only linear, but also quadratic terms remain, nevertheless, correct in the asymptotic expressions for the moments of distribution function in this case.

УДК 535.2-4+535.854

ВЛИЯНИЕ РЕГЕНЕРАТИВНОГО УСИЛИТЕЛЯ НА ЧУВСТВИТЕЛЬНОСТЬ РЕЗОНАНСНОГО ВОЛОКОННОГО КОЛЬЦЕВОГО ИНТЕРФЕРОМЕТРА К ВРАЩЕНИЮ

И. А. Андронова, П. А. Шилягин

Институт прикладной физики РАН, г. Нижний Новгород, Россия

Рассмотрена возможность использования регенеративного усилителя на активном волокне, размещённого внутри резонатора, для повышения чувствительности к вращению схемы резонансного волоконного кольцевого интерферометра (РВКИ) с наружным отражателем и широкополосным источником излучения. Рассмотрено влияние шумов усилителя на чувствительность схмеы РВКИ к вращению. Показано, что использование регенеративного усилителя позволяет получить максимальную чувствительность при меньшей мощности источника за счёт компенсации потерь в резонаторе. Проведено сравнение с чувствительностью традиционной нерезонансной схемы волоконного кольцевого интерферометра.

В основе волоконных оптических гироскопов лежит эффект Саньяка, который состоит в том, что разность фаз $\varphi_{\rm s}$ встречных волн, распространяющихся в замкнутом контуре, связана со скоростью вращения $\Omega_{\rm s}$ контура соотношением

$$\varphi_{\rm s} = \frac{8\pi NS}{\lambda c}\,\Omega_{\rm s} = \frac{4\pi RL}{\lambda c}\,\Omega_{\rm s} = \frac{2L^2}{\lambda cN}\,\Omega_{\rm s}, \tag{1}$$

где $S = \pi R^2$ — площадь контура, R — его радиус, $L = 2\pi RN$ — длина волокна, N — число витков, λ и c — длина волны и скорость света соответственно. К настоящему времени физические исследования в области традиционной волоконной гироскопии на основе эффекта Саньяка в однопроходных волоконных кольцевых интерферометрах (ВКИ) в общих чертах завершены. Однако проблема повышения чувствительности в оптической гироскопии остаётся актуальной. Решение этой проблемы с использованием резонансных волоконных кольцевых интерферометров (РВКИ), амплитудные и фазовые характеристики которых при высокой добротности весьма чувствительны к появлению саньяковской разности фаз [1–8], встретило ряд трудностей. Эти трудности были связаны с использованием лазерных источников излучения и необходимостью подстройки частоты генерации лазера к резонансной частоте волоконного резонатора, а также с влиянием когерентных рассеянных сигналов на измеряемую разность фаз и др. Поэтому исследование РВКИ проходило менее активно, чем исследование традиционных однопроходных гироскопов. Идея использования в РВКИ широкополосных источников излучения [9, 10], которые значительно раньше стали применяться в традиционных ВКИ для уменьшения влияния эффектов обратного рассеяния, позволила надеяться на реализацию потенциальных возможностей резонансных схем с существенно большей чувствительностью к вращению, чем в традиционных ВКИ. К настоящему времени в литературе появился целый ряд работ по исследованию различных схем РВКИ с широкополосными источниками излучения и внутрирезонаторными усилителями [11–15]. Однако проблема повышения чувствительности к вращению в волоконной гироскопии за счёт использования различных схем РВКИ, в том числе с применением волоконных усилителей, к настоящему времени далека от завершения как в экспериментальном, так и в теоретическом плане.

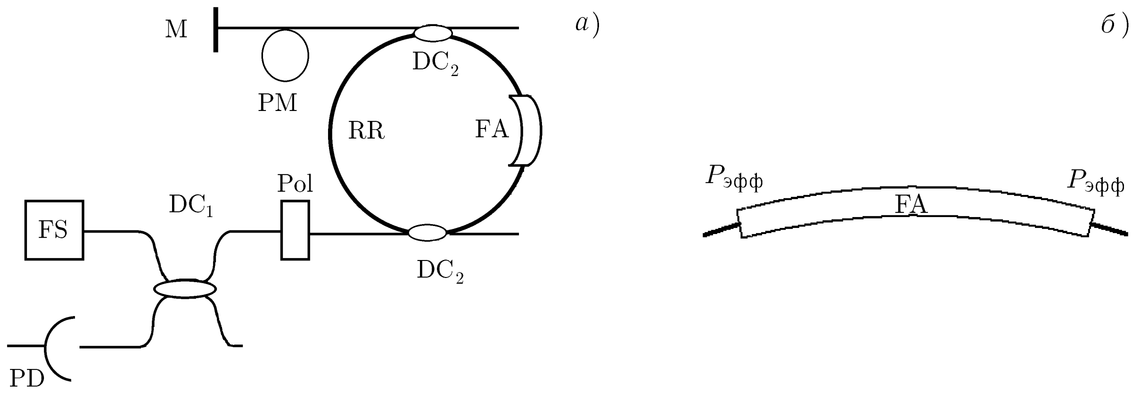

Рис. 1. (*а*) Схема резонансного волоконного кольцевого интерферометра (амплитудный вариант): FS — широкополосный источник излучения, RR — многовитковый волоконный контур, DC_1 — светоделитель, DC_2 — светоделители-отражатели, M — наружный отражатель, FA — волоконный усилитель, PM — модулятор, PD — фотодетектор, Pol — поляризатор. (*б*) Волоконный усилитель

Цель данной работы — исследование возможности увеличения чувствительности сигнала на выходе схемы резонансного волоконного интерферометра с наружным отражателем и модулятором (рис. 1*а*) к вращению при использовании внутрирезонаторного усилителя на активном волокне, а также оптимизация параметров схемы с учётом шумов источника и усилителя для получения максимальной чувствительности и сравнение с характеристиками традиционной нерезонансной схемы волоконного кольцевого интерферометра (ВКИ).

Рассматриваемая схема РВКИ представлена на рис. 1*а*. В состав схемы входят широкополосный источник (FS), многовитковый волоконный контур (RR), волоконный усилитель (FA), два светоделителя-отражателя (DC_2), наружный отражатель (M) и модулятор (PM). Усилитель представляет собой вваренный в волоконный контур отрезок активного волокна с собственной накачкой.

Специфика этой схемы состоит в том, что сигнал вращения на выходе определяется не фазовыми, а амплитудными характеристиками резонатора, т. к. представляет собой сигнал, прошедший интерферометр последовательно в прямом, а после отражения от наружного зеркала и прохождения модулятора — во встречном направлении. По мере увеличения вращения интенсивность сигнала на выходе будет падать из-за смещения резонансных кривых по частоте для волн встречных направлений относительно положения покоя. При малых скоростях вращения выходной сигнал, как и в традиционной схеме ВКИ, оказывается чётной функцией скорости и не зависит от направления вращения. Поэтому необходимо использовать методику, позволяющую определять направление вращения и получать линейную или близкую к линейной зависимость от скорости вращения. Данная схема ранее рассматривалась в работе [16], где был предложен компенсационный метод, в котором скорость вращения определяется по смещению частоты сигнала, отражённого от наружного зеркала, необходимому для восстановления первоначального (без вращения) уровня интенсивности сигнала на выходе схемы. Компенсационный метод привлекателен тем, что позволяет измерять большой диапазон скоростей вращения, ограниченный только возможностями элемента для смещения частоты, но этот метод усложняет и существенно удорожает схему. В работе [16] было продемонстрировано измерение величины и знака скорости вращения в РВКИ с использованием более простого и дешёвого метода модуляции отражённого сигнала. Однако вопросы оптимизации параметров схемы и возможность использования внутрирезонаторного (регенеративного) усилителя для повышения предельной чувствительности модуляционного метода в этой работе не затрагивались.

В данной работе расчёт сигнала на выходе схемы РВКИ проводится для одной поляризационной моды, что означает отсутствие разъюстировки осей анизотропии элементов схемы и отсутствие связи поляризационных мод. Размещение в резонаторе усилителя позволяет скомпенсировать потери в элементах схемы, увеличить добротность, не уменьшая коэффициент пропускания зеркала, и за счёт этого повысить мощность сигнала на выходе и чувствительность схемы к вращению. Однако введение усилителя в резонатор приводит к появлению дополнительных шумовых источников, и их влияние на чувствительность схемы к вращению требует отдельного рассмотрения. В соответствии с [17] шум усилителя может быть представлен шумовым источником на входе усилителя со спектральной плотностью мощности на частоте ν

$$S_{\text{эфф}}(\nu) = Kh\nu, \tag{2}$$

где K — коэффициент, который зависит от инверсии населённостей и коэффициента усиления активной среды волоконного усилителя, h — постоянная Планка. С ростом инверсии и коэффициента усиления $K \to 1$. Таким образом, с введением усилителя в схему внутри резонатора появляются два дополнительных широкополосных источника излучения (рис. 1*б*).

Для выяснения принципиальных физических закономерностей будем рассматривать случай, когда спектры источника и усилителя равномерны и совпадают по ширине, а усилитель работает в линейном режиме. Выделим одну спектральную компоненту источника излучения на входе интерферометра $E_0(\omega)\exp(i\omega t)$ на частоте ω и проследим за её распространением в схеме РВКИ (рис. 1*а*) от входа до выхода. В результате сигнал на выходе схемы в соответствии с порядком прохождения элементов в прямом (+) и обратном (−) направлении запишется в виде

$$E(\omega) = E_0(\omega)\exp(i\omega t)T^+(\omega)\exp[i\varphi^+(\omega)]\, r \exp[i\Phi\sin(\Omega t)]T^-(\omega)\exp[i\varphi^-(\omega)], \tag{3}$$

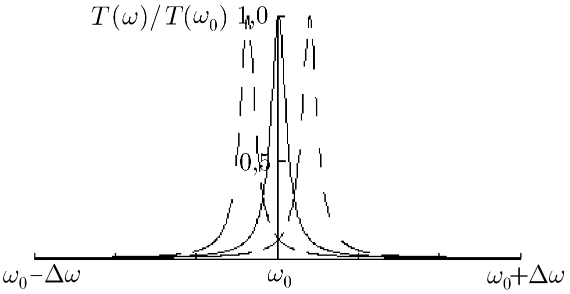

Рис. 2. Зависимость пропускания резонатора T от частоты ω в пределах одного резонанса без вращения (сплошная линия) и при наличии вращения для встречных волн (пунктир). Графики нормированы на $T(\omega_0)$

где $T^+(\omega)$, $T^-(\omega)$ и $\varphi^+(\omega)$, $\varphi^-(\omega)$ — амплитудные и фазовые коэффициенты прохождения через кольцевой резонатор на частоте ω, различие которых для волн встречных направлений связано с эффектом Саньяка, r — коэффициент отражения наружного зеркала, Φ и Ω — амплитуда и частота фазовой модуляции соответственно. Коэффициенты $T^+(\omega)$, $T^-(\omega)$ выразим через параметры резонатора в соответствии с [18]:

$$T^\pm(\omega)\exp[i\varphi^\pm(\omega)] = \frac{t_0\sqrt{R_\text{f}}\,\exp(i\,\delta^\pm/2)}{1 - R_\text{f}\exp(i\,\delta^\pm)}\ , \tag{4}$$

$$T^\pm(\omega) = \frac{t_0\eta}{2}\,\frac{1}{\sqrt{1+\eta^2\sin^2(\delta^\pm/2)}}\ , \tag{5}$$

где $\delta^\pm(\omega) = kL \pm \varphi_\text{s} = Ln\omega/c \pm \varphi_\text{s}$ — набег фазы за один обход резонатора, n — показатель преломления волокна, $c/(Ln)$ — межмодовый интервал, $R_\text{f} = \alpha^2 t_\text{c}^2 t_\text{f} g$ — ослабление поля за один обход резонатора (условие отсутствия генерации $R_\text{f} < 1$), α — коэффициент отражения светоделителя по амплитуде, t_c, t_f и $g(\omega)$ — потери в светоделителях, волокне и усиление за один обход контура соответственно, $t_0 = t_\text{c}\,(1-\alpha^2)/\alpha$, $\Delta\nu_\text{R} = c\,(1-R_\text{f})/(\pi Ln\sqrt{R_\text{f}})$ — полоса резонатора, $\eta = 2\sqrt{R_\text{f}}\,/(1-R_\text{f}) = 2c/(\pi Ln\,\Delta\nu_\text{R})$ — резкость резонансной кривой (обозначения соответствуют работам [19–21]). Характеристики пропускания резонатора для волн встречных направлений представлены на рис. 2 в отсутствие (сплошная кривая) и при наличии (пунктир) вращения. Разность резонансных частот для волн встречных направлений $\nu^+ - \nu^- = \varphi_{\rm s} c/(2L) = 2\pi R\Omega_{\rm s}/\lambda$ не зависит от L. Однако чувствительность к вращению, определяемая по смещению резонансов, будет расти с ростом L и уменьшением полосы резонатора $\Delta\nu \sim c/L$.

Полагая амплитуду фазовой модуляции Φ не слишком большой, при расчётах будем учитывать только три компоненты выходного сигнала $E(\omega) = d_1 + d_2 + d_3$ на частотах ω, $\omega + \Omega$ и $\omega - \Omega$, которые имеют вид

$$d_1 = \frac{1}{2}\, E_0(\omega) r J_0(\Phi) T^+(\omega) T^-(\omega) \cos[\omega t + \varphi^+(\omega) + \varphi^-(\omega)],$$

$$d_2 = \frac{1}{2}\, E_0(\omega) r J_1(\Phi) T^+(\omega) T^-(\omega + \Omega) \cos[(\omega + \Omega)\, t + \varphi^+(\omega) + \varphi^-(\omega + \Omega)],$$

$$d_3 = \frac{1}{2}\, E_0(\omega) r J_1(\Phi) T^+(\omega) T^-(\omega - \Omega) \cos[(\omega - \Omega)\, t + \varphi^+(\omega) + \varphi^-(\omega - \Omega)], \tag{6}$$

где J_0 и J_1 — функции Бесселя первого рода.

Полезный сигнал интерференции на выходе фотоприёмника на частоте модуляции Ω от одной спектральной компоненты источника на частоте $\omega \propto d_1 d_2^* + d_1 d_3^* + d_2 d_1^* + d_3 d_1^*$ в соответствии с (6) запишется в виде

$$I_\Omega(\omega) = \frac{1}{8}\, S(\omega) r^2 J_0(\Phi) J_1(\Phi)\, [T^+(\omega)]^2\, T^-(\omega)\, [T^-(\omega + \Omega) - T^-(\omega - \Omega)]. \tag{7}$$

В (7) и ниже опускаем коэффициент, определяющий связь между фототоком и оптической мощностью на выходе интерферометра, который при определении порогового сигнала сокращается. Постоянная составляющая сигнала интерференции в фототоке на выходе интерферометра от одной спектральной компоненты источника на частоте $\omega \propto d_1 d_1^* + d_2 d_2^* + d_3 d_3^*$ в соответствии с (6) запишется в виде

$$I_=(\omega) = \frac{1}{8}\, S(\omega) r^2 J_0^2(\Phi)\, [T^+(\omega)]^2 \left\{[T^-(\omega)]^2 + [T^-(\omega + \Omega)]^2 + [T^-(\omega - \Omega)]^2\right\}. \tag{8}$$

В дальнейшем для простоты расчётов выбираем Φ таким образом, что $J_0(\Phi) = J_1(\Phi)$, что соответствует $\Phi \approx 1{,}45$. Результирующий полезный и шумовой сигналы на выходе РВКИ находятся интегрированием сигнала от одной спектральной компоненты $I(\omega)$ по спектру источника излучения:

$$I_\sim(\Omega) = \int\limits_0^\infty I_\Omega(\omega)\, {\rm d}\omega, \qquad I_= = \int\limits_0^\infty I_=(\omega)\, {\rm d}\omega. \tag{9}$$

В рассматриваемом случае на ширине спектра источника укладывается много резонансов кольцевого резонатора, поэтому $I_\sim(\Omega)$ и $I_=$ представляют собой выражения вида $\int_0^\infty S(\omega) F(\omega)\, {\rm d}\omega$, где $S(\omega)$ — спектр источника, $F(\omega)$ — периодическая функция пропускания резонатора с периодом, равным межмодовому интервалу $\Delta\omega = c/(Ln)$. Поскольку в данной задаче спектр источника принят равномерным: $S(\omega) = S_0$, перейдём от интегрирования по спектру к сумме интегралов по отдельному резонансу и, учитывая число этих резонансов на ширине спектра источника, получим

$$\int\limits_0^\infty S(\omega) F(\omega)\, {\rm d}\omega \approx N S_0 \int\limits_{\omega_0 - \Delta\omega}^{\omega_0 + \Delta\omega} F(\omega)\, {\rm d}\omega,$$

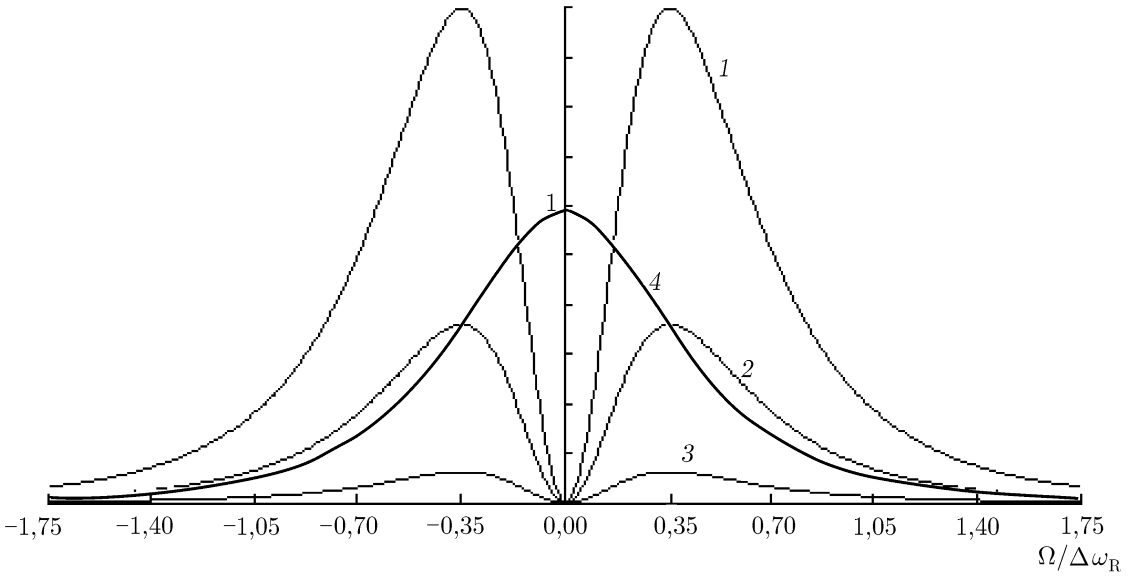

Рис. 3. Зависимость интенсивности сигнала вращения (кривые *1–3*) и спектральной плотности шума (кривая *4*) в относительных единицах от частоты модуляции при $R_{\mathrm{f}} = 0{,}89$ ($\Delta\nu_{\mathrm{R}} = 72{,}4$ кГц): кривая *1* соответствует $\varphi_{\mathrm{s}} = 6\cdot 10^{-8}$, *2* — $\varphi_{\mathrm{s}} = 4{,}8\cdot 10^{-8}$, *3* — $\varphi_{\mathrm{s}} = 2\cdot 10^{-8}$

где S_0 и Π — спектральная амплитуда и ширина полосы источника излучения, ω_0 — одна из резонансных частот, $N = \Pi Ln/c$ — число резонансов в спектре источника. Введём новую переменную интегрирования $\delta = kL = \omega\,(Ln/c)$, тогда подынтегральная функция $F(\omega)$ заменяется новой функцией $F(\delta)$ с периодом 2π, и искомое выражение перепишется в виде

$$\int\limits_0^\infty S(\omega)F(\omega)\,\mathrm{d}\omega \approx NS_0 \int\limits_{(m-1)\,\pi}^{(m+1)\,\pi} F(\delta)\,\mathrm{d}\delta\ \frac{c}{Ln} = NS_0 \int\limits_{-\pi}^{\pi} F(\delta)\,\mathrm{d}\delta\ \frac{c}{Ln}\,. \tag{10}$$

По этой формуле и будут проводиться дальнейшие расчёты.

Для определения пороговой чувствительности в схеме необходимо на выходе фотоприёмника рассчитать полезную составляющую сигнала на частоте модуляции Ω и шум на этой частоте в полосе приёма. Интегральная форма полезного сигнала на выходе фотоприёмника на частоте модуляции Ω после замены ω на δ записывается в виде

$$I_\sim(\Omega) = N\ \frac{c}{Ln}\ \frac{1}{8}\ Sr^2 J_0^2(\Phi) \int\limits_{-\pi}^{\pi} [T^+(\delta)]^2\, T^-(\delta) \left[T^-\Big(\delta + \frac{Ln}{c}\ \Omega\Big) - T^-\Big(\delta - \frac{Ln}{c}\ \Omega\Big)\right] \mathrm{d}\delta. \tag{11}$$

Можно показать, что $I_\sim(\Omega) = 0$ при $\varphi_{\mathrm{s}} = 0$ и $I_\sim(\Omega)$ — нечётная функция φ_{s}. Это позволяет определять направление вращения, что невозможно при отсутствии модуляции. Результаты численного интегрирования (11) представлены на рис. 3, где приведены зависимости величины полезного сигнала от частоты модуляции при трёх различных значениях скорости вращения $\Omega_{\mathrm{s}} \propto \varphi_{\mathrm{s}}$, что позволяет определить оптимальную частоту модуляции как частоту, на которой интенсивность сигнала при заданной величине φ_{s} будет наибольшей. Из рис. 3 видно, что при малых значениях φ_{s}, которые представляют основной интерес, шум не зависит от φ_{s}. Оптимальная частота модуляции Ω_{opt} связана с полосой резонатора соотношением

$$\Omega_{\mathrm{opt}} = 0{,}35\,\Delta\omega_{\mathrm{R}} = 0{,}7\ \frac{c}{Ln}\ \frac{1-R_{\mathrm{f}}}{\sqrt{R_{\mathrm{f}}}}\ , \qquad \Delta\omega_{\mathrm{R}} = 2\pi\,\Delta\nu_{\mathrm{R}}. \tag{12}$$

При достаточно малых скоростях вращения ($c\,\Delta\varphi_{\rm s}/L \ll \Delta\nu_{\rm R}$) интенсивность полезного сигнала после интегрирования можно представить в виде аналитического выражения:

$$I_\sim(\Omega) = k_1 J_0^2(\Phi) \left(P_0 + \frac{1}{t_0}\ P_{\text{эфф}} \right) t_0^4\ \frac{(2+\eta^2)\,\eta^7}{(1+\eta^2)^{5/2}}\ \varphi_{\rm s}, \tag{13}$$

где $k_1 = 7 \cdot 10^{-3}$, P_0 — мощность источника на входе схемы, $P_{\text{эфф}}$ — мощность эффективного дополнительного источника на входе усилителя, которая находится путём интегрирования (2) в пределах полосы усиления.

Интегральное выражение для интенсивности постоянной составляющей имеет вид

$$I_= = N\ \frac{c}{Ln}\ \frac{1}{8}\ Sr^2 J_0^2(\Phi) \int\limits_{-\pi}^{\pi} [T^+(\delta)]^2 \left\{ [T^-(\delta)]^2 + \left[T^-\left(\delta + \frac{Ln}{c}\ \Omega\right)\right]^2 + \left[T^-\left(\delta - \frac{Ln}{c}\ \Omega\right)\right]^2 \right\} {\rm d}\delta. \tag{14}$$

В результате интегрирования при малых $\varphi_{\rm s}$ получим аналитическое выражение для интенсивности постоянной составляющей:

$$I_= = k_2 J_0^2(\Phi) \left(P_0 + \frac{2}{t_0}\ P_{\text{эфф}} \right) t_0^4\ \frac{(2+3\eta^2+\eta^4)\,\eta^4}{(1+\eta^2)^{5/2}}\,, \qquad k_2 = 7 \cdot 10^{-2}. \tag{15}$$

Перейдём к рассмотрению дополнительных («избыточных») шумов, возникающих в результате детектирования широкополосного излучения, прошедшего волоконный резонатор. В соответствии с [22], избыточный шум каждой из трёх компонент сигнала на выходе фотоприёмника может быть записан в виде свёртки по частоте квадратов элементов (6):

$$\overline{D_i^2}(\nu) = \int\limits_0^\infty \overline{d_i^2(\omega)}\,\overline{d_i^2(\omega+\nu)}\,{\rm d}\omega, \tag{16}$$

где черта обозначает статистическое усреднение. Полный избыточный шум схемы представляет собой сумму:

$$\overline{D_\nu^2} = \sum_{i=1}^{3} \overline{D_i^2}(\nu). \tag{17}$$

При оптимальной частоте модуляции (12) и малых $\varphi_{\rm s}$ в результате интегрирования выражения (17) получаем аналитическое выражение для величины избыточных шумов в полосе 1 Гц:

$$\overline{D_\nu^2} = k_3 J_0^4(\Phi) \left[P_0^2 + \left(1 + \frac{1{,}7}{t_0^2}\right) \frac{1}{t_0^2}\ P_{\text{эфф}}^2 \right] t_0^8\ \frac{1}{\Pi}\ \frac{(8+8\eta^2+5\eta^4)\,(2+\eta^2)\,\eta^8}{(1+\eta^2)^{7/2}}\ , \tag{18}$$

где $k_3 = 5 \cdot 10^{-5}$.

Пороговая чувствительность определяется из условия равенства единице отношения сигнала $I_\sim(\Omega)$ к квадратному корню из спектральной плотности суммарной шумовой составляющей фототока на выходе РВКИ на частоте модуляции Ω, которое записывается в виде

$$\frac{I_\sim(\Omega)}{\sqrt{N(\Omega)\,\Delta\nu}} = 1, \tag{19}$$

где спектральная плотность шума $N(\Omega)$ представляет сумму двух компонент — «избыточной» и дробовой: $N(\Omega) = \overline{D_\Omega^2} + 2eI_=$, e — элементарный заряд, $\Delta\nu$ — полоса приёма. В дальнейшем пороговый сигнал рассматривается на оптимальной частоте модуляции, что соответствует на рис. 3 пересечению кривых интенсивности сигнала и спектральной плотности шума в максимуме по-

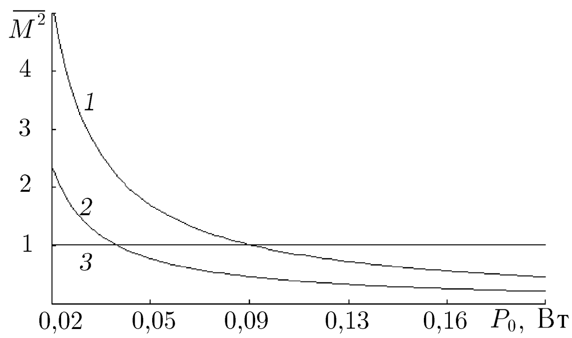

Рис. 4. Зависимость глубины модуляции дробового (кривые *1, 2*) и избыточного (кривая *3*) шума на выходе РВКИ от мощности излучения на выходе: *1* — без усиления ($\alpha = 0{,}896$), *2* — с усилением 12,2 % ($\alpha = 0{,}846$) при $\eta = 6$

лезного сигнала. Следует отметить, что глубина модуляции избыточной компонентой шума $\overline{M_\Pi^2} = \overline{D_\Omega^2}/I_=^2$ не зависит от мощности излучения на выходе РВКИ, в то время как глубина модуляции дробовой компоненты $\overline{M_=^2} = 2e/I_=$ обратно пропорциональна указанной величине, что иллюстрирует рис. 4.

Таким образом, с увеличением интенсивности сигнала на выходе схемы за счёт мощности источника или внутрирезонаторного усилителя можно достичь интенсивности, для которой глубина модуляции дробового шума меньше избыточного: $\overline{M}_\Pi \gg \overline{M}_=$. В этом случае на выходе можно ограничиться рассмотрением только избыточных шумов источника, что в значительной мере упрощает задачу нахождения порогового сигнала вращения. Определим мощность $P_{0\,\text{пор}}$, начиная с которой можно не учитывать дробовой шум, соотношением $\overline{M}_\Pi/\overline{M}_= = 5$. Величину $P_{0\,\text{пор}}$ при малых $\varphi_{\rm s}$ можно получить в аналитической форме из (16) и (19):

$$P_{0\,\text{пор}} = k_4\, \frac{e\Pi}{J_0^2(\Phi)t_0^4} \left[\frac{(\eta^4+3\eta^2+2)\,(1+\eta^2)}{\eta^4\,(5\eta^4+8\eta^2+8)\,(2+\eta^2)}\right], \qquad k_4 = 1{,}2\cdot 10^4~\text{Вт/А}. \tag{20}$$

Из (20) можно показать, что при заданных характеристиках светоделителя, определяемых параметром t_0, и при $\eta \gg 1$ мощность, при которой влиянием дробовых шумов можно пренебречь, пропорциональна η^{-4}, т. е. при росте резкости резонансной кривой за счёт усиления мощность источника, при которой можно пренебречь дробовым шумом, уменьшается. Пороговое значение скорости вращения в диапазоне выходных мощностей, для которых $\overline{M}_\Pi \gg \overline{M}_=$, аналитически можно представить в следующем виде:

$$\Omega_{\rm sp}^{\rm R} = \frac{\lambda c}{4\pi RL}\,\frac{1}{\sqrt{\Pi}}\,F(\eta) G\biggl(\frac{P_\text{эфф}}{P_0}\biggr), \tag{21}$$

где в соответствии с (15) и (19)

$$F(\eta) = \frac{(1+\eta^2)^{3/4}\sqrt{8+8\eta^2+5\eta^4}}{\sqrt{2+\eta^2}\;\eta^3}, \tag{22}$$

$$G\biggl(\frac{P_\text{эфф}}{P_0}\biggr) = \frac{1+P_\text{эфф}/(t_0P_0)}{\sqrt{1+\left(1+\dfrac{1{,}7}{t_0^2}\right)\dfrac{1}{t_0^2}\left(\dfrac{P_\text{эфф}}{P_0}\right)^2}}\,. \tag{23}$$

Нетрудно показать, что $F(\eta)$ при больших значениях η убывает пропорционально $\eta^{-1/2}$, и $\Omega_{\rm sp}^{\rm R}$ пропорциональна $\sqrt{\Delta\nu_{\rm R}}$; функция G растёт с ростом аргумента. При $P_0 > P_{0\,\text{пор}}$ и $P_\text{эфф} = 0$ имеем $G(0) = 1$, пороговая скорость вращения $\Omega_{\rm sp}^{\rm R}$ не зависит от мощности источника P_0.

Для сравнения с чувствительностью традиционной (нерезонансной) схемы ВКИ запишем в соответствии с [23] выражение для пороговой скорости вращения:

$$\Omega_{\rm sp}^{\rm T} = \frac{\lambda c\sqrt{e/I_= + 1/\Pi}\;\Delta\nu}{2\pi RLJ_1[2\Phi\sin(\Omega\tau/2)]}, \tag{24}$$

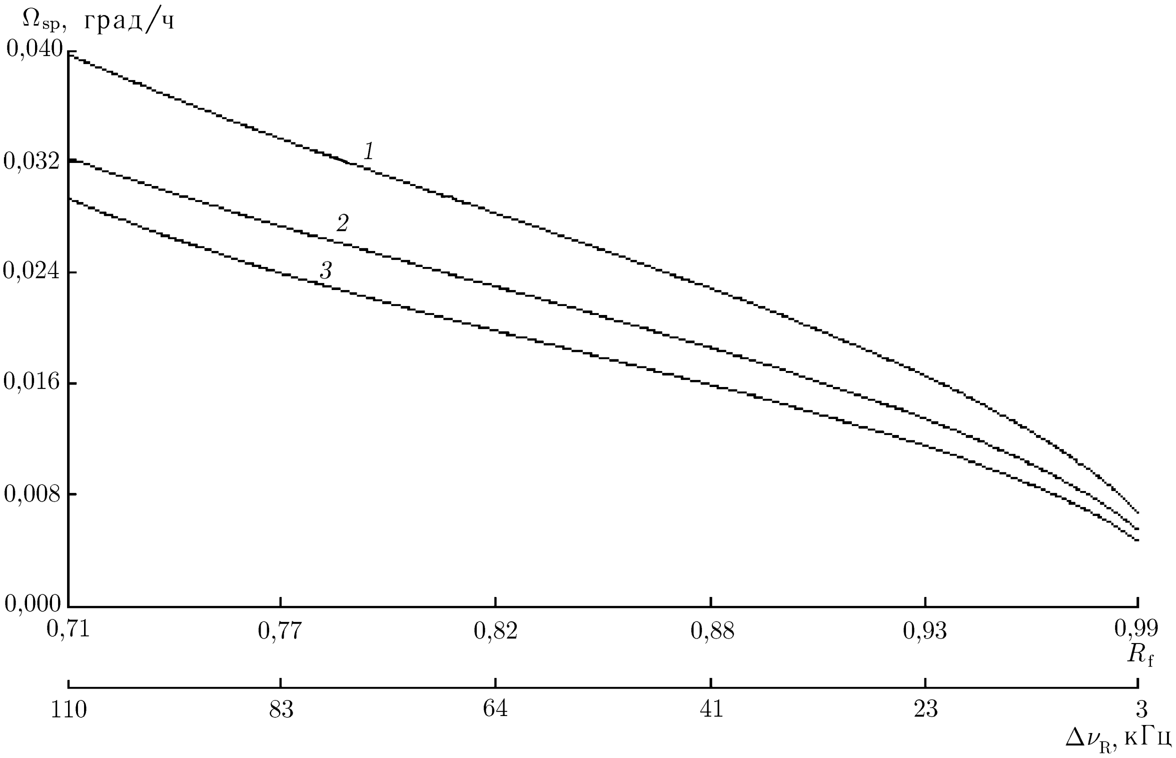

Рис. 5. Зависимость пороговой скорости вращения от $R_{\rm f}$ (или от ширины полосы резонатора) при частоте модуляции $\Omega = \Omega_{\rm opt} = 0{,}35\,\Delta\omega_{\rm R}$: кривая *1* соответствует $P_{\text{эфф}}/P_0 = 0{,}157$ ($P_0 = 3$ мВт), *2* — $P_{\text{эфф}}/P_0 = 0{,}094$ ($P_0 = 5$ мВт), *3* — $P_{\text{эфф}} = 0$ (нешумящий усилитель)

где $\Delta\nu$ — полоса приёма сигнала, $\tau = Ln/c$. Из (24) видно, что при $e/I_= \ll 1/\Pi$ (т. е. при $\overline{M}_\Pi \gg \gg \overline{M}_=$) $\Omega_{\rm sp}^{\rm T} \sim \sqrt{1/\Pi}$. Отношение порогового сигнала для традиционной схемы ВКИ и резонансной схемы (РВКИ) получим, используя (21) и (24) при одних и тех же параметрах схемы (длине волокна, радиусе намотки, ширине полосы и мощности источника и $\overline{M}_\Pi \gg \overline{M}_=$). Это соотношение имеет вид

$$\frac{\Omega_{\rm sp}^{\rm R}}{\Omega_{\rm sp}^{\rm T}} \approx F(\eta) G(P_{\text{эфф}}/P_0)\, J_1[2\Phi \sin(\Omega\tau/2)]. \qquad (25)$$

Анализ (25) показывает, что выигрыш в чувствительности для резонансной схемы определяется параметром резкости η и отношением мощности источника широкополосного излучения P_0 к эффективной мощности усилителя $P_{\text{эфф}}$. На рис. 5 представлены зависимости порогового сигнала при трёх различных аргументах функции G. Как видно из рис. 5, учёт шумов усилителя, величина которых определялась из аппроксимации к малым усилениям результатов экспериментальной работы [24], снижает пороговую чувствительность в зависимости от мощности источника излучения. Так, при мощности 5 мВт шумы усилителя понижают чувствительность на 14 %, а при мощности 3 мВт — на 35 %.

Приведённая ниже табл. 1 позволяет проанализировать влияние ряда параметров на оптимальную частоту модуляции Ω, область линейной зависимости выходного сигнала $\Delta\Omega_{\rm lin}$ и минимальную обнаруживаемую скорость вращения резонансного кольцевого интерферометра с усилением ($\Omega_{\rm sp}^{\rm R}$) и без усиления ($\tilde{\Omega}_{\rm sp}^{\rm R}$) и провести сравнение с традиционной нерезонансной схемой ВКИ. Таблица 1 получена с использованием выражения (25) для следующих параметров схемы: $R = 0{,}05$ м, $L = 200$ м, $\lambda = 0{,}8$ мкм, полоса длин волн источника излучения $\Delta\lambda = 20$ нм, $\Phi = 1{,}45$, мощность источника $P_0 = 1$ мВт.

Таблица 1

g	(R_f)	η	Ω_opt, кГц	$\Delta\Omega_\mathrm{lin}$, град/ч	φ_sp, рад	$\Omega_\mathrm{sp}^\mathrm{R}$, град/ч	$\tilde{\Omega}_\mathrm{sp}^\mathrm{R}$, град/ч
1,58	(0,75)	6,9	64,3	$25{,}6\cdot10^{3}$	$1{,}37\cdot10^{-7}$	$28{,}0\cdot10^{-3}$	0,35
1,68	(0,80)	8,9	49,8	$20{,}1\cdot10^{3}$	$10{,}10\cdot10^{-8}$	$22{,}0\cdot10^{-3}$	0,49
1,79	(0,85)	12,3	36,3	$15{,}4\cdot10^{3}$	$8{,}85\cdot10^{-8}$	$18{,}0\cdot10^{-3}$	0,96
1,87	(0,89)	17,2	25,3	$11{,}8\cdot10^{3}$	$7{,}43\cdot10^{-8}$	$15{,}0\cdot10^{-3}$	10,70
2,00	(0,95)	39,0	11,4	$7{,}9\cdot10^{2}$	$4{,}86\cdot10^{-8}$	$10{,}0\cdot10^{-3}$	—
2,04	(0,97)	65,7	5,0	$3{,}9\cdot10^{2}$	$3{,}74\cdot10^{-8}$	$7{,}71\cdot10^{-3}$	—
Нерезонансный интерферометр					$4{,}46\cdot10^{-7}$	0,176	0,176

В табл. 1 g — усиление по полю за один проход, Ω_opt — оптимальная частота модуляции, $\Delta\Omega_\mathrm{lin}$ — область линейной зависимости выходного сигнала от скорости вращения, η — резкость резонансной кривой, φ_sp — пороговое значение φ_s, $\Omega_\mathrm{sp}^\mathrm{R}$ — пороговая угловая скорость вращения с усилением, $\tilde{\Omega}_\mathrm{sp}^\mathrm{R}$ — в отсутствие усилителя ($g = 1$) при росте R_f за счёт отражения α.

Сравнение пороговых значений $\Omega_\mathrm{sp}^\mathrm{R}$ и $\tilde{\Omega}_\mathrm{sp}^\mathrm{R}$ при одинаковой добротности резонатора, приведённых в табл. 1, показывает, что при постоянной мощности на входе интерферометра (1 мВт) пороговый сигнал существенно выше в случае, когда увеличение добротности осуществляется за счёт роста отражения светоделителя (α), чем за счёт введения в резонатор усиления (g). В первом случае происходит уменьшение мощности на выходе, возрастание дробовой составляющей шума и рост $\tilde{\Omega}_\mathrm{sp}^\mathrm{R}$. Во втором случае рост добротности сопровождается ростом мощности на выходе, уменьшением дробовой составляющей шума и значительным уменьшением $\Omega_\mathrm{sp}^\mathrm{R}$. Из табл. 1 также видно, что пороговая скорость вращения в традиционной схеме ВКИ ($\Omega_\mathrm{sp}^\mathrm{T} = 0{,}176$ град/ч) заметно выше, чем в резонансной схеме с усилением ($\Omega_\mathrm{sp}^\mathrm{T} = (7{,}71\div28{,}0)\cdot10^{-3}$ град/ч). Без усиления чувствительность рассмотренной резонансной схемы при заданной мощности $P_0 = 1$ мВт ниже, чем традиционной. Таким образом, уменьшение порогового сигнала при неизменных параметрах схемы РВКИ возможно только за счёт уменьшения дробовой составляющей шума либо путём увеличения мощности источника излучения P_0, либо за счёт внутрирезонаторного усилителя. Нетрудно показать из (20), что для получения выходной мощности, необходимой для выхода из области влияния дробовых шумов и получения максимальной чувствительности, внутрирезонаторное размещение усилителя требует существенно меньших коэффициентов усиления, чем при его размещении непосредственно за источником.

Следует отметить, что весь вышеприведённый анализ РВКИ с усилителем относится к линейному режиму усилителя. Вопросы насыщения усиливающей среды в этой работе не рассматриваются, однако приближение к порогу генерации может приводить к нелинейным эффектам в активной среде усилителя и к изменению его частотных и шумовых характеристик. Кроме того, увеличение добротности резонатора за счёт введения усилителя налагает ограничения на равномерность амплитудно-частотной характеристики последнего, поскольку в частотной области максимального усиления может начаться генерация до достижения оптимального усиления для других частотных областей.

Анализ приведённого материала позволяет сформулировать методы повышения предельной чувствительности РВКИ. Они делятся на методы, аналогичные методам повышения чувствительности ВКИ, и методы, присущие только РВКИ. К первым относится увеличение коэффициента пропорциональности между скоростью вращения и разностью фаз встречных волн $\varphi_\mathrm{s} = 4\pi RL\Omega_\mathrm{s}/(\lambda c)$ за счёт увеличения длины волокна, радиуса намотки, укорочения длины волны и увеличения ширины полосы источника излучения П; ко вторым — увеличение добротности резонатора с одновременным увеличением мощности излучения на выходе интерферометра, что при наличии потерь в резонаторе возможно только за счёт введения усилителя или повышения мощности источника.

Результаты работы состоят в следующем.

1) Рассмотрена схема резонансного волоконного кольцевого интерферометра с отражателем и широкополосным источником излучения, внутрирезонаторным усилителем для измерения скорости вращения с использованием модуляционного метода. Показано, что при реальных потерях в резонаторе увеличение чувствительности рассмотренной резонансной схемы по сравнению с традиционной (нерезонансной) схемой волоконного кольцевого интерферометра при отсутствии усилителя возможно только за счёт увеличения мощности источника.

2) Показано, что размещение усилителя внутри резонатора (регенеративный режим) приводит к увеличению пороговой чувствительности резонансного кольцевого интерферометра при меньшем коэффициенте усиления, чем при его размещении непосредственно на выходе источника.

3) При увеличении добротности резонансного кольцевого интерферометра уменьшается область линейной зависимости сигнала от скорости вращения.

4) Показано, что в той области выходных мощностей, для которых дробовой шум меньше шума источника излучения и усилителя, пороговый сигнал резонансного кольцевого интерферометра не зависит от выходной мощности излучения, а зависит только от полосы резонатора и ширины спектра источника излучения.

5) Показано, что шумы усилителя, обусловленные спонтанным излучением, могут снижать выигрыш в чувствительности к вращению, полученный за счёт усиления.

Работа частично поддержана грантом № 00–1596732 РФФ.

СПИСОК ЛИТЕРАТУРЫ

1. Eziekiel S., Balsamo S. R. // Appl. Phys. 1977. V. 30. P. 478.
2. Shupe D. M. // Appl. Optics. 1981. V. 20, No. 2. P. 286.
3. Iwatsuku K., Hotate K., Nigashiguchi M. // Appl. Optics. 1984. V. 23, No. 21. P. 3 916.
4. Iwatsuku K., Hotate K., Nigashiguchi M. // Appl. Optics. 1986. V. 25, No. 15. P. 2 602.
5. Байбородин Ю. В., Мащенко А. И. // Зарубежная радиоэлектроника. 1988. № 3. С. 95.
6. Lefevre H. The fiber-optic gyroscope. Boston, London: Artech House Inc., 1993. P. 160.
7. Kringlebotn J. T., Blotekjaer K., Pannel C. N. // Fiber and Integration Optics. 1995. V. 14. P. 265.
8. Ohtsuka Y. // J. Lightwave Techn. 1985. V. 9. P. 378.
9. Farhadiroushan I. M., Giles I. P., Youngquist R. C. // SPIE. 1986. V. 719. P. 178.
10. Новиков М. А., Иванов В. В. // Письма в ЖТФ. 1998. Т. 24, № 17. С. 24.
11. Rosenthal A. P. // J. Opt. Soc. Am. 1962. V. 52, No. 10. P. 1 143.
12. Hotate K. // SPIE. 1996. V. 2 895. P. 68.
13. Иванов В. В., Новиков М. А., Геликонов В. М. // Письма в ЖТФ. 1999. Т. 25, вып. 23.
14. Hu Z., Ma X., Li F., et al. // SPIE. 1999. V. 3 860. P. 348.
15. Алексеев Э. И., Базаров Е. Н., Губин В. П. и др. // Квантовая электроника. 2004. Т. 34, № 1. С. 62.
16. Иванов В. В., Новиков М. А., Геликонов В. М. // Квантовая электроника. 2000. Т. 30, № 2. С. 119.
17. Карлов Н. В. Лекции по квантовой электронике М.: Наука, 1988. 50 с.
18. Андронова И. А., Токман М. М. Модуляционный метод измерения скорости вращения в резонансном волоконном кольцевом интерферометре с низкокогерентным источником излучения: Препринт № 618 ИПФ РАН. Нижний Новгород, 2002.

19. Yu A., Siddiqui A. S. // Electron. Lett. 1992. V. 28, № 19. P. 1778.
20. Yu A., Siddiqui A. S. // IEE Proc. J. 1993. V. 140, No. 2. P. 150.
21. Hu Z., Ma X., Li F., et al. // SPIE. 1999. V. 3860. P. 348.
22. Берштейн И. Л. // ЖТФ. 1941. Т. 11, вып. 4. С. 302.
23. Андронова И. А., Берштейн И. Л. // Изв. вузов. Радиофизика. 1989. Т. 32, № 4. С. 426.
24. Laming R. I., Payne D. N. // IEEE Photon. Technol. Lett. 1990. V. 2, No. 6. P. 418.

Поступила в редакцию 12 июля 2004 г.; принята в печать 25 апреля 2005 г.

THE INFLUENCE OF A REGENERATIVE AMPLIFIER TO SENSITIVITY TO ROTATION OF A RESONANT FIBRE RING INTERFEROMETER

I. A. Andronowa and P. A. Shilyagin

The possibility of using a regenerative amplifier based on an active fiber inside a resonator to enhance the sensitivity to rotation one of the resonant fibre ring interferometer scheme (RFRI) with a low coherence source and external reflector is considered. The influence of amplifier noises on the RFRI sensitivity to rotation is discussed. It is shown that the presence of the regenerative amplifier allows to realize maximum sensitivity with a smaller source intensity due to compensation the dissipation in resonator. The comparison with the sensitivity of the conventional nonresonant fibre ring interferometer (FRI) is made.

УДК 517.9

РОСТ И НАСЫЩЕНИЕ ФЛУКТУАЦИЙ В НЕЛИНЕЙНОМ ОСЦИЛЛЯТОРЕ НА ПОРОГЕ БИФУРКАЦИИ СПОНТАННОГО НАРУШЕНИЯ СИММЕТРИИ

Е. Д. Суровяткина

Московский педагогический государственный университет
и Институт космических исследований РАН, г. Москва, Россия

Рассмотрено явление усиления шума в нелинейном осцилляторе на пороге бифуркации спонтанного нарушения симметрии. Исследования проводились на модели нелинейного осциллятора, в котором потенциальный рельеф трансформировался от моностабильного к симметричному бистабильному, а действующий на систему шум полагался гауссовым и короткокоррелированным. Исследована зависимость дисперсии флуктуаций от режима системы и скорости приближения к бифуркационному порогу. Приведены аналитические оценки как линейного роста, так и нелинейного насыщения флуктуаций, которые удовлетворительно согласуются с результатами численного моделирования. Отмечено, что при быстрых бифуркационных переходах в нелинейном осцилляторе наблюдаются шумозависимый гистерезис и нарушение вероятностной симметрии устойчивых конечных состояний.

ВВЕДЕНИЕ

Как известно, на пороге возникновения автоколебаний в нелинейных системах наблюдается предгенерационное усиление шума. Это явление наблюдается как в радиофизических системах, так и в оптических автоколебательных системах [1]. Усиление шума по мере приближения к порогу генерации обусловлено здесь уменьшением потерь в осцилляторе, в результате чего вещественная часть одного из ляпуновских показателей системы из отрицательной становится положительной. Исходное состояние системы в этом случае теряет устойчивость, а усиленный предгенерационный шум служит эффективной затравкой для возникновения автоколебаний.

Предгенерационное усиление шума является частным случаем более общего явления, а именно предбифуркационного усиления шума и слабых сигналов, обусловленного уменьшением (вплоть до обращения в нуль в критической точке) декремента затухания [2–5]. Развитая в работах [2–5] линейная теория предсказывает неограниченный рост флуктуаций при приближении к точке бифуркации.

Нелинейный анализ предбифуркационного усиления шума был выполнен в работе [6] для случая бифуркации удвоения периода в системе, описываемой отображением логистического типа. В этом случае, как и при возникновении генерации, ляпуновский показатель из отрицательного становится положительным, но смена знака вещественной части $\mathrm{Re}\,\lambda$ ляпуновского показателя λ ведёт теперь не к возникновению генерации, а к переходу системы из неустойчивого состояния в одно из двух возможных новых устойчивых состояний.

Целью данной работы является рассмотрение явления предбифуркационного усиления шума в нелинейном осцилляторе, подверженном бифуркации спонтанного нарушения симметрии. Как известно, при такой бифуркации вместо одного устойчивого состояния равновесия появляются два новых устойчивых состояния, а первоначальное состояние теряет устойчивость.

В разделе 1 описана динамическая модель нелинейного осциллятора, в котором может происходить удвоение числа устойчивых состояний равновесия. Примером системы такого рода могут служить одномерные поперечные колебания стержня, напряжённого вдоль оси (см. раздел 1). В разделах 2–4 приведены аналитические и численные оценки уровня флуктуаций при быстрых и медленных изменениях управляющего параметра. Раздел 5 содержит результаты численного моделирования. В разделе 6 описано явление шумозависимого гистерезиса в рассматриваемой системе. Наконец, в разделе 7 показано, что при быстрых бифуркационных переходах в нелинейном осцилляторе наблюдается нарушение вероятностной симметрии устойчивых постбифуркационных состояний.

1. ДИНАМИЧЕСКАЯ МОДЕЛЬ

Рассмотрим колебания в нелинейном осцилляторе, описываемом уравнением второго порядка

$$\frac{\mathrm{d}x^2}{\mathrm{d}t^2} + 2\gamma\,\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial U(x)}{\partial x} = \eta(t). \tag{1}$$

Здесь γ — показатель затухания, $\eta(t)$ — шумовое воздействие на систему, а $U(x)$ — потенциальная энергия.

Как известно, число минимумов потенциального рельефа $U(x)$ в нелинейном осцилляторе определяет число состояний равновесия. При бифуркации спонтанного нарушения симметрии происходит переход от потенциального профиля с одним минимумом к профилю с двумя минимумами и соответствующее удвоение числа устойчивых состояний. Бифуркации спонтанного нарушения симметрии соответствует модельный потенциал

$$U(x) = Ax^4 + Bx^2, \tag{2}$$

который при $B > 0$ имеет единственный минимум $U_{\min} = 0$ (при $x = 0$), а при $B < 0$ приобретает два отрицательных и одинаковых по величине минимума, разделённых максимумом $U_{\max} = 0$, находящимся при $x = 0$.

Переход от одного (при $B > 0$) к двум (при $B < 0$) устойчивым состояниям показан на рис. 1. Пусть параметр B зависит от времени: $B = B(t)$, принимая положительные значения при $t < t^*$ и отрицательные при $t > t^*$. Такое поведение демонстрирует, например, функция

$$B(t) = -\omega_0^2 \operatorname{arctg}[\beta\,(t - t^*)]. \tag{3}$$

Здесь коэффициент β характеризует скорость изменения управляющего параметра.

Физическим прототипом системы, описываемой потенциалом (2), могут служить поперечные одномерные (по оси x) колебания плоского стержня (линейки), вдоль оси которого действует нарастающая во времени сдавливающая сила. При критическом сдавливающем напряжении, при котором коэффициент B обращается в нуль, стержень испытывает бифуркацию спонтанного нарушения симметрии и принимает выгнутую форму, отвечающую одному из двух устойчивых состояний x^+ и x^- [7].

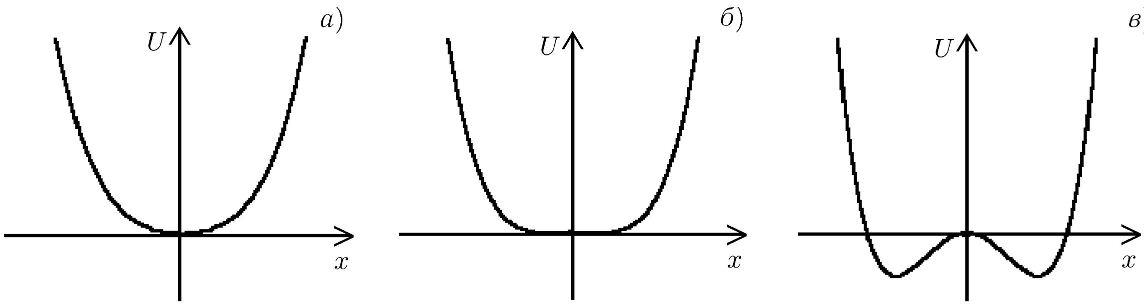

Рис. 1. Потенциальный рельеф $U(x)$ в нелинейном осцилляторе, допускающем бифуркации удвоения числа устойчивых состояний равновесия: (*а*) $B > 0$, (*б*) $B = 0$, (*в*) $B < 0$

Задача данной статьи состоит в том, чтобы определить дисперсию $\langle (x(t) - \bar{x})^2 \rangle \equiv \sigma_x^2$ отклика $x(t)$ на флуктуационное воздействие $\eta(t)$ (здесь $\bar{x}$ — среднее значение, совпадающее со стационарным значением) и тем самым описать как рост, так и нелинейное насыщение предбифуркационного усиления шума в моменты времени $t < t^*$, предшествующие бифуркации.

Относительно флуктуационного воздействия $\eta(t)$ примем, что $\eta(t)$ описывает стационарный случайный процесс с корреляционной функцией

$$\langle \eta(t')\eta(t'') \rangle = \sigma_\eta^2 \Psi_\eta(t' - t''), \tag{4}$$

где σ_η^2 — дисперсия, а $\Psi_\eta(\tau)$ — коэффициент корреляции (нормированная корреляционная функция процесса, спадающая до уровня 0,5 на временах порядка времени корреляции τ_η). Ниже мы ограничимся рассмотрением короткокоррелированных процессов, приближающихся по своим свойствам к белому шуму. В этом случае время корреляции τ_η мало́ по сравнению с периодом свободных колебаний $2\pi/\omega_0$, так что $\omega_0 \tau_\eta \ll 2\pi$.

2. ОЦЕНКИ ФЛУКТУАЦИЙ ПРИ МЕДЛЕННОМ ИЗМЕНЕНИИ ПАРАМЕТРОВ СИСТЕМЫ: ПРИБЛИЖЕНИЕ ВКБ

При малой амплитуде колебаний и $B(t) > 0$ уравнение (1) можно линеаризовать, записав его в виде

$$\ddot{x} + 2\gamma\dot{x} + \omega^2(t)x = \eta(t), \tag{5}$$

где $\omega(t) = \sqrt{B(t)}$. При достаточно медленном уменьшении $B(t)$, а именно при выполнении неравенства $\beta \ll \omega_0$, решение $x(t)$ удовлетворительно описывается приближением Вентцеля—Крамерса—Брюллюэна (ВКБ):

$$x(t) = \frac{A_x}{\sqrt{\omega(t)}} \exp(-\gamma t) \cos[\Phi(t) + \varphi_0], \tag{6}$$

где A_x и φ_0 — некоторые постоянные,

$$\Phi(t) = \int\limits_0^t \omega(t')\,\mathrm{d}t'.$$

В приближении ВКБ легко записать и функцию Грина уравнения (5):

$$g(t, t') = \begin{cases} 0, & t < t'; \\ \dfrac{\exp[-\gamma\,(t - t')]}{\sqrt{\omega(t)\omega(t')}} \sin[\Phi_g(t, t')], & t > t', \end{cases} \tag{7}$$

где

$$\Phi_g(t, t') = \int\limits_{t'}^t \omega(t'')\,\mathrm{d}t''.$$

Используя функцию Грина (7), решение неоднородного линейного уравнения (5) можно записать в виде

$$x(t) = \int\limits_{-\infty}^t g(t, t')\eta(t')\,\mathrm{d}t', \tag{8}$$

а дисперсию $\sigma_x^2 = \langle x^2(t) \rangle$ — в виде

$$\sigma_x^2 = \int\limits_{-\infty}^{t} \mathrm{d}t' \int\limits_{-\infty}^{t} \mathrm{d}t''\, g(t,t') g(t',t'')\, \langle \eta(t') \eta(t'') \rangle. \tag{9}$$

Для короткокоррелированного шума (4) и для функции Грина (7) выражение (9) принимает вид

$$\sigma_x^2 = \tau_\eta \sigma_\eta^2 \int\limits_{-\infty}^{t} g^2(t,t')\, \mathrm{d}t' = \tau_\eta \sigma_\eta^2\, \frac{1}{\omega(t)} \int\limits_{-\infty}^{t} \frac{\exp[-\gamma\,(t-t')]}{\omega(t')}\, \sin^2[\Phi_g(t,t')]\, \mathrm{d}t'. \tag{10}$$

Заменяя величину $\sin^2[\Phi_g(t,t')]$ её средним значением $1/2$, получаем оценку

$$\sigma_x^2(t) = \sigma_\eta^2\, \frac{\tau_\eta}{2\gamma\omega(t)\Omega(t)}\ , \tag{11}$$

где через $1/\Omega(t)$ обозначен интеграл

$$\frac{1}{\Omega(t)} = \gamma \int\limits_{-\infty}^{t} \frac{\exp[-\gamma\,(t-t')]}{\omega(t')}\, \mathrm{d}t', \tag{12}$$

представляющий собой величину $1/\omega(t')$, усреднённую по времени с весом $\gamma \exp[-\gamma\,(t-t')]$.

Полезно отметить, что при постоянном значении частоты $\omega = \omega_0$ соотношение (11) переходит в известное выражение для установившихся флуктуаций в линейном осцилляторе под действием короткокоррелированной флуктуационной силы $\eta(t)$ [1]:

$$\sigma_0^2 = \sigma_\eta^2\, \frac{\tau_\eta}{2\gamma\omega_0^2}\ . \tag{13}$$

По мере приближения к точке бифуркации $t = t^*$, в которой $B(t)$ обращается в нуль, линейная оценка (11) описывает возрастание дисперсии флуктуаций, что отвечает явлению предбифуркационного усиления шума. Величины $1/\omega(t)$ и $1/\Omega(t)$, входящие в выражение (11), при $t \to t^*$ обращаются в бесконечность. Вслед за ними стремится к бесконечности и линейная оценка (11). Разумеется, оценка дисперсии (11), основанная на приближении ВКБ, в окрестности точки бифуркации неприменима, но, как будет показано в следующем разделе, при $t \to t^*$ удаётся получить нелинейную оценку дисперсии σ_x^2, принимающую конечное значение даже непосредственно в точке бифуркации.

Особенность бифуркационного сценария для рассматриваемой системы состоит в том, что по мере приближения к точке бифуркации вещественная часть ляпуновских показателей $\gamma = \mathrm{Re}\,\lambda$ не уменьшается до нуля, как при бифуркации Ландау—Хопфа и бифуркации удвоения периода, поэтому рост флуктуаций не связан здесь с уменьшением потерь. Вместо этого мы имеем дело с превращением пары комплексно-сопряжённых ляпуновских показателей $\lambda_{1,2} = -\gamma \pm i\omega$ в пару вещественных показателей, один из которых положителен. В результате предбифуркационное усиление шума связано теперь с уменьшением частоты $\omega = \mathrm{Im}\,\lambda$, поскольку при $\omega \to 0$ амплитуда свободных колебаний увеличивается как $1/\sqrt{\omega(t)}$.

3. НЕЛИНЕЙНЫЕ ОЦЕНКИ ИНТЕНСИВНОСТИ ФЛУКТУАЦИЙ В ОКРЕСТНОСТИ ТОЧКИ БИФУРКАЦИИ

Как уже было сказано выше, выражение (11), основанное на линейной теории, в непосредственной окрестности точки бифуркации теряет силу, поскольку при неограниченном росте флуктуаций начинают сказываться нелинейные эффекты. Нелинейными эффектами можно пренебречь, пока слагаемое четвёртой степени в (2) мало́ по сравнению с квадратичным слагаемым:

$$A\,\langle x^4\rangle \ll |B(t)|\,\langle x^2\rangle. \tag{14}$$

Вблизи бифуркационного порога, когда потенциальный рельеф ещё имеет единственное состояние равновесия, вполне допустимо считать распределение вероятности вынужденных флуктуаций x одномодовым (одногорбым). В этих условиях для оценок можно воспользоваться гауссовой моделью флуктуаций x (см. приведённые ниже вычисления, которые вполне удовлетворительно согласуются с результатами численного моделирования). Разумеется, в постбифуркационном режиме с двумя равновесными состояниями, когда возникает бимодальное (двугорбое) вероятностное распределение, целесообразно аппроксимировать распределение вероятности «бигауссовой» моделью, как это предложено в работе [8]. В данной работе, однако, мы ограничимся аналитическими оценками только для предбифуркационных флуктуаций, тогда как для постбифуркационных флуктуаций воспользуемся результатами численного моделирования, не анализируя модальную структуру распределения флуктуаций x.

На основании вышеизложенного, считая флуктуации η и x гауссовыми, при оценках положим

$$\langle x^4\rangle = 3\,\langle x^2\rangle^2 = 3\sigma_x^4$$

и перепишем неравенство (14) в форме

$$|B(t)| \gg 3A\sigma_x^2 = \frac{3A\sigma_\eta^2\tau_\eta}{2\gamma\omega(t)\Omega(t)}\,. \tag{15}$$

Если предположить дополнительно, что характерное время затухания колебаний $1/\gamma$ меньше характерного времени $1/\beta$ изменения частоты $\omega(t)$, из (12) получим

$$\frac{1}{\Omega(t)} \approx \frac{1}{\omega(t)}\ , \tag{16}$$

тогда условие (15) примет вид

$$B^2(t) = \omega^4(t) \gg \frac{3A\sigma_\eta^2\tau_\eta}{2\gamma}\ . \tag{17}$$

Из (17) следует оценка

$$B_{\min} \sim \left(\frac{3A\sigma_\eta^2\tau_\eta}{2\gamma}\right)^{1/2} \tag{18}$$

допустимой отстройки от точки бифуркации, а из (11) с учётом (16) — оценка

$$\sigma^2_{x\,\max} \sim \sigma_\eta\sqrt{\frac{\tau_\eta}{6\gamma A}} \tag{19}$$

максимальной интенсивности флуктуаций. Таким образом, при $B < B_{\min}$ линейные эффекты роста флуктуаций уступают место нелинейному насыщению, которое достигается при интенсивности флуктуаций $\sigma^2_{x\,\max}$. Подобное насыщение интенсивности флуктуаций обнаружено также в случае бифуркации удвоения периода [6], а общий подход к оценке уровня насыщения описан в работе [9].

Отношение $\sigma^2_{x\,\max}$ к интенсивности флуктуаций (13) при $\omega = \omega_0$ удобно назвать фактором предбифуркационного усиления шума:

$$K = \frac{\sigma^2_{x\,\max}}{\sigma^2_0} = \frac{1}{\sigma_\eta}\left(\frac{2\gamma}{3A\tau_\eta}\right)^{1/2}\omega_0^2. \tag{20}$$

Эта величина показывает, во сколько раз интенсивность флуктуаций в зоне насыщения превышает стационарные флуктуации в осцилляторе.

4. ФЛУКТУАЦИИ ПРИ ОЧЕНЬ БЫСТРОМ ИЗМЕНЕНИИ ПАРАМЕТРОВ ОСЦИЛЛЯТОРА

Выше мы рассмотрели флуктуации при достаточно медленном ($\beta < \gamma$) изменении параметров осциллятора. При быстром прохождении точки бифуркации, т. е. при $\beta > \gamma$, можно ожидать некоторого уменьшения $\sigma^2_{x\,\max}$ по сравнению со случаем $\beta < \gamma$.

Тенденцию к уменьшению $\sigma^2_{x\,\max}$ с ростом скорости β прохождения точки бифуркации можно проиллюстрировать рассмотрением предельного случая $\beta \to \infty$, когда при $t < t^*$ величина $B(t)$ постоянна и равна ω_0^2, а при $t > t^*$ тоже постоянна, но равна $-\omega_0^2$. В этом случае интенсивность флуктуаций при $t < t^*$ постоянна и равна σ_0^2, так что фактор усиления флуктуаций обращается в единицу: $K = 1$. Разумеется, при $t > t^*$ начнётся экспоненциальный рост флуктуаций вследствие потери устойчивости состояния равновесия $x = 0$, но это никак не отразится на поведении флуктуаций при $t < t^*$.

5. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

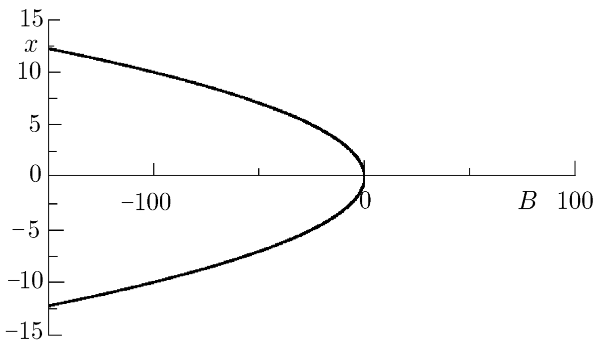

Рис. 2. Бифуркационная диаграмма нелинейного осциллятора

Описанный выше нелинейный осциллятор анализировался численно в случае, когда $A = 0{,}5$, а параметр B принимал значения от $B_0 = 100$ до $B_{\mathrm{f}} = 10^{-8}$, т. е. уменьшался практически до нуля, при этом изменения параметра B считались достаточно медленными (квазистационарный режим). При расчётах показатель затухания γ принимался равным 0,1, так что для выполнения условия квазистационарности требовалось, чтобы характерное время изменения B было малым по сравнению с 0,1.

Указанный диапазон изменения параметра B позволяет определить дисперсию флуктуаций как в непосредственной близости к бифуркационной точке $B_{\mathrm{c}} = 0$, так и вдали от неё (напомним, что каждому значению B на удалении от точки бифуркации отвечает частота колебаний $\omega_0 = \sqrt{B}$). При $B > 0$, когда система имеет одно устойчивое решение (см. бифуркационную диаграмму на рис. 2), начальным значением служила устойчивая точка $x(0) = 0$, начальное значение производной $x'(0)$ полагалось равным нулю.

Генератор случайных чисел обеспечивал нормально распределённые значения $\eta(t)$ с нулевым средним ($\langle\eta(t)\rangle = 0$), при этом среднеквадратичное отклонение σ_η менялось в диапазоне от 10^{-6}

до 10^{-1}. Корреляционная функция случайного процесса характеризовалась временем τ_η, лежащим в интервале от 10^{-2} до 10^{-3}, т. е. малым по сравнению с $2\pi/\omega_0$. В этом случае результаты не зависели от вида корреляционной функции Ψ_η, как это и должно быть для процессов, приближающихся к белому шуму. Использование иных генераторов случайных чисел, обеспечивающих, например, равномерно распределённые величины $\eta(t)$, дало результаты, качественно подобные случаю нормально распределённого шумового воздействия. Численное решение уравнения (5) проводилось с использованием метода Рунге—Кутты четвёртого порядка с фиксированным шагом.

Результаты численного моделирования представлены на рис. 3, где крестиками показана зависимость дисперсии флуктуаций σ_x^2 от параметра B. Численный эксперимент проводился при достаточно медленном изменении параметра B (квазистационарный режим) и дисперсии шума $\sigma_\eta^2 = 10^{-8}$. Наклонная штриховая линия соответствуют оценке дисперсии флуктуаций σ_0^2 в линейном режиме (13), которая описывает возрастание флуктуаций при приближении к бифуркационному порогу. Как видно из графика, результаты численного моделирования соответствуют линейной оценке до значений $B > B_{\min}$, что согласуется с оценкой (18). При дальнейшем приближении к точке бифуркации $B = 0$ дисперсия флуктуаций достигает насыщения при значении $\sigma_{x\,\max}^2$, которое отмечено горизонтальной штрихпунктирной линией и соответствует нелинейной оценке (19).

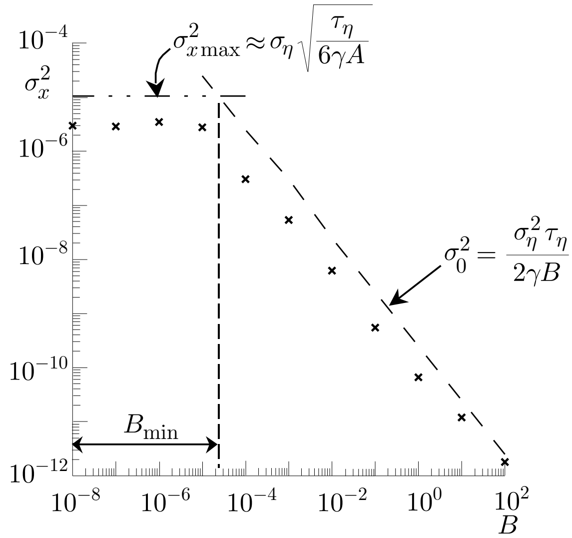

Рис. 3. Предбифуркационное усиление шума для бифуркации спонтанного нарушения симметрии. Зависимость дисперсии флуктуаций σ_x^2 от параметра B в квазистационарном режиме ($\sigma_\eta^2 = 10^{-8}$; $\tau_\eta = 3{,}5 \cdot 10^{-3}$; $\gamma = 0{,}1$; $A = 0{,}5$)

В рассматриваемом квазистационарном случае при дисперсии шума $\sigma_\eta^2 = 10^{-8}$ коэффициент усиления флуктуаций $K_{\max}$ составил $1{,}55 \cdot 10^6$. Это значение по порядку величины согласуется с теоретической оценкой $K_{\max}^{\rm teor} = 6{,}17 \cdot 10^6$, полученной на основе (20).

Описанные результаты качественно согласуются с данными, полученными ранее для бифуркации удвоения периода [6]: в обоих случаях средний квадрат флуктуаций σ_x^2 пропорционален среднеквадратичному отклонению шума σ_η, а коэффициент усиления $K_{\max}$ обратно пропорционален σ_η.

6. ШУМОЗАВИСИМЫЙ ГИСТЕРЕЗИС В ОКРЕСТНОСТИ ТОЧКИ БИФУРКАЦИИ

Явление шумозависимого гистерезиса состоит в том, что после прохождения бифуркационного значения система задерживается в окрестности вновь возникшего неустойчивого состояния равновесия и лишь спустя некоторое время переходит в одно из двух возможных устойчивых состояний [10, 11]. Чем быстрее меняется управляющий параметр, тем отчётливее выражено явление гистерезиса.

В квазистационарном режиме, когда параметр B изменяется достаточно медленно, бифуркация в системе наступает при критическом значении $B = B_{\rm c} = 0$. При быстром же изменении пара-

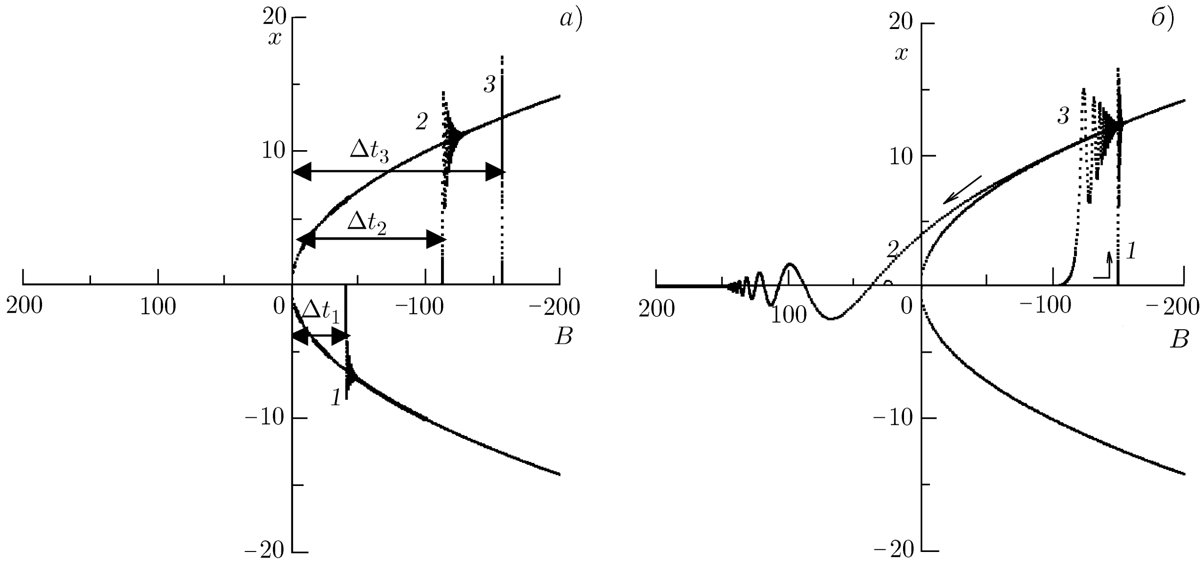

Рис. 4. Бифуркационная диаграмма и решения дифференциального уравнения (1): (*а*) при медленном прохождении через точку бифуркации ($\beta = 0{,}03$, график *1*), при быстром прохождении ($\beta = 0{,}3$, график *2*) и при очень быстром прохождении ($\beta = 300$, график *3*); (*б*) при прямом и обратном прохождении через точку бифуркации со скоростями $\beta = 3$ (график *1*) и $\beta = -3$ (график *2*), график *3* отражает результат воздействия шума с $\sigma_\eta^2 = 1{,}87 \cdot 10^{-7}$ при $\beta = 3$

метра B бифуркация удвоения числа устойчивых состояний наступает лишь спустя некоторое время после прохождения критического значения $B_{\text{c}} = 0$, при этом время задержки зависит от скорости изменения параметра β.

На рис. 4*а* представлены результаты численного моделирования бифуркационного перехода в нелинейном осцилляторе в условиях, когда параметр B изменятся по закону (3). Для наглядности на рисунке совмещены бифуркационная диаграмма исследуемой модели, т. е. установившиеся значения при постоянном значении B, и зависимость $x(B)$ с изменяющимся параметром B. Как видно из рис. 4*а*, при изменении параметра B система, преодолев значение $B = B_{\text{c}}$, ещё некоторое время пребывает в окрестности неустойчивой ветви (это время существенно зависит от скорости β) и лишь затем переходит в одно из двух возможных устойчивых состояний равновесия. При высокой скорости перехода $\beta = 300$ (график *3*) время запаздывания Δt_3 существенно больше, чем время запаздывания Δt_2 при $\beta = 0{,}3$ (график *2*) и время Δt_1 при очень медленном переходе со скоростью $\beta = 0{,}03$ (график *1*).

Как известно, при прямом и обратном прохождении бифуркационной точки система задерживается в окрестности прежних устойчивых точек, при этом явление затягивания приводит к образованию гистерезисной петли, размер которой существенно зависит от шума. Следует отметить, что при прямом прохождении бифуркационной точки система более чувствительна к воздействию шума, чем при обратном ходе. Рисунок 4*б* демонстрирует явление гистерезиса в нелинейном осцилляторе при $\beta = 3$ (прямой ход) и $\beta = -3$ (обратный ход). Как видно из рисунка, с ростом шума в системе размер гистерезисной петли сокращается. Этот эффект может быть использован для измерения слабых шумов в нелинейных системах, как это было предложено нами ранее для системы с бифуркациями удвоения периода [11].

7. НАРУШЕНИЕ ВЕРОЯТНОСТНОЙ СИММЕТРИИ В НЕЛИНЕЙНОМ ОСЦИЛЛЯТОРЕ ПРИ БИФУРКАЦИИ УДВОЕНИЯ ЧИСЛА УСТОЙЧИВЫХ СОСТОЯНИЙ

Как было показано в работах [11, 12] на примере бифуркации удвоения периода, при малых шумах динамическим бифуркациям присуще явление нарушения вероятностной симметрии: если при наличии заметного шума вероятности попадания в равноправные конечные состояния одинаковы и составляют 50 %, то в отсутствие шума конечное состояние системы полностью предсказуемо и определяется только начальными условиями и скоростью изменения управляющего параметра. Явление нарушения вероятностной симметрии наблюдается и при бифуркации спонтанного нарушения симметрии. В отсутствии шума система с вероятностью 100 % попадает в одно из двух возможных конечных состояний, определяемое скоростью бифуркационного перехода и начальными условиями. При воздействии шума на нелинейный осциллятор с изменяющимися параметрами вероятности попадания в одно из двух возможных устойчивых состояний выравниваются.

На рис. 5 иллюстрируется явление нарушения вероятностной симметрии в нелинейном осцилляторе. В отсутствие шума при начальном значении $x(0) = 0$, начальном значении производной $x'(0) = 1$ и скорости $\beta = 3$ зависимость решения x от параметра B показана графиком *1* (рис. 5*а*). В результате воздействия шума система может попасть как в «верхнее» (график *2*), так и в «нижнее» (график *3*) состояние, но существенно раньше, чем в «бесшумном» переходе (при численных расчётах дисперсия шума полагалась равной $\sigma_\eta^2 = 1{,}87 \cdot 10^{-6}$).

На рис. 5*б* показана зависимость вероятности попадания в конечные состояния в результате бифуркационного перехода от начальных значений x_0. Как видно из рисунка, при $\sigma_\eta^2 = 1{,}87 \cdot 10^{-7}$ шум «размывает» границы зон притяжения конечных состояний. В отличие от дискретных отображений [13], картина зон притяжения конечных состояний в случае нелинейного осциллятора будет зависеть не только от начальных значений и скорости бифуркационного перехода, но и от

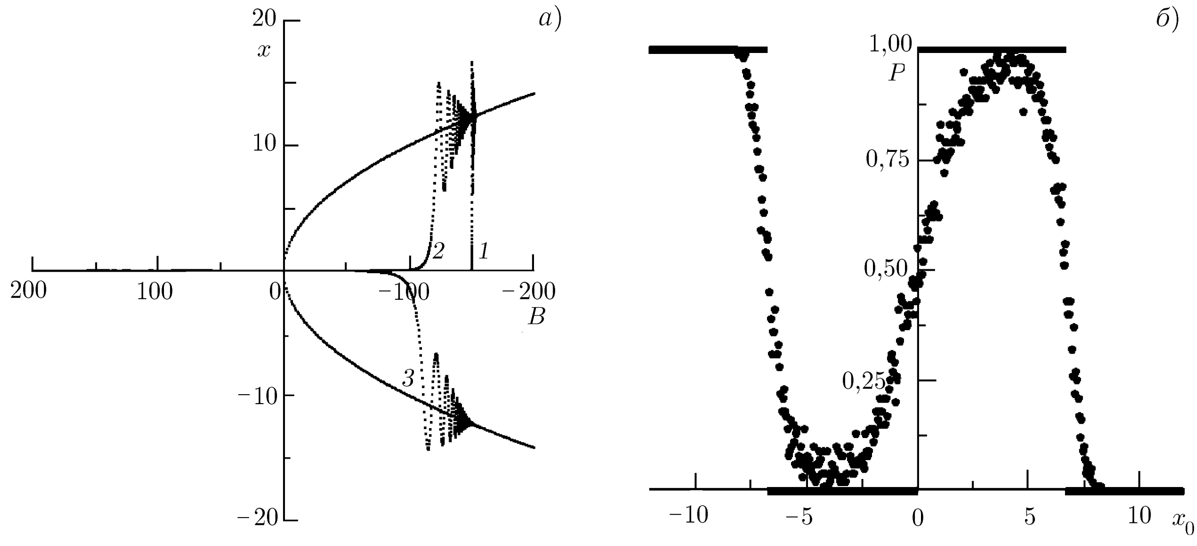

Рис. 5. Нарушение вероятностной симметрии в нелинейном осцилляторе: (*а*) бифуркационная диаграмма и решение уравнения (1) при $\beta = 3$ (график *1*), решения при наличии шума с дисперсией $\sigma_\eta^2 = 1{,}87 \cdot 10^{-7}$ (графики *2* и *3*); (*б*) зоны притяжения конечных состояний (сплошные линии) и размывание границ бассейнов притяжения шумом с дисперсией $\sigma_\eta^2 = 1{,}87 \cdot 10^{-7}$ при скорости $\beta = 3$ (точки)

начального значения производной. В случае малых скоростей зоны притяжения будут дробиться, что приводит к ещё большей чувствительности системы к шуму.

При замедлении прохождения точки бифуркации под влиянием предбифуркационного усиления шума вероятности попадания в одно из двух возможных состояний выравниваются. В то же время при быстром прохождении точки бифуркации флуктуации в окрестности бифуркационной точки снижаются, а предсказуемость попадания в определённое конечное состояние повышается.

8. ЗАКЛЮЧЕНИЕ

В данной работе исследованы флуктуации в нелинейном осцилляторе, испытывающем бифуркации спонтанного нарушения симметрии при наличии шума. Установлено, что в результате воздействия на систему гауссового короткокоррелированного шума дисперсия вынужденных флуктуаций σ_x^2 в режиме насыщения (вблизи точки бифуркации) достигнет уровня, пропорционального среднеквадратичному значению шумовой силы σ_η: $\sigma_x^2 \propto \sigma_\eta$, тогда как в линейном режиме (вдали от бифуркационного порога) дисперсия вынужденных флуктуаций σ_x^2 пропорциональна дисперсии шума σ_η^2: $\sigma_x^2 \propto \sigma_\eta^2$. Аналитические оценки удовлетворительно согласуются с результатами численного моделирования.

В работе показано, что предбифуркационное усиление флуктуаций способствует установлению вероятностной симметрии конечных состояний равновесия. При достаточно медленном изменении параметра осциллятора воздействие слабых шумов приводит к выравниванию вероятностей попадания в каждое из двух возможных конечных состояний. Однако при быстром бифуркационном переходе эффект предбифуркационного усиления шумов ослабляется, система становится менее чувствительной к воздействию шумов, а конечные состояния — более предсказуемыми (явление нарушения вероятностной симметрии).

Показано, что бифуркация спонтанного нарушения симметрии в нелинейном осцилляторе при изменяющемся управляющем параметре сопровождается явлением затягивания, параметры которого существенно зависят от уровня шума (явление шумозависимого гистерезиса).

Автор признателен РФФИ за поддержку данной работы в рамках гранта 02–02–17418.

СПИСОК ЛИТЕРАТУРЫ

1. Ахманов С. А., Дьяков Ю. Е., Чиркин А. С. Введение в статистическую радиофизику и оптику. М.: Наука, 1981.
2. Wiesenfeld K. // Phys. Rev. A. 1985. V. 32. P. 1 744.
3. Wiesenfeld K. // J. Stat. Phys. 1985. V. 38. P. 1 071.
4. Wiesenfeld K., McNamara B. // Phys. Rev. A. 1986. V. 33. P. 629.
5. Wiesenfeld K., Pedersen N. // Phys. Rev. A. 1987. V. 36. P. 1 440.
6. Кравцов Ю. А., Бильчинская С. Г., Бутковский О. Я. и др. // ЖЭТФ. 2001. Т. 120, №. 6. С. 1 527.
7. Физическая энциклопедия. Т. 4. М.: Изд-во Большая Российская энциклопедия, 1994. С. 652.
8. Музычук О. В. // Актуальные проблемы статистической радиофизики. Н. Новгород: ТАЛАМ, 2003.
9. Kravtsov Yu. A., Surovyatkina E. D. // Phys. Lett. A. 2003. V. 319. P. 348.
10. Baesens C. // Physica D. 1991. V. 53. P. 319.
11. Бутковский О. Я., Кравцов Ю. А., Суровяткина Е. Д. // ЖТФ. 1997. Вып. 9. С. 128.
12. Бутковский О. Я., Браш Дж. С., Кравцов Ю. А., Суровяткина Е. Д. // ЖЭТФ. 1996. Т. 109, № 6. С. 2 201.

13. Бутковский О. Я., Кравцов Ю. А., Суровяткина Е. Д. // ЖЭТФ. 1998. Т. 113, вып. 1. С. 369.

Поступила в редакцию 19 марта 2004 г.; принята в печать 25 февраля 2005 г.

GROWTH AND NONLINEAR SATURATION OF FLUCTUATIONS IN NONLINEAR OSCILLATOR ON THE THRESHOLD OF THE BIFURCATION OF SPONTANEOUS SYMMETRY BREAKING

E. D. Surovyatkina

The phenomenon of the prebifurcation noise amplification in nonlinear oscillator experiencing bifurcation of the spontaneous symmetry breaking. Our theoretical estimates have proved to be in good agreement with the result of numerical simulations. It is shown that in saturation regime fluctuation variance is proportional to the standard deviation of external noise, whereas in linear regime is proportional to the noise variance. It is shown that the phenomenon of the prebifurcation noise amplification is more pronounced in the case of slow transition through the bifurcation point. The fluctuation amplification favors to the establishment of the probability symmetry of the final equilibrium states. In contrast, under quick transition through the bifurcation point, when the probability symmetry of final equilibrium states is broken and the oscillator reaches definite final state, the effect of amplification is much less pronounced. Under backward and forward passages through the bifurcation point a loop of the noise-dependent hysteresis arises.

УДК 621.315.592+539.293.4

ГЕНЕРАЦИЯ АКУСТИЧЕСКИХ ФОНОНОВ В ПОЛУПРОВОДНИКОВОЙ СВЕРХРЕШЁТКЕ ПРИ ВНУТРИЗОННОМ ПОГЛОЩЕНИИ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ

М. В. Вязовский, Г. А. Сыродоев

Волгоградский государственный педагогический университет, г. Волгоград, Россия

Найден коэффициент затухания звуковой волны при внутризонном многофотонном поглощении электромагнитной волны в сверхрешётке. Для двух направлений распространения звука, поперечного и продольного относительно оси сверхрешётки, найдены области изменения знака коэффициента затухания звука. Проведено численное суммирование ряда для коэффициента усиления звуковой волны при типичных параметрах сверхрешётки. Дана численная оценка коэффициента усиления. Отмечается, что многофотонное поглощение будет оказывать влияние на коэффициент усиления звука в сверхрешётке при значительно меньшей величине поля, чем в обычном полупроводнике.

Генерация неравновесных акустических фононов в однородных полупроводниках при внутризонном поглощении электромагнитной волны была рассмотрена в работах [1–3]. Там же была определена область неустойчивости этих фононов, в которой может происходить усиление звуковой волны. В работах [4, 5] аналогичная задача была решена и для оптических фононов. В полупроводниковых сверхрешётках влияние электромагнитной волны на затухание звука рассматривалось в [6], но область неустойчивости фононов в этой работе не находилась. Отметим также работу [7], в которой была найдена область смены знака коэффициента затухания звука в сверхрешётке в постоянном электрическом поле.

В настоящей работе находится коэффициент затухания продольных акустических фононов в одномерной сверхрешётке при внутризонном поглощении интенсивной электромагнитной волны и определяется интервал волновых векторов фононов $\mathbf{q}$, в котором коэффициент затухания меняет знак. Для решения этой задачи будем исходить из эффективного гамильтониана взаимодействия электрона в n-й минизоне проводимости с электромагнитной волной:

$$\hat{\mathrm{H}}_1 = \varepsilon_n\Big[\hat{\mathbf{p}} - \frac{e}{c}\,\mathbf{A}(t)\Big] - \varepsilon_n(\hat{\mathbf{p}}), \tag{1}$$

где

$$\varepsilon_n(\hat{\mathbf{p}}) = \frac{\hat{p}_\perp^2}{2\mu} + \frac{\Delta_n}{2}\left[1 - \cos\!\left(\frac{\hat{p}_z d}{\hbar}\right)\right] \tag{2}$$

— энергия электрона в n-й минизоне, z — ось сверхрешётки, Δ_n — ширина n-й минизоны, μ — эффективная масса электрона, d — постоянная сверхрешётки, $\hat{\mathbf{p}} = -i\hbar\nabla$, $\mathbf{A}(t)$ — вектор-потенциал электромагнитной волны, $\hbar$ — постоянная Планка, e — заряд электрона, c — скорость света. Гамильтониан электрон-фононного взаимодействия для продольных акустических фононов возьмём в виде

$$\hat{\mathrm{H}}_2 = i\sum_{\mathbf{q}}\big\{B_{\mathbf{q}}\exp[i\,(\mathbf{qr} - \omega_{\mathbf{q}}t)]\hat{b}_{\mathbf{q}} - B_{\mathbf{q}}^*\exp[-i\,(\mathbf{qr} - \omega_{\mathbf{q}}t)]\hat{b}_{\mathbf{q}}^+\big\}, \tag{3}$$

где для случая деформационного потенциала взаимодействия $B_{\mathbf{q}} = \sqrt{\hbar/(2\rho\omega_{\mathbf{q}}V)}\,q\Xi$, Ξ — константа деформационного потенциала, ρ и V — плотность и объём кристалла соответственно, $\omega_{\mathbf{q}}$ — частота фонона, $\hat{b}_{\mathbf{q}}$ и $\hat{b}_{\mathbf{q}}^+$ — операторы уничтожения и рождения фононов соответственно.

Пусть плоская монохроматическая электромагнитная волна распространяется перпендикулярно оси сверхрешётки (оси z), тогда в дипольном приближении

$$A_z = A_0 \cos(\omega t), \qquad E_z = -\frac{1}{2}\,\frac{\partial A_z}{\partial t}\,. \tag{4}$$

Считая нижнюю минизону занятой, а все верхние свободными, ограничимся приближением одной минизоны. Учитывая известное разложение $\exp(ia\sin\varphi) = \sum_{m=-\infty}^{+\infty} J_m(a)\exp(im\varphi)$, где $J_m(a)$ — функция Бесселя m-го порядка, для $\hat{\mathrm{H}}_1$ получим выражение

$$\hat{\mathrm{H}}_1 = \sum_{m=-\infty}^{+\infty} \left[J_{2m}(a)\exp(2mi\varphi)\hat{C}_{2m} + J_{2m+1}(a)\exp[-(2m+1)\,i\varphi]\hat{C}_{2m+1}\right], \tag{5}$$

где

$$\hat{C}_{2m+1} = (-1)^m\,\frac{\Delta}{2}\sin(\hat{p}_z d/\hbar), \qquad \hat{C}_{2m} = \frac{\Delta}{2}\left[\delta_{0\,2m}/J_0(a) - (-1)^m\right]\cos(\hat{p}_z d/\hbar),$$

$\varphi = \omega t$, $a = eEd/(\hbar\omega)$, E — амплитуда электрического поля волны, $\delta_{0\,2m}$ — символ Кронекера, Δ — ширина низшей минизоны. Во втором порядке теории возмущений находим вероятность электронного перехода в единицу времени в нижней минизоне из состояния с квазиимпульсом $\mathbf{p} = \hbar\mathbf{k}$ в состояние $\mathbf{p}' = \hbar\mathbf{k}'$ с поглощением фотонов и излучением (поглощением) фонона:

$$w^{-}_{\mathbf{k}',\mathbf{k}}(\mathbf{q}) = \frac{2\pi B^2_{\mathbf{q}}}{\hbar^4}\,N_{\mathbf{k}}\,(1 - N_{\mathbf{k}'}) \sum_{m=1}^{\infty} \left\{\frac{C^2_m(k_z - q_z)}{(\omega_{\mathbf{k}+\mathbf{q},\mathbf{k}} + \omega_{\mathbf{q}})^2} + \frac{C^2_m(k_z)}{(m\omega)^2}\right\} \delta_{\mathbf{k}'\,\mathbf{k}-\mathbf{q}}\delta(\omega_{\mathbf{k}'\,\mathbf{k}} - m\omega + \omega_{\mathbf{q}})\,(n_{\mathbf{q}} + 1),$$

$$w^{+}_{\mathbf{k}'\,\mathbf{k}}(\mathbf{q}) = \frac{2\pi B^2_{\mathbf{q}}}{\hbar^4}\,N_{\mathbf{k}}\,(1 - N_{\mathbf{k}'}) \sum_{m=1}^{\infty} \left\{\frac{C^2_m(k_z + q_z)}{(\omega_{\mathbf{k}+\mathbf{q}\,\mathbf{k}} - \omega_{\mathbf{q}})^2} + \frac{C^2_m(k_z)}{(m\omega)^2}\right\} \delta_{\mathbf{k}'\,\mathbf{k}+\mathbf{q}}\delta(\omega_{\mathbf{k}'\,\mathbf{k}} - m\omega - \omega_{\mathbf{q}})n_{\mathbf{q}}, \tag{6}$$

где $\omega_{\mathbf{k}'\,\mathbf{k}} = (\varepsilon_{\mathbf{k}'} - \varepsilon_{\mathbf{k}})/\hbar$, $\varepsilon_{\mathbf{k}}$ — энергия в состоянии $\mathbf{k}$, $n_{\mathbf{q}}$ и $N_{\mathbf{k}}$ — числа заполнения фононов и электронов. Символы Кронекера и δ-функции выражают законы сохранения квазиимпульса и энергии электрона, фонона и фотонов при переходах, $C^2_m(k_z) = \Delta^2 J^2_m(a)\cos^2(k_z d)/4$ при чётном m, $C^2_m(k_z) = \Delta^2 J^2_m(a)\sin^2(k_z d)/4$ при нечётном m, величина $w^{-}_{\mathbf{k}'\,\mathbf{k}}(\mathbf{q})$ в (6) относится к излучению фонона, $w^{+}_{\mathbf{k}'\,\mathbf{k}}(\mathbf{q})$ — к поглощению. Полагая, что $\exp[-m\hbar\omega/(k_0 T)] \ll 1$, где k_0 — постоянная Больцмана, T — температура образца, можно пренебречь переходами с излучением фотонов, поэтому в сумме (6) оставлены члены с $m > 0$ (член с $m = 0$ не вызывает переходов с поглощением фотонов). Как видно, выражения (6) представляют собой сумму вероятностей переходов с поглощением m фотонов и излучением (поглощением) фонона. Вклад многофотонных переходов, как следует из свойств функций Бесселя, существенен лишь при $a = eEd/(\hbar\omega) \geq 1$.

Скорость изменения числа заполнения фононами состояния $\mathbf{q}$ будет определяться уравнением

$$\frac{\mathrm{d}n_{\mathbf{q}}}{\mathrm{d}t} = -\Gamma_{\mathbf{q}} n_{\mathbf{q}} + G_{\mathbf{q}}, \tag{7}$$

где коэффициент затухания $\Gamma_{\mathbf{q}} = n^{-1}_{\mathbf{q}} \sum_{\mathbf{k}'}\sum_{\mathbf{k}} \left(w^{-}_{\mathbf{k}'\,\mathbf{k}} - w^{+}_{\mathbf{k}'\,\mathbf{k}}\right)$. Первое слагаемое в скобках — вероятность перехода с поглощением фонона, второе — с вынужденным испусканием фонона; член $G_{\mathbf{q}}$ в (7) — вероятность перехода со спонтанным испусканием фонона. Здесь мы учитываем только электронное затухание звука. Рассмотрим два случая: когда вектор $\mathbf{q}$ перпендикулярен оси сверхрешётки и когда он параллелен ей. В первом случае для невырожденной сверхрешётки, взяв для $N_{\mathbf{k}}$ распределение Больцмана и учитывая, что $\omega_{\mathbf{k}'\,\mathbf{k}} = \mp\hbar\mathbf{k}\mathbf{q}/\mu + \hbar^2\mathbf{q}^2/(2\mu)$, после интегрирования получим

$$\Gamma_q = g_1 \sum_{m=1}^{\infty} \left[1 + (-1)^m\,\frac{I_2(\beta)}{I_0(\beta)}\right] \frac{J^2_m(a)}{m^2}\exp(-\gamma)\mathrm{sh}\left[\frac{\mu\omega v_{\mathrm{s}}}{k_0 T q}\left(\frac{\hbar q^2}{2\mu\omega} - m\right)\right], \tag{8}$$

где $I_s(\beta)$ — функция Бесселя мнимого аргумента s-го порядка, $\beta = \Delta/(2k_0T)$,

$$\gamma = \frac{\mu\omega^2}{2q^2k_0T}\left[m^2 + \left(\frac{\omega_q}{\omega}\right)^2 + \left(\frac{\hbar q^2}{2\mu\omega}\right)^2 - m\,\frac{\hbar q^2}{\mu\omega}\right], \qquad g_1 = \frac{\sqrt{\pi}\;\Xi^2 n_0\,\sqrt{\mu}\;\Delta^2}{2\,\sqrt{2}\;\rho v_{\mathrm{s}}\hbar^3\omega^2\,\sqrt{k_0T}}\,,$$

n_0 — концентрация электронов в минизоне проводимости, v_{s} — скорость звука. Каждый из членов суммы (8) меняет знак при $q = \sqrt{2m\mu\omega/\hbar}$, становится отрицательным при $q < \sqrt{2m\mu\omega/\hbar}$ и при $q \to 0$ стремится к нулю. Следовательно, члены суммы имеют минимум в интервале $0 < q < \sqrt{2m\mu\omega/\hbar}$, что подтверждается численным анализом. Если электронный коэффициент нарастания будет больше коэффициента затухания из-за фонон-фононного взаимодействия, то система фононов становится неустойчивой (звуковая волна, частота которой попадёт в этот интервал, будет усиливаться до тех пор, пока нелинейные эффекты не ограничат нарастание её амплитуды). На рис. 1a представлены графики зависимости членов суммы (8) от волнового числа q, а также график суммы всех членов при $\Xi = 10$ эВ, $\rho = 5$ г/см^3, $\Delta = 0{,}05$ эВ, $v_{\mathrm{s}} = 10^5$ см/с, $T = 25$ К, $n_0 = 10^{14}$ см$^{-3}$, $\hbar\omega = 0{,}005$ эВ, $E = 2\cdot 10^4$ В/см, $\beta = 10$. Из рис. 1a видно, что вклады первых трёх гармоник (при $eEd/(\hbar\omega) > 1$) сравнимы по величине, и область неустойчивости соответствует интервалу $0 < q \leq 0{,}5\cdot 10^6$ см$^{-1}$. Численная оценка даёт коэффициент усиления в этой области $\alpha_q = \Gamma_q/v_{\mathrm{s}} = 10^2$ см$^{-1}$. При $q > \sqrt{2m\mu\omega/\hbar}$ все члены суммы (8) положительны и определяют затухание фононов при многофононном поглощении электромагнитной волны. Во втором случае (вектор $\mathbf{q}$ направлен вдоль оси сверхрешётки), используя распределение Больцмана, находим

$$\Gamma_q = \frac{g_2}{|\sin(qd/2)|}\sum_{m=2}^{\infty}\left\{\frac{J_m^2(a)}{m^2}\left[\Theta(1-b_m^2)\exp[\beta\cos(k_md)]\,\frac{\cos^2[(k_m+q)\,d] + \cos^2(k_md)}{\sqrt{1-b_m^2}} - \right.\right.$$
$$\left.\left. - \Theta(1-\bar b_m^2)\exp[\beta\cos(\bar k_md)]\,\frac{\cos^2[(\bar k_m-q)\,d] + \cos^2(\bar k_md)}{\sqrt{1-\bar b_m^2}}\right]\right\}$$

при чётных m,

$$\Gamma_q = \frac{g_2}{|\sin(qd/2)|}\sum_{m=1}^{\infty}\left\{\frac{J_m^2(a)}{m^2}\left[\Theta(1-b_m^2)\exp[\beta\cos(k_md)]\,\frac{\sin^2[(k_m+q)\,d] + \sin^2(k_md)}{\sqrt{1-b_m^2}} - \right.\right.$$
$$\left.\left. - \Theta(1-\bar b_m^2)\exp[\beta\cos(\bar k_md)]\,\frac{\sin^2[(\bar k_m-q)\,d] + \sin^2(\bar k_md)}{\sqrt{1-\bar b_m^2}}\right]\right\} \tag{9}$$

при нечётных m. Здесь

$$g_2 = \frac{\Xi^2 n_0 q\Delta}{4\rho v_{\mathrm{s}}\hbar^2\omega^2 I_0(\beta)}\,, \qquad k_m = -\,\frac{q}{2} + \frac{1}{d}\arcsin b_m, \qquad \bar k_m = -\,\frac{q}{2} + \frac{1}{d}\arcsin \bar b_m,$$

$$b_m = \frac{m\hbar\omega + \hbar\omega_q}{\Delta\sin(qd/2)}\,, \qquad \bar b_m = \frac{m\hbar\omega - \hbar\omega_q}{\Delta\sin(qd/2)}\,, \qquad q = q_z, \qquad \Theta(x) = \begin{cases}1, & x > 0;\\ 0, & x < 0.\end{cases}$$

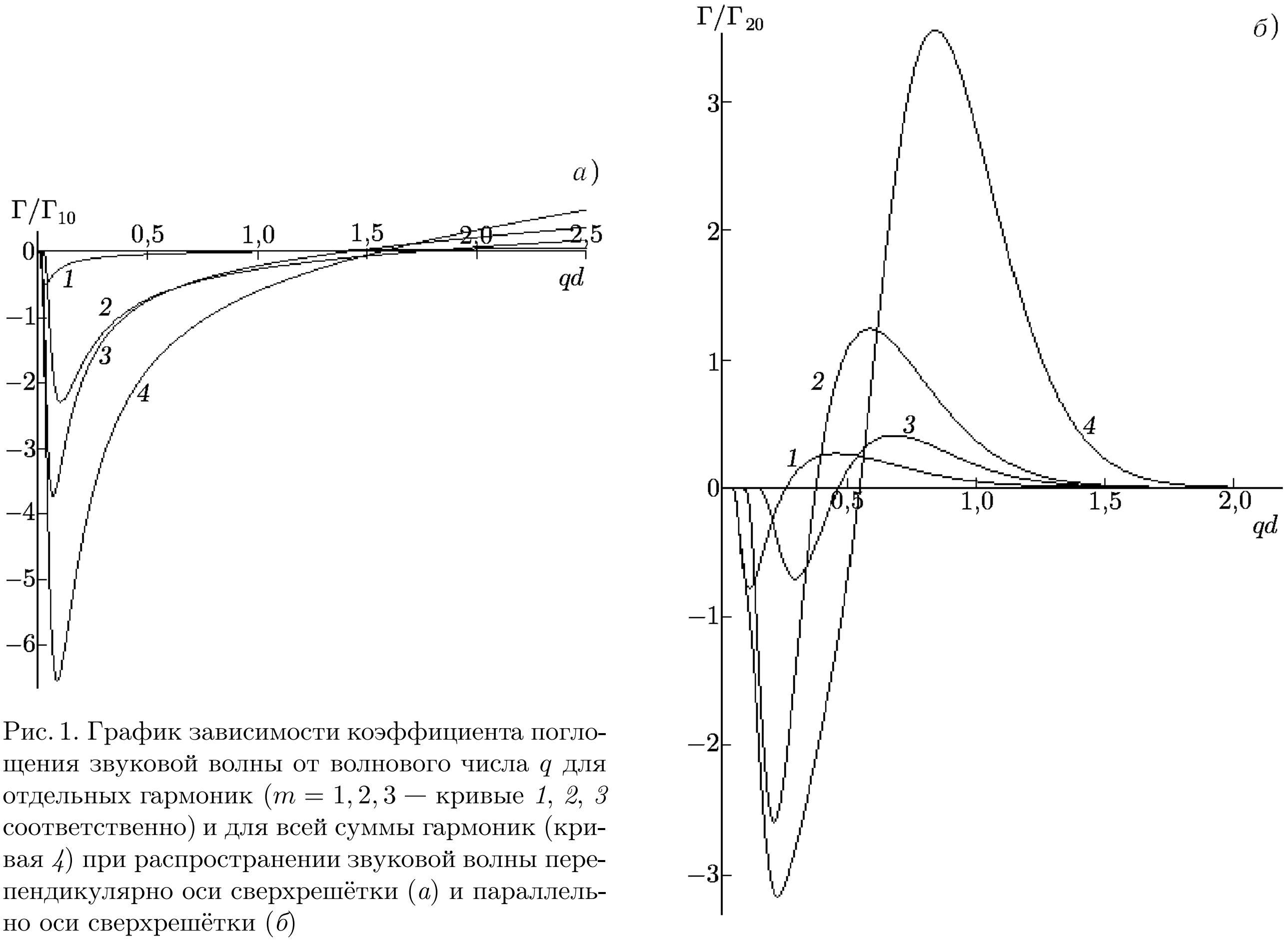

Рис. 1. График зависимости коэффициента поглощения звуковой волны от волнового числа q для отдельных гармоник ($m = 1, 2, 3$ — кривые *1*, *2*, *3* соответственно) и для всей суммы гармоник (кривая *4*) при распространении звуковой волны перпендикулярно оси сверхрешётки (*а*) и параллельно оси сверхрешётки (*б*)

Учитывая, что $\hbar\omega_q/\Delta \ll 1$, проводим разложение в ряд по этой величине членов сумм (9). В результате для чётных членов ряда (m — чётное) получим

$$\Gamma_q = \frac{g_2}{|\sin(qd/2)|} \sum_{m=2}^{\infty} \Theta(\kappa_m^2)\, \frac{\hbar\omega_q}{\Delta}\, \frac{J_m^2(a)}{\kappa_m m^2} \exp\bigl\{\Delta[\kappa_m \cos(qd/2) + m\hbar\omega/\Delta]/(k_0T)\bigr\} \times$$
$$\times \left[\left(\frac{\Delta}{k_0T} - \frac{m\hbar\omega\cos(qd/2)}{k_0T\kappa_m \sin^2(qd/2)} + \frac{m\hbar\omega}{\Delta\kappa_m^2 \sin^2(qd/2)}\right)\left(\cos^2(qd/2) - \frac{m\hbar\omega}{\Delta\sin(qd/2)}\cos(qd)\right) - \right.$$
$$\left. - \frac{2\hbar\omega\cos(qd)}{\Delta\sin^2(qd/2)}\right], \qquad (10)$$

где $\kappa_m = \sqrt{1 - (m\hbar\omega)^2/[\Delta\sin(qd/2)]^2}$. Для нечётных m в (10) нужно заменить член $\cos^2(qd/2)$ во второй скобке ряда на $\sin^2(qd/2)$ и знак перед последним слагаемым с минуса на плюс. Предполагая выполнение неравенств $\Delta/(k_0T) \gg m\hbar\omega/\Delta$, $\hbar\omega/\Delta \ll 1$, $(m\hbar\omega)^2/[\Delta\sin(qd/2)]^2 \ll 1$, находим, что волновые числа q_m, при которых каждый из членов суммы (10) меняет знак, пропорциональны $\sqrt{m}$. Это подтверждается также и графиками зависимости от q отдельных членов суммы (10) на рис. 1*б*, которые построены для следующих параметров: $\Delta = 0{,}05$ эВ, $\hbar\omega = 0{,}05\Delta$, $T = 20$ К, $E = 2\cdot10^4$ В/см, $d = 10^{-6}$ см, $n_0 = 10^{15}$ см$^{-3}$. Здесь же построен и график всей суммы (с учётом и нечётных членов) при тех же значениях параметров. Из графиков видно, что при

$eEd/(\hbar\omega) > 1$ необходимо учитывать многофотонные члены. Отметим также, что многофотонное поглощение нужно учитывать при значительно меньшей величине электромагнитного поля, чем в однородном полупроводнике. Это объясняется тем, что аргумент функции Бесселя пропорционален постоянной сверхрешётки, которая значительно больше постоянной основной решётки. Как видно из (10), при $m\hbar\omega/[\Delta\sin(qd/2)] = 1$ имеем $\kappa_m = 0$, и те члены, в знаменателях которых стоит κ_s, обращаются в бесконечность. Чтобы избавиться от расходимостей, необходимо феноменологически ввести время затухания квазичастицы-электрона τ, определяемое процессами релаксации, и заменить в (5) δ-функции лоренцевской кривой с полушириной τ^{-1}. Здесь мы ограничимся численным определением выражения (10) при $(m\hbar\omega)^2/[\Delta\sin(qd/2)]^2 \ll 1$. Область смены знака коэффициента затухания Γ_q находится в интервале 10^4 см$^{-1}$ $< q < 10^5$ см$^{-1}$. Оценка величины Γ_q при приведённых выше параметрах даёт $\Gamma_q/v_{\rm s} = 10^3$ см$^{-1}$.

СПИСОК ЛИТЕРАТУРЫ

1. Эпштейн Э. М. // Письма в ЖЭТФ. 1971. Т. 13, № 9. С. 511; ФТП. 1977. Т. 11, № 2. С. 421.
2. Эпштейн Э. М. // ФТП. 1977. Т. 11, № 2. С. 421.
3. Васько Ф. Т. // ФТТ. 1977. Т. 19, № 11. С. 3 274.
4. Гуревич В. Л., Паршин Д. А. // ЖЭТФ. 1977. Т. 72, № 4. С. 1 589.
5. Эпштейн Э. М., Шмелёв Г. М., Цуркан Г. И. Фотостимулированные процессы в полупроводниках. Кишинев: Штиница, 1987. 168 с.
6. Mensah S. Y., Allotey F. K., Adypongs K. // Superlattices and Microstructures. 1997. V. 22, No. 4. P. 453.
7. Shmelev G. M., Mensah S. Y., Tsurkan G. I. // J. Phys. C. 1988. V. 21. P. 1 073.

Поступила в редакцию 30 апреля 2004 г.; принята в печать 23 ноября 2004 г.

ACOUSTIC-PHONON GENERATION IN A SEMICONDUCTOR SUPERLATTICE DUE TO INTRABAND ABSORPTION OF AN ELECTROMANGETIC WAVE

M. V. Vyazovsky and G. A. Syrodoev

We find the decay rate of an acoustic wave in the case of intraband multiphoton absorption of an electromagnetic wave in a superlattice. The domains of sign reversal of the acoustic decay rate are found for two directions, namely, longitudinal and transverse with respect to the superlattice axis, of acoustic-wave propagation. The series describing the acoustic-wave growth rate for typical superlattice parameters is evaluated numerically. The numerical estimate for the growth rate is given. It is pointed out that multiphoton absorption should affect the acoustic-wave growth rate even if the field is much weaker than in an ordinary semiconductor.

УДК 621.3+535.2

ОБОБЩЁННЫЙ РЕЛЕЕВСКИЙ КРИТЕРИЙ БИНАРНОГО РАЗРЕШЕНИЯ

Г. Н. Бочков, А. Т. Гаврилин, К. В. Горохов

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

Предлагается критерий разрешения однотипных некогерентных сигналов по двум скалярным параметрам и оцениваются соответствующие ему пределы разрешения.

Под критерием разрешения Релея в радиолокации понимают следующее соглашение: два источника сигналов уверенно разрешаются по некоторому скалярному параметру, если суммарная сигнальная функция этого параметра на выходе устройства первичной обработки имеет форму двугорбой кривой, при одногорбой кривой разрешения нет [1]. В случае, когда параметром разрешения является время прихода сигналов, а линейные тракты приёмника узкополосны, в качестве сигнальной функции чаще берут огибающую сигналов этих трактов, предполагая некогерентное сложение сигналов разных источников.

Для большинства возникающих или используемых в оптике и радиотехнике сигналов суммарная сигнальная функция становится двугорбой, когда сдвиг сигналов по параметру сравнивается с шириной парциальной сигнальной функции. Обычно в качестве этой ширины принимают либо её среднеквадратичное значение, либо ширину функции на уровне 0,5 от максимума. В то же время достаточно типична ситуация, когда оптик с определённым опытом наблюдений в состоянии «на глазок» разрешить изображения двух точечных источников, расстояния между которыми меньше радиуса диска Эйри [2]. Этот пример показывает, что отождествление пределов разрешения с шириной сигнальной функции не вполне корректно.

Ф. М. Вудворд [3], по-видимому, первым отметил различие задач оптимального оценивания параметра сигнала на фоне шума и оптимального разрешения двух сигналов по этому параметру. Если дисперсия оптимальной оценки при больших отношениях сигнал/шум обратно пропорциональна квадрату второй производной сигнальной функции по параметру, вычисленной в точке максимума [4], то для анализа разрешающей способности параболическая аппроксимация сигнальной функции в окрестности упомянутой точки не вполне адекватна: графическое сложение двух одинаково ориентированных парабол может дать только параболу. Эти рассуждения Вудворда можно распространить и на двумерный случай, заменив в них параболы на эллиптические параболоиды.

В данной работе предлагается математически формализованный критерий разрешения по двумерному параметру $(x,\, y)$ и оцениваются соответствующие этому критерию пределы разрешения.

Модуль сигнальной функции параметра $(x,\, y)$ назовём функцией неопределённости и будем обозначать $C(x,y)$. Не ограничивая общности, считаем, что существует конечная окрестность точки $(x=0, y=0)$, в которой функция $C(x,y)$ имеет единственный максимум, равный единице. Тогда предлагаемый критерий формулируется следующим образом: два однотипных и равномощных сигнала уверенно разрешаются, если сумма их функций неопределённости $C(x-x_1, y-y_1)+ + C(x-x_2, y-y_2)$ в упомянутой окрестности имеет седловую точку. Напомним, что седловой точкой поверхности называется точка, вблизи которой поверхность лежит по разные стороны от своей касательной плоскости (рис. 1).

Вынесенное в заголовок статьи название сформулированного критерия продиктовано двумя причинами. Во-первых, критерий разрешения по Релею в его оптической версии относится именно к двумерным рельефам освещённости, хотя и достаточно специального вида (лоренцианов и гауссианов). Во-вторых, сечение графика функции вертикальной плоскостью, проходящей через седловую точку графика в направлении максимального возрастания функции, имеет в этой точке характерный провал в согласии с одномерным критерием Релея.

Сформулированный выше критерий наследует основные достоинства и ограничения классического релеевского критерия, и мы не будем сейчас на них останавливаться, а сосредоточим внимание на конструктивных возможностях критерия в части, касающейся синтеза сигналов с высокой разрешающей способностью и оценки пределов разрешения. С этой точки зрения приведённое выше геометрическое определение седловой точки является малоконструктивным, поэтому мы несколько усилим его, предположив, что исследуемая поверхность гладкая (задаётся дважды непрерывно дифференцируемыми функциями), и отождествим седловую точку с точкой, в окрестности которой поверхность имеет отрицательную гауссову кривизну [5].

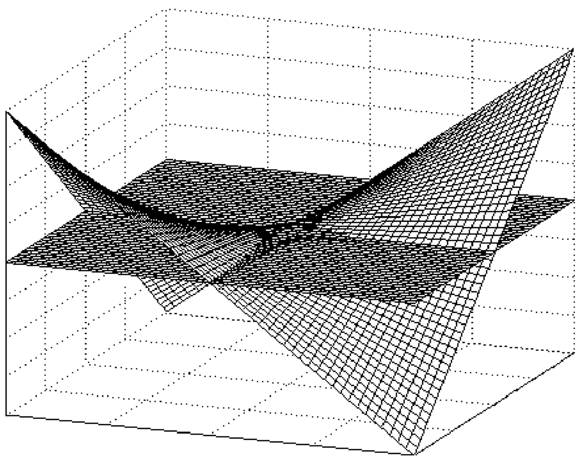

Рис. 1

Для замкнутости изложения приведём вкратце необходимые сведения по данному вопросу. Известно [5], что внутреннюю геометрию поверхности $\mathbf{r}\colon \mathrm{R}^2 \to \mathrm{R}^3$ характеризует 2-я квадратичная форма:

$$\Pi = -(\mathrm{d}\mathbf{r}, \mathrm{d}\mathbf{n}) = (\mathrm{d}^2\mathbf{r}, \mathbf{n}) = (\mathbf{r}_{uu}, \mathbf{n})\,\mathrm{d}u^2 + 2\,(\mathbf{r}_{uv}, \mathbf{n})\,\mathrm{d}u\,\mathrm{d}v + (\mathbf{r}_{vv}, \mathbf{n})\,\mathrm{d}v^2 \equiv L\,\mathrm{d}u^2 + 2M\,\mathrm{d}u\,\mathrm{d}v + N\,\mathrm{d}v^2,$$

где $\mathbf{r}(u, v)$ — радиус-вектор текущей точки поверхности, $\mathbf{n} = [\mathbf{r}_u, \mathbf{r}_v]/|[\mathbf{r}_u, \mathbf{r}_v]|$ — единичная нормаль к поверхности, $\mathbf{r}_u$ и $\mathbf{r}_v$ — касательные векторы к поверхности, ориентированные вдоль координатных линий: $\mathbf{r}(u, v = \mathrm{const})$ и $\mathbf{r}(u = \mathrm{const}, v)$ соответственно.

Седловая (гиперболическая) точка, характеризуется тем, что в ней форма Π знакопеременна, т. е. в зависимости от отношения $\mathrm{d}u/\mathrm{d}v$ может принимать как положительные, так и отрицательные значения. Гауссова кривизна K поверхности, определяемая отношением

$$K = \begin{vmatrix} L & M \\ M & N \end{vmatrix} \Big/ |[\mathbf{r}_u, \mathbf{r}_v]|^2,$$

в седловой точке отрицательна [5].

В случае явного задания поверхности, когда $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + C(x, y)\mathbf{k}$, имеем

$$\Pi = \frac{\partial^2 C}{\partial x^2}\,\mathrm{d}x^2 + 2\,\frac{\partial^2 C}{\partial x \partial y}\,\mathrm{d}x\,\mathrm{d}y + \frac{\partial^2 C}{\partial y^2}\,\mathrm{d}y^2,$$

$$K = \frac{\dfrac{\partial^2 C}{\partial x^2}\,\dfrac{\partial^2 C}{\partial y^2} - \left(\dfrac{\partial^2 C}{\partial x \partial y}\right)^2}{\sqrt{1 + \left(\dfrac{\partial C}{\partial x}\right)^2 + \left(\dfrac{\partial C}{\partial y}\right)^2}}\,.$$

Обратимся вновь к суммарной функции неопределённости, возникающей при некогерентном сложении двух сигналов:

$$C_\Sigma(x,y) = C(x-x_1, y-y_1) + C(x-x_2, y-y_2).$$

Из соображений симметрии понятно, что седловая точка у $C_\Sigma(x,y)$ может появиться в середине отрезка, соединяющего точки (x_1, y_1) и (x_2, y_2). Перенесём начало координат в эту срединную точку и обозначим $a = (x_2 - x_1)/2$, $b = (y_2 - y_1)/2$. Тогда суммарная функция неопределённости запишется в виде

$$C_\Sigma(x,y) = C(x-a, y-b) + C(x+a, y+b). \tag{1}$$

Нас интересуют условия возникновения седла в начале координат (рис. 2). Полагая в (1) $x = 0$, $y = 0$, имеем

$$C_\Sigma(0,0) \equiv C_0(a,b) = C(-a,-b) + C(a,b). \tag{2}$$

Если функция неопределённости обладает центральной симметрией, т. е.

$$C(-x,-y) = C(x,y),$$

то выражение (2) упрощается:

$$C_0(a,b) = 2C(a,b). \tag{3}$$

Равенство (3) сводит задачу исследования суммарной функции неопределённости к рассмотрению расположения гиперболических точек на графике одной из парциальных функций неопределённости.

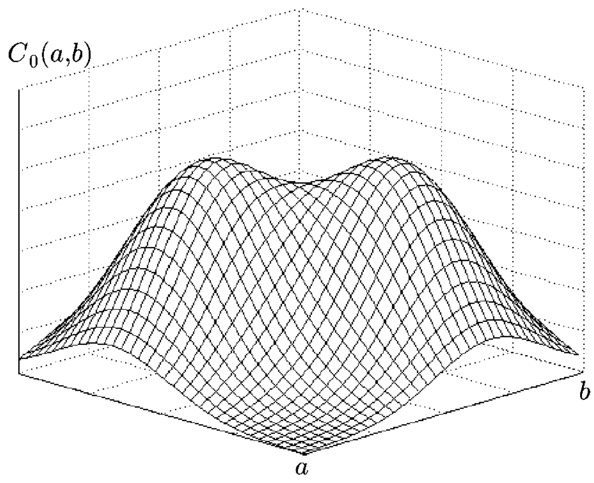

Рис. 2

Рассмотрим частный, но общепринятый в радиолокации случай, когда в качестве функции неопределённости выступает модуль классической функции Вудворда [3]

$$\chi(x,y) = \int\limits_{-\infty}^{+\infty} \dot s(t - x/2)\dot s^*(t + x/2)\exp(jyt)\,\mathrm{d}t \tag{4}$$

узкополосного сигнала с комплексной амплитудой $\dot s(t)$. Тогда для большинства используемых в оптике и радиотехнике сигналов функция $C(x,y) = |\chi(x,y)|^2$ в окрестности локального максимума обладает положительной гауссовой кривизной ($K(x,y) > 0$), а на некотором удалении от максимума $K(x,y)$ меняет знак. При этом точки на поверхности с $K(x,y) = 0$ образуют замкнутую линию — «параболический пояс».

Таким образом, для обеспечения высокой разрешающей способности по предложенному критерию нужно стремиться к минимизации размеров упомянутого пояса.

Нормированная на энергию сигнала функция (4) обладает следующим замечательным свойством:

$$\frac{1}{2\pi}\iint\limits_{-\infty}^{+\infty} |\chi(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = 1, \tag{5}$$

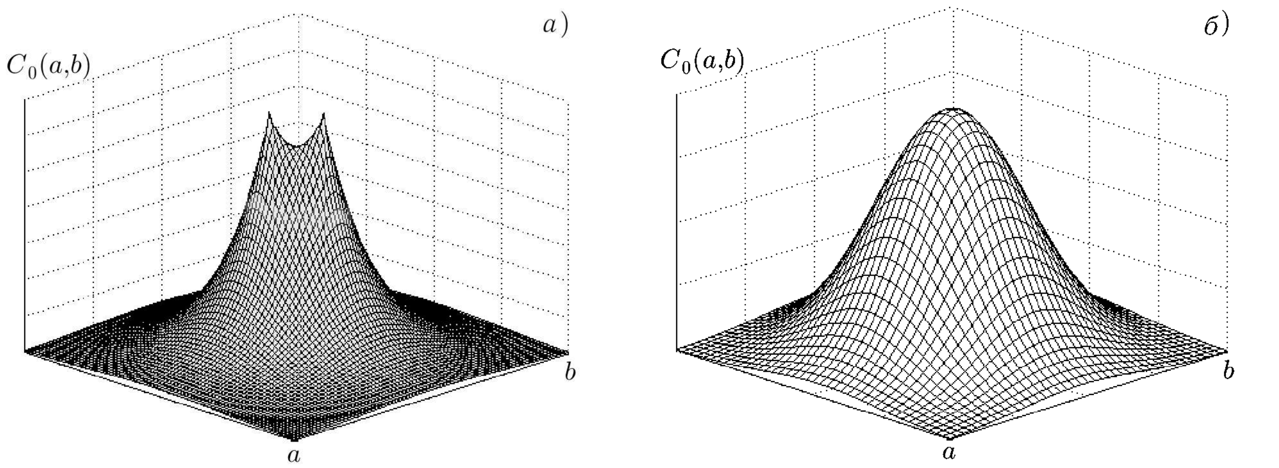

Рис. 3

которое часто трактуют как невозможность совместного высокого разрешения по x и y. Между тем само по себе равенство (5) не препятствует достижению сколь угодно малых пределов разрешения по двум параметрам. Проиллюстрируем это на следующем примере.

Пусть функция неопределённости описывается выражением $C(x,y) = \exp\left[-\frac{1}{2}\sqrt{x^2+y^2}\right]$ и, как нетрудно убедиться, удовлетворяет условию (5). Эта функция не дифференцируема в начале координат, а во всех остальных точках имеет отрицательную гауссову кривизну, что согласно нашему критерию обеспечивает совместное сверхразрешение по обоим параметрам (рис. 3*а*). Для сравнения на рис. 3*б* изображена сумма сигнальных функций гауссовой формы при аналогичных параметрах разнесения источников. Здесь мы намеренно отвлекаемся от рассмотрения влияния шумов, чтобы подчеркнуть принципиальную сторону дела. Это тем более важно, что в качестве функции неопределённости может выступать не только абсолютная величина функции Вудворда, но и, например, смешанные двумоментные кумулянты [6] (в частности, функции асимметрии и эксцесса) стационарных случайных сигналов или двумерные сечения полиспектров [6] (в частности, биспектры). Указанные характеристики нечувствительны к аддитивным гауссовским помехам, что делает предложенный критерий наиболее адекватным применительно к разрешению сигналов по статистикам высших порядков.

Тем не менее рассмотрим вкратце классический вариант, когда сигнальная функция линейна по отношению к входному сигналу. В этом случае она входит в достаточную для разрешения сигналов статистику [4] аддитивно наряду с шумовой функцией. Последняя как функция временно́й задержки и доплеровского смещения представляет собой гауссовское случайное поле, являясь откликом семейства линейных фильтров на гауссовский шум. Статистика локальных максимумов и минимумов, а также седловых точек случайных полей не поддаётся строгому аналитическому исследованию [7], однако в первом приближении понятно, что для надёжного разрешения достаточно обеспечить глубину седла, которая бы превосходила среднеквадратическое значение флуктуаций указанного поля. На самом деле требования к интенсивности помех можно ослабить, если учесть, что при дельта-коррелированном по времени и частоте входном шуме корреляционная функция поля по форме совпадает с функцией неопределённости [4]. Последнее означает, что характерные масштабы изменчивости случайного поля сравнимы с масштабами функции неопределённости, поэтому крутые (в этих масштабах) всплески и впадины шумовой функции крайне маловероятны в пределах технически реализуемых апертур. Таким образом, предложенный критерий обеспечивает высокую надёжность разрешения не только в случае предельно больших отношений сигнал/шум, но и в большинстве практически значимых ситуаций.

Работа выполнена при поддержке Совета при Президенте РФ по поддержке ведущих научных школ (проект НШ-1729.2003.2) и НП «Университеты России» (проект УР.01.01.020).

СПИСОК ЛИТЕРАТУРЫ

1. Васин В. В. Радиолокационные устройства (теория и техника построения). М.: Сов. радио, 1970. 680 с.
2. Ландсберг Г. С. Оптика: 5 изд. М., 1976. 926 с.
3. Вудворд Ф. М. Теория вероятностей и теория информации с приложениями в радиолокации. М.: Сов. радио, 1955. 132 с.
4. Куликов Е. И., Трифонов А. П. Оценка параметров сигналов на фоне помех. М.: Сов. радио, 1978. 296 с.
5. Дубровин Б. А., Новиков С. П., Фоменко А. Т. Современная геометрия: Методы и приложения. М.: Наука, 1986. 760 с.
6. Малахов А. Н. Кумулянтный анализ случайных процессов и их преобразований. М.: Сов. радио, 1978. 520 с.
7. Лонге-Хиггинс М. С. Гидродинамическая неустойчивость. М.: Изд-во иностр. лит-ры, 1964. С. 124.

Поступила в редакцию 16 мая 2004 г.; принята в печать 27 декабря 2004 г.

GENERALIZED RAYLEIGH CRITERION OF TWO-POINT RESOLUTION

G. N. Bochkov, A. T. Gavrilin, and K. V. Gorokhov

The criterion of two-dimensional incoherent signals resolution is suggested. Corresponding limits are estimated.

УДК 621.396.67.01

ХАРАКТЕРИСТИКИ ОБНАРУЖЕНИЯ ПРОСТРАНСТВЕННЫХ СИГНАЛОВ ДЛЯ СТАТИСТИК ОБОБЩЁННОГО ОТНОШЕНИЯ ПРАВДОПОДОБИЯ В СЛУЧАЕ КОРОТКИХ ВЫБОРОК

О. В. Болховская, А. А. Мальцев, К. В. Родюшкин

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

В работе исследуются характеристики обнаружения пространственных сигналов антенной решёткой в случае коротких выборок. Все исследуемые решающие статистики получены на основе обобщённого отношения правдоподобия для выборок произвольного объёма. В качестве примера построены рабочие характеристики обнаружения полезного сигнала, представляющего собой одну или две плоских когерентных волны на фоне пространственно-однородного и неоднородного шумов. Проведено сравнение эффективности используемых статистик с χ^2-критерием для обнаружения шумоподобных сигналов.

ВВЕДЕНИЕ

Проблема обнаружения пространственных сигналов с неизвестными параметрами в случае коротких выборок является актуальной в гидроакустике, радиолокации и связи. При этом задача обнаружения, как правило, решается путём обработки выходных сигналов многоэлементной антенной решётки с использованием критерия обобщённого отношения правдоподобия (generalized likelihood ratio (GLR)) [1–6]. Хорошо известно, что в случае большого объёма выборки этот подход является асимптотически оптимальным [6, 7]. Однако теоретически обосновать оптимальность GLR-метода для коротких выборок достаточно сложно. Поэтому применение обобщённого отношения правдоподобия в случае коротких выборок требует проведения дополнительных исследований. В работах [8, 9] была оценена эффективность использования тест-статистик на основе обобщённого отношения правдоподобия в случае коротких выборок для критерия Неймана—Пирсона. В настоящей работе приводятся результаты сравнительного анализа рабочих характеристик GLR-статистик, полученных при различном объёме априорной информации о принимаемом полезном сигнале и шуме.

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим p-элементную узкополосную приёмную антенную решётку с произвольным расположением датчиков. Будем считать, что сигналы с элементов антенны образуют p-мерный вектор $\mathbf{z}$, являющийся комплексным случайным гауссовским вектором. Предполагается, что осуществляется N выборок выходного сигнала антенной решётки $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N$, которые являются статистически независимыми, одинаково распределёнными случайными векторами с нулевым средним значением и пространственной ковариационной матрицей $\mathbf{\Sigma}$.

Задача обнаружения узкополосного пространственно-коррелированного полезного сигнала антенной решёткой формулируется как классическая двухальтернативная задача различения двух гипотез: гипотезы H_0, состоящей в том, что принимается только шум, и гипотезы H_1, когда принимается аддитивная смесь сигнала и шума.

В рассматриваемой постановке задачи обнаружения в зависмости от имеющейся априорной информации рассматривалось шесть вариантов задания характеристик шума гипотезами

H_{0i} (пространственными ковариационными матрицами $\mathbf{\Sigma}_{0i}$) и полезного сигнала гипотезами H_{1i} (пространственными ковариационными матрицами $\mathbf{\Sigma}_{1i}$), где i — номер варианта. Получаемые при этом решающие GLR-статистики приведены в табл. 1, где a_{ij} — элементы ненормированной ковариационной матрицы

$$\mathbf{A} = \sum_{n=1}^{N} \mathbf{z}_n \mathbf{z}_n^{\dagger}, \tag{1}$$

$\mathbf{z}_n$ — n-ое выборочное значение вектора выходного сигнала антенной решётки, $\mathrm{sp}(\mathbf{A})$ и $|\mathbf{A}|$ обозначают соответственно след и детерминант матрицы $\mathbf{A}$.

Как видно, в первых трёх вариантах предполагалось, что полезным сигналом является любой сигнал, характеристики которого отличны от характеристик шума. В трёх последних вариантах дополнительно к характеристикам шума задавалась и определённая априорная информация о полезном сигнале. Характеристики шума в различных вариантах задавались в виде сложной или простой гипотезы.

В первом варианте (сложная гипотеза H_0, статистика V_1) предполагалось отсутствие любой априорной информации о шуме кроме условия независимости его отсчётов в различных элементах антенны.

Во втором варианте (сложная гипотеза H_0, статистика V_2) дополнительно к независимости шума в элементах антенны предполагалось наличие априорной информации о его однородности (одинаковой мощности в разных элементах антенны: $\sigma_{ii}^2 = \sigma^2 = \mathrm{const}$).

В третьем варианте (простая гипотеза H_0, статистика V_3) ковариационная матрица шума предполагалась полностью известной (диагональной, с известной дисперсией σ^2). Очевидно, что в этом случае с помощью нормировки ковариационная матрица шума всегда может быть сведена к единичной.

В четвёртом варианте (простая гипотеза H_0, статистика V_4) кроме полной априорной информации о шуме в антенных элементах (гауссовость, независимость, однородность, единичная мощность) дополнительно предполагается, что полезный сигнал полностью пространственно-когерентен. В этом случае статистика V_4 определяется максимальным собственным числом $\hat{\lambda}_1$ выборочной ковариационной матрицы сигнала

$$\hat{\mathbf{\Sigma}} = \mathbf{A}/N. \tag{2}$$

В пятом варианте (сложная гипотеза H_0, статистика V_5) рассматривалась задача обнаружения шумового пространственно-неоднородного сигнала на фоне однородного собственного шума антенны с априори неизвестной мощностью.

В шестом варианте (простая гипотеза H_0, статистика V_6) в случае гипотезы H_0 ковариационная матрица шума равна единичной, однако в случае альтернативной гипотезы H_1 ковариационная матрица сигнала не произвольная (как для статистики V_3), а представляет собой сумму единичной матрицы и произвольной матрицы $\mathbf{M}$:

$$\mathbf{\Sigma} = \mathbf{I} + \mathbf{M}. \tag{3}$$

В этом случае нахождение статистики в целом происходит так же, как и в третьем варианте, однако для максимально правдоподобной оценки матрицы $\mathbf{M}$ получается следующее выражение:

$$\hat{\mathbf{M}} = \mathbf{A}/N - \mathbf{I}, \tag{4}$$

где

$$\mathbf{A} = N\,(\mathbf{I} + \hat{\mathbf{M}}) = N\hat{\mathbf{\Sigma}}. \tag{5}$$

Очевидно, что матрица $\mathbf{M}$ должна быть эрмитовской и положительно-определённой. Из этого следует, что каждое собственное число матрицы $\mathbf{\Sigma}$, являющейся суммой двух матриц (3), должно быть не меньше единицы. Поэтому все собственные числа оценки этой матрицы $\hat{\mathbf{\Sigma}}$, меньшие единицы, следует заменить на единичные. Матрицу, удовлетворяющую этому требованию и получаемую из $\hat{\mathbf{\Sigma}}$ путём такой замены, обозначим $\tilde{\mathbf{\Sigma}}$, а соответствующую ей выборочную ненормированную матрицу обозначим $\tilde{\mathbf{A}}$.

Вывод точных аналитических выражений для GLR-статистик V_1, V_2, V_3 и V_4, приведённых в табл. 1, дан в работах [8, 10, 11]. Приведённые в табл. 1 выражения для статистик V_5 и V_6 легко могут быть получены аналогичным образом. Выражения, найденные для решающих статистик V_1, V_2, V_3, V_5 и V_6 (см. табл. 1) позволяют найти точные аналитические формулы для моментов любого порядка и представить их функции распределения в виде ряда по ортогональным полиномам Якоби [11–13]. Точное выражение для функции распределения максимального собственного числа $\hat{\lambda}_1$, через которое выражается статистика V_4, было получено в работе [10]. Это дало возможность провести сравнение рабочих характеристик всех приведённых в табл. 1 решающих статистик в различных случаях.

2. ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

Было проведено численное моделирование схем обнаружения пространственного сигнала, работающих на основе статистик V_1–V_6. Для обнаружения пространственных сигналов использовалась 5-элементная линейная эквидистантная антенная решётка ($p = 5$) с расстоянием между соседними элементами, равным половине длины волны.

В первом варианте схемы обнаружения в соответствии с нулевой гипотезой H_{01} собственный шум моделировался как неоднородный (имеющий разные мощности в антенных элементах) с диагональной ковариационной матрицей

$$\Sigma_{01} = \begin{pmatrix} \sigma_{11}^2 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp}^2 \end{pmatrix}. \tag{6}$$

При этом мощности собственных шумов в элементах антенной решётки брались равными 0,2; 0,2; 1; 1,8 и 1,8 соответственно, что соответствует достаточно сильно неоднородному шуму. Во всех остальных вариантах схемы обнаружения шум моделировался как однородный, с одинаковой мощностью в антенных элементах и ковариационной матрицей

$$\mathbf{\Sigma}_{0i} = \sigma^2 \mathbf{I}, \tag{7}$$

где $i = 2, \dots, 6$. Однако во втором и пятом вариантах дисперсия (мощность) шума предполагалась неизвестной, а в третьем, четвёртом и шестом вариантах известной и равной единице:

$$\mathbf{\Sigma}_{0i} = \mathbf{I}, \tag{8}$$

где $i = 3, 4, 6$.

Во всех рассматриваемых случаях след ковариационной матрицы шума оставался равным числу элементов антенной решётки:

Таблица 1

Номер варианта i	GLR-статистика	Физический смысл гипотезы H_{0i}	Физический смысл гипотезы H_{1i}
1	$V_1 = \dfrac{\lvert\mathbf{A}\rvert}{\prod\limits_{i=1}^{p} a_{ii}}$	независимость шумов	сигнал с любой ковариационной матрицей, отличной от $\mathbf{\Sigma}_{01}$
2	$V_2 = \dfrac{\lvert\mathbf{A}\rvert}{[\mathrm{sp}(\mathbf{A})/p]^p}$	независимость и однородность шумов	сигнал с любой ковариационной матрицей, отличной от $\mathbf{\Sigma}_{02}$
3	$V_3 = \left(\dfrac{e}{N}\right)^p \lvert\mathbf{A}\rvert \exp[-\mathrm{sp}(\mathbf{A})/N]$	независимость, однородность шумов и нормировка их мощности на единицу	сигнал с любой ковариационной матрицей, отличной от $\mathbf{\Sigma}_{03}$
4	$V_4 = e\,(\hat{\lambda}_1^{-1} \exp \hat{\lambda}_1)^{-1}$	независимость, однородность шумов и нормировка их мощности на единицу	сигнал в виде когерентной волны с априори неизвестным волновым фронтом
5	$V_5 = \dfrac{\prod\limits_{i=1}^{p} a_{ii}}{[\mathrm{sp}(\mathbf{A})/p]^p}$	независимость и однородность шумов	сигналом является шум с независимыми отсчётами
6	$V_6 = \left(\dfrac{e}{N}\right)^p \lvert\tilde{\mathbf{A}}\rvert \exp[-\mathrm{sp}(\mathbf{A})/N]$	независимость, однородность шумов и нормировка их мощности на единицу	ковариационная матрица сигнала является суммой единичной и произвольной положительно-определённой матриц

$$\mathrm{sp}(\mathbf{\Sigma}_{0i}) = p. \tag{9}$$

Рассматривались три модели полезного сигнала. В первой, простейшей модели полезный сигнал представлял собой плоскую когерентную волну, падающую нормально к плоскости апертуры антенной решётки. В этом случае ковариационная матрица $\hat{\mathbf{\Sigma}}$ сигналов антенной решётки имеет одно максимальное собственное число («сигнальное»), равное сумме мощности сигнала и единицы, а все остальные собственные числа («шумовые») равны единице.

Во второй модели полезный сигнал представлял собой сумму двух независимых плоских волн, одна из которых падала нормально, а другая под углом (порядка 0,41 рад), соответствующим первому нулю диаграммы направленности для 5-элементной антенной решётки из полуволновых вибраторов [14]. В этом случае ковариационная матрица сигналов антенной решётки имеет два одинаковых максимальных собственных числа (равных сумме единицы и мощности каждого из сигналов), а все остальные собственные числа равны единице.

В третьей модели полезный сигнал задавался как однородный или неоднородный шум неизвестной мощности.

Было проведено сравнение эффективности использования статистик V_1–V_6 для обнаружения описанных выше полезных сигналов и шумов. Исследовались рабочие характеристики обнаружителей (вероятности правильного обнаружения в зависимости от вероятности ложной тревоги) при заданном отношении сигнал/шум (отношение мощности полезного сигнала к мощности шума в одном антенном элементе). Отношение сигнал/шум во всех экспериментах задавалось равным -5 дБ (как наиболее показательное для 5-элементной антенной решётки при используемом объёме выборки $N = 15$).

На рис. 1*а*, *б* представлены рабочие характеристики обнаружения когерентного сигнала в виде плоской волны на фоне соответственно однородного и неоднородного собственного шума для всех исследуемых статистик. Как видно из приведённых результатов, в первом случае (при однородном шуме) наиболее эффективной является статистика V_4, вероятность правильного обнаружения P_{RD} для которой составляет около 0,9 при вероятности ложной тревоги $P_{\mathrm{FA}} = 0{,}1$. При таком же значении вероятности ложной тревоги статистики V_6, V_3, V_1 и V_2 имеют вероятности правильного обнаружения $P_{\mathrm{RD}} = 0{,}8$; 0,7; 0,59 и 0,57 соответственно; статистика V_5 вообще непригодна для обнаружения в данном случае, поскольку условия эксперимента не соответствуют условиям её применимости.

В случае, когда шумовой фон становится неоднородным, статистика V_1, предназначенная для обнаружения на фоне неоднородного шума, становится самой эффективной. При вероятности ложной тревоги $P_{\mathrm{FA}} = 0{,}1$ вероятности правильного обнаружения для статистик V_1, V_6 и V_4 равны 0,9; 0,8 и 0,6 соответственно, а статистики V_2, V_3 и V_5 перестают быть пригодными для обнаружения.

На рис. 2 представлены рабочие характеристики статистик V_1–V_6 при обнаружении частично когерентного сигнала в виде суммы двух плоских волн от независимых источников на фоне однородного и неоднородного шума. Из представленных результатов видно, что наиболее эффективным для обнаружения суммы двух плоских волн на фоне однородного шума (см. рис. 2*а*) является использование статистик V_4 или V_6 (их эффективность в этом случае практически одинакова). При вероятности ложной тревоги $P_{\mathrm{FA}} = 0{,}1$ вероятность правильного обнаружения для этих статистик примерно равна $P_{\mathrm{RD}} = 0{,}8$. При этой вероятности ложной тревоги для статистики V_3 вероятность правильного обнаружения составляет $P_{\mathrm{RD}} = 0{,}5$, а для статистик V_2 и V_1 — порядка 0,3. Статистика V_5 в данном случае непригодна для использования. Если же аддитивный шум на антенных элементах становится неоднородным, то реально работают только статистики V_1, V_4 и V_6, причём статистика V_6 при вероятности ложной тревоги $P_{\mathrm{FA}} = 0{,}1$ имеет вероятность правильного обнаружения $P_{\mathrm{RD}} = 0{,}9$, в то время как для статистик V_1 и V_4 имеем $P_{\mathrm{RD}} = 0{,}7$ и $P_{\mathrm{RD}} = 0{,}45$ соответственно.

На рис. 3 показаны рабочие характеристики обнаружения шумового сигнала на фоне однородного по элементам антенны шума. Сигнал при этом моделировался как однородный (рис. 3*а*) и неоднородный (рис. 3*б*) по элементам антенной решётки шум неизвестной мощности. Для сравнения на приведённых графиках добавлена рабочая характеристика классического χ^2-критерия, который является оптимальным для обнаружения однородного гауссовского сигнала на фоне такого же шума.

Из приведённых на рис. 3*а* графиков видно, что для однородного шума рабочие характеристики статистики V_6 и χ^2-критерия практически совпадают. Вероятности правильного обнаружения при $P_{\mathrm{FA}} = 0{,}1$ для статистик V_6, χ^2, V_4 и V_3 равны 0,9; 0,9; 0,7 и 0,3 соответственно. Статистики V_1, V_2 и V_5 в данном случае фактически непригодны для обнаружения шумовых сигналов, поскольку они приспособлены для приёма пространственно-когерентных сигналов.

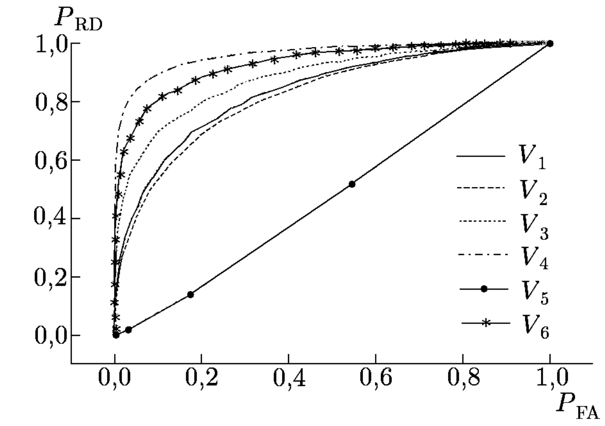

Рис. 1*а*. Рабочие характеристики для статистик V_1–V_6 при обнаружении плоской волны на фоне однородного шума

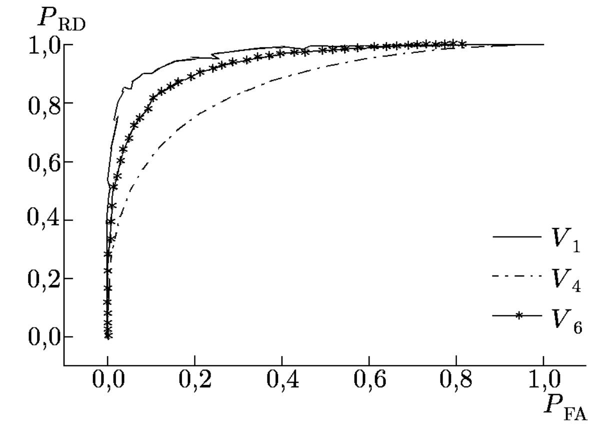

Рис. 1*б*. Рабочие характеристики для статистик V_1, V_4 и V_6 при обнаружении плоской волны на фоне неоднородного шума

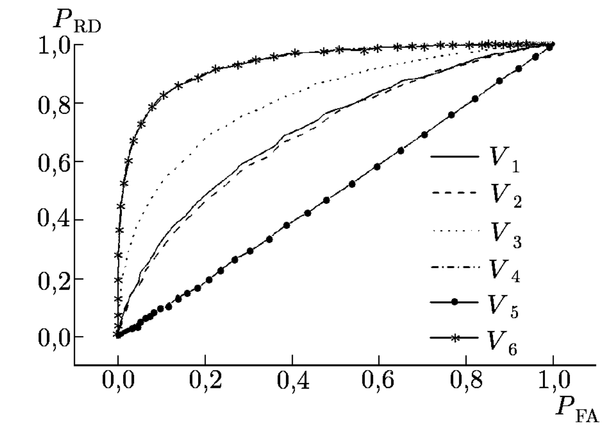

Рис. 2*а*. Рабочие характеристики для статистик V_1–V_6 при обнаружении суммы двух плоских волн на фоне однородного шума

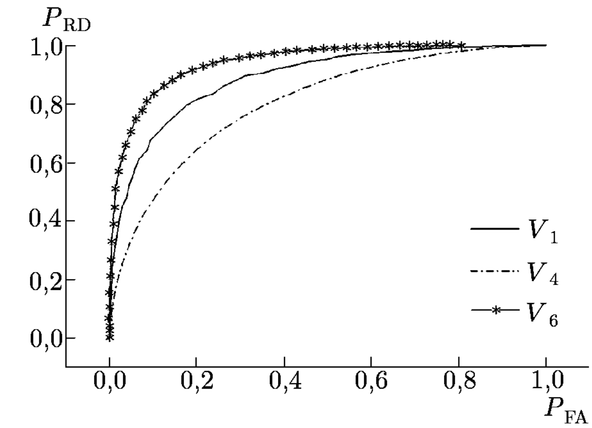

Рис. 2*б*. Рабочие характеристики для статистик V_1, V_4, V_6 при обнаружении суммы двух плоских волн на фоне неоднородного шума

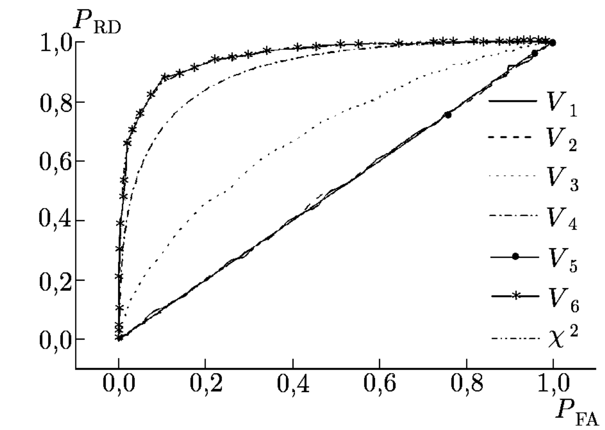

Рис. 3*а*. Рабочие характеристики для статистик V_1–V_6 и χ^2 для однородного шумового сигнала на фоне однородного шума

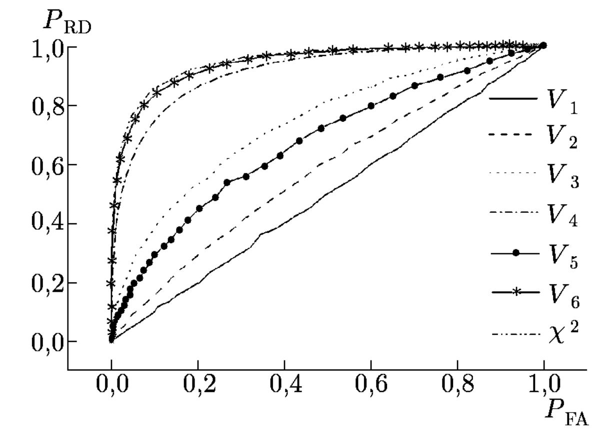

Рис. 3*б*. Рабочие характеристики для статистик V_1–V_6 и χ^2 для неоднородного шумового сигнала на фоне однородного шума

На рис. 3*б* аналогичные кривые построены для обнаружения неоднородного шумового полезного сигнала на фоне однородного шума.

Из приведённых графиков видно, что рабочие характеристики статистик χ^2, V_4 и V_6 практически не меняются по сравнению со случаем, когда сигналом являлся однородный шум (ср. рис. 3*а*). Статистика V_3 увеличивает вероятность правильного обнаружения от 0,3 до 0,4 при $P_{\mathrm{FA}} = 0{,}1$, статистики V_2 и V_5 также несколько улучшают свои характеристики, однако всё равно остаются практически непригодными для обнаружения.

3. ЗАКЛЮЧЕНИЕ

В работе проведено сравнение эффективности использования различных решающих статистик (алгоритмов обработки), полученных на основе обобщённого отношения правдоподобия, для обнаружения пространственных сигналов в случае коротких выборок. На основании анализа приведённых выше рабочих характеристик можно сделать следующие выводы.

Статистика V_6 является наиболее универсальной из всех рассмотренных статистик. Хотя статистика V_6 предназначена для обнаружения сигнала на фоне однородного шума, её рабочие характеристики малочувствительны (робастны) к достаточно сильной неоднородности аддитивного шума (порядка 9 дБ). Кроме того, эта статистика малочувствительна к характеристикам полезного сигнала. Например, в задаче обнаружения полезного сигнала, являющегося когерентной волной, на фоне однородного шума проигрыш эффективности статистики V_6 по сравнению с наилучшей (асимптотически оптимальной) для этого случая статистикой V_4 оказывается незначительным (порядка 12 %).

Для обнаружения пространственно-когерентного сигнала на фоне сильно неоднородного шума можно также использовать статистику V_1, которая имеет в этом случае наилучшие рабочие характеристики, превосходящие характеристики статистик V_6 и V_4 приблизительно на 10 и 45 % процентов соответственно. Интересно отметить, что вычисление значений статистики V_1 существенно проще, чем статистик V_6 и V_4, поскольку для неё не требуется решение задачи на собственные значения выборочной ковариационной матрицы входных сигналов.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 03–02–17141 и Совета при Президенте Российской Федерации по поддержке ведущих научных школ (проект № НШ-1729.2003.2).

СПИСОК ЛИТЕРАТУРЫ

1. Вайнштейн Л. А., Зубаков В. Д. Выделение сигналов на фоне случайных помех. М.: Советское радио, 1960.
2. Тихонов В. И. Статистическая радиотехника. М.: Советское радио, 1966.
3. Шестов Н. С. Выделение оптических сигналов на фоне случайных помех. М.: Советское радио, 1967.
4. Левин Б. Р. Теоретические основы статистической радиотехники. М.: Советское радио, 1974.
5. Теория обнаружения сигналов / Под ред. П. А. Бакута. М.: Радио и связь, 1984.
6. Трифонов А. П., Шинаков Ю. С. Совместное различение сигналов и оценка их параметров на фоне помех. М.: Радио и связь, 1986.
7. Anderson T. W. An introduction to multivariate statistical analysis. New York: John Wiley and Sons, 1960.
8. Болховская О. В., Мальцев А. А. // Изв. вузов. Радиофизика. 2002. Т. 45, № 12. С. 1 077.
9. Болховская О. В., Мальцев А. А., Родюшкин К. В. // Изв. вузов. Радиофизика. (в печати).

10. Родюшкин К. В. Обнаружение, разрешение и оценивание числа источников сигналов антенной решёткой в случае коротких выборок и неизвестных волновых фронтов: Дис. ... канд. физ.-мат. наук. Нижний Новгород, 2001.
11. Bolkhovskaya O. V., Maltsev A. A., Lo Presti L., Sellone F. // Int. Conf. on Electromagnetics in Advanced Applications (ICEAA 01), September 2001, Torino, Italy. P. 655.
12. Справочник по специальным функциям с формулами, графиками и математическими таблицами. М.: Наука, 1979.
13. Luke Y. L. Mathematical functions and their approximations. New York: Academic Press Inc., 1975.
14. Горелик Г. С. Колебания и волны. М.: Гос. изд-во физ.-мат. лит-ры, 1959.

Поступила в редакцию 7 сентября 2004 г.; принята в печать 19 ноября 2004 г.

CHARACTERISTICS OF SPATIAL SIGNALS DETECTION FOR THE GENERALIZED LIKELIHOOD RATIO STATISTICS FOR THE SHORT SAMPLE CASE

O. V. Bolkhovskaya, A. A. Maltsev, and K. V. Rodyushkin

In this paper we study the characteristics of spatial-signal detection by an antenna array in the case of short samples. All the studied deciding statistics are obtained on the basis of the generalized likelihood ratio for a sample of an arbitrary volume. Performance data of detection of useful signal comprising one or two plane coherent waves on the background of spatially homogeneous and inhomogeneous noise are constructed as an example. The efficiency of the used statistics is compared with the χ^2 criterion applied for the detection of noise-like signals.