

С о д е р ж а н и е

Алимов В. А., Рахлин А. В. О некоторых особенностях декаметровая радиоастрономии	95
Егорченков Р. А., Кравцов Ю. А. Модифицированное приближение геометрической оптики: учёт экспоненциально малых рассеянных полей	106
Маненков А. Б. Влияние размерности на решения задач дифракции в открытых волноводах	115
Беляев В. В., Маюнов А. Т., Михайлов Г. Д., Разиньков С. Н. Оценка радиолокационной заметности рупорной антенны с нелинейным контактом при наблюдении "нелинейной" радиолокационной станцией	126
Метелёв С. А., Шишкин Ю. В. Принцип построения двухканального пространственного сепаратора сигнала и помехи с предварительным ортонормированием входных процессов	130
Цвык А. И., Воробьёв Г. С., Нестеренко А. В., Желтов В. Н. Влияние локальной неоднородности магнитостатического поля на стартовые и выходные характеристики генератора дифракционного излучения	144
Авербах В. С., Власов С. Н., Заславский Ю. М. Движение капли жидкости в капилляре под действием статического и акустического полей	155
Белокопытов Г. В., Журавлёв М. В. Электротепловая колебательная неустойчивость в сегнетоэлектрических резонаторах. Двухмодовый режим	162
Брюханов Ю. А., Мясников Е. А. Предельные циклы нелинейной рекурсивной цифровой системы второго порядка с усечением по модулю после сложения	174
Опечатки к статье Т. М. Заборонковой, А. В. Кудрина, Е. Ю. Петрова "О распределении тока вдоль цилиндрической антенны в магнитоактивной плазме в диапазоне очень низких частот" (Известия ВУЗов. Радиофизика. 1999. Т. 42, № 8, С. 750-764)	182

УДК 533.951, 537.868

О НЕКОТОРЫХ ОСОБЕННОСТЯХ ДЕКАМЕТРОВОЙ РАДИОАСТРОНОМИИ

В. А. Алимов, А. В. Рахлин

Рассмотрены некоторые эффекты распространения коротких радиоволн в случайно-неоднородной околоземной плазме применительно к основным задачам декаметрового радиоастрономии. Показано, в частности, что при работе систем КВ радиоастрономии на средних широтах возможны значительные (десятки—сотни процентов) ошибки в определении истинного значения интенсивности КВ радиоизлучения внеземного источника из-за рассеяния радиоволн и фокусировок (дефокусировок) излучения в ионосфере. Установлено, что в условиях развитой ионосферной турбулентности угловое разрешение дискретных радиоисточников методом КВ радиоинтерферометрии или известным методом мерцаний может быть весьма невысоким (единицы угловых градусов). Предложен модифицированный метод мерцаний, основанный на спектральном анализе радиоизлучения дискретного источника с выделением полезного сигнала в области высоких спектральных частот, соответствующего дифракции КВ излучения на слабых неоднородностях межпланетной плазмы, на фоне сильных, но сравнительно низкочастотных флуктуаций принимаемого излучения, обусловленных влиянием развитой турбулентной структуры ионосферы. Показано, что этот метод позволяет существенно повысить угловое разрешение внеземных дискретных радиоисточников (до единиц угловых секунд).

За последние несколько десятилетий в радиоастрономии достигнут определенный прогресс, в основном за счет исследований, выполненных в метровом—сантиметровом диапазонах длин волн [1]. В то же время декаметровый диапазон остался наименее освоенным в радиоастрономических исследованиях. Это обстоятельство обусловлено не только техническими трудностями в построении современных больших радиоастрономических комплексов и сложной помеховой обстановкой в КВ диапазоне, но и определенными принципиальными ограничениями, которые вносит в эти исследования околоземная плазма. Следует заметить, что отдельные аспекты проблемы влияния околоземной плазмы на радиоастрономические исследования рассматривались в целом ряде работ (см., например, [2] и цитированную там литературу). Но применительно к декаметровой радиоастрономии эти вопросы практически не анализировались.

Ниже мы рассмотрим некоторые аспекты распространения коротких радиоволн в случайно-неоднородной ионосферной плазме применительно к задачам декаметрового радиоастрономии, прежде всего для выяснения роли неоднородной структуры ионосферы в радиоастрономических исследованиях и возможных путях решения основных задач декаметрового радиоастрономии.

1. ИЗМЕРЕНИЯ ИНТЕНСИВНОСТИ РАДИОИЗЛУЧЕНИЯ ВНЕЗЕМНЫХ ИСТОЧНИКОВ

Как известно, для определения спектральной плотности потока радиоизлучения внеземных источников или их яркостной температуры проводят измерения интенсивности излучения этих радиоисточников. При измерениях методом наземной декаметровой радиоастрономии радиоизлучение от внеземных источников проходит через неоднородную ионосферу и принимается направленным антенным устройством. Верхняя ионосфера, как правило, характеризуется развитой турбулентной структурой с размерами неоднородностей электронной концентрации от нескольких десятков метров до нескольких десятков километров и широким спектром квазирегулярных крупномасштабных неоднородностей типа перемещающихся ионосферных возмущений с размерами в десятки и сотни километров [2]. Распространение коротких радиоволн в такой многомасштабной неоднородной среде в возмущенных условиях может быть описано с помощью метода рефракционного рассеяния радиоволн [3]. При этом слой

неоднородной ионосферной плазмы заменяется эквивалентным фазовым экраном, свойства которого адекватно описывают неоднородную структуру среды распространения, а дифракция радиоволн на нем определяет основные статистические характеристики принимаемого на Земле радиоизлучения [3]. Причем с использованием теоремы взаимности [4] задачу об интенсивности принимаемого на направленную антенну на Земле радиоизлучения внеземного КВ источника целесообразно свести к задаче об интенсивности КВ излучения наземного передатчика с направленным антенным устройством при приеме его в области расположения внеземного источника.

Итак, рассмотрим следующую задачу.* Пусть на область ионосферы — фазовый экран, расположенный на расстоянии z_u от земной поверхности, падает КВ излучение $E_0(x, t)$ от направленного наземного устройства. Экран вносит в проходящее излучение только фазовые вариации (регулярные $\varphi_u(x, t)$ и случайные $s(x, t)$). После дифракции на экране излучение принимается всенаправленным антенным устройством, расположенным на оси z на расстоянии z_b ($z_b \gg z_u$) от экрана.

Тогда поле КВ сигнала в точке приема можно записать в следующем виде [3, 5]:

$$E(t) = \left(\frac{ik}{2\pi z_u} \right)^{1/2} \int_{-\infty}^{+\infty} E_0(x, t) \exp \left[-\frac{\Gamma(t)}{2} + i(\varphi_u(x, t) + s(x, t)) - \frac{ikx^2}{2z_u} \right] dx, \quad (1)$$

где $\Gamma(t)$ — показатель ослабления интенсивности радиоволн в ионосферной плазме [2, 4], $k = 2\pi/\lambda$ — волновое число, λ — длина волны. Ось x перпендикулярна к оси z .

Пусть случайные фазовые флуктуации, вносимые в проходящее КВ излучение, велики ** ($\overline{s^2} \gg 1$, где черта сверху означает усреднение по ансамблю неоднородностей) и имеют характерный пространственный масштаб $l_E \simeq L_0/\sqrt{\overline{s^2}}$, где L_0 — внешний масштаб ионосферной турбулентности [2], а масштаб квазирегулярных фазовых возмущений волны значительно больше масштаба l_E . Тогда с учетом [3, 5] и (1) для средней интенсивности принимаемого радиоизлучения можно записать следующее соотношение:

$$\begin{aligned} \bar{I}(t) \simeq \frac{k}{2\pi z_u} e^{-\Gamma(t)} \iint_{-\infty}^{+\infty} dx_1 dx_2 E_0(x_1, t) E_0^*(x_2, t) \exp [i\varphi'_u(x_1 + x_2/2, t)(x_1 - x_2)] \times \\ \times \exp \left[-\frac{(x_1 - x_2)^2}{l_E^2} \right] \exp \left[-\frac{ik}{2z_u}(x_1^2 - x_2^2) \right], \end{aligned} \quad (2)$$

где $\varphi'_u(x, t)$ — производная по переменной $\bar{x} = (x_1 + x_2)/2$ от функции $\varphi_u(\bar{x}, t)$.

Без ограничения общности положим для определенности регулярное распределение освещенности экрана $E_0(x, t)$ гауссовским с характерным масштабом a_u и регулярным смещением максимума освещенности $x_0(t)$ из-за вращения Земли или за счет электронного сканирования диаграммы направленности наземной приемно-передающей антенны по координате x :

$$E_0(x, t) = A_0 \exp \left[-\frac{[x - x_0(t)]^2}{a_u^2} \right]. \quad (3)$$

Здесь $a_u \simeq 2z_u\vartheta_a$, $\vartheta_a \simeq 1/(ka)$ — полуширина диаграммы направленности приемно-передающей антенны, a — характерный горизонтальный размер этой антенны ($\vartheta_a \ll 1$), A_0 — амплитуда падающего поля, пропорциональная амплитуде излучения источника.

*Для простоты и большей наглядности результатов ограничимся анализом двумерной задачи.

**Напомним, что в формирование амплитудных вариаций принимаемого КВ излучения, прошедшего слой ионосферной плазмы, могут вносить существенный вклад и крупномасштабные неоднородности с размерами от единиц до десятков километров, для которых $\overline{s^2} \gg 1$ [3].

С учетом (3), проводя необходимые преобразования в соотношении (2), получаем следующее выражение для функции $\bar{I}(t)$:

$$\bar{I}(t) \simeq I_0 e^{-\Gamma(t)} \frac{l_E}{\sqrt{\pi} a} \int_{-\infty}^{+\infty} d\tilde{x} \exp\left[-2(\tilde{x} - x'_0(t))^2\right] \exp\left[-\frac{l_E^2}{a^2}(\tilde{x} - \varphi'_u(x, t)\frac{a}{2})^2\right], \quad (4)$$

где введены обозначения $I_0 = A_0^2$, $\tilde{x} = x/a_u$, $\tilde{x}_0(t) = x_0(t)/a_u$.

Дальнейшие упрощения выражения (4) связаны с заданием конкретного вида функции $\varphi_u(x, t)$ — регулярного набега фазы волны в ионосферном слое. Мы будем полагать, что

$$\varphi_u(x, t) = \varphi(t) + \alpha_u(t)x + \varphi_0(t) \exp[-x^2/b^2(t)].$$

Здесь первое слагаемое описывает набег фазы волны в плоскостной ионосфере, множитель $\alpha_u(t)$ — линейный градиент фазового набега вдоль оси x , обуславливающий квазирегулярную рефракцию радиоволн, третье слагаемое описывает крупномасштабное возмущение $\varphi_0(t)$ фазы волны с масштабом $b(t)$, ответственное за фокусировку радиоволн [2, 5].

Подставляя $\varphi_u(x, t)$ в соотношение (4), с учетом результатов работ [5, 6] можно получить общее выражение для искомой функции $\bar{I}(t)$. В ряде наиболее интересных предельных случаев выражение для средней интенсивности принимаемого КВ радиоизлучения имеет следующий вид (при $\alpha_u(t) \equiv 0$):

$$\bar{I}(t) \simeq \begin{cases} I_0 \frac{l_E(t)}{a} \exp(-\Gamma(t)) & \text{при } l_E/a < 1; \\ I_0 \left[\frac{l_E(t)}{b(t)} \varphi_0(t)\right]^{2/3} \exp(-\Gamma(t) - 2\tilde{x}_0^2(t)) & \text{при } l_E/a > 1; \quad \varphi_0(t) < 0; \quad |z/z_\Phi| \simeq 1; \\ I_0 \exp(-\Gamma(t) - 2\tilde{x}_0^2(t)) & \text{при } l_E/a > 1; \quad \varphi_0(t) < 0; \quad |z/z_\Phi| \ll 1; \\ \frac{I_0}{|1+z/z_\Phi|} \exp(-\Gamma(t) - 2\tilde{x}_0^2(t)) & \text{при } l_E/a > 1; \quad \varphi_0(t) > 0, \end{cases} \quad (5)$$

где $z_\Phi = kb^2(t)/2|\varphi_0(t)|$ — фокусное расстояние линзоподобной крупномасштабной неоднородности.

Из (5) видно, что средняя интенсивность принимаемого радиоизлучения воспроизводит кривую прохождения дискретного источника через диаграмму направленности приемно-передающей антенны ($I(t) \simeq I_0 \exp[-2|x'_0(t)|^2]$, см. (3)) только в случае, когда ширина диаграммы направленности антенны $\vartheta_a = 1/(ka)$ больше ширины углового спектра рассеяния радиоволн в ионосфере $\vartheta_s \simeq 1/(kl_E)$ и когда практически отсутствуют крупномасштабные линзоподобные неоднородности ($|z/z_\Phi| \ll 1$). Во всех других случаях искажения кривой прохождения могут быть весьма значительны. Причем возможно как существенное ослабление интенсивности радиоизлучения дискретного источника I_0 (при $\vartheta_a < \vartheta_s$ и (или) во время присутствия в ионосфере крупномасштабных дефокусирующих ($\varphi_0(t) > 0$) неоднородностей электронной концентрации), так и значительное увеличение I_0 при наличии в ионосфере крупномасштабных фокусирующих неоднородностей ($\varphi_0(t) < 0$ и $|z/z_\Phi| \simeq 1$).

Для характерных значений основных параметров случайных неоднородностей ионосферной плазмы (внешний масштаб ионосферной турбулентности $L_0 \simeq 30$ км, относительные флуктуации электронной концентрации $\Delta N/N \simeq 10^{-2} \div 10^{-1}$, толщина слоя с неоднородностями $L \simeq 150$ км [2, 3]) средний квадрат набега фазы волны [3] $\overline{s^2} \simeq \frac{k^2 f_{\text{кр}}^4}{2 f^4} \frac{\Delta N}{N} L_0 L$ на частоте $f = 10$ МГц (при отношении критической частоты слоя F_2 к рабочей частоте $f_{\text{кр}}/f = 0,5$) составляет $\overline{s^2} \simeq 4 \cdot (10^2 \div 10^4)$.* При этом характерный масштаб неоднородности поля на выходе ионосферного слоя $l_E = L_0/\sqrt{\overline{s^2}} \simeq 1,5 \cdot (10^{-1} \div 1)$ км. Таким образом, для антенных КВ устройств с умеренной направленностью (при

*В дальнейшем конкретные оценки статистических характеристик КВ сигналов и соответствующих параметров дискретных источников делаются для частоты $f = 10$ МГц при отношении $f_{\text{кр}}/f = 0,5$.

$a \simeq 150$ м, $\vartheta_a \simeq 10^{-1}$) могут реализовываться самые различные ситуации в определении величины $\bar{I}(t)$, предусмотренные соотношениями (5), в зависимости от конкретных ионосферных условий, в том числе случаи ослабления интенсивности радиоизлучения из-за рассеяния (при $l_E/a < 1$) и случаи значительного (в несколько раз) увеличения интенсивности принимаемого радиоизлучения из-за фокусировки коротких волн на крупномасштабных ионосферных неоднородностях [7]. Для остронаправленных наземных КВ антенн ($\vartheta_a \lesssim 10^{-2}$) типичной становится потеря усиления приемных антенн ($l_E/a \ll 1$, см. (5)) с резко выраженным эффектом занижения истинного значения интенсивности радиоизлучения внеземного источника.

Итак, в возмущенных условиях среднеширотной ионосферы, когда существует развитая турбулентность ионосферной плазмы и квазирегулярная крупномасштабная структура неоднородностей электронной концентрации в верхней ионосфере, возможны значительные (десятки—сотни процентов) ошибки в определении истинного значения интенсивности КВ радиоизлучения внеземного источника за счет рассеяния радиоволн и фокусировок (дефокусировок) излучения в ионосфере.

При этом, кстати, возможна и сильная деполяризация КВ радиоизлучения при его трансionoсферном распространении, так что степень линейной поляризации излучения на выходе из ионосферного слоя [2, 8]

$$p \simeq \exp\left(-2s^2\left[\frac{f_H^2}{f^2} \cos^2 \alpha + \frac{1}{2}\left(\frac{L}{L_0} \frac{f_{кр}^2 f_H}{f^3} \sin \alpha\right)^2\right]\right), \quad (6)$$

где f_H — гирочастота электронов в земной ионосфере [2, 4], α — угол между волновым вектором \mathbf{k} и направлением магнитного поля Земли \mathbf{H} , может быть практически равна нулю, т. к. $s^2 \gtrsim 4 \cdot 10^2$ (см. выше). Последнее обстоятельство может отрицательно сказаться на работе больших радиоастрономических КВ антенн, использующих поляризационно разнесенный прием.

Следует заметить, однако, что и в обычных (невозмущенных) условиях среднеширотной ионосферы, когда развитая турбулентная структура отсутствует, возможны заметные искажения в кривых прохождения дискретных КВ радиоисточников и ошибки в измерениях интенсивности их КВ излучения.

Вначале рассмотрим случай полного отсутствия неоднородностей электронной концентрации в ионосфере, который соответствует довольно редкому особо спокойному состоянию дневной среднеширотной ионосферы [2]. Для этого случая ($\varphi_0(t) \equiv 0$, $s(t) \equiv 0$) исходя из соотношений (1), (3) несложно получить следующее выражение для интенсивности принимаемого радиоизлучения:

$$I(t) \simeq I_0 e^{-\Gamma(t)} \exp\left(-\frac{1}{2} \left[\frac{t}{t_0} - \frac{\vartheta_{\text{реф}}(t)}{\vartheta_a}\right]^2\right), \quad (7)$$

где $t_0 = \vartheta_a/\alpha_a$, $\alpha_a t$ — линейное угловое смещение максимума диаграммы направленности приемной КВ антенны, обусловленное либо вращением Земли (в этом случае $\alpha_a \simeq 4 \cdot 10^{-3}$ мин $^{-1}$), либо системой электронного слежения антенного устройства за космическим объектом; $\vartheta_{\text{реф}}(t)$ — временная зависимость угла рефракции проходящего ионосферу радиоизлучения (для КВ диапазона и среднеширотной ионосферы характерные значения угла $\vartheta_{\text{реф}} \simeq 1 \div 3^\circ$ [2]).

Для умеренно направленных приемных антенн ($\vartheta_{\text{реф}}/\vartheta_a \ll 1$) функция $I(t)$ воспроизводит теоретическую кривую прохождения радиоисточника (см. (7)):

$$I(t) \simeq I_0 e^{-\Gamma(t)} e^{-t^2/(2t_0^2)}. \quad (8)$$

Но для остронаправленных антенн ($\vartheta_{\text{реф}}/\vartheta_a \gtrsim 1$) вид функции $I(t)$ будет существенно зависеть от временной зависимости функции $\vartheta_{\text{реф}}(t)$ и может значительно отличаться от теоретической кривой прохождения (8).

Рассмотрим теперь случай сравнительно небольших флуктуаций коротких радиоволн в ионосфере. И в этом случае среднее значение интенсивности принимаемого излучения \bar{I} может заметно отличаться

от истинного значения I_0 . Действительно, из соотношения (1) путем несложных преобразований можно получить следующее приближенное выражение для средней интенсивности принимаемого излучения (при $\overline{s^2} \ll 1$, $\varphi_u(t) \equiv 0$):

$$\overline{I}(t) \simeq (1 - \overline{s^2})I_{\text{рег}} + \overline{s^2} \overline{I}_{\text{рас}}. \quad (9)$$

Здесь $I_{\text{рег}}$, $\overline{I}_{\text{рас}}$ — интенсивность соответственно регулярной и диффузной компонент КВ сигнала.

Легко показать, что для остронаправленных антенн ($\vartheta_a = 1/(ka) \ll \vartheta_s = 1/(kl)$, где l — характерный пространственный масштаб слабых неоднородностей электронной концентрации; в наших условиях $l \lesssim 1$ км [3, 9]) величина $\overline{I}_{\text{рас}}$ будет пренебрежимо мала. Тогда искомая функция $\overline{I}(t) = (1 - \overline{s^2}) e^{-\Gamma} I_0$ и ошибка в определении истинного значения интенсивности радиоизлучения внеземного источника, возникающая вследствие влияния слабых неоднородностей ионосферы, будет составлять $\delta I \simeq \overline{s^2}$ (при $\overline{s^2} \simeq 10^{-1}$ ошибка определения $\delta I \simeq 10\%$).

Здесь еще следует заметить, что и при практически полном отсутствии турбулентной структуры в верхней ионосфере возможны значительные вариации интенсивности принимаемого космического КВ радиоизлучения, которые обусловлены присутствием в среднеширотной ионосфере квазирегулярных крупномасштабных неоднородностей типа перемещающихся ионосферных возмущений. Согласно расчетам [5, 6] в этом случае интенсивность принимаемого КВ радиоизлучения может меняться в широких пределах от $I_0 e^{-\Gamma(t)} / (1 + z/z_\phi)$ до $I e^{-\Gamma(t)} \sqrt{2\varphi_0}$ (см. (5)), так что ошибка определения величины I_0 может составить сотни процентов (см. выше).

Итак, по-видимому, единственным критерием истинности измерений интенсивности радиоизлучения от внеземных КВ радиоисточников является практически полное отсутствие искажений в кривой прохождения радиоисточника и флуктуаций в принимаемом излучении [10]. Лишь жесткий отбор записей КВ сигналов, свободных от указанных недостатков, в сочетании с текущим риометрическим контролем за поглощением радиоволн в ионосфере может гарантировать необходимую надежность измерений потоков радиоизлучения дискретных источников методом КВ радиоастрономии. В противном случае ошибки в определении истинного значения величины I_0 могут достигать сотен процентов.

Это заключение остается в силе и при определении спектров излучения дискретных источников в диапазоне коротких радиоволн. Поскольку даже при одновременных измерениях радиояркости одного и того же источника регистрируемые величины \overline{I}_1 и \overline{I}_2 для двух частот в КВ диапазоне могут существенно отличаться из-за дисперсии в эффектах рассеяния и фокусировок проходящего КВ радиоизлучения в ионосфере (см. (5)).*

2. ИЗМЕРЕНИЯ КООРДИНАТ ДИСКРЕТНЫХ ИСТОЧНИКОВ РАДИОИЗЛУЧЕНИЯ И ИХ УГЛОВЫХ РАЗМЕРОВ

Наиболее распространенным методом измерения координат внеземных источников КВ излучения и их угловых размеров является интерферометрический метод [1, 2].

Для средней интенсивности радиоизлучения, принимаемого от дискретного радиоисточника с угловым распределением радиояркости $\mathcal{I}_0(\vartheta)$, на выходе двухэлементного интерферометра с базой d можно записать следующее соотношение (см. также [1, 2]):

$$\overline{I}_{\text{инт}}(\vartheta) \simeq \int \mathcal{I}_0(\vartheta - \vartheta') [1 + M(d) \cos(kd\vartheta')] d\vartheta', \quad (10)$$

где $M(d)$ — коэффициент модуляции средней интерференционной картины на Земле. В нашем случае при дифракции КВ радиоизлучения в случайно-неоднородной ионосфере функция $M(d)$ равна [2]

$$M(d) \simeq \exp[-D_s(d)/2], \quad (11)$$

*Вопросы методики абсолютных измерений спектральных плотностей КВ радиоизлучения космических дискретных источников заслуживают специального рассмотрения.

где $D_s(d)$ — структурная функция флуктуаций фазы волны в ионосферном слое. В случае сильных фазовых возмущений (см. выше) $D_s(d) \simeq d^2/l_E^2$.

Аппроксимируя распределение \mathcal{I}_0 гауссовой кривой $\mathcal{I}_0 = \mathcal{I}_0 \exp[-\vartheta^2/(2\vartheta_0^2)]$, где ϑ_0 — угловой размер радиосточника, из соотношения (10) находим

$$\bar{I}(\vartheta) \simeq P_0 [1 + M_{\text{эфф}}(d) \cos(kd\vartheta)], \quad (12)$$

где $M_{\text{эфф}}(d) \simeq M(d) \exp(-\vartheta_0^2/\vartheta_{\text{инт}}^2)$, $\vartheta_{\text{инт}} = \sqrt{2}/(kd)$ — угловой размер главного лепестка диаграммы направленности интерферометра, P_0 — интенсивность излучения, принимаемого одной антенной интерферометра.

В случае $\overline{s^2} \gg 1$

$$M_{\text{эфф}}(d) \simeq \exp[-(\vartheta_s^2 + \vartheta_0^2)/\vartheta_{\text{инт}}^2]. \quad (13)$$

Из соотношений (12), (13) следует, что угловое разрешение дискретных радиосточников интерферометрическим методом ограничено величиной углового спектра рассеяния радиоволн в ионосфере $\vartheta_s = 1/(kl_E)$. В условиях развитой турбулентности верхней ионосферы, когда $l_E \simeq 1, 5 \cdot (10^{-1} \div 1)$ км (см. выше), величина $\vartheta_s = 3 \cdot 10^{-2}$. Такое низкое угловое разрешение интерферометрического метода в КВ диапазоне, естественно, ставит вопрос о возможностях других методов радиоастрономических измерений угловых характеристик внеземных КВ радиосточников.

Как известно [2], перспективным методом измерений угловых характеристик внеземных радиосточников является метод мерцаний радиоволн на неоднородностях межпланетной плазмы. Этот метод основан на измерении индекса амплитудных флуктуаций принимаемого излучения. Однако в КВ диапазоне в условиях сильных ионосферных флуктуаций интенсивности проходящего излучения этот метод, как легко показать, имеет сравнительно низкое угловое разрешение порядка величины углового спектра рассеянных коротких радиоволн в ионосфере [11].

Тем не менее, можно предложить некий модифицированный метод мерцаний (МММ), который будет свободен от указанного недостатка. К изложению этого метода радиоастрономических измерений мы сейчас и перейдем.

Рассмотрим следующую задачу. Пусть излучение от бесконечно удаленного дискретного источника, имеющего угловое распределение интенсивности излучения $\mathcal{I}_0(\vartheta)$, проходит два слоя крупномасштабных неоднородностей электронной концентрации и принимается на некотором расстоянии z_u от второго слоя. Первый слой характеризуется средним квадратом фазовых флуктуаций проходящего излучения $\overline{s_{\text{св}}^2}$ ($\overline{s_{\text{св}}^2} \ll 1$), френелевским масштабом в спектре неоднородностей электронной концентрации $l_{\text{св}}$ (см. ниже) и поперечной (к лучу зрения на источник) скоростью их дрейфа $\mathbf{V}_{\text{св}}$. Второй слой вносит сильные фазовые искажения в проходящее излучение ($\overline{s_{\text{н}}^2} \gg 1$), имеет развитую турбулентную структуру с характерным внешним масштабом турбулентности L_0 (соответственно эффективный масштаб флуктуаций комплексного поля $l_E \simeq L_0/\sqrt{\overline{s_{\text{н}}^2}}$) и поперечной скоростью дрейфа неоднородностей $\mathbf{V}_{\text{н}}$. Протяженность первого слоя $L_{\text{св}}$, второго — $L_{\text{н}}$.

С учетом указанных особенностей распространения радиоволн в такой двухслойной среде мы будем рассматривать режим насыщенных мерцаний в плоскости наблюдения, когда пространственно-временную корреляционную функцию флуктуаций интенсивности можно записать в следующем виде (ср. [11]):

$$\Gamma_I(\boldsymbol{\rho}, \tau) = \iint |\Gamma_{EE^*}(\mathbf{r}_{01}, \mathbf{r}_{02}, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)|^2 \mathcal{I}_I(\boldsymbol{\vartheta}_1) \mathcal{I}_I(\boldsymbol{\vartheta}_2) d\boldsymbol{\vartheta}_1 d\boldsymbol{\vartheta}_2, \quad (14)$$

где $\Gamma_{EE^*}(\tau = t_1 - t_2) = \overline{E(\mathbf{r}_{01}, t_1, \boldsymbol{\vartheta}_1) E^*(\mathbf{r}_{02}, t_2, \boldsymbol{\vartheta}_2)}$ — пространственно-временная корреляционная функция флуктуаций комплексных полей плоских волн, приходящих с направлений $\boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2$; $\boldsymbol{\rho} = \mathbf{r}_{01} - \mathbf{r}_{02}$.

Расчет функции $\Gamma_{EE^*}(\boldsymbol{\rho}, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$ будем проводить в рамках метода рефракционного рассеяния радиоволн [11]:

$$\Gamma_{EE^*}(\mathbf{r}_{01}, \mathbf{r}_{02}, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2) \simeq \frac{k^2}{(2\pi)^2 z_u^2} \iint_{-\infty}^{+\infty} d\mathbf{r}_1 d\mathbf{r}_2 \Gamma_{EE^*}(\mathbf{r}_1, \mathbf{r}_2, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2) \times \overline{\exp\{i[s_u(\mathbf{r}_1, t_1, \boldsymbol{\vartheta}_1) - s_u(\mathbf{r}_2, t_2, \boldsymbol{\vartheta}_2)]\}} \exp\left[-\frac{ik}{2z_u}(|\mathbf{r}_1 - \mathbf{r}_{01}|^2 - |\mathbf{r}_2 - \mathbf{r}_{02}|^2)\right]. \quad (15)$$

Здесь $\Gamma_{EE^*}(\mathbf{r}_1, \mathbf{r}_2, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2)$ — корреляционная функция флуктуаций комплексного поля на входе второго статистически однородного слоя плазмы. С учетом [11]

$$\Gamma_{EE^*}(\mathbf{r}_1, \mathbf{r}_2, \tau, \boldsymbol{\vartheta}_1, \boldsymbol{\vartheta}_2) \simeq \frac{k^2}{(2\pi)^2 z_{cb}^2} \iint_{-\infty}^{+\infty} d\mathbf{r}'_1 d\mathbf{r}'_2 \exp\{ik[\mathbf{r}'_1 \boldsymbol{\vartheta}_1 - \mathbf{r}'_2 \boldsymbol{\vartheta}_2]\} \times \overline{\exp\{i[s_{cb}(\mathbf{r}'_1, t_1, \boldsymbol{\vartheta}_1) - s_{cb}(\mathbf{r}'_2, t_2, \boldsymbol{\vartheta}_2)]\}} \exp\left[-\frac{ik}{2z_{cb}}(|\mathbf{r}'_1 - \mathbf{r}_1|^2 - |\mathbf{r}'_2 - \mathbf{r}_2|^2)\right], \quad (16)$$

где $z_{cb} = L_{cb}/2$ — расстояние от первого эквивалентного фазового экрана, заменяющего первый статистически однородный слой плазмы [11], до второго неоднородного слоя.

Для входящих в соотношение (15), (16) пространственно-временных корреляционных функций комплексных полей в отдельных неоднородных слоях с учетом [2, 11] можно получить следующие выражения:

$$\overline{\exp\{i[s_u(\mathbf{r}_1, t_1, \boldsymbol{\vartheta}_1) - s_u(\mathbf{r}_2, t_2, \boldsymbol{\vartheta}_2)]\}} \simeq \exp\left[-\frac{|\mathbf{r}_1 - \mathbf{r}_2 + \mathbf{V}_u \tau|^2 + \vartheta^2 L_u^2/3}{2l_E^2}\right], \quad (17)$$

$$\overline{\exp\{i[s_{cb}(\mathbf{r}'_1, t_1, \boldsymbol{\vartheta}_1) - s_{cb}(\mathbf{r}'_2, t_2, \boldsymbol{\vartheta}_2)]\}} \simeq (1 - 2\overline{s_{cb}^2}) + 2\overline{s_{cb}^2} \rho_s(\mathbf{r}'_1 - \mathbf{r}'_2 + \mathbf{V}_{cb} \tau, \boldsymbol{\vartheta}) = (1 - 2\overline{s_{cb}^2}) + 2\overline{s_{cb}^2} \exp\left[-\frac{|\mathbf{r}'_1 - \mathbf{r}'_2 + \mathbf{V}_u \tau|^2}{l_{cb}^2}\right]. \quad (18)$$

В соотношении (18) для простоты и большей наглядности дальнейших расчетов коэффициент корреляции фазовых флуктуаций излучения в первом слое $\rho_s(\boldsymbol{\rho}', \tau, \boldsymbol{\vartheta})$ был принят в гауссовом виде ($\boldsymbol{\vartheta} = \boldsymbol{\vartheta}_1 - \boldsymbol{\vartheta}_2$, $\boldsymbol{\rho}' = \mathbf{r}'_1 - \mathbf{r}'_2$). С этой же целью распределение радиояркости источника излучения также будем аппроксимировать гауссовой кривой с нормировкой $\int_{-\infty}^{\infty} \mathcal{I}_0(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta} \equiv 1$:

$$\mathcal{I}_0(\boldsymbol{\vartheta}) = \frac{1}{2\pi\vartheta_0^2} \exp(-\vartheta^2/2\vartheta_0^2). \quad (19)$$

Используя соотношения (14)–(19) и проводя необходимые преобразования, в результате получаем выражение для временной корреляционной функции флуктуаций интенсивности принимаемого излучения ($\boldsymbol{\rho} \equiv 0$):

$$\Gamma_I(\tau) \simeq (1 - 4\overline{s_{cb}^2}) \frac{1}{1 + 4\vartheta_0^2/\vartheta_u^2} \exp\left[-\frac{V_u^2 \tau^2}{l_E^2} \frac{1}{1 + 4\vartheta_0^2/\vartheta_u^2}\right] + 4\overline{s_{cb}^2} \frac{1}{1 + 2\vartheta_0^2/\vartheta_{cb}^2} \exp\left[-\left(\frac{V_u^2 \tau^2}{l_E^2} + \frac{V_{cb}^2 \tau^2}{2l_{cb}^2} \frac{1}{1 + 2\vartheta_0^2/\vartheta_{cb}^2}\right)\right]. \quad (20)$$

При выводе этого соотношения были использованы следующие приближения:

$$\overline{s_{cb}^2} \ll 1; \quad \overline{s_u^2} \gg 1; \quad z_{cb} \gg z_u; \quad |\mathbf{V}_{cb}| \gg |\mathbf{V}_u|; \quad \vartheta_{cb} = \frac{l_{cb}}{z_{cb}} \ll \vartheta_u = \frac{l_E}{z_u}. \quad (21)$$

Из формулы (20) следует, что индекс мерцаний радиоволн

$$\Gamma_I(0) \simeq [1 + 4\vartheta_0^2/\vartheta_u^2]^{-1}.$$

Это выражение совпадает с соответствующим выражением для индекса мерцаний радиоволн в однослойной модели неоднородной среды с сильными фазовыми флуктуациями проходящего излучения [2]. Из него, в частности, следует, что угловое разрешение для дискретных радиоисточников в общепринятом методе мерцаний в условиях сильных флуктуаций радиоволн ограничивается угловым размером неоднородностей поля волны ϑ_u , которое в силу принятых выше приближений значительно больше угловых размеров неоднородностей первого слоя среды распространения ($\vartheta_u \gg \vartheta_{св}$).

Используя соотношение (20), легко получить выражение для частотного спектра флуктуаций интенсивности принимаемого радиоизлучения:

$$I(\Omega, \vartheta_0) = \int_{-\infty}^{+\infty} \Gamma_I(\tau) \exp(i\Omega\tau) d\tau = \frac{(1 - 4\overline{s_{св}^2}) \sqrt{\pi}}{\Omega_u \sqrt{1 + 4\vartheta_0^2/\vartheta_u^2}} \exp \left[-\frac{\Omega^2}{4\Omega_u^2} \left(1 + 4\frac{\vartheta_0^2}{\vartheta_u^2} \right) \right] + \\ + \frac{4\overline{s_{св}^2} \sqrt{\pi}}{\sqrt{\Omega_u^2 + \frac{\Omega_{св}^2}{2(1+2\vartheta_0^2/\vartheta_{св}^2)}}} \frac{1}{(1 + 2\vartheta_0^2/\vartheta_{св}^2)} \exp \left[-\frac{\Omega^2}{4 \left(\Omega_u^2 + \frac{\Omega_{св}^2}{2(1+2\vartheta_0^2/\vartheta_{св}^2)} \right)} \right], \quad (22)$$

где введены обозначения $\Omega_u = |\mathbf{V}_u/l_E|$ и $\Omega_{св} = |\mathbf{V}_{св}/l_{св}|$.

Мы будем интересоваться нормированными спектрами флуктуаций $I(\Omega) = I(\Omega, \vartheta_0)/I(\Omega, 0)$ на отдельных выделенных частотах Ω_u и $\Omega_{св}$ в зависимости от отношений ϑ_0/ϑ_u и $\vartheta_0/\vartheta_{св}$. При этом будем полагать, что выполнено неравенство $\Omega_{св} \gg \Omega_u$. Легко видеть, что на частоте Ω_u величина $I(\Omega_u)$ уменьшается примерно в 2 раза при $\vartheta_0 = \vartheta_u$ и становится пренебрежимо малой при $\vartheta_0 > \vartheta_u$. В то же время $I(\Omega_{св})$ уменьшается до значений порядка e^{-1} при $\vartheta_0 \simeq \vartheta_{св}$ и становится пренебрежимо малой при $\vartheta_0 > \vartheta_{св}$.

Таким образом, анализируя спектр мерцаний радиоволн от дискретного источника, прошедших два слоя с резко отличными статистическими свойствами неоднородностей (см.(21)), на различных спектральных частотах, можно существенно повысить угловое разрешение классического метода мерцаний [2] и фактически обеспечить измерение углового распределения радиояркости источника $\mathcal{I}(\vartheta)$ в широком диапазоне углов $\vartheta_{св} \lesssim \vartheta \lesssim \vartheta_u$.

Такой модифицированный метод мерцаний (МММ), по-видимому, при определенных условиях может быть успешно реализован в декаметровых радиоастрономических исследованиях внеземных дискретных источников радиоизлучения. Действительно, в декаметровом диапазоне радиоволн излучение от удаленных космических источников, как правило, проходит два неоднородных слоя с резко отличными статистическими свойствами: слой межпланетной плазмы ($L_{св} \simeq 10^8$ км) со сравнительно небольшими фазовыми флуктуациями коротких волн ($\overline{s_{св}^2} \lesssim 10^{-1}$) на быстро дрейфующих неоднородностях электронной концентрации ($V_{св} = 300 \div 400$ км/с) с размерами порядка соответствующей зоны Френеля ($l_{св} \simeq \sqrt{z_{св}/k} \simeq 10^3$ км) и возмущенный ионосферный слой ($\overline{s_u^2} \gtrsim 10^2$, $l_E \simeq 1,5 \cdot 10^{-1}$ км, $L_u = 150$ км, высота неоднородного слоя $z_H = 300$ км, см. выше) с относительно медленным дрейфом неоднородностей ($V_H = 100$ м/с)[2]. При этом характерные значения собственных частот мерцаний коротких радиоволн при дифракции излучения от внеземных дискретных источников в отдельных слоях могут отличаться лишь незначительно ($\Omega_u \simeq \Omega_{св} \simeq 5 \cdot 10^{-1}$ Гц). Заметим, что эти оценки согласуются с общими представлениями о роли ионосферы в эффектах межпланетных мерцаний коротких радиоволн, отмеченных в [12], но не согласуются с оценками собственных частот мерцаний в ионосфере и межпланетной плазме, изложенными в [13, 14], в которых неоправданно занижена частота ионосферных мерцаний коротких радиоволн (см., например, [2]).

Итак, не все необходимые условия (см. (21) и ниже) для успешной реализации МММ в декаметровый радиоастрономии внеземных дискретных источников, вообще говоря, выполнены, поскольку $\Omega_u \simeq \Omega_{св}$. Но эта ситуация может быть исправлена, если при приеме флуктуирующего излучения от удаленных космических источников применяются остронаправленные КВ антенны (ср. [14]). При этом при соответствующих вычислениях спектральных характеристик флуктуаций принимаемого излучения на выходе направленной антенны следует использовать соотношения (14)–(19) и методику вычислений [10]. В результате для нормированной автокорреляционной функции флуктуаций интенсивности принимаемого излучения на выходе остронаправленной приемной антенны получается следующее соотношение (ср. (20)):

$$\Gamma_I(\tau) \simeq \frac{1 - 4\overline{s_{св}^2}}{1 + 4\vartheta_0^2/\vartheta_u^2} \exp\left[-\frac{V_u^2 \tau^2}{d^2}\right] + \frac{4\overline{s_{св}^2}}{1 + 2\vartheta_0^2/\vartheta_{св}^2} \exp\left[-\frac{V_{св}^2 \tau^2}{2l_{св}^2} \frac{1}{1 + 2\vartheta_0^2/\vartheta_{св}^2}\right], \quad (23)$$

где d — характерный размер приемной антенны ($l_E \ll d \ll l_{св}$).

Этот результат следовало ожидать, поскольку такая остронаправленная антенна усредняет мелко-масштабные пространственные флуктуации излучения, обусловленные дифракцией его в возмущенной ионосфере, но практически не искажает (а точнее подчеркивает) крупномасштабные неоднородности в принимаемом излучении, обусловленные распространением его в межпланетной плазме. При этом в результирующем спектре мерцаний радиоволн должны наиболее рельефно выделяться быстрые флуктуации принимаемого КВ излучения, связанные с его дифракцией на неоднородностях солнечного ветра, на фоне сглаженных ионосферных флуктуаций, сосредоточенных вблизи нулевой частоты.

В заключение следует заметить, что выше мы рассматривали возможности угловых измерений внеземных дискретных радиоисточников методами КВ радиоастрономии в условиях развитой ионосферной турбулентности. Но, к счастью, на средних широтах условия сильной геофизической возмущенности в ночной ионосфере реализуются хотя и сравнительно часто, но далеко не всегда [2]. При этом в обычных (невозмущенных) для среднеширотной ионосферы условиях ионосферный слой достаточно тонкий ($L \simeq 30$ км [2]), а возмущения ΔN электронной концентрации N мелкомасштабных неоднородностей ($L \lesssim 1$ км) слабы ($\Delta N/N \lesssim 10^{-3}$ [9]). При этом фазовые возмущения, вносимые в проходящее КВ излучение, малы ($\overline{s^2} \ll 1$), и сравнительно интенсивные крупномасштабные неоднородности с размерами в десятки километров не вносят сколько-нибудь значительный вклад в нарушение когерентности принимаемых на разнесенные антенны КВ радиоизлучения от внеземных дискретных радиоисточников. В этих условиях коэффициент модуляции средней интерференционной картины на Земле $M(d) \simeq 1$ (см. (11)), так что указанное выше ограничение на работоспособность систем КВ радиоинтерферометрии практически снимается.

ЗАКЛЮЧЕНИЕ

Выше были рассмотрены некоторые эффекты распространения коротких радиоволн в случайно-неоднородной околоземной плазме применительно к основным задачам декаметровый радиоастрономии — измерению интенсивности КВ радиоизлучения внеземных источников, их координат и угловых размеров. Показано, в частности, что при работе систем КВ радиоастрономии даже в условиях среднеширотной ионосферы возможны значительные (в десятки—сотни процентов) ошибки в определении истинного значения интенсивности КВ радиоизлучения внеземного источника из-за рассеяния и фокусировок (дефокусировок) излучения в ионосфере. Единственным критерием истинности измерений интенсивности радиоизлучения внеземных КВ радиоисточников является, по-видимому, практически полное отсутствие искажений в кривых прохождения радиоисточников и флуктуаций в принимаемом излучении. Лишь жесткий отбор записей КВ сигналов, свободных от указанных недостатков, может

гарантировать необходимую надежность измерений потоков радиоизлучения дискретных источников методом КВ радиоастрономии.

В условиях развитой ионосферной турбулентности угловое разрешение $\Delta\vartheta$ дискретных радиоисточников методом КВ радиоинтерферометрии или известным методом мерцаний может быть весьма невысоким ($\Delta\vartheta \gtrsim 1 \div 2^\circ$). В то же время можно предложить для этих целей модифицированный метод мерцаний (МММ), который позволит повысить угловое разрешение внеземных дискретных КВ радиоисточников на несколько порядков. В КВ радиоастрономии МММ основан на спектральном анализе мерцаний радиоизлучения дискретного источника и выделении полезного сигнала на высоких спектральных частотах, соответствующего дифракции излучения на слабых неоднородностях межпланетной плазмы (солнечного ветра), на фоне сильных, но низкочастотных флуктуаций принимаемого наземной остроуправленной КВ антенной излучения, обусловленных влиянием развитой турбулентной структуры ионосферы.

Предложенный модифицированный метод мерцаний может оказаться весьма эффективным методом КВ радиоастрономических исследований. Поэтому для оценки реальных возможностей этого метода в ближайшее время целесообразно провести соответствующие экспериментальные исследования. Однако при этом следует иметь в виду, что необходимым условием успешной реализации предложенного метода являются насыщенные ионосферные мерцания коротких радиоволн в сочетании со слабыми межпланетными мерцаниями излучения дискретных радиоисточников при приеме его на остроуправленную КВ антенну. Это обстоятельство должно быть особо учтено при методической проработке соответствующего эксперимента.

Работа выполнена в рамках проекта ИНТАС № 97–1964.

ЛИТЕРАТУРА

1. Кисляков А. Г., Разин В. А., Цейтлин Н. М. Введение в радиоастрономию, Ч. I. — М.: Наука, 1995.
2. Гершман Б. Н., Ерухимов Л. М., Яшин Ю. Я. Волновые явления в ионосфере и космической плазме. — М.: Наука, 1984.
3. Алимов В. А., Рахлин А. В., Выборнов Ф. И. // Изв. вуз. Радиофизика. 1997. Т. 40, № 11. С. 1323.
4. Гинзбург В. Л. Распространение электромагнитных волн в плазме. — М.: Наука, 1967.
5. Алимов В. А. // Изв. вуз. Радиофизика. 1989. Т. 32, № 7. С. 795.
6. Алимов В. А. // Изв. вуз. Радиофизика. 1995. Т. 38, № 6. С. 551.
7. Токарев Ю. В. и др. // Изв. вуз. Радиофизика. 1999. Т. 42, № 8. С. 799.
8. Алимов В. А., Рахлин А. В., Тамойкин В. В. // Изв. вуз. Радиофизика. 1996. Т. 39, № 8. С. 1007.
9. Алимов В. А., Ерухимов Л. М. // Изв. вуз. Радиофизика. 1995. Т. 38, № 12. С. 1227.
10. Токарев Ю. В. // Изв. вуз. Радиофизика. 1996. Т. 39, № 8. С. 945.
11. Алимов В. А., Рахлин А. В. // Изв. вуз. Радиофизика. 1997. Т. 40, № 5. С. 586.
12. Власов В. И., Чашей И. В., Шишов В. И., Шишова Т. Д. // Геомаг. и аэроном. 1979. Т. 19, № 3. С. 401.
13. Жук И. Н. // Изв. вуз. Радиофизика. 1980. Т. 23, № 8. С. 893.
14. Vovkoop V. P., Zhouck I. N. // Astrophys. Space Sci. 1981. V. 79. P. 165.
15. Алимов В. А., Ерухимов Л. М. // Изв. вуз. Радиофизика. 1967. Т. 10, № 10. С. 620.

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

Поступила в редакцию
18 мая 1999 г.

ON SOME FEATURES OF DECAMETER RADIO ASTRONOMY*V. A. Alimov and A. V. Rakhlin*

We consider some effects of the VHF radio wave propagation in randomly irregular plasma near the Earth. Applications of these effects to main problems of decameter radio astronomy are discussed. In particular, we show that significant measurement errors of about tens to hundreds percent for the intensity of VHF radio emission from an extra-terrestrial source may occur due to scattering and focusing/defocusing of the radiation in the ionosphere if VHF radio astronomical facility operates at middle latitudes. We find that the angular resolution of discrete radio sources observed using radio interferometry and the well-known scintillation methods cannot be better than about a degree due to the effect of the developed ionospheric turbulence. We propose a modified scintillation method. It is based on the spectral analysis of radio emission from discrete sources, which allows the useful high-frequency signal corresponding to diffraction of VHF emission at weak inhomogeneities of interplanetary plasma to be separated against strong, relatively low-frequency fluctuations of the received radiation due to the effect of the developed turbulent structure of the ionosphere. We show that such a method allows the angular resolution of extra-terrestrial radio sources to be improved up to arcsecond level.

УДК 514:535:551.501.793

МОДИФИЦИРОВАННОЕ ПРИБЛИЖЕНИЕ ГЕОМЕТРИЧЕСКОЙ ОПТИКИ: УЧЁТ ЭКСПОНЕНЦИАЛЬНО МАЛЫХ РАССЕЯННЫХ ПОЛЕЙ

Р. А. Егорченков¹, Ю. А. Кравцов²

Предложено модифицированное приближение геометрической оптики (МПГО), основанное на представлении решения уравнения Гельмгольца в виде ряда с нулевым геометрооптическим приближением в качестве главного члена. Предлагаемый метод уточняет геометрооптическое приближение, предоставляя возможность учёта экспоненциально малого рассеяния назад в среде с плавными неоднородностями. При рассеянии вперёд решение переходит в обычный геометрооптический ряд по обратным степеням волнового числа. Приведены оценки рассеяния назад с привлечением метода возмущений для лучей. Установлены условия, при которых рассеянное назад поле согласуется с борновским приближением теории рассеяния.

ВВЕДЕНИЕ

Как известно [1, 2], приближение геометрической оптики дает решение уравнения Гельмгольца

$$\Delta u + k^2 \varepsilon u = 0, \quad (1)$$

где $\varepsilon = \varepsilon(\mathbf{r})$ — диэлектрическая проницаемость среды, k — волновое число в вакууме, u — волновое поле, в виде асимптотического ряда по обратным степеням волнового числа

$$u = \left(A_0 + \frac{A_1}{ik} + \frac{A_2}{(ik)^2} + \dots \right) e^{ik\psi} \quad (2)$$

(дебаевская процедура [3]) или, что фактически то же самое, в виде ряда по малому безразмерному параметру $\mu = 1/(kl)$ (рытовская процедура [4]), где l — характерный масштаб неоднородности.

Первый член ряда (2) (нулевое приближение геометрической оптики $u = A_0 e^{ik\psi}$) описывает эволюцию поля в “бездифракционном” приближении, которое отвечает сохранению энергии в лучевой трубке, тогда как последующие члены ряда учитывают слабые (малоугловые) дифракционные эффекты. Отклонения луча на большие углы и, тем более, рассеяние назад приближением геометрической оптики (2) не охватываются. В плавно неоднородной среде ($kl \gg 1$) такие эффекты экспоненциально малы, порядка $\exp(-kl)$, и в принципе не могут быть описаны геометрооптическим рядом по обратным степеням k , поскольку экспоненциальные множители вида $\exp(-kl)$ при $k \rightarrow \infty$ убывают быстрее любой степени $1/k$.

В данной работе предлагается модифицированное приближение геометрической оптики (МПГО), которое позволяет оценить экспоненциально малые поправки к нулевому приближению геометрической оптики. Метод обобщает процедуру Нортовера [5], который рассматривал одномерную волновую задачу, описываемую уравнением

$$\frac{\partial^2 u}{\partial z^2} + k^2 \varepsilon u = 0. \quad (3)$$

Нортовер заметил, что решение уравнения (3) в нулевом приближении ВКБ

$$u_0 = C \varepsilon^{-1/4} \exp(ik \int \varepsilon^{1/2} dz), \quad (4)$$

где $C = \text{const}$, отвечающем нулевому приближению геометрической оптики, строго удовлетворяет уравнению

$$\frac{\partial^2 u}{\partial z^2} + (k^2 \varepsilon + m) u = 0,$$

в котором $m = -\varepsilon^{-1/4} \partial^2(\varepsilon^{-1/4}) / \partial z^2$ представляет собой малую по сравнению с $k^2 \varepsilon$ величину. Если добавить малое слагаемое mu одновременно в левую и правую части уравнения (3) и искать решение в виде $u = u_0 + w$, то для поправки w получится уравнение

$$\frac{\partial^2 w}{\partial z^2} + (k^2 \varepsilon + m) w = m(w + u_0), \quad (5)$$

решение которого может быть получено методом возмущений по малому параметру m . Если неоднородность $\varepsilon(z)$ сосредоточена в ограниченной области $-l < z < l$, вне которой $\varepsilon = 1$ и $m = 0$, а первичная волна $u_\infty = B \exp(ikz)$ падает на неоднородность со стороны отрицательных z , то слева от неоднородности, т. е. при $z \rightarrow -\infty$, поправка первого порядка дается выражением

$$w_1 = ikB \exp(-ikz) \int_{-\infty}^{+\infty} dz' m(z') \exp[2ik \int_{-\infty}^{z'} \sqrt{\varepsilon(z'')} dz'']. \quad (6)$$

Это выражение описывает экспоненциально малое (при гладких крупномасштабных неоднородностях среды) обратное рассеяние в одномерных задачах. Идея обобщения процедуры Нортонера на случай трехмерно-неоднородной среды была предложена в работе [6].

В данной работе развивается детальная версия модифицированного приближения геометрической оптики. Рассмотрение проводится в два этапа. Сначала строится теория возмущений для функции Грина (разд. 1), а затем и для самого волнового поля (разд. 2). Анализ полученного разложения проводится в последующих разделах. В разд. 3 устанавливается связь полученного разложения с дебаевским разложением (2), в разд. 4 — с борновским приближением.

1. РАЗЛОЖЕНИЕ ФУНКЦИИ ГРИНА В РЯД С ИСПОЛЬЗОВАНИЕМ НУЛЕВОГО ПРИБЛИЖЕНИЯ ГЕОМЕТРИЧЕСКОЙ ОПТИКИ В КАЧЕСТВЕ ГЛАВНОГО ЧЛЕНА

Поле точечного монохроматического источника $g(\mathbf{r}, \mathbf{r}_s)$ (функция Грина волновой задачи) удовлетворяет уравнению

$$\Delta g(\mathbf{r}, \mathbf{r}_s) + k^2 \varepsilon(\mathbf{r}) g(\mathbf{r}, \mathbf{r}_s) = \delta(\mathbf{r} - \mathbf{r}_s), \quad (7)$$

где \mathbf{r}_s — радиус-вектор источника. Геометрооптическое приближение для функции Грина записываем в форме, аналогичной (2):

$$g_{\text{GO}}(\mathbf{r}, \mathbf{r}_s) = \left(G_0 + \frac{1}{ik} G_1 + \dots \right) e^{ik\Psi}, \quad (8)$$

где Ψ удовлетворяет уравнению эйконала

$$(\nabla \Psi)^2 = \varepsilon, \quad (9)$$

а амплитуды G_0, G_1, \dots — уравнениям переноса

$$\begin{aligned} 2\nabla G_0 \nabla \Psi + G_0 \Delta \Psi &= 0, \\ 2\nabla G_1 \nabla \Psi + G_1 \Delta \Psi &= -\Delta G_0, \\ 2\nabla G_2 \nabla \Psi + G_2 \Delta \Psi &= -\Delta G_1, \end{aligned} \quad (10)$$

Добавив к левой части уравнения (7) слагаемое $-Mg$, где

$$M(\mathbf{r}, \mathbf{r}_s) = \Delta G_0(\mathbf{r}, \mathbf{r}_s)/G_0(\mathbf{r}, \mathbf{r}_s), \quad (11)$$

мы получим уравнение, которому подчиняется функция Грина

$$g_{GO}^0 = G_0 e^{ik\Psi}, \quad (12)$$

вычисленная в нулевом приближении геометрической оптики:

$$\Delta g_{GO}^0 + (k^2\varepsilon - M) g_{GO}^0 = \delta(\mathbf{r} - \mathbf{r}_s). \quad (13)$$

Добавив к левой и правой частям исходного уравнения (7) величину $-Mg$, получим уравнение

$$\Delta g + (k^2\varepsilon - M) g = \delta(\mathbf{r} - \mathbf{r}_s) - Mg, \quad (14)$$

решение которого представим в виде

$$g = g_{GO}^0 + \gamma. \quad (15)$$

С учетом (13) уравнение для поправки γ можно представить в форме

$$\Delta\gamma + (k^2\varepsilon - M)\gamma = -M(g_{GO}^0 + \gamma) \quad (16)$$

или в эквивалентной интегральной форме

$$\gamma(\mathbf{r}, \mathbf{r}_s) = - \int M(\mathbf{r}', \mathbf{r}_s) \left[g_{GO}^0(\mathbf{r}', \mathbf{r}_s) + \gamma(\mathbf{r}', \mathbf{r}_s) \right] g_{GO}^0(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}'. \quad (17)$$

Формальное решение уравнения (17) можно записать в виде ряда по степеням малой величины M :

$$\gamma = \sum_{n=1}^{\infty} \hat{M}^n g_{GO}^0 = \sum_{n=1}^{\infty} \gamma_n, \quad (18)$$

где через \hat{M} обозначен интегральный оператор, действующий по правилу

$$\hat{M}f(\mathbf{r}, \mathbf{r}_s) = - \int M(\mathbf{r}', \mathbf{r}_s) f(\mathbf{r}', \mathbf{r}_s) g_{GO}^0(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}'.$$

Первый член ряда (18) представляется интегралом

$$\gamma_1 = - \int M(\mathbf{r}', \mathbf{r}_s) g_{GO}^0(\mathbf{r}', \mathbf{r}_s) g_{GO}^0(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}'. \quad (19)$$

2. РАЗЛОЖЕНИЕ РЕШЕНИЯ В РЯД ДЛЯ ВОЛНЫ ПРОИЗВОЛЬНОГО ВИДА

Располагая функцией Грина (15), нетрудно построить разложение для волнового поля $u(\mathbf{r})$, которое подчиняется уравнению Гельмгольца $\Delta u + k^2\varepsilon u = 0$. Представим решение этого уравнения в виде суммы нулевого приближения геометрической оптики u_{GO}^0 и поправки w :

$$u = u_{GO}^0 + w = A_0 e^{ik\Psi} + w. \quad (20)$$

Учитывая, что ψ удовлетворяет уравнению эйконала $(\nabla\psi)^2 = \varepsilon$, а амплитуда нулевого приближения A_0 — уравнению переноса $2(\nabla A_0 \nabla \psi) + A_0 \Delta \psi = 0$ (см. (10)), для поля w получим неоднородное уравнение

$$\Delta w + k^2 \varepsilon w = -\Delta A_0 e^{ik\psi}. \quad (21)$$

Решение этого уравнения может быть выражено через функцию Грина (15):

$$w(\mathbf{r}) = - \int \Delta A_0(\mathbf{r}') e^{ik\psi(\mathbf{r}')} g(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}', \quad (22)$$

которое с учетом (18) может быть представлено в виде

$$w(\mathbf{r}) = - \int \Delta A_0(\mathbf{r}') e^{ik\psi(\mathbf{r}')} g_{\text{ГО}}^0(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}' - \sum_{n=1}^{\infty} \int \Delta A_0(\mathbf{r}') e^{ik\psi(\mathbf{r}')} \gamma_n(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}', \quad (23)$$

или, если пренебречь поправками γ_n , через геометрическую функцию Грина (12):

$$w_1(\mathbf{r}) = - \int \Delta A_0(\mathbf{r}') e^{ik\psi(\mathbf{r}')} g_{\text{ГО}}^0(\mathbf{r}, \mathbf{r}') d^3\mathbf{r}'. \quad (24)$$

Рассмотрим некоторые частные случаи использования интегральной оценки (24).

3. РАССЕЯНИЕ ВПЕРЕД

Рассмотрим вначале рассеяние волнового поля вперед, предполагая, что точка наблюдения C находится на луче ABC , который проходит через неоднородную среду (рис. 1). Учитывая выражение (12) для геометрической функции Грина и переходя в (24) к интегрированию по лучевым координатам ξ', η', τ' , перепишем (24) в виде

$$w(\boldsymbol{\xi}) = - \int \Delta A_0(\boldsymbol{\xi}') G_0(\boldsymbol{\xi}', \boldsymbol{\xi}) \exp[ik(\psi(\boldsymbol{\xi}') + \Psi(\boldsymbol{\xi}', \boldsymbol{\xi}))] D(\boldsymbol{\xi}') d\xi' d\eta' d\tau',$$

где $\boldsymbol{\xi} = (\xi, \eta, \tau)$, а

$$D(\boldsymbol{\xi}) = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \tau)}$$

— якобиан перехода от декартовых координат x, y, z к лучевым координатам ξ, η, τ (положим, что координата τ отсчитывается вдоль луча, η и ξ — поперек).

Зафиксируем параметр τ' на луче и вычислим интеграл по поперечным к лучу координатам ξ' и η' по методу стационарной фазы. Для этого разложим суммарную фазу $\psi(\boldsymbol{\xi}) + \Psi(\boldsymbol{\xi}', \boldsymbol{\xi})$ в ряд по ξ' и η' :

$$\begin{aligned} \psi(\boldsymbol{\xi}') + \Psi(\boldsymbol{\xi}', \boldsymbol{\xi}) &= \psi(\boldsymbol{\xi}) + \frac{1}{2} \frac{\partial^2}{\partial \xi'^2} (\psi + \Psi)|_{(\xi, \eta)} (\xi - \xi')^2 + \\ &+ \frac{\partial^2}{\partial \xi' \partial \eta'} (\psi + \Psi)|_{(\xi, \eta)} (\xi - \xi') (\eta - \eta') + \frac{1}{2} \frac{\partial^2}{\partial \eta'^2} (\psi + \Psi)|_{(\xi, \eta)} (\eta - \eta')^2. \end{aligned} \quad (25)$$

Здесь ξ и η — координаты стационарной точки, в которой первые производные от суммарного эйконала $\psi + \Psi$ по ξ' и η' равны нулю в силу экстремальных свойств луча. Через $\psi(\boldsymbol{\xi})$ в (25) обозначен полный эйконал ψ в точке наблюдения C , складывающийся из эйконала первичной волны ψ на участке AB и эйконала функции Грина Ψ на участке BC . Интеграл по поперечным переменным ξ' и η' приближенно равен

$$\int \exp[ik(\psi + \Psi)] d\xi' d\eta' \simeq \frac{\pi i}{k} \exp[ik\psi(\boldsymbol{\xi})] E^{-1/2}(\tau, \tau'), \quad (26)$$

где

$$E = \left\{ \frac{\partial^2(\psi + \Psi)}{\partial \xi'^2} \frac{\partial^2(\psi + \Psi)}{\partial \eta'^2} - \left[\frac{\partial^2(\psi + \Psi)}{\partial \xi' \partial \eta'} \right]^2 \right\} \Big|_{(\xi, \eta)},$$

так что поправка w_1 в точке наблюдения C принимает вид

$$w_1(C) = -\frac{\pi i}{k} \exp[ik\psi(\tau)] \int_A^C \Delta A_0(\tau') G_0(\tau', \tau) D(\tau') E^{-1/2} d\tau'. \tag{27}$$

В (27) во всех аргументах опущены координаты опорного луча ξ и η .

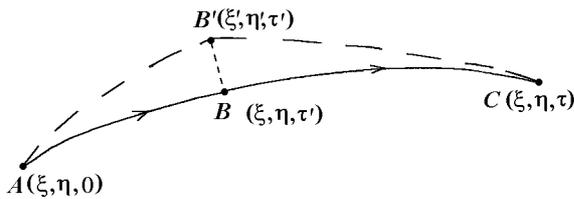


Рис. 1. К вычислению первой поправки к геометрикооптическому решению при рассеянии вперед. A — точка выхода луча, B' — точка интегрирования, B — точка интегрирования на луче после вычисления интеграла по поперечным координатам ξ' и η' , C — точка наблюдения

В этом интеграле усматривается сходство с первым приближением геометрической оптики w_{GO}^1 , которое согласно [2] имеет вид

$$w_{GO}^1 = -\frac{e^{ik\psi(\tau)}}{2ik} \int_A^C \Delta A_0 \left[\frac{D(\tau')}{D(\tau)} \right]^{1/2} d\tau'. \tag{28}$$

При рассеянии вперед поправка (27) асимптотически переходит в первое приближение геометрической оптики.

4. РАССЕЯНИЕ НАЗАД

Рассматривая рассеяние назад, поместим точку наблюдения D рядом с точкой A на первичном фронте волны (рис. 2).

Если расстояние L от точки наблюдения D до точки B на луче удовлетворяет условию $\sqrt{\lambda L} \ll l$, при котором поперечник первой зоны Френеля мал по сравнению с размером неоднородности l , то рассеянное поле можно вычислить, используя метод стационарной фазы при интегрировании по поперечным переменным ξ' и η' , при этом

$$w_1(D) = -\frac{\pi i}{k} \int \exp[ik\psi_D(\tau')] \Delta A_0(\tau') G_0(\tau, \tau') E^{-1/2}(\tau, \tau') d\tau'. \tag{29}$$

Значения всех величин в (29) берутся в стационарной точке (ξ, η) . В отличие от выражения (27) эйконал ψ_D складывается здесь из эйконала $\psi(\mathbf{r}'_B)$ на прямом пути от точки A до точки B и из эйконала Ψ на обратном пути. Поэтому при смещении точки τ' суммарный эйконал ψ_D будет меняться с удвоенной скоростью по сравнению с $\psi(\mathbf{r}')$:

$$d\psi_D = 2 d\psi = 2\varepsilon(\tau') d\tau'.$$

В силу предполагаемой медленности изменения остальных множителей интеграл (29) имеет экспоненциально малую величину: если характерный масштаб изменения свойств среды обозначить через l , то поправка в точке наблюдения D

$$w_1(D) \sim e^{-kl}. \tag{30}$$

Если же точка наблюдения D находится в зоне Фраунгофера по отношению к области неоднородности, то интеграл (24) можно привести к виду

$$w_1 \approx \frac{e^{ik\varphi}}{4\pi R} \int \Delta A_0 \sqrt[4]{\varepsilon} \exp(i\mathbf{q}\rho) d^3\rho = \frac{e^{ik\varphi}}{R} 2\pi^2 \Phi(\mathbf{q}), \tag{31}$$

где $\mathbf{q} = \nabla'\psi + \nabla'\Psi$ — вектор рассеяния, соответствующий центру неоднородности, R — расстояние от точки наблюдения до центра неоднородности, $\varphi = \psi(\mathbf{r}_0) + \Psi(\mathbf{r}_0, \mathbf{r}_D)$ — суммарный оптический путь первичной волны до центра неоднородности и рассеянной волны от центра неоднородности до точки наблюдения D , $\boldsymbol{\rho} = \mathbf{r}' - \mathbf{r}_0$ — смещение точки рассеяния от центра неоднородности \mathbf{r}_0 . Наконец, через $\Phi(\mathbf{q})$ в (31) обозначено преобразование Фурье от $\Delta A_0 \sqrt[4]{\varepsilon}$:

$$\Phi(\mathbf{q}) = \frac{1}{8\pi^3} \int \Delta A_0(\boldsymbol{\rho}) \sqrt[4]{\varepsilon(\boldsymbol{\rho})} \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\rho. \tag{32}$$

В силу предполагаемой медленности изменения $\Delta A_0 \sqrt[4]{\varepsilon}$ по сравнению с $\exp(i\mathbf{q}\boldsymbol{\rho})$ преобразование Фурье (32) характеризуется экспоненциальным множителем, подобным (30). Это справедливо при рассеянии назад, когда вектор рассеяния \mathbf{q} вдвое превышает волновой вектор \mathbf{k}_0 первичного поля, $\mathbf{q} = 2\mathbf{k}_0 = 2\nabla\psi(\mathbf{r}_0)$, но не при рассеянии вперед, когда $\mathbf{q} = 0$.

В формуле (31) неявно предполагается, что на больших расстояниях от \mathbf{r}_0 амплитуда A_0 убывает как $1/R$, при этом ΔA_0 стремится к нулю. Наличие каустик, на которых амплитуда A_0 и лапласиан ΔA_0 имеют особенности, усложняют проведение расчетов. Однако особенности на каустиках можно устранить переходом от приближения геометрической оптики к интегральному представлению Маслова [7, 8].

В заключение оценим поле, рассеянное назад, в случае слабой неоднородности, при

$$\varepsilon(\mathbf{r}) = 1 + \nu(\mathbf{r}), \tag{33}$$

где $|\nu| \ll 1$.

Сделаем дополнительное упрощающее предположение, что на локальную неоднородность с поперечным размером l падает плоская волна, распространяющаяся вдоль оси z .

Решая уравнения геометрической оптики методом возмущений, для относительного изменения амплитуды $\delta A_0/A_0$, получим следующее выражение:

$$\frac{\delta A_0}{A_0} = -\frac{1}{4}\nu - \frac{1}{4}\Delta_{\perp} \int_0^z dz' \int_0^{z'} \nu(x, y, z'') dz'', \tag{34}$$

которое несколько уточняет формулу (35.7) из [9], сохраняя слагаемое $-\nu/4$, которым в статистических задачах обычно пренебрегают (здесь $\Delta_{\perp} \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$ — лапласиан по поперечным к оси z координатам). Слагаемое $-\nu/4$ в (34) обусловлено изменением свойств среды вдоль луча (оно отражает изменение амплитуды по закону $A_0 \sim \varepsilon^{-1/4}$, реализующееся в одномерной задаче), тогда как второе слагаемое связано с искривлением лучей внутри неоднородности.

Подставляя (34) в интеграл (31) и учитывая соотношение (33), находим

$$w_1 = -\frac{A_0 e^{ik\varphi}}{R} \frac{\pi^2}{2} [\Phi_1(\mathbf{q}) + \Phi_2(\mathbf{q}) + \Phi_3(\mathbf{q})], \tag{35}$$

где

$$\Phi_1(\mathbf{q}) = \frac{1}{8\pi^3} \int \Delta \nu \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\rho, \quad \Phi_2(\mathbf{q}) = \frac{1}{8\pi^3} \int \Delta_{\perp} \nu \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\rho,$$

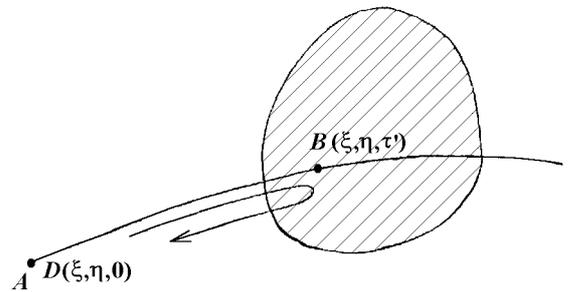


Рис. 2. К вычислению экспоненциально малой поправки к геометрическому решению при рассеянии назад. Точка наблюдения D совпадает с точкой выхода луча A , текущая точка интегрирования B находится внутри заштрихованной неоднородности. Оптические пути $\psi = \psi_{AB}$ и $\Psi = \psi_{BD}$ равны, и мы получаем удвоенный эйконал $2ik\psi$ в экспоненте подынтегрального выражения поправки w_1

$$\Phi_3(\mathbf{q}) = \frac{1}{8\pi^3} \int \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\boldsymbol{\rho} \int_0^z dz' \int_0^{z'} \Delta_{\perp} \nu dz''.$$

Согласно свойству преобразования Фурье имеем

$$\Phi_1 = \mathbf{q}^2 \Phi_{\nu}, \quad \Phi_2 = \mathbf{q}_{\perp}^2 \Phi_{\nu}, \quad \Phi_3 = \frac{1}{8\pi^3} \mathbf{q}_{\perp}^4 \int \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\boldsymbol{\rho} \int_0^z dz' \int_0^{z'} \nu(x, y, z'') dz'',$$

$$\Phi_{\nu} = \frac{1}{8\pi^3} \int \nu \exp(i\mathbf{q}\boldsymbol{\rho}) d^3\boldsymbol{\rho},$$

где $\mathbf{q}_{\perp} = (q_x, q_y)$.

При рассеянии строго назад поперечная компонента вектора рассеяния обращается в нуль ($\mathbf{q}_{\perp} = 0$), а продольная компонента q_z равна $2k$. В результате

$$w_1 = -\frac{A_0 e^{ik\varphi}}{R} 2\pi^2 k^2 \Phi_{\nu}(0, 0, 2k). \quad (36)$$

Этот результат согласуется с борновским приближением теории рассеяния

$$w_B = -\frac{A_0 e^{ik\varphi}}{R} 2\pi^2 k^2 \Phi_{\nu}(\mathbf{q}). \quad (37)$$

В случае плавных возмущений борновское приближение (37) описывает экспоненциально малое рассеяние. Например, в случае неоднородности с гауссовой формой, $\nu = \nu_{\max} \exp(-\rho^2/(2l^2))$, при рассеянии назад получаем

$$w_B = -\frac{A_0 e^{ik\varphi}}{2R} l^3 k^2 \nu_{\max} \sqrt{2\pi} \exp(-2k^2 l^2).$$

Отметим, что борновское приближение применимо, строго говоря, при условии

$$k \int_{-\infty}^{+\infty} \nu dz \ll 1, \quad (38)$$

которое означает малость набега фазы на длине неоднородности. Между тем, вывод соотношения (37) требует только малости отклонения лучей по сравнению с размером l неоднородности [2], т. е. выполнения условия вида

$$\int z |\nabla_{\perp} \nu| dz \ll l. \quad (39)$$

Это условие заведомо слабее неравенства (38), поскольку вместо $k\nu$ содержит комбинацию $z |\nabla_{\perp} \nu|/l$, которая по порядку величины сравнима с $1/l$, так что $z |\nabla_{\perp} \nu|/l \sim \nu_{\max}/l \ll k\nu_{\max}$. Таким образом, опирающаяся на МПГО формула (35) оправдывает применение борновского приближения (37) даже в условиях, когда теория возмущений, строго говоря, неприменима.

Отметим, что при рассеянии в стороны, когда поперечная составляющая вектора рассеяния \mathbf{q}_{\perp} отлична от нуля, согласие между выражением (35) и борновским приближением (37) нарушается. Это означает, что ряд (23) для поля w устроен иначе, чем борновское разложение, и что равенство полей (35) и (37) при рассеянии назад, увы, не сохраняется при рассеянии вбок и, тем более, вперед.

В средах с плавно меняющимися параметрами экспоненциально слабые рассеянные поля играют пренебрежимо малую роль в силу малости множителя $\exp(-kl)$ при $l \gg \lambda$. Исключение может составить разве что случай локализованной плазменной неоднородности в окрестности плазменной частоты $\omega_{\text{пл}}$. При $\omega = \omega_{\text{пл}}$ диэлектрическая проницаемость плазмы

$$\varepsilon = 1 - \frac{4\pi e^2 N(\mathbf{r})}{m\omega^2} \equiv 1 - \frac{\omega_{\text{пл}}^2}{\omega^2},$$

где $N(\mathbf{r})$ — концентрация частиц с массой m и зарядом e , обращается в нуль, так что при $\omega < \omega_{\text{пл}}$ экспоненциально малое рассеяние назад может смениться вполне ощутимым отражением в силу малости ε . В этой области частот фактор $\exp[ik(\psi + \Psi)]$ перестает быть быстроосциллирующим, способствуя тем самым резкому увеличению обратного рассеяния. Вблизи плазменной частоты локальная длина волны $\lambda = \lambda_0/\sqrt{\varepsilon}$ увеличивается. Когда λ сравнивается с масштабом неоднородности l , условия применимости геометрической оптики нарушаются. Таким образом, оценки вида (31) применимы при условии $\lambda/\sqrt{\varepsilon} \ll l$.

5. ЗАКЛЮЧЕНИЕ

Можно констатировать, что предлагаемое разложение (23) обладает следующими свойствами: 1) при рассеянии назад поле w_1 , представляющее собой поправку к нулевому приближению геометрической оптики, имеет экспоненциально малую величину; 2) при рассеянии вперед поле w_1 приобретает структуру, подобную первому приближению геометрической оптики; 3) при выполнении условия (39), которое является более слабым, чем требование (38) применимости борновского приближения, рассеянное назад поле (35) совпадает с борновским приближением (37), хотя при рассеянии вбок между этими полями появляются различия.

Разложение (23) следовало бы отнести к так называемым непertурбативным методам, поскольку в нем отсутствует явный малый параметр. Как известно, непertурбативные методы иногда дают хорошее (и даже очень хорошее) приближение, но в силу отсутствия явного малого параметра условия хорошей аппроксимации обычно трудно обосновать. В данном случае трудным для анализа является поведение геометрикооптических полей в окрестности каустик. Каустические сингулярности лучевого поля, в принципе, можно устранить путем использования асимптотических методов — метода Маслова или метода эталонных интегралов, но оценки качества аппроксимации в этом случае существенно усложняются.

В заключение авторы выражают свою признательность Ю. Н. Черкашину за обсуждения проблемы.

ЛИТЕРАТУРА

1. Борн М., Вольф Э. Основы оптики. — М.: Наука, 1970.
2. Кравцов Ю. А., Орлов Ю. И. Геометрическая оптика неоднородных сред. — М.: Наука, 1980.
3. Debye P. A note to Sommerfeld A., Runge J. // Anwendung der Vektorrechnung auf die Grundlagen der Geometrischen Optik. 1911. V. 35. P. 277.
4. Рытов С. М. // ДАН СССР. 1938. Т. 18, № 2. С. 263.
5. Northover F. N. // J. Math. Phys. 1969. V. 10, № 4. P. 715.
6. Кравцов Ю. А., Орлов Ю. И. // Современные проблемы распространения и рассеяния волн. — М.: ИРЭ АН СССР, 1979.
7. Маслов В. П., Федорюк М. В. Квазиклассическое приближение для уравнений квантовой механики. — М.: Наука, 1976.
8. Kravtsov Yu. A., Orlov Yu. I. Caustics Catastrophes and Wave Fields. 2nd ed. — Berlin, Heidelberg: Springer Verlag, 1998.
9. Рытов С. М., Кравцов Ю. А., Татарский В. И. Введение в статистическую радиофизику. Ч. II. — М.: Наука, 1978.

¹ Московский педагогический гос.
университет и Орский
гуманитарно-технологический
институт (филиал Оренбургского
гос. университета), Россия

² Институт космических
исследований РАН, Москва,
Россия и Центр космических
исследований Польской академии
наук, Варшава, Польша

Поступила в редакцию
12 мая 1999 г.

**MODIFIED GEOMETRICAL-OPTICS APPROXIMATION:
TAKING EXPONENTIALLY SMALL SCATTERED FIELDS INTO ACCOUNT**

R. A. Egorchenkov and Yu. A. Kravtsov

We propose a modified geometrical-optics approximation (MGOA). It is based on the solution of the Helmholtz equation in the form of a series in which the main term describes zero-order geometrical-optics approximation. The proposed method allows for taking into account exponentially small backscattering in smoothly inhomogeneous medium and thus improves the geometrical-optics approximation. In the case of forward scattering, the obtained solution is reduced to conventional geometrical-optics series over negative powers of wave number. We estimate the backscattering using perturbation technique for rays and determine conditions under which the backscattered field agrees with the Born approximation of the scattering theory.

ВЛИЯНИЕ РАЗМЕРНОСТИ НА РЕШЕНИЯ ЗАДАЧ ДИФРАКЦИИ В ОТКРЫТЫХ ВОЛНОВОДАХ

А. Б. Маненков

Сравниваются характеристики отражения поверхностных (волноводных) мод от обрыва планарного диэлектрического волновода и оптического волокна с вытянутым поперечным сечением. В области малых замедлений выведены аналитические соотношения для коэффициентов отражения основных мод. Анализ проведен в скалярном приближении с помощью метода интегральных уравнений и вариационных принципов. Показано, что различие характеристик рассматриваемых задач тесно связано с аналитическими свойствами системы собственных мод волноводов.

ВВЕДЕНИЕ

Задачи дифракции волноводных (поверхностных) мод на скачкообразных нерегулярностях открытых диэлектрических волноводов являются предметом интенсивных исследований последние несколько десятков лет [1–11]. Геометрия реальных систем такого типа достаточно сложна, поэтому их анализ часто проводится с помощью упрощенных моделей. В частности, исследование трехмерных систем с вытянутым поперечным сечением (см. ниже) заменяется анализом двумерных структур, причем такая замена обычно делается без должного обоснования.

Цель работы заключается в том, чтобы на одном частном примере определить, как влияет размерность геометрии волноводной системы на дифракционные эффекты в нерегулярных открытых волноводах. В качестве такого примера выбрана достаточно часто встречающаяся задача об отражении основной поверхностной моды (ПМ) от обрыва двумерного диэлектрического слоя и от обрыва трехмерного волновода. Подобные системы являются удобной моделью для анализа различных устройств оптоэлектроники, антенной техники и для изучения распространения света в диэлектрических и металлических пленках [1–12]. Несмотря на частный характер рассматриваемой геометрии, большинство полученных выводов может быть отнесено к более широкому классу проблем (см. раздел 5). Отметим, что характеристики двумерных и трехмерных систем кроме размерности зависят также от поляризации волн [1–5, 10]. В работе исследовано только влияние размерности; для этого решаются скалярные задачи, в которых поляризационные эффекты отсутствуют.

Геометрия задачи для трехмерной проблемы изображена на рис. 1. Для простоты предполагаем, что поперечное сечение волновода (оптического волокна) имеет две оси симметрии — x и y . Постоянные показатели преломления сердцевины волокна и его оболочки (при $z < 0$) обозначим через n_1 и n_0 (причем $n_1 > n_0$), а показатель преломления правого полупространства — через n_f . Далее предполагаем, что все показатели преломления близки друг к другу, т. е. $|n_1 - n_0| \ll n_0$ и $|n_1 - n_f| \ll n_f$. Эти условия достаточны для применимости скалярного приближения [8]. Для таких структур одна из поперечных компонент электрического поля является доминирующей, так что остальными компонентами можно пренебречь. Двумерная система (полубесконечный диэлектрический слой) получается из геометрии, изображенной на рис. 1, предельным переходом при бесконечном увеличении поперечного размера волновода вдоль оси x и постоянном размере вдоль оси y . В дальнейшем будем, как обычно, предполагать, что для двумерной задачи поля не зависят от координаты x . Предполагаем также, что волноводы одномодовые. Временной множитель $\exp(-ikt)$, где k — волновое число в вакууме, c — скорость света, во всех соотношениях опущен.

1. МЕТОД РАСЧЕТА

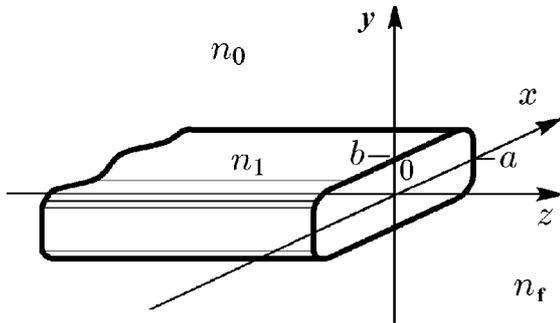


Рис. 1

Решение описанных выше задач получено на основе интегральных уравнений (ИУ), которые выводятся с помощью спектральных разложений полей по собственным модам диэлектрического волновода и свободного пространства. Вывод ИУ во многом повторяет известную технику, широко применяемую для анализа нерегулярностей металлических волноводов [13–15]. Однако обобщение ее на открытые системы имеет ряд особенностей (в основном для трехмерной геометрии), поэтому опишем коротко этот подход, опуская детали.

Предполагаем, что на обрыв волновода (волокна) набегаёт слева (рис. 1) основная поверхностная (волноводная) мода, которую в соответствии со стандартной системой обозначений [16] будем называть E_{11} -модой. Далее обозначим через $E_s(1)$ электрическое поле ПМ, через $E_{m\kappa}(1)$ и $E_{m\kappa}(2)$ — поля радиационных мод (РМ) соответственно волновода ($z < 0$) и правого полупространства ($z > 0$); поперечные волновые числа поверхностной и радиационных мод равны соответственно κ_s и κ , m — дискретный индекс [17–19]. Числа 1 и 2 в скобках отмечают величины, относящиеся соответственно к левой и правой областям пространства (рис. 1). При $z < 0$ спектральное разложение поля имеет вид

$$E = [\exp(i\beta_s z) + \rho \exp(-i\beta_s z)] E_s(1) + \sum_m \int_0^\infty \rho_m(\kappa) E_{m\kappa}(1) \exp(-i\beta_\kappa z) d\kappa, \tag{1}$$

а справа от обрыва

$$E = \sum_m \int_0^\infty \tau_m(\kappa) E_{m\kappa}(2) \exp(i\gamma_\kappa z) d\kappa. \tag{2}$$

В этих формулах ρ — коэффициент отражения волноводной моды, $\rho_m(\kappa)$ и $\tau_m(\kappa)$ — коэффициенты разложений (амплитуды), β_s , β_κ и γ_κ — продольные волновые числа собственных мод: $\beta_s^2 = k^2 n_0^2 - \kappa_s^2$, $\beta_\kappa^2 = k^2 n_0^2 - \kappa^2$, $\gamma_\kappa^2 = k^2 n_f^2 - \kappa^2$ ($0 < \kappa < \infty$). Аналогичные разложения можно написать для магнитного поля H . При выводе этих формул следует учесть, что в скалярной теории удобно определить это поле соотношением $H = (ik)^{-1} \partial E / \partial z$.

Система собственных мод может быть построена методом S -оператора [17–19]. Поля этих мод ортонормированы:

$$\begin{aligned} \langle E_s(1), E_s(1) \rangle &= 1, & \langle E_s(1), E_{m\kappa}(1) \rangle &= 0, \\ \langle E_{m\kappa}(1), E_{m'\kappa'}(1) \rangle &= \delta_{mm'} \delta(\kappa - \kappa'), & \langle E_{m\kappa}(2), E_{m'\kappa'}(2) \rangle &= \delta_{mm'} \delta(\kappa - \kappa'), \end{aligned} \tag{3}$$

где $\delta_{mm'}$ — символ Кронекера, $\delta(\kappa - \kappa')$ — дельта-функция,

$$\langle E_1, E_2 \rangle = \int_{z=0} E_1 E_2 dx dy \tag{4}$$

— интеграл по плоскости $z = 0$ от произведения функций E_1 и E_2 . Используя эти соотношения, все коэффициенты разложений можно выразить через поле $\mathcal{E} = E(x, y, 0)$ в плоскости $z = 0$; например, $\rho = \langle \mathcal{E}, E_s(1) \rangle - 1$. Сшивая поля в плоскости обрыва, нетрудно получить ИУ для поля \mathcal{E} :

$$\hat{\Xi}_e[\mathcal{E}] = 2\beta_s E_s(1), \tag{5}$$

где интегральный оператор $\hat{\Xi}_e$ равен

$$\hat{\Xi}_e[\mathcal{E}] = \langle \mathcal{E}, E_s(1) \rangle \beta_s E_s(1) + \sum_m \int_0^\infty \langle \mathcal{E}, E_{m\kappa}(1) \rangle \beta_\kappa E_{m\kappa}(1) d\kappa + \\ + \sum_m \int_0^\infty \langle \mathcal{E}, E_{m\kappa}(2) \rangle \gamma_\kappa E_{m\kappa}(2) d\kappa. \quad (6)$$

Отметим, что ядро уравнения (5) имеет особенность при совпадении аргументов (координат) [2].

С помощью стандартной техники [6, 13–15] из ИУ (5) выводятся стационарные функционалы для различных характеристик задачи. В частности, для коэффициента отражения ρ основной моды имеем

$$\frac{1 - \rho}{1 + \rho} = \frac{\langle \mathcal{E}, \hat{\Xi}_{er}[\mathcal{E}] \rangle}{\beta_s \langle \mathcal{E}, E_s(1) \rangle^2}, \quad (7)$$

где $\hat{\Xi}_{er}[\mathcal{E}] = \hat{\Xi}_e[\mathcal{E}] - \langle \mathcal{E}, E_s(1) \rangle \beta_s E_s(1)$. Соотношение (7) является электрической формулировкой вариационного принципа. Аналогичным образом можно вывести ИУ и стационарные функционалы, в которые входит магнитное поле $\mathcal{H} = H(x, y, 0)$ в плоскости стыка (магнитную формулировку вариационного принципа) [13]. Отметим, что в общем случае эти функционалы не совпадают с электрическими.

В силу стационарности выражения (7) для величины ρ можно получить достаточно точные результаты, если подставить в эту формулу приближенное распределение поля в плоскости обрыва $z = 0$. Ниже будем использовать поле, пропорциональное полю падающей моды: $\mathcal{E} = A E_s(1)$, где A — некоторая константа. Эта константа сокращается после подстановки в функционал (7), поэтому формально можно считать, что $A = 1 + \rho$. При таком выборе функции \mathcal{E} в плоскости обрыва пренебрегаем РМ волновода (см. ниже). После простых преобразований с учетом (3) получаем

$$\rho = \frac{\sum_m \int_0^\infty (\beta_s - \gamma_\kappa) Q_{m\kappa}^2 d\kappa}{\sum_m \int_0^\infty (\beta_s + \gamma_\kappa) Q_{m\kappa}^2 d\kappa}, \quad (8)$$

где $Q_{m\kappa} = \langle E_s(1), E_{m\kappa}(2) \rangle$. При выводе этой формулы было использовано равенство Парсевала [20]

$$\langle E_s(1), E_s(1) \rangle = \sum_m \int_0^\infty \langle E_s(1), E_{m\kappa}(2) \rangle^2 d\kappa, \quad (9)$$

которое позволяет внести β_s под знак интеграла. С помощью формулы Грина выражения для функций $Q_{m\kappa}$ можно преобразовать к виду

$$Q_{m\kappa} = \frac{k^2 (n_1^2 - n_0^2)}{\kappa^2 + |\kappa_s|^2} \iint_{\Omega_1} E_s(1) E_{m\kappa}(2) dx dy, \quad (10)$$

в котором интегралы вычисляются только по поперечному сечению волновода Ω_1 . Аналогичные, но несколько более простые функционалы получаются для двумерной задачи о рассеянии на обрыве плоского диэлектрического слоя [20].

Из последних соотношений можно вывести ряд аналитических оценок. Например, в области низких частот ($k \rightarrow 0$) имеем

$$\rho \rightarrow \rho_0 = (n_0 - n_f)/(n_0 + n_f), \quad (11)$$

а при $k \rightarrow \infty$

$$\rho \rightarrow \rho_\infty = (n_1 - n_f)/(n_1 + n_f). \quad (12)$$

2. СОБСТВЕННЫЕ МОДЫ

Поля РМ в правом полупространстве ($z > 0$) имеют простое аналитическое представление. Для трехмерной задачи четные моды равны

$$E_{m\kappa}(2) = (\kappa\xi_m/\pi)^{1/2} J_{2m}(\kappa r) \cos(2m\varphi), \quad (13)$$

где $m = 0, 1, 2, \dots$ — целый индекс, r и φ — полярные координаты, J_{2m} — функции Бесселя 1-го рода, $\xi_0 = 1/2$, $\xi_m = 1$ для $m > 0$. Для задачи о рассеянии на обрыве диэлектрической пластины поля четных мод при $z > 0$ пропорциональны $\cos(\kappa y)$:

$$E_{m\kappa}(2) = (1/\pi)^{1/2} \cos(\kappa y).$$

При указанном выше выборе функции \mathcal{E} в силу соотношений (3) поля РМ волновода не вошли в (8), поэтому не будем рассматривать их подробно. Отметим лишь, что свойства ядер ИУ для задач с открытыми волноводами отличаются от свойств ИУ для задач с закрытыми системами. Это отличие связано в основном с тем, что в открытых структурах спектр собственных мод включает непрерывную часть [17–19], причем РМ определены неоднозначно. В частности, в качестве РМ свободного пространства вместо функций (13) можно использовать систему плоских волн, тогда разложение (2) будет иметь вид двойного интеграла Фурье. Как обычно, преимущества одних разложений перед другими определяются геометрией задачи.

При вычислении интегралов в (8) и (10) приходится многократно рассчитывать поле волноводной моды $E_s(1)$, поэтому эффективность метода определяется в основном свойствами процедур вычисления этой функции. Для определения характеристик ПМ использовался модифицированный метод интегральных уравнений, который позволяет быстро и точно рассчитывать ПМ в волноводе с гладкой формой поперечного сечения [11, 21]. Поле $E_s(1)$ во внутренней области Ω_1 волокна аппроксимируется конечным рядом Релея; с учетом симметрии задачи это разложение имеет вид [16]

$$E_s(r, \varphi) = \sum_{m=0}^{M-1} A_m J_{2m}(ur) \cos(2m\varphi), \quad u^2 = k^2 (n_1^2 - n_0^2) - w^2, \quad (14)$$

где $w = |\kappa_s|$ — модуль поперечного волнового числа моды. Внутри области Ω_1 поле $E_s(1)$ удовлетворяет волновому уравнению, а на граничном контуре поперечного сечения Λ — интегральному уравнению [21]

$$\oint_{\Lambda} \left[K_0(wr_{12}) \frac{\partial}{\partial \nu_2} E_s(\mathbf{r}_2) + w K_1(wr_{12}) \nu_2 \frac{\mathbf{r}_2 - \mathbf{r}_1}{r_{12}} E_s(\mathbf{r}_2) \right] dl_2 = 0, \quad (15)$$

где \mathbf{r}_1 и \mathbf{r}_2 — радиус-векторы соответственно точки коллокации с координатами (x_1, y_1) и произвольной точки (x_2, y_2) на контуре Λ , $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$, K_0 и K_1 — модифицированные цилиндрические функции, dl_2 — элемент дуги контура Λ , ν_2 — внешняя нормаль к этому контуру в точке \mathbf{r}_2 . Алгебраизацию

задачи проводим, используя метод коллокации [11, 16]: левую часть (15) приравняем нулю в конечном числе точек (узлов) $\mathbf{r}_l = (x_l, y_l)$, $l = 1, 2, \dots, M$. Для симметричных структур эти точки удобно расположить в первом квадранте плоскости xy .

Вычислительный процесс, основанный на описанном коллокационном алгоритме, сходится, если использовать адаптивное распределение узлов [11, 16, 21]. Адаптивные точки лежат на контуре Λ , и их предельная плотность распределения (при $M \rightarrow \infty$) должна быть пропорциональна $|\Psi'(\zeta)|$, где $\Psi(\zeta)$ — аналитическая функция комплексной переменной $\zeta = x + iy$, осуществляющая конформное отображение области, расположенной вне Λ , на область $|\Psi(\zeta)| > 1$. Эти точки можно найти численно с помощью простой процедуры Фекете [22]. Обозначим через ζ_j комплексные координаты узлов на контуре Λ , где j — номера точек. В число этих узлов надо включить точки, которые расположены во всех четырех квадрантах, так что $j = 1, 2, \dots, 4M$. Далее введем функцию $W(\zeta_1, \dots, \zeta_{4M}) = \prod_{j \neq l} |\zeta_j - \zeta_l|$.

Узлы Фекете определяются как система точек, в которых эта функция достигает максимума; координаты точек максимума находились стандартными программами оптимизации. В пределе $M \rightarrow \infty$ распределение точек $\{\zeta_j\}$ пропорционально $|\Psi'(\zeta)|$.

3. РЕЗУЛЬТАТЫ РАСЧЕТОВ

Для вычисления коэффициентов отражения ρ использовались соотношения (8) и (10). Расчеты проводились для волокон с суперэллиптической формой поперечного сечения. В поперечной плоскости граница Λ такого волновода определяется выражением

$$(x/a)^{2N_c} + (y/b)^{2N_c} = 1, \quad (16)$$

где N_c — целое число, a и b — полуоси суперэллипса. Меняя параметры в формуле (16), легко получить целый ряд кривых, которые описывают форму поперечного сечения волноводов, довольно широко применяющихся в оптоэлектронике. Например, при $N_c = 1$ формула (16) определяет обычный эллипс, а при $N_c \gg 1$ этот контур близок к прямоугольнику (со слегка сглаженными углами). При фиксированном размере b и $a \rightarrow \infty$ суперэллиптический волновод переходит в двумерную структуру — диэлектрическую пластину.

Рассмотрим теперь результаты расчетов коэффициента отражения основной ПМ в таких системах. На рис. 2 представлены зависимости $|\rho|$ от размера $2b$ для двух задач: сплошная кривая 1 построена для задачи о рассеянии основной моды на обрыве эллиптического волновода, а штриховая кривая 2 — на обрыве пластины. Параметры рассматриваемых систем были следующими: длина волны в свободном пространстве $\lambda = 2\pi/k = 1,3$ мкм, $n_1 = 1,5$, $n_0 = 1,455$, $n_f = 1,425$. Размер малой оси эллиптического волокна $2b$ совпадал с толщиной пластины (которую также обозначаем через $2b$). Для эллиптического волновода отношение полуосей постоянно и равно $a/b = 4$, а для пластины формально $a/b = \infty$. На рис. 2 горизонтальные пунктирные прямые показывают предельные значения ρ_0 и ρ_∞ (см. выше). Как видно из рисунка, характер зависимостей для двумерной (обрыв пластины) и трехмерной (обрыв эллиптического волокна) систем в области малых b заметно отличается.

На рис. 2 вертикальная штриховая линия указывает границу одномодового режима волокна. Критическое значение высоты эллиптического волновода $2b_c \approx 1,5$ мкм; при больших значениях $2b$ на обрыве может возбуждаться высшая мода * E_{31} . При указанных выше параметрах диэлектрический слой становится многомодовым при сравнительно большой толщине $2b_c = 3,564$ мкм; это значение лежит за пределами графика. При $b > b_c$ использованное в работе простое приближение для поля \mathcal{E} становится менее точным, так как не выполняется условие одномодовости волокна. В области высоких частот,

* При $a > b$ эта мода является второй высшей модой волокна [13]. Из-за симметрии рассматриваемой задачи первая высшая мода E_{21} на таком обрыве не возбуждается.

когда справедливо приближение геометрической оптики, указанная выше зависимость решений задач от их размерности, становится, по-видимому, незначительной. Тем не менее, геометрические свойства задач могут влиять на их решения и в этом диапазоне, например в узких областях вблизи критических частот высших ПМ, величина которых зависит от формы поперечного сечения волновода. Однако для слабонаправляющих систем, когда $n_1 \sim n_0$, эти эффекты невелики (см. ниже).

Отмеченные различия сохраняются и при другой форме поперечного сечения волокна. В качестве примера на рис. 3 изображены зависимости $|\rho|$ от размера $2b$ для задач о рассеянии на обрыве суперэллиптического волновода (кривая 1) и на обрыве пластины (кривая 2). Геометрические параметры суперэллиптического волокна были следующими: $a/b = 4, N_c = 4$; остальные параметры волокна и параметры пластины были те же, что и для рис. 2. Как и на рис. 2, вертикальная штриховая линия указывает критическое значение высоты суперэллиптического волновода $2b_c \approx 1,44$ мкм, которое определяет границу одномодового режима.

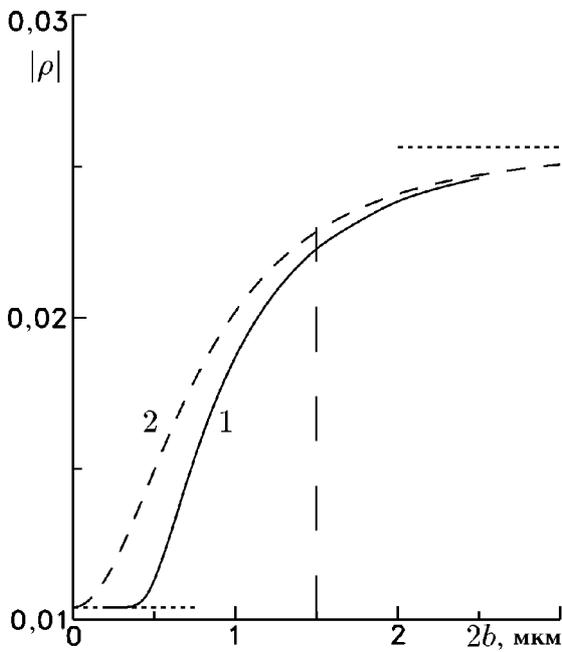


Рис. 2

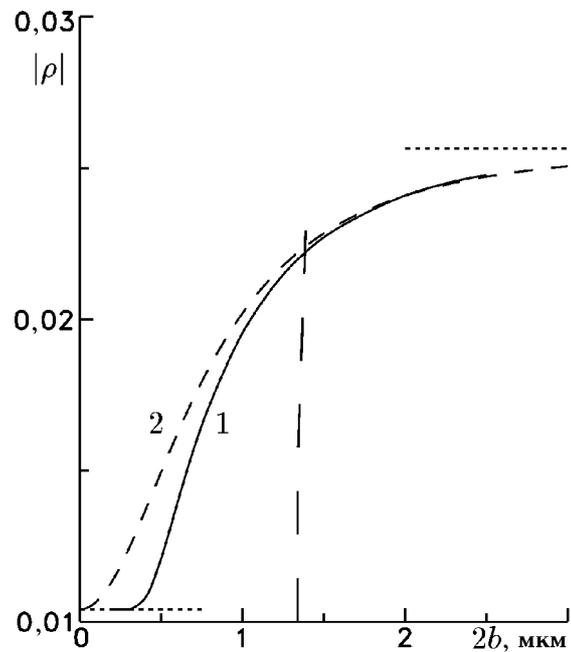


Рис. 3

В рассмотренных выше случаях показатель преломления $n_f < n_0$. Для систем, у которых $n_0 < n_f < n_1$, при некотором размере $2b$ коэффициент отражения ρ может быть близок к нулю. Такой режим согласования волноводной моды интересен для различных приложений [2, 4, 5]. Результаты расчета $|\rho|$ для подобных систем изображены на рис. 4. Как и выше, кривая 1 построена для суперэллиптического волокна ($a/b = 4, N_c = 4$), а кривая 2 — для диэлектрической пластины. Для обеих задач параметры были следующими: $\lambda = 1,3$ мкм, $n_1 = 1,5, n_0 = 1,455, n_f = 1,485$. Для рассматриваемых структур значения размера $2b$, при которых $\rho \approx 0$,

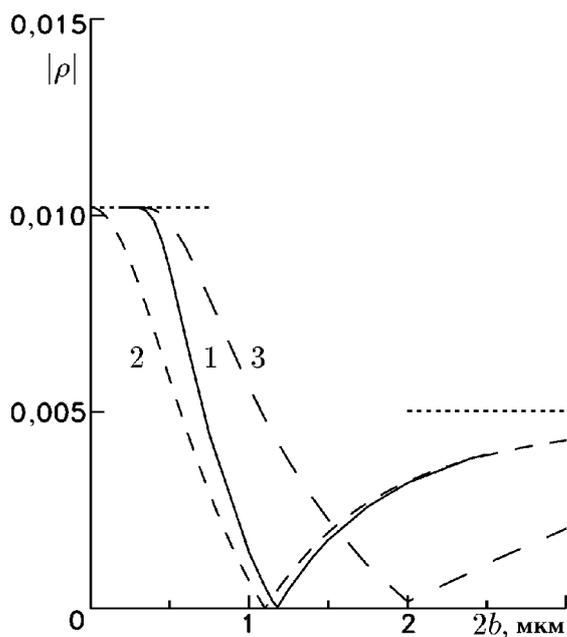


Рис. 4

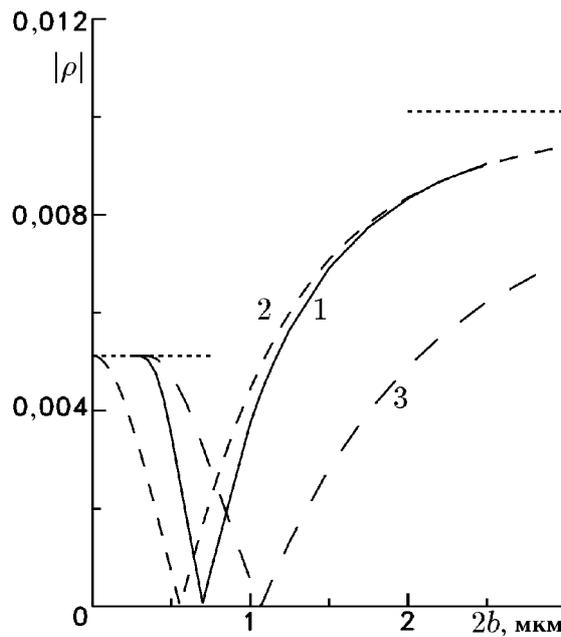


Рис. 5

отличаются.** Сдвиг точек согласования для этих задач становится больше, если n_f приближается к n_0 . Такой случай иллюстрирует рис. 5, приведенные на котором кривые вычислены при показателе преломления правого полупространства $n_f = 1,47$; остальные параметры были те же, что и для рис. 4. Отметим, что критические размеры, которые определяют границы одномодовых режимов, для рассматриваемых волноводов здесь те же, что и для рис. 3, поэтому они не показаны. Для сравнения на рис. 4 и 5 приведены также данные (кривые 3), построенные с помощью формулы, полученной методом “эффективного показателя преломления” [20]:

$$\rho_{\text{eff}} = (\beta_s - kn_f) / (\beta_s + kn_f). \quad (17)$$

Результаты сравнения показывают, что такое приближение является достаточно грубым.

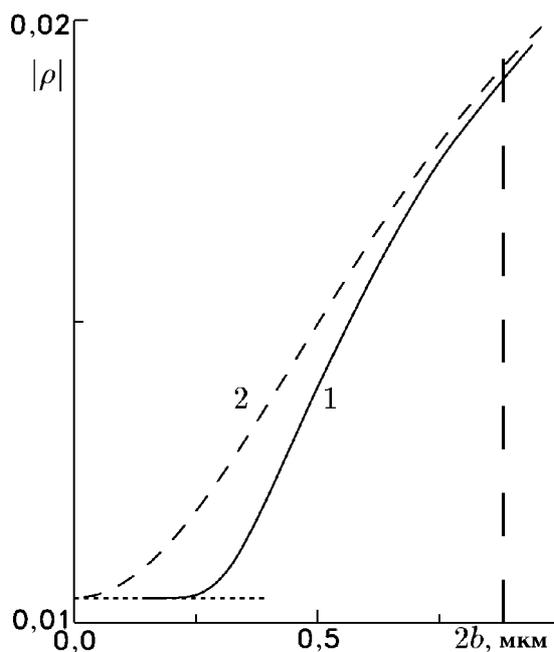


Рис. 6

На рис. 6 представлены значения $|\rho|$ при изменении размера $2b$ для задач о рассеянии на обрыве суперэллиптического волновода (кривая 1) с более вытянутым сечением и на обрыве пластины (кривая 2). Геометрические параметры суперэллиптического волокна были следующими: $a/b = 8$, $N_c = 2$; остальные параметры волокна и параметры пластины были те же, что и для рис. 2, 3 (в частности $n_f = 1,425$). Как и следовало ожидать, при увеличении отношения a/b зависимости $|\rho(2b)|$ для

** Поскольку используемая методика является приближенной, то, естественно, нельзя утверждать, что в отмеченных точках ρ обращается в ноль точно.

рассматриваемых трехмерных и двумерных задач сближаются. Как и выше, граница одномодового режима справа отмечена вертикальной штриховой линией. Поскольку соотношение (8) применимо только для одномодовых режимов, а при вытянутой форме волновода ($a \gg b$) с увеличением b спектр волноводных мод быстро сгущается, то расчеты ограничены анализом области сравнительно небольших замедлений моды E_{11} суперэллиптического волокна.

4. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Приведенные выше данные показывают, что различие рассматриваемых задач проявляется в наибольшей степени в области малых замедлений, поэтому рассмотрим эту область подробнее. Предположим, что для трехмерных систем безразмерный параметр $V_3 = k [ab (n_1^2 - n_0^2)]^{1/2} \ll 1$. В этом случае пространственный спектр РМ свободного пространства узкий ($\kappa \ll kn_f$) и все соотношения могут быть упрощены. В частности, во всех разложениях полей доминируют основные цилиндрические гармоники, причем внутри области Ω_1 поля можно считать постоянными [17, 23]. Используя эти свойства собственных мод, для трехмерных систем получим

$$\rho \approx \frac{n_0 - n_f}{n_0 + n_f} + \frac{w_3^2}{k^2 n_0 (n_0 + n_f)} \ln \left(\frac{kn_0}{w_3} \right), \quad (18)$$

где w_3 — модуль поперечного волнового числа волноводной моды E_{11} (индекс 3 обозначает размерность задачи). Для задачи о рассеянии на обрыве диэлектрической пластины аналогичные рассуждения при условии $V_2 = kb (n_1^2 - n_0^2)^{1/2} \ll 1$ дают другой результат

$$\rho \approx \frac{n_0 - n_f}{n_0 + n_f} + \frac{w_2^2}{k^2 n_0 (n_0 + n_f)}, \quad (19)$$

где w_2 — модуль поперечного волнового числа основной моды пластины (индекс 2 также соответствует размерности задачи). Приведенные выше соотношения показывают, что для обеих задач коэффициенты отражения ρ зависят от параметров волновода через поперечные волновые числа мод (т. е. в конечном счете через их фазовые скорости). При малых поперечных размерах волноводов ($b \rightarrow 0$) для рассматриваемых систем поперечные волновые числа ПМ существенно различны [11, 16, 17]: для трехмерных структур $w_3 b \sim \exp[-c_3/(k^2 ab)]$, где c_3 — некоторая константа, тогда как для двумерных задач $w_2 b \sim (kb)^2$. Как следует из этих оценок, при одинаковых размерах b поперечное волновое число суперэллиптического волновода существенно меньше поперечного волнового числа пластины (если $V_2 \ll 1$ и $V_3 \ll 1$), поэтому при малых b для волокна коэффициент отражения ПМ практически постоянен, т. е. $\rho \approx \rho_0$, тогда как для пластины величина $|\rho - \rho_0|$ пропорциональна b^2 (см. рис. 2–6). Численные данные показывают, что значения $|\rho|$ для рассматриваемых задач различны вплоть до значений $V_3 \sim 1$.

Из формул (18), (19) следует, что в области малых замедлений зависимость ρ от параметров задачи определяется в основном дисперсионными характеристиками ПМ, которые различны для двумерных и трехмерных задач. В свою очередь, эти характеристики можно связать с аналитическими свойствами собственных значений $s_m(\kappa)$ оператора рассеяния \hat{S} , который используется для построения РМ волновода [17, 18], поскольку поперечные волновые числа ПМ являются полюсами функций $s_m(\kappa)$ в комплексной плоскости κ . Анализ показывает, что связь между свойствами оператора рассеяния и решением задачи на самом деле является более глубокой. Например, аналитические свойства оператора \hat{S} определяют вид диаграммы направленности сферической (или цилиндрической) волны, а также пространственной волны внутри волновода (при $z < 0$ и $\mathbf{r} \in \Omega_1$), которые возбуждаются на обрыве.

Используя результаты работ [11, 17, 18], нетрудно показать, что в трехмерной задаче поле пространственной волны E_{sp} при $z \rightarrow -\infty$ спадает как $\exp(-ikn_0z)/|z \ln^2 z|$ (при $|z| \gg ka^2$). Логарифмический характер зависимости поля от координаты определяется логарифмической особенностью фазы собственного значения для основной ветви РМ: при $\kappa \rightarrow 0$ имеем $s_0(\kappa) \approx \ln(-\kappa/\kappa_s)/\ln(\kappa/\kappa_s)$. Отметим, что точка ветвления $\kappa = 0$ в комплексной плоскости поперечного волнового числа присутствует у подынтегральных функций в разложении (1) только для трехмерных систем. Для двумерной геометрии $\kappa = 0$ — регулярная точка, в этом случае поле пространственной волны убывает по другому закону: $E_{sp} \sim \exp(-ikn_0z)/|z|^{3/2}$.

Вывод основного соотношения (8) и оценок (18), (19) основан на предположении, что поля РМ волновода возбуждаются с малой амплитудой. Для того, чтобы обосновать это предположение, исследуем подробнее структуру поля \mathcal{E} . Для этого преобразуем ИУ первого рода к уравнению второго рода с малым ядром. Добавляя и вычитая из (6) спектральные разложения функций $\bar{\beta} \delta(\mathbf{r}_1 - \mathbf{r}_2)$ и $\bar{\gamma} \delta(\mathbf{r}_1 - \mathbf{r}_2)$ по полям собственных мод [20], после преобразований рядов получим

$$(\bar{\beta} + \bar{\gamma}) \mathcal{E} = 2\beta_s E_s(1) - (\beta_s - \bar{\beta}) E_s(1) \langle \mathcal{E}, E_s(1) \rangle - \sum_m \int_0^\infty (\beta_\kappa - \bar{\beta}) E_{m\kappa}(1) \langle \mathcal{E}, E_{m\kappa}(1) \rangle d\kappa - \sum_m \int_0^\infty (\gamma_\kappa - \bar{\gamma}) E_{m\kappa}(2) \langle \mathcal{E}, E_{m\kappa}(2) \rangle d\kappa. \quad (20)$$

Вспомогательные параметры $\bar{\beta}$ и $\bar{\gamma}$ можно подобрать так, чтобы ядро нового ИУ было мало. В этом случае (20) можно решить методом итераций. В частности, можно выбрать $\bar{\beta} = \beta_s$ и $\bar{\gamma} = \sum_m \int_0^\infty \gamma_\kappa Q_{m\kappa}^2 d\kappa$.

После двух итераций получаем приближенное выражение для поля на обрыве:

$$\mathcal{E} = \frac{2\beta_s E_s(1)}{\beta_s + \bar{\gamma}} - \frac{2\beta_s}{(\beta_s + \bar{\gamma})^2} \sum_m \int_0^\infty (\gamma_\kappa - \bar{\gamma}) Q_{m\kappa} E_{m\kappa}(2) d\kappa. \quad (21)$$

Из (21) нетрудно вывести формулу для ρ , которая в этом приближении совпадает с (8). Следующие итерации получаются аналогичным образом. Отметим, что этот подход можно применить также к волноводам, которые могут направлять несколько ПМ, а также к системам с более сложной геометрией.

Оценка второго слагаемого в правой части формулы (21) показывает, что при условии $n_1 \approx n_0$ оно существенно меньше первого [20]. Почти для всех точек поперечной плоскости $z = 0$ слагаемые, полученные при следующих итерациях, будут еще меньше, поэтому в разложении (1) доминирующим является первое слагаемое, которое пропорционально $E_s(1)$. Эта оценка неточна в очень малой окрестности контура Λ в плоскости $z = 0$, т. е. на диэлектрическом ребре, где соответствующий ряд итераций сходится медленно. Указанный факт объясняется тем, что на диэлектрических ребрах вторая производная скалярного поля E может иметь слабую (логарифмическую) сингулярность [20, 24]. Однако влияние этого эффекта на величину ρ незначительно, поскольку сингулярность весьма слабая, а площадь области, где она может сказываться, мала. Анализ подобных эффектов изложен в работах [20, 24]. Различные тесты показали, что при условии $|n_1/n_0 - 1| < 0,03$ относительная погрешность расчета ρ по формуле (8) по порядку величины равна 10^{-3} (или меньше). Для скалярных задач магнитные формулировки вариационного принципа [20] менее точны (обычно в несколько раз), поэтому выше при расчетах использовался только электрический функционал (8).

ЗАКЛЮЧЕНИЕ

Рассмотренный в работе пример показывает, что в общем случае результаты анализа двумерной открытой волноводной системы * нельзя переносить на трехмерную систему. В основном сильное различие дифракционных характеристик проявляется в области малых замедлений. Это различие можно объяснить установленной выше связью между коэффициентом отражения мод и их фазовой скоростью. Оценки показывают, что указанные выводы сохраняются и для систем с более сложной геометрией, например, в случае, когда на торцах волноводов находятся просветляющиеся слоистые покрытия [4, 5]. Заметим также, что с учетом поляризационных эффектов различие характеристик двумерных и трехмерных задач может быть больше [25], чем это следует из скалярной теории.

Вопрос о соответствии решений задач дифракции в открытых двумерных и трехмерных системах был рассмотрен в работе на примере только одной частной геометрии волноводной системы. Тем не менее, основные выводы статьи являются достаточно общими. В качестве примера можно указать большой класс задач о рассеянии ПМ на стыке двух диэлектрических волноводов разного поперечного сечения. Большинство дифракционных характеристик этих систем определяются фазовыми скоростями ПМ, распространяющимися в таких волноводах [23], поэтому, как и в рассмотренных в работе задачах, в области малых замедлений характеристики двумерных и трехмерных систем отличаются. Граница, ниже которой это различие существенно, определяется соотношением $k[\Omega(n_1^2 - n_0^2)]^{1/2} \sim 1$, где Ω — минимальная из площадей поперечного сечения волноводов. Отметим, что для этих задач остаются справедливыми все выводы, касающиеся аналитических свойств подынтегральных функций в спектральных разложениях, которые, в свою очередь, определяют различия в структуре рассеянных полей [17, 23].

Работа выполнена при частичной поддержке РФФИ (грант № 97–02–16334).

ЛИТЕРАТУРА

1. Ikegami T. // IEEE J. Quantum Electron. 1972. V. QE8, № 6. P. 470.
2. Rozzi T. E., In't Veld G. H. // IEEE Trans. Microwave Theory. Tech. 1980. V. MTT28, № 2. P. 61.
3. Buus J. // IEEE J. Quantum Electron. 1981. V. QE17, № 12. P. 2256.
4. Vassallo C. // J. Opt. Soc. Am. 1988. V. A5, № 11. P. 1918.
5. Kendall P. C., Roberts D. A., Robson P. N., Adams M. J., Robertson M. J. // IEEE Proc., pt. J. 1993. V. 140, № 1. P. 49.
6. Маненков А. Б. // Изв. вуз. Радиофизика. 1982. Т. 25, № 12. С. 1484.
7. Uzunoglu N. K., Capsalis C. N., Tigelis I. // J. Opt. Soc. Am. 1987. V. A4, № 11. P. 2150.
8. Manenkov A. B. // Optical Quantum Electron. 1991. V. 23, № 1. P. 81.
9. Tigelis I., Uzunoglu N. K., Capsalis C. N. // J. Electr. Waves Appl. 1991. V. 5, № 12. P. 447.
10. Sewell P., Reed M., Benson T. M., Kendall P. C. // IEEE J. Quantum Electron. 1997. V. QE33, № 12. P. 2311.
11. Маненков А. Б. // Изв. вуз. Радиофизика. 1997. Т. 40, № 8. С. 1004.
12. Jamid H. A., Al-Bader S. J. // IEEE Photonics Technol. Lett. 1997. V. 9, № 2. P. 220.
13. Левин Л. Теория волноводов. Методы решения волноводных задач. — М.: Радио и связь, 1981. — С. 291.
14. Швингер Ю. // Зарубежная радиоэлектроника. 1970. №3. С. 6.
15. Хенл Х., Мауэ А., Вестпфаль К. Теория дифракции. — М.: Мир, 1964. — С. 213.

* Для сравнения укажем, что в закрытых металлических волноводах зависимость решений от размерности обычно менее значительна.

16. Рожнев А. Г., Маненков А. Б. // Радиотехника и электроника. 1997. Т. 42, № 7. С. 785.
17. Маненков А. Б. // Изв. вуз. Радиофизика. 1981. Т. 24, № 1. С. 84.
18. Manenkov A. B. // Optical Quantum Electron. 1991. V. 23, № 5. P. 621.
19. Вайнштейн Л. А., Маненков А. Б. // Лекции по электронике СВЧ и радиофизике. Кн. 1. — Саратов: СГУ, 1986. — С. 141.
20. Маненков А. Б., Тигелис И. Г. // Изв. вуз. Радиофизика. 1999. Т. 42, № 1. С. 73.
21. Клеев А. И., Маненков А. Б., Рожнев А. Г. // Радиотехника и электроника. 1993. Т. 38, № 11. С. 1938.
22. Гайер Д. Лекции по теории аппроксимации в комплексной области. — М.: Мир, 1986. С. 82.
23. Маненков А. Б. // Изв. вуз. Радиофизика. 1982. Т. 25, № 11. С. 1329.
24. Васильев А. Д., Маненков А. Б. // Изв. вуз. Радиофизика. 1987. Т. 30, № 3. С. 405.
25. Reed V., Sewell P., Benson T. M., Kendall P. C. // Microwave Optical Technol. Lett. 1997. V. 15, № 4. P. 196.

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

Поступила в редакцию
26 июня 1999 г.

EFFECT OF DIMENSIONALITY ON THE SOLUTIONS OF OPEN-WAVEGUIDE DIFFRACTION PROBLEMS

A. B. Manenkov

We compare reflection characteristics of the guided (surface) modes from an abruptly terminated planar dielectric waveguide and an optical fiber with a prolate cross section. Analytical expressions for the reflection coefficients of the dominant modes are derived for the small-slowing (near-cutoff) region. The analysis is made in the scalar approximation using integral-equation method and variational principles. It is shown that the difference in features of the considered problems is closely related to analytical properties of eigenmode systems of the waveguides.

УДК 621.396.969

ОЦЕНКА РАДИОЛОКАЦИОННОЙ ЗАМЕТНОСТИ РУПОРНОЙ АНТЕННЫ С НЕЛИНЕЙНЫМ КОНТАКТОМ ПРИ НАБЛЮДЕНИИ "НЕЛИНЕЙНОЙ" РАДИОЛОКАЦИОННОЙ СТАНЦИЕЙ

В. В. Беляев, А. Т. Маюнов, Г. Д. Михайлов, С. Н. Разиньков

Теоретически и экспериментально исследовано рассеяние электромагнитных волн рупорной антенной с нелинейным контактом $Al-Al_2O_3-Al$ на третьей гармонике облучающего поля в дециметровом диапазоне длин волн. Оценено влияние нелинейности контакта на радиолокационную заметность рупорной антенны при наблюдении "нелинейной" радиолокационной станцией.

Антенные системы современных радиоэлектронных устройств содержат большое число нелинейных электрических контактов "металл—диэлектрик—металл". К ним относятся микротрещины и заклепки в рефлекторах апертурных антенн, стыки в местах соединения элементов антенно-фидерных трактов и т. д. [1]. Под воздействием электромагнитного излучения нелинейные контакты генерируют вторичные поля на гармониках и комбинационных частотах облучающих волн [1, 2]. Как показывают экспериментальные данные [1, 3, 4], наибольшую мощность имеет излучение на третьей гармонике и комбинационных частотах третьего порядка. Избирательный прием рассеянного поля на гармониках и комбинационных частотах позволяет обнаруживать радиоэлектронные средства "нелинейными" радиолокационными станциями (РЛС) в условиях фоновых помех, а также при использовании противорадиолокационной маскировки [3, 5]. Устранение нелинейного вторичного излучения на этапе конструирования радиоэлектронных средств актуально для снижения радиолокационной заметности.

В предлагаемой работе исследовано рассеяние поля рупорной антенной с нелинейным контактом (на примере контакта $Al-Al_2O_3-Al$) на третьей гармонике облучающей волны и оценена дальность обнаружения данной антенны "нелинейной" радиолокационной станцией, использующей в качестве зондирующего сигнала импульсный сигнал с частотой 3 ГГц и скважностью импульсов $Q = 1000$.

В качестве показателя радиолокационной заметности антенны выбиралась усредненная по направлениям рассеяния, лежащим в переднем полупространстве антенны, "нелинейная" эффективная площадь рассеяния (ЭПР) [1] на третьей гармонике падающей волны $\sigma_{зср}$.

Рупорная антенна (рис. 1) выполнена в виде соединения пирамидального рупора и прямоугольного волновода, имеющего активную нагрузку. Рупор и волновод являются идеально проводящими. Размеры поперечного сечения волновода равны $a = 0,72 \lambda_0$ и $b = 0,34 \lambda_0$, где λ_0 — рабочая длина волны антенны; характерные размеры рупора в H - и E -плоскостях равны соответственно $a_p = 1,48 \lambda_0$ и $b_p = 0,94 \lambda_0$, длина грани рупора $L = 2,10 \lambda_0$. Спротивление активной нагрузки 50 Ом. Нагрузка антенны является согласованной на частоте облучения и на ее третьей гармонике. Нелинейный элемент представляет собой тонкий стержень, содержащий в центре контакт $Al-Al_2O_3-Al$ со статической вольт-амперной характеристикой $i = 0,0183 V + 0,003 V^3$, где i измеряется в миллиамперах, а приложенная к контакту разность потенциалов V — в вольтах [4]. Стержень размещается в E -плоскости прямоугольного волновода на расстоянии $l_0 = \lambda_0/4$ от активной нагрузки и на расстоянии λ_0 от границы соединения волновода с рупором. Контакт выполнен в виде сквозной щели шириной 50 ангстрем, лежащей в плоскости поперечного сечения стержня и расположенной в его центре.

Стержень в E -плоскости прямоугольного волновода является моделью детекторной секции диапазона сверхвысоких частот (СВЧ) для анализа рассеяния электромагнитных волн. Контакт металл—диэлектрик—металл можно рассматривать как модель микротрещины в антенной системе.

При расчете вторичного поля антенны будем полагать, что раскрыв рупора облучается плоской монохроматической волной. Длина волны облучающего поля λ равна рабочей длине волны рупорной антенны λ_0 , поэтому коэффициент отражения падающей волны от раскрыва рупора равен нулю. Поле в рупорной антенне с нелинейным контактом складывается из облучающего и вторичного поля с частотой падающей волны, а также вторичных полей с частотами гармоник. Будем полагать, что напряженность падающего поля E_0 удовлетворяет условию квазилинейного приближения [1], поэтому электромагнитные волны на гармониках не оказывают взаимного влияния друг на друга.

Расчет вторичного поля осуществляется путем решения системы интегральных уравнений, связывающих электрические и магнитные поля на поверхности стержня S :

$$\frac{1}{2} (\mathbf{n} \times \mathbf{E}_p(\mathbf{r})) + jk_p W_0 \int_S ((\mathbf{n} \times \mathbf{G}_p(\mathbf{r}, \mathbf{r}')) \times \mathbf{r}') \mathbf{H}_p(\mathbf{r}') d\mathbf{S}' + j \int_S ((\mathbf{n} \times \text{rot } \mathbf{G}_p(\mathbf{r}, \mathbf{r}')) \times (\mathbf{n} \times \mathbf{E}_p(\mathbf{r}')) d\mathbf{S}' = (\mathbf{n} \times \mathbf{E}_0(\mathbf{r})) \delta_{1p}, \quad (1)$$

где \mathbf{E}_p и \mathbf{H}_p — электрическая и магнитная компоненты рассеянного поля на p -й гармонике падающей волны, W_0 — волновое сопротивление свободного пространства, \mathbf{n} — внутренняя нормаль к поверхности S , δ_{1p} — символ Кронекера, $k_p = 2\pi p/\lambda$ — волновое число p -й гармоники в свободном пространстве, $\mathbf{G}_p(\mathbf{r}, \mathbf{r}')$ — соленоидальная функция Грина прямоугольного волновода для длины волны, соответствующей p -й гармонике падающего излучения в свободном пространстве. В явном виде $\mathbf{G}_p(\mathbf{r}, \mathbf{r}')$ получена в [6].

Уравнения (1) получены из граничных условий на поверхности стержня и внутренней поверхности волновода с использованием векторной теоремы Грина.

Для тонкого стержня можно считать, что тангенциальные составляющие полей $\mathbf{E}_p(\mathbf{r})$ и $\mathbf{H}_p(\mathbf{r})$ не зависят от φ (см. рис. 1) и содержат только одну компоненту: $\mathbf{E}_{p\tau}(\mathbf{r}) = E_{p\tau}(y)\mathbf{e}_y$, $\mathbf{H}_{p\tau}(\mathbf{r}) = H_{p\tau}(y)\mathbf{e}_y$, \mathbf{e}_y — орт декартовой системы координат на рис. 1. При этом $E_{p\tau}(y) = U_p \delta(y - b/2)$, где U_p — напряжение на нелинейном контакте, создаваемое p -й гармоникой, $\delta(y - b/2)$ — дельта-функция Дирака. Плотность тока на поверхности стержня определяется выражением $J_p(y) = H_{p\tau}(y)$.

Таким образом, в приближении тонкого стержня система интегральных уравнений (1) приводится к виду

$$\int_0^1 G_p(\xi, \xi') J_p(\xi') d\xi' = j \frac{2\pi}{W_0} \left[U_p \delta(\xi) - \frac{V}{b} \left(1 - \exp \left\{ -2j \sqrt{k_1^2 - (\pi/a)^2} l_0 \right\} \right) \delta_{1p} \right], \quad (2)$$

где $\xi = y/b$.

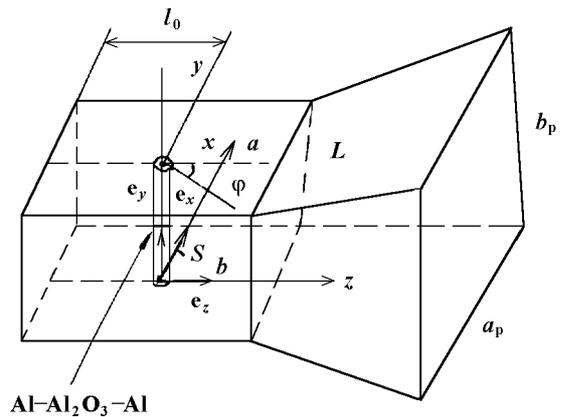


Рис. 1. Рупорная антенна с контактом Al-Al₂O₃-Al

В результате решения (2) методом Галеркина найдено амплитудно-фазовое распределение плотности тока на поверхности стержня на третьей гармонике облучающей волны. По известному распределению плотности тока на поверхности S путем решения уравнения Гельмгольца вычислено вторичное поле на раскрыве антенны. При расчетах учитывалось преобразование плоской волны в сферическую при переходе из волновода в рупор. При заданном распределении поля на раскрыве рупора методом Пойгенса—Кирхгофа найдено поле рассеяния в дальней зоне на третьей гармонике. Результаты расчетов вторичного поля позволили найти “нелинейную”ЭПР антенны.

Вычисленная угловая зависимость $\sigma_3(\theta)$ “нелинейной”ЭПР рупорной антенны с контактом Al—Al₂O₃—Al на третьей гармонике при плотности потока энергии облучающего поля на раскрыве рупора 0,1 Вт/м² представлена на рис. 2 (сплошная линия). Угол θ отсчитывается от внешней нормали к раскрыву антенны.

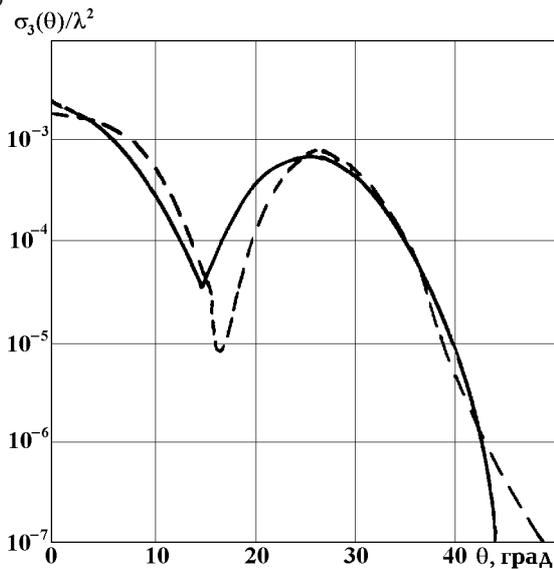


Рис. 2. Зависимость “нелинейной”ЭПР рупорной антенны с контактом Al—Al₂O₃—Al на третьей гармонике облучающего поля от угла прихода облучающей волны (сплошная линия — теория; пунктирная — эксперимент)

Для проверки точности вычисления “нелинейной”ЭПР проведено сравнение теоретической зависимости $\sigma_3(\theta)$ с результатами измерения. Для измерения ЭПР рупорной антенны на третьей гармонике облучающего поля использовалась амплифазометрическая радиолокационная измерительная установка [7]. Для устранения паразитных излучений от непреднамеренных нелинейных контактов в месте соединения волновода с рупором наносилась клеевая пленка толщиной 0,5 мм. Нелинейный контакт создавался путем действия паров воды на соприкасающиеся поверхности распиленного алюминиевого штыря в течение 1 минуты.

Экспериментальная угловая зависимость $\sigma_3(\theta)$ представлена на рис. 2 (пунктирная линия) вместе с теоретическими данными. Максимальное различие вычисленных и измеренных значений “нелинейной”ЭПР (при $\theta = 15^\circ$) не превышает систематической погрешности измерений (5,5 дБ).

Средняя “нелинейная”ЭПР σ_{3cp} , найденная в результате усреднения теоретической угловой зависимости $\sigma_3(\theta)$ по направлениям рассеяния, лежащим в

переднем полупространстве антенны, равна $7,9 \cdot 10^{-5}$ м².

На основании полученного результата из уравнения нелинейной радиолокации для импульсного сигнала [3] определена максимальная дальность R_{max} обнаружения рассматриваемой рупорной антенны по вторичному излучению поля на третьей гармонике облучающего поля “нелинейной”РЛС. При этом использовались следующие параметры типовой “нелинейной”РЛС: пиковая мощность излучаемого радиолокационного сигнала 50 кВт, чувствительность приемного устройства РЛС 10^{-12} Вт, коэффициенты усиления передающей и приемной антенн 30 дБ. Затухание радиолокационных сигналов диапазона СВЧ при распространении вблизи земной поверхности 0,02 дБ/м [3]. Расчет R_{max} осуществлялся с учетом того, что применение импульсного сигнала приводит к возрастанию дальности обнаружения по сравнению с использованием монохроматического излучения не менее чем в $Q^{(p-1)/(2p+2)}$ раз.

Расчеты показали, что исследуемая в данной работе рупорная антенна может быть обнаружена РЛС с указанными выше характеристиками на расстоянии 190 м.

Наличие N одинаковых контактов Al—Al₂O₃—Al в рупорной антенне вследствие аддитивности мощ-

ности вторичного излучения на гармониках [8] может приводить к возрастанию ее средней “нелинейной” ЭПР в N раз по сравнению с антенной, содержащей один нелинейный контакт. При этом максимальная дальность обнаружения антенны средствами “нелинейной” радиолокации возрастает в $N^{1/(2p+2)}$ раз [8] (в частности при приеме информационного сигнала на третьей гармонике — в $N^{1/8}$ раз).

Устранение рассеяния поля на гармониках обеспечивается путем нанесения в местах точечных металлических контактов изолирующих (клеевых) пленок толщиной $50 \div 100$ ангстрем, лакокрасочных и анодированных покрытий [5], а также за счет включения в тракты антенных систем полосовых СВЧ фильтров [7]. Изолирующие покрытия повышают потенциальный барьер диэлектрика, препятствуя проникновению электронов через контакт. Вследствие этого вторичное поле на гармониках снижается. Нанесение клеевой пленки на исследуемый контакт позволяет уменьшить “нелинейную” ЭПР антенны на третьей гармонике облучающей волны примерно на $40 \div 45$ дБ·мВт. Включение в тракт рупорной антенны волноводного фильтра, работа которого основана на поглощении гармоник композитным гиромангнитным материалом, нанесенным на стенки волновода, снижает паразитное излучение поля на 40 дБ·мВт.

Таким образом, на основе математического моделирования рассеяния поля рупорной антенной с нелинейным контактом и экспериментальной проверки результатов оценено влияние вторичного излучения на гармониках облучающей волны на радиолокационную заметность рупорной антенны в дециметровом диапазоне длин волн.

ЛИТЕРАТУРА

1. Штейншлегер В. Б. // УФН. 1984. Т. 142, № 1. С. 131.
2. Горбачев А. А., Ларцов С. В., Тараканов С. П., Чигин Е. П. // Радиотехника и электроника. 1997. Т. 42, № 7. С. 782.
3. Harger R. // IEEE Trans. 1976. V. AES-12, № 2. P. 230.
4. Хайга В. // ТИИЭР. 1975. Т. 63. С. 67.
5. Вернигоров Н. С., Харин В. Б. // Радиотехника и электроника. 1997. Т. 42, № 12. С. 1447.
6. Бравер И. М., Гарб Х. Л. // Радиотехника и электроника. 1982. Т. 27, № 2. С. 253.
7. Маюнов А. Т., Михайлов Г. Д., Разиньков С. Н. // Измерительная техника. 1997. № 12. С. 35.
8. Васенкова Л. В., Горбачев А. А. // Изв. вуз. Радиофизика. 1995. Т. 38, № 7. С. 743.

5 Центральный
научно-исследовательский
институт МО РФ, г. Воронеж,
Россия

Поступила в редакцию
2 августа 1999 г.

THE ESTIMATE OF DETECTABILITY OF THE HORN ANTENNA WITH THE NONLINEAR JUNCTION BY “NONLINEAR”RADAR

V. V. Belyaev, A. T. Mayunov, G. D. Mikhailov, and S. N. Razin'kov

We study theoretically and experimentally the scattering of decimeter electromagnetic waves at the third harmonic by horn antenna with the nonlinear junction $Al-Al_2O_3-Al$. The effect of the junction nonlinearity on the detectability of the horn antenna by “nonlinear” radar.

УДК 621.396:621.391.82

ПРИНЦИП ПОСТРОЕНИЯ ДВУХКАНАЛЬНОГО ПРОСТРАНСТВЕННОГО СЕПАРАТОРА СИГНАЛА И ПОМЕХИ С ПРЕДВАРИТЕЛЬНЫМ ОРТОНОРМИРОВАНИЕМ ВХОДНЫХ ПРОЦЕССОВ

С. А. Метелев, Ю. В. Шишкин

В работе предложен и исследован алгоритм работы адаптивного пространственного сепаратора (разделителя) сигнала и помехи с ортогонализацией и нормированием входных процессов. Показано, что весовые коэффициенты второго адаптивного сумматора могут быть выражены через весовые коэффициенты, формирующие первый выход сепаратора. Аналитическими и численными методами исследована помехоустойчивость такого устройства. Полученные выводы подтверждены тестовым экспериментом с реальными радиосигналами.

Для борьбы со случайными и преднамеренными сигналоподобными помехами, не имеющими априорно известных различий с полезным сигналом, для адаптивных антенных систем (ААС) предложены алгоритмы сепарации, или разделения, сигнально-помеховой смеси на составные компоненты [1–3]. Алгоритмы [1, 3] основаны на разложении входных процессов по собственным векторам корреляционной матрицы и эффективно работают лишь при значительных различиях в мощности парциальных колебаний.

Большинство реальных радиоканалов распространения помехи и сигнала являются каналами с переменными параметрами, что вызывает флуктуации уровней принимаемых радиоволн. Это приводит к снижению качества пространственно-временной обработки сигнала (ПВОС) по алгоритмам [1, 3].

Оптимальный сепаратор [2] построен по критерию средней квадратичной ошибки (СКО) и реализует градиентный алгоритм. Такой сепаратор предназначен для сигналов с угловой манипуляцией (что позволяет достаточно просто сформировать эталонный сигнал), которые наиболее распространены в дискретных каналах радиосвязи. Эффективность разделения в таком сепараторе практически не зависит от типа помех (шумоподобные или сигналоподобные) и их мощности.

Процедура обработки в сепараторе [2] начинается с ортогонализации входных процессов и их нормирования в схемах автоматической регулировки усиления (АРУ). Ортонормированные колебания далее поступают на два блока, работающие по критерию СКО — алгоритмы Уидроу—Хоффа с двумя ортогональными эталонными сигналами, формируемыми путем ограничения выходных колебаний. Алгоритмы Уидроу—Хоффа обеспечивают асимптотически оптимальную обработку сигналов, минимизирующую среднеквадратичную ошибку и максимизирующую отношения сигнал/(помеха+шум) на одном выходе и отношение помеха/(сигнал+шум) на другом выходе. При этом второй эталонный сигнал формируется в процессе компенсации первого эталонного сигнала из ограниченного второго выхода сепаратора при помощи ортогонализатора эталонных колебаний. Наличие ортогонализатора в цепи второго эталона несколько искажает эталонное колебание, что снижает помехоустойчивость второго выхода сепаратора.

Вместе с тем, наличие предпроцессора ортонормировки перед оптимальным алгоритмом минимизации СКО (МСКО) обеспечивает возможность упростить построение второго адаптивного процессора и исключить второй блок Уидроу—Хоффа. Принципу построения такого упрощенного варианта сепаратора и исследованию эффективности его работы посвящена данная статья.

1. ФУНКЦИОНАЛЬНАЯ СВЯЗЬ ВЕСОВЫХ КОЭФФИЦИЕНТОВ СЕПАРАТОРА

Рассмотрим приведенную на рис. 1 функциональную схему сепаратора, которая в обобщенном виде изображает алгоритм разделения двух колебаний в двух адаптивных сумматорах. Перед адаптивными сумматорами стоит ортогонализатор и схемы АРУ, обеспечивающие ортонормирование антенных колебаний. Вычисление весовых коэффициентов осуществляется в адаптивном процессоре по критерию СКО, причем в данном случае не имеет значения, какой конкретно алгоритм используется — градиентный или непосредственного обращения корреляционной матрицы.

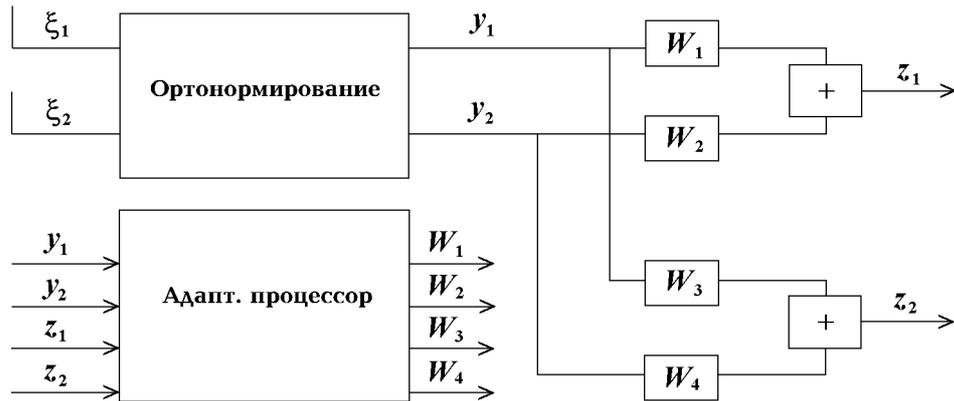


Рис. 1. Блок-схема двухканального сепаратора с ортогонализацией и нормированием входных процессов

Предположим, что комплексные амплитуды антенных колебаний $\xi_1(t)$ и $\xi_2(t)$ представляют из себя сумму колебаний сигнала $s(t) = ac_s(t)$ и помехи $p(t) = bc_p(t)$, различающихся в ветвях разнесения по амплитудам (вследствие неидентичности диаграмм направленности приемных антенн) и по фазам (вследствие пространственного разнесения):

$$\begin{aligned}\xi_1(t) &= a_1 c_s(t) e^{i\beta_s} + b_1 c_p(t) e^{i\beta_p} + n_1(t) e^{i\eta_1(t)}, \\ \xi_2(t) &= a_2 c_s(t) e^{i(\beta_s + \Delta\beta_s)} + b_2 c_p(t) e^{i(\beta_p + \Delta\beta_p)} + n_2(t) e^{i\eta_2(t)}.\end{aligned}\quad (1)$$

Здесь $c_s(t)$, $c_p(t)$ — модулирующие функции источников соответственно полезного сигнала и помехи, β_s и β_p — фазы колебаний сигнала и помехи на первой антенне, $\Delta\beta_s$ и $\Delta\beta_p$ — разность фаз колебаний сигнала и помехи между второй и первой антеннами, n_1 , n_2 и η_1 , η_2 — соответственно амплитуды и фазы некоррелированных шумов в ветвях разнесения.

Процедура ортогонализации осуществляется линейным преобразованием антенных колебаний и имеет множество решений. Одна из возможностей такого преобразования, с помощью алгоритма квадратурного автокомпенсатора (или двухканального преобразования Грама—Шмидта), будет рассмотрена ниже при анализе помехоустойчивости сепаратора. Для анализа функциональных связей весовых коэффициентов сепаратора конкретный вид ортогонализующего преобразования значения не имеет. Комплексные процессы $x_1(t)$ и $x_2(t)$ после ортогонализации по-прежнему являются суммой двух колебаний: полезного сигнала $s(t)$ и помехи $p(t)$ (влиянием тепловых шумов в каналах сначала пренебрежем):

$$\begin{aligned}x_1(t) &= s_1(t) e^{i\varphi_s} + p_1(t) e^{i\varphi_p}, \\ x_2(t) &= s_2(t) e^{i(\varphi_s + \Delta\varphi_s)} + p_2(t) e^{i(\varphi_p + \Delta\varphi_p)},\end{aligned}\quad (2)$$

где выделены начальные фазы φ_s , φ_p и разности фаз $\Delta\varphi_s$, $\Delta\varphi_p$ между колебаниями сигнала и помехи в ветвях разнесения после ортогонализации,

$$s_{1,2}(t) = A_{s1,s2}c_s(t), \quad p_{1,2}(t) = A_{p1,p2}c_p(t),$$

$$\overline{c_{s,p}(t)c_{s,p}(t)^*} = 1, \quad \overline{c_{s,p}(t)c_{p,s}(t)^*} = 0.$$

Здесь горизонтальная черта сверху обозначает усреднение по времени, A_{s1} и A_{s2} — амплитуды полезного сигнала, A_{p1} и A_{p2} — амплитуды помехи после ортогонализатора. Сигнал и помеха предполагаются некоррелированными.

Ниже для краткости в обозначениях процессов будем опускать обозначение зависимости от времени t .

Ортонормированные процессы обладают следующими замечательными свойствами:

1) Из условия ортогональности $\overline{x_1x_2^*} = 0$ следует соотношение

$$\overline{s_1s_2^*}e^{-i\Delta\varphi_s} + \overline{p_1p_2^*}e^{-i\Delta\varphi_p} = 0, \quad (3)$$

которое выполняется при одновременном выполнении условий

$$\overline{s_1s_2^*} = \overline{p_1p_2^*}, \quad \text{или} \quad A_{s1}A_{s2} = A_{p1}A_{p2}, \quad (4)$$

и

$$\Delta\varphi_s = \Delta\varphi_p + \pi. \quad (5)$$

Из (4) видно, что отношение амплитуд сигнала и помехи в первой ветви разнесения после ортогонализации становится равно отношению амплитуд помехи и сигнала во второй ветви разнесения (инверсия амплитуд или мощностей), а разности фаз полезного и помехового колебаний отличаются на π (колебания противофазны) согласно (5).

2) После ортогонализации осуществляется нормирование колебаний:

$$y_1 = x_1 / \sqrt{\overline{x_1x_1^*}}, \quad y_2 = x_2 / \sqrt{\overline{x_2x_2^*}},$$

$$\sqrt{\overline{x_1x_1^*}} = \sqrt{A_{s1}^2 + A_{p1}^2}, \quad \sqrt{\overline{x_2x_2^*}} = \sqrt{A_{s2}^2 + A_{p2}^2}.$$

Нормированные амплитуды сигнала S_1 , S_2 и помехи P_1 , P_2 в соответствующих ветвях разнесения с учетом (4) оказываются связанными между собой:

$$S_1 = A_{s1} / \sqrt{A_{s1}^2 + A_{p1}^2} = 1 / \sqrt{1 + A_{p1}^2/A_{s1}^2} = 1 / \sqrt{1 + A_{s2}^2/A_{p2}^2} = A_{p2} / \sqrt{A_{s2}^2 + A_{p2}^2} = P_2, \quad (6)$$

т. е.

$$S_1 = P_2, \quad S_2 = P_1. \quad (7)$$

Амплитуды (или мощности) сигнала и помехи после ортонормирования становятся не только инвертированы, но и перекрестно равны, при этом противофазность нормированных суммарных колебаний y_1 и y_2 сохраняется. Эти два фундаментальных свойства ортонормированных колебаний определяют связь между весовыми коэффициентами адаптивных сумматоров. Действительно, первый выход сепаратора представляет из себя сумму

$$z_1 = W_1y_1 + W_2y_2 = W_1S_1c_s e^{i\varphi_s} + W_1P_1c_p e^{i\varphi_p} + W_2S_2c_s e^{i(\varphi_s + \Delta\varphi_s)} + W_2P_2c_p e^{i(\varphi_p + \Delta\varphi_p)}. \quad (8)$$

Минимальное значение среднеквадратичной ошибки после настройки весовых коэффициентов в этом случае равно нулю и достигается при полной компенсации помехи и когерентном сложении сигнала.

Предположим, что в адаптивном процессоре найдены веса W_1 и W_2 , обеспечивающие компенсацию помехи на первом выходе сепаратора:

$$W_1 P_1 c_p e^{i\varphi_p} + W_2 P_2 c_p e^{i(\varphi_p + \Delta\varphi_p)} = e^{i\varphi_p} (W_1 P_1 c_p + W_2 P_2 c_p e^{i\Delta\varphi_p}) = 0. \quad (9)$$

Из (9) следует

$$|W_2| / |W_1| = P_1 / P_2, \quad (10)$$

$$\Delta\phi_{21} = \pi - \Delta\varphi_p, \quad (11)$$

где $\Delta\phi_{21} = \arg W_2 - \arg W_1 = \phi_2 - \phi_1$ — разность аргументов весовых коэффициентов.

Проводя аналогичные выкладки для второго выхода $z_2 = W_3 y_1 + W_4 y_2$, можно показать, что для компенсации полезного сигнала на этом выходе необходимо обеспечить выполнение условий

$$|W_4| / |W_3| = S_1 / S_2, \quad (12)$$

$$\Delta\phi_{43} = \pi - \Delta\varphi_s, \quad (13)$$

откуда с использованием (5), (7), (10), (11) следует

$$|W_4| / |W_3| = |W_1| / |W_2|, \quad (14)$$

$$\Delta\phi_{43} = \pi + \Delta\phi_{21}. \quad (15)$$

При таких весовых коэффициентах на втором выходе сепаратора происходит минимизация сигнала и максимизация помехи вследствие их противофазности.

Таким образом, для того чтобы осуществить разделение двух компонент в двух ортонормированных колебаниях, весовые коэффициенты в сепараторе должны быть связаны соотношением (14), (15). Иначе говоря, достаточно вычислить в адаптивном процессоре весовые коэффициенты W_1 , W_2 для первого адаптивного сумматора, после чего по ним можно определить весовые коэффициенты для формирования второго выхода W_3 , W_4 по следующему правилу:

$$W_3 = W_2^*, \quad W_4 = -W_1^*. \quad (16)$$

2. ПОМЕХОУСТОЙЧИВОСТЬ СЕПАРАТОРА

На первом выходе сепаратора происходит вычитание колебаний помехи и когерентное сложение сигнала при взвешенном суммировании с коэффициентами (10), (11):

$$\begin{aligned} z_1 &= W_1 S_1 c_s e^{i\varphi_s} + W_2 S_2 c_s e^{i(\varphi_s + \Delta\varphi_s)} = |W_1| e^{i\phi_1} S_1 c_s e^{i\varphi_s} + P_1 / P_2 |W_1| e^{i(\phi_1 + \Delta\phi_{21})} S_2 c_s e^{i(\varphi_s + \Delta\varphi_s)} = \\ &= |W_1| e^{i\phi_1} S_1 c_s e^{i\varphi_s} \left(1 + S_2^2 / S_1^2 e^{i(\pi - \Delta\varphi_p + \Delta\varphi_s)} \right) = |W_1| e^{i\phi_1} S_1 c_s \left(1 + S_2^2 / S_1^2 \right) e^{i\varphi_s}. \end{aligned} \quad (17)$$

При построении ортогонализатора по алгоритму квадратурного автокомпенсатора легко проанализировать три предельных случая.

Если сигнал в антенных колебаниях много меньше помехи, то он выделяется на выходе ортогонализатора и $S_2 \ll S_1$, $|W_2| \ll |W_1| \approx 1$, т. е. $|z_1| \approx S_1$.

Если на выходе антенн сигнал много больше помехи, то $S_2 \gg S_1$, $|W_1| \ll |W_2| \approx 1$, а

$$z_1 = |W_1| S_2^2 / S_1 e^{i\phi_1} c_s e^{i\varphi_s} = |W_2| e^{i\phi_1} S_2 c_s e^{i\varphi_s} = |W_2| e^{i\phi_2} S_2 c_s e^{i(\varphi_s + \Delta\varphi_s)}, \quad (18)$$

откуда следует, что $|z_1| \approx S_2$, т. е. на выходе выделяется сигнал из второй ветви разнесения.

Если сигналы в ветвях разнесения равны, то на выходе сепаратора происходит когерентное сложение сигналов равной амплитуды с двух ветвей разнесения с равными весовыми коэффициентами, причем

$$|W_2| \approx |W_1| \approx 1/\sqrt{2} \quad S_2 \approx S_1 \approx 1/\sqrt{2}, \quad (19)$$

и после взвешивания и суммирования выходной сигнал снова будет примерно равен единице. Таким образом обеспечивается и минимум среднеквадратичной ошибки, и максимум отношения сигнал/помеха.

Известно, что оптимальная схема обработки входных процессов в отсутствие помехи обеспечивает повышение отношения сигнал/(помеха+шум) (ОСПШ) на 3 дБ (т. е. увеличение ОСШП вдвое). Из проведенного рассмотрения видно, что амплитуда сигнала на выходе не возрастает. Поэтому причиной повышения ОСШШ является уменьшение шума на выходе оптимального компенсатора по сравнению с его входом. Для доказательства этого нетривиального факта необходимо рассмотреть работу схемы при наличии некоррелированных (тепловых) шумов в ветвях разнесения.

Рассмотрим ААС, состоящую из двух идентичных пространственно разнесенных антенн с изотропными диаграммами направленности. В этом случае амплитуды сигнала и помехи в ветвях разнесения будут равны, и антенные колебания (1) могут быть представлены в виде

$$\begin{aligned} \xi_1 &= ac_s e^{i\beta_s} + bc_p e^{i\beta_p} + n_1 e^{i\eta_1}, \\ \xi_2 &= ac_s e^{i(\beta_s + \Delta\beta_s)} + bc_p e^{i(\beta_p + \Delta\beta_p)} + n_2 e^{i\eta_2}. \end{aligned} \quad (20)$$

Двухканальный ортогонализатор Грама—Шмидта имеет один весовой коэффициент

$$W_o = \frac{\overline{\xi_1 \xi_2^*}}{\xi_2 \xi_2^*} = \frac{a^2 e^{-i\Delta\beta_s} + b^2 e^{-i\Delta\beta_p}}{a^2 + b^2 + |n_2|^2} \quad (21)$$

и обеспечивает преобразование колебания только в первой ветви разнесения:

$$x_1 = \xi_1 - W_o \xi_2 = \frac{b^2 (1 - e^{i\Delta\beta}) + |n_2|^2}{a^2 + b^2 + |n_2|^2} ac_s e^{i\beta_s} + \quad (22)$$

$$+ \frac{a^2 (1 - e^{-i\Delta\beta}) + |n_2|^2}{a^2 + b^2 + |n_2|^2} bc_p e^{i\beta_p} - \frac{a^2 e^{-i\Delta\beta_s} + b^2 e^{-i\Delta\beta_p}}{a^2 + b^2 + |n_2|^2} n_2 e^{i\eta_2} + n_1 e^{i\eta_1},$$

$$x_2 = \xi_2 = ac_s e^{i(\beta_s + \Delta\beta_s)} + bc_p e^{i(\beta_p + \Delta\beta_p)} + n_2 e^{i\eta_2}, \quad (23)$$

где $\Delta\beta = \Delta\beta_s - \Delta\beta_p$ — разность разностей фаз антенных колебаний сигнала и помехи, $|n_{1,2}|^2 = n_{1,2} n_{1,2}^*$ — мощность шумов в соответствующей ветви разнесения.

После нормирования колебания принимают вид

$$y_1 = \frac{b^2 (1 - e^{i\Delta\beta}) + |n_2|^2}{Q} ac_s e^{i\beta_s} + \frac{a^2 (1 - e^{-i\Delta\beta}) + |n_2|^2}{Q} bc_p e^{i\beta_p} - \quad (24)$$

$$- \frac{a^2 e^{-i\Delta\beta_s} + b^2 e^{-i\Delta\beta_p}}{Q} n_2 e^{i\eta_2} + \frac{a^2 + b^2 + |n_2|^2}{Q} n_1 e^{i\eta_1},$$

$$y_2 = \frac{1}{\sqrt{a^2 + b^2 + |n_2|^2}} \left\{ ac_s e^{i(\beta_s + \Delta\beta_s)} + bc_p e^{i(\beta_p + \Delta\beta_p)} + n_2 e^{i\eta_2} \right\}, \quad (25)$$

$$Q = \left\{ \left(2a^2b^2(a^2 + b^2 + 2|n_2|^2)(1 - \cos(\Delta\beta)) + (a^2 + b^2)|n_2|^4 + \right. \right. \\ \left. \left. + (a^4 + b^4 + 2a^2b^2 \cos(\Delta\beta))|n_2|^2 + (a^2 + b^2 + |n_2|^2)^2|n_1|^2 \right\}^{1/2}. \quad (26)$$

Оптимальные весовые коэффициенты для каждого из выходов сепаратора находятся из уравнения Винера—Хопфа $\mathbf{w} = \mathbf{R}_{yy}^{-1} \mathbf{r}_{yd}$, где \mathbf{R}_{yy}^{-1} — обратная корреляционная матрица, \mathbf{r}_{yd} — вектор корреляции процессов с эталонным колебанием. Для сепаратора этот вектор и весовые коэффициенты являются матрицами, поэтому уравнение принимает вид

$$\mathbf{W} = \mathbf{R}_{yy}^{-1} \mathbf{R}_{yd}.$$

Для ортонормированных колебаний обратная корреляционная матрица становится единичной, и элементы матрицы весовых коэффициентов определяются корреляцией этих колебаний с эталонными колебаниями:

$$\mathbf{W} = \mathbf{R}_{yd}.$$

В качестве последних можно использовать сигналы источников $c_{s,p}$:

$$W_{1,2} = \overline{y_{1,2}^* c_s}, \quad (27)$$

$$W_{3,4} = \overline{y_{1,2}^* c_p}. \quad (28)$$

Для первой ветви разнесения получим

$$W_1 = \frac{a [b^2 (1 - e^{-i\Delta\beta}) + |n_2|^2]}{Q} e^{-i\beta_s}, \quad (29)$$

$$W_2 = \frac{a}{\sqrt{a^2 + b^2 + |n_2|^2}} e^{-i(\beta_s + \Delta\beta_s)}. \quad (30)$$

Оптимальные весовые коэффициенты для второй ветви разнесения имеют вид

$$W_3 = \frac{b [a^2 (1 - e^{i\Delta\beta}) + |n_2|^2]}{Q} e^{-i\beta_p}, \quad (31)$$

$$W_4 = \frac{b}{\sqrt{a^2 + b^2 + |n_2|^2}} e^{-i(\beta_p + \Delta\beta_p)}. \quad (32)$$

Перед тем, как перейти к анализу помехоустойчивости сепаратора, проверим соответствие точных решений (29)–(32) для весовых коэффициентов критерию СКО при наличии некоррелированных шумов функциональной связи (14), (15), найденной в приближении отсутствия шумов.

Детальный анализ выражений (29)–(32), проведенный численными методами, показал:

1) соотношения (14), (15) выполняются точно при любом уровне сигнала и помехи лишь при отличии разности фаз антенных колебаний сигнала и помехи $\Delta\beta = 180^\circ$, т. е. при разности азимутов сигнала и помехи 90° (для пространственного разнеса антенн $d = \lambda/2$, где λ — длина волны излучения);

2) при сближении азимутов сигнала и помехи соотношения (14), (15) перестают быть точными, но вплоть до разности азимутов 30° возникающая погрешность не превышает 1 % при приемлемых уровнях сигнала (не менее 10 дБ по сравнению с некоррелированным шумом приемных трактов);

3) заметные отличия соотношений (14), (15) от точных решений (более чем на 10 %) возникают при сближении азимутов менее чем на 10° .

Таким образом, функциональная связь весовых коэффициентов (14), (15), найденная в пренебрежении шумами, имеет место в достаточно широком диапазоне изменения параметров сигнально-помеховой обстановки и нарушается лишь в тех условиях, когда и полный алгоритм оптимального сепаратора [2] снижает свою помехоустойчивость (при малых угловых разнесах и уровнях сигнала и помехи).

Вначале проанализируем помехоустойчивость сепаратора (под которой здесь будем понимать ОСПШ) для некоторых частных случаев.

2.1. СИЛЬНЫЙ ПОЛЕЗНЫЙ СИГНАЛ В ОТСУТСТВИЕ ПОМЕХИ

В случае одного полезного сигнала и в предположении слабых шумов ($a^2 \gg |n_{1,2}|^2$)

$$\begin{aligned} y_1 &= 1 / \sqrt{|n_1|^2 + |n_2|^2} \left\{ n_1 e^{i\eta_1} - n_2 e^{i\eta_2 - i\Delta\beta_s} + ac_s e^{i\beta_s} |n_2|^2 / a^2 \right\}, \\ y_2 &= 1 / \sqrt{a^2 + |n_2|^2} \left\{ ac_s e^{i(\beta_s + \Delta\beta_s)} + n_2 e^{i\eta_2} \right\}. \end{aligned} \quad (33)$$

Видно, что в первой ветви выделяются в основном шумовые колебания со слабой сигнальной составляющей, причем попавший в первую ветвь шум второго канала испытал фазовый сдвиг и, естественно, остался коррелированным с шумом второй ветви.

Весовые коэффициенты для первого сумматора имеют вид

$$W_1 = \frac{a |n_2|^2 / a^2}{\sqrt{|n_1|^2 + |n_2|^2}} e^{-i\beta_s}, \quad W_2 = \frac{a}{\sqrt{a^2 + |n_2|^2}} e^{-i(\beta_s + \Delta\beta_s)}. \quad (34)$$

Из (34) следует, что весовой коэффициент в первой ветви настолько меньше второго весового коэффициента, насколько мощность сигнала больше мощности шума.

В сигнале z_1 на выходе первого сумматора шумовая компонента, коррелированная с шумом в сигнале z_2 , будет иметь амплитуду, равную половине амплитуды шума на выходе второго сумматора, и вычтется из него. В результате в суммарном колебании останутся две шумовые составляющие из первой и второй ветви с одинаковыми, но уменьшенными вдвое по сравнению с шумом на входе амплитудами за счет нормировки в первом канале и компенсации коррелированного шума второй ветви на адаптивном сумматоре:

$$z_1 = c_s + \frac{n_1}{2a} e^{i(\eta_1 - \beta_s)} + \frac{n_2}{2a} e^{i(\eta_2 - (\beta_s + \Delta\beta_s))}.$$

Далее для упрощения предположим, что мощности тепловых шумов в ветвях разнесения равны, и вычислим ОСПШ на выходе:

$$\frac{1}{z_1 z_1^*} = 1 + \frac{|n_1|^2 + |n_2|^2}{4a^2} = 1 + \frac{|n|^2}{2a^2},$$

где первое слагаемое является мощностью сигнала, $|n|^2/(2a^2)$ — мощность шумов; при этом ОСПШ равно отношению первого слагаемого ко второму:

$$\text{ОСПШ}_z = \frac{1}{|n|^2/2a^2} = 2 \frac{a^2}{|n|^2} = 2 \text{ОСПШ}_\xi,$$

т. е. выходное ОСПШ_z вдвое выше входного ОСПШ_ξ.

2.2. СИЛЬНЫЕ СИГНАЛ И ПОМЕХА (ПРЕНЕБРЕЖИМО МАЛЫЕ ШУМЫ)

При $\{a^2, b^2\} \gg |n_{1,2}|^2$ из (25)–(27), (30), (31) получим

$$y_1 = \frac{b^2(1 - e^{i\Delta\beta})}{Q} a c_s e^{i\beta_s} + \frac{a^2(1 - e^{-i\Delta\beta})}{Q} b c_p e^{i\beta_p}, \quad (35)$$

$$y_2 = \frac{1}{\sqrt{a^2 + b^2}} \left\{ a c_s e^{i(\beta_s + \Delta\beta_s)} + b c_p e^{i(\beta_p + \Delta\beta_p)} \right\}, \quad (36)$$

$$Q = \sqrt{2a^2b^2(a^2 + b^2)(1 - \cos(\Delta\beta))}, \quad (37)$$

$$W_1 = \frac{ab^2(1 - e^{-i\Delta\beta})}{Q} e^{-i\beta_s}, \quad W_2 = \frac{a}{\sqrt{a^2 + b^2}} e^{-i(\beta_s + \Delta\beta_s)}, \quad (38)$$

и колебание на первом выходе сепаратора является суммой двух когерентных колебаний сигнала:

$$z_1 = W_1 y_1 + W_2 y_2 = \frac{b^2}{a^2 + b^2} c_s + \frac{a^2}{a^2 + b^2} c_s = c_s.$$

Помеховое колебание на первом выходе полностью скомпенсировано, на втором выходе сепаратора ситуация противоположная: происходит когерентное сложение помехи и компенсация сигнала.

2.3. ОБЩИЙ СЛУЧАЙ

При произвольных отношениях уровней сигнала и помехи к некоррелированным шумам в трактах приема необходимо найти мощности сигнала (С), помехи (П) и результирующего шума (Ш) после суммирования колебаний (24), (25) с весовыми коэффициентами (29), (30). После этого вычисляется ОСПШ:

$$\text{ОСПШ} = \frac{С}{П + Ш}.$$

При вычислении ОСПШ предположим равенство мощностей некоррелированных шумов в антенных колебаниях ($|n_1|^2 = |n_2|^2 = |n|^2$) и отнормируем входные амплитуды сигнала и помехи на среднеквадратичную амплитуду шумов. Мощность сигнала на выходе при этом находится достаточно просто и может быть выражена через веса:

$$С = (|W_1|^2 + |W_2|^2)^2 = \frac{4a^4}{D^2} [b^2(1 - \cos(\Delta\beta)) + 1]^2. \quad (39)$$

Выражения для некомпенсированных остатков помехи и некоррелированных шумов принимают вид

$$П = \frac{2a^2b^2}{D^2} (1 + \cos(\Delta\beta)), \quad (40)$$

$$Ш = \frac{2a^2}{D^2} \left\{ b^2(b^2 + 1)(1 - \cos(\Delta\beta)) + 1 \right\}, \quad (41)$$

$$D = 2a^2b^2(1 - \cos(\Delta\beta)) + 2(a^2 + b^2) + 1. \quad (42)$$

Помеха и шум на выходе зависят от трех параметров: уровней сигнала a и помехи b на входе, отнормированных на уровень шумов, и разности углов прихода, определяющих разность $\Delta\beta$ разностей фаз антенных колебаний сигнала и помехи. На рис. 2 приведены зависимости $\Pi(b)$, $\text{Ш}(b)$, $\text{ОСПШ}(b)$ при фиксированном уровне сигнала на входе $a = 10$ (ОСПШ на входе составляло 100, или 20 дБ), ОСПШ на входе с ростом уровня помехи падало от 100 до 0,01 (эта зависимость приведена на правом графике). Угловой разнос между сигналом и помехой на этих графиках являлся параметром, амплитуда помехи на входе менялась от 0 до 100.

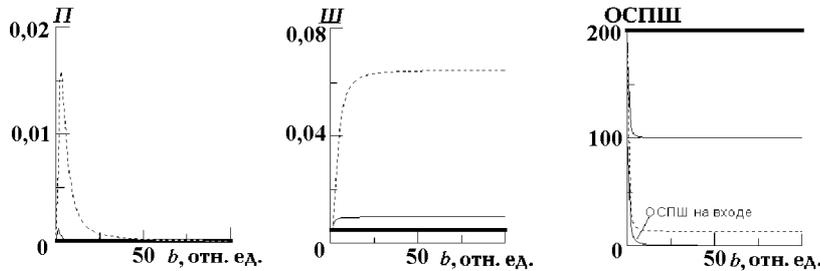


Рис. 2. Зависимости мощности нескомпенсированных остатков помехи (слева), шумов (в центре) и ОСПШ (справа) на выходе сепаратора от амплитуды помехи на входе при трех величинах углового разноса сигнала и помехи: 10° (пунктирная линия), 30° (тонкая линия) и 90° (жирная линия)

Из этих зависимостей следует: 1) в отсутствие помехи ($b = 0$) на выходе сепаратора ОСПШ равно 200, т. е. достигается выигрыш 3 дБ; 2) с появлением помехи ОСПШ на выходе сепаратора сначала падает так же, как и входное ОСПШ — скомпенсировать слабую помеху не позволяют некоррелированные шумы; по мере увеличения b уровень помехи на входе превысит уровень шумов, после чего алгоритм начинает на нее реагировать и происходит почти полная компенсация помехи; 3) выигрыш в 3 дБ в ОСПШ при этом достигается только в случае ортогональных углов прихода сигнала и помехи. Сближение азимутов приводит к резкому возрастанию шумов и нескомпенсированных остатков помехи, при этом уровень сигнала меняется слабо и остается близким к 1 (график $C(b)$ здесь не приведен из-за своей тривиальности).

Помехоустойчивость второго выхода сепаратора для второго колебания (помехи), естественно, ведет себя точно таким же образом вследствие полной симметрии задачи.

Теперь рассчитаем помехоустойчивость второго выхода, получаемого с использованием функциональной связи (16), и сравним его с точным решением по критерию МСКО. Для этого будем считать, что помеха выделяется на первом выходе сепаратора, сигнал — на втором выходе. При суммировании колебаний y_1, y_2 с весовыми коэффициентами (16) мощность помехи на втором выходе сепаратора равна нулю для любых уровней сигнала и помехи по отношению к уровню шумов. Мощности сигнала и шумов имеют вид

$$C = \frac{2a^2b^2}{D} [1 - \cos(\Delta\beta)], \tag{43}$$

$$\text{Ш} = \frac{2b^2}{D}, \tag{44}$$

$$\text{ОСПШ} = a^2 [1 - \cos(\Delta\beta)]. \tag{45}$$

Поведение этих зависимостей от уровня помехи на входе при фиксированной амплитуде входного сигнала $a = 10$ показано на рис. 3.

Основной вывод, который можно сделать из анализа зависимостей на рис. 2, заключается в том, что второй выход сепаратора, построенного по упрощенной схеме с использованием соотношений (16), реализует критерий максимума подавления помехи (МПП), а не МСКО. ОСПШ алгоритма МПП не

зависит от уровня помехи, выделяемой на первом выходе сепаратора, и даже в отсутствие помехи не происходит когерентного сложения сигнала на втором выходе и выигрыша в 3 дБ нет. Исключением является случай ортогональных углов прихода сигнала и помехи, когда алгоритмы МСКО и МПП полностью совпадают по помехоустойчивости. Поэтому данный алгоритм МПП является субоптимальным алгоритмом.

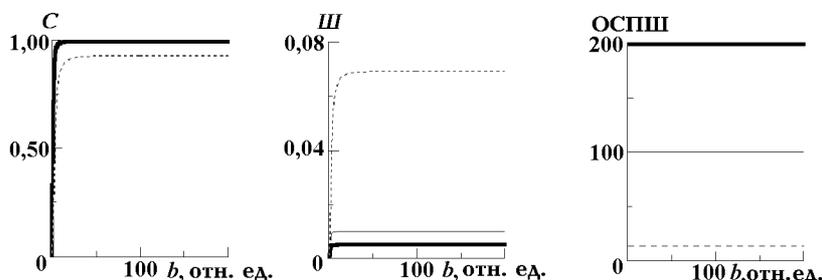


Рис. 3. Зависимости мощности сигнала (слева), шумов (в центре) и ОСПШ (справа) на втором выходе упрощенного алгоритма сепаратора от амплитуды помехи на входе при тех же величинах углового разнеса сигнала и помехи, что и на рис. 2

Область параметров, в которой алгоритм МПП проигрывает алгоритму МСКО по величине ОСПШ, относительно невелика, поэтому данный алгоритм по своим характеристикам близок к оптимальному. Кроме того, необходимо учитывать, что при слабой помехе ($b \leq a$) в алгоритме с формированием эталонного колебания в петле обратной связи [2] полезный сигнал будет выделяться в первом канале, оптимальном по критерию СКО.

3. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Проведенный в предыдущем разделе анализ выполнен при бесконечно большом времени усреднения в корреляторах (или при бесконечно узкой полосе эквивалентного фильтра низкой частоты (ФНЧ)). Такое усреднение при нахождении элементов корреляционной матрицы и вектора корреляции с эталонными сигналами можно осуществлять лишь в стационарных каналах связи. Реальные каналы связи являются каналами с переменными параметрами, что резко ограничивает допустимое время усреднения. Задача анализа помехоустойчивости сепаратора с конечным временем усреднения сводится к нелинейным интегродифференциальным уравнениям, поэтому вследствие сложности аналитического рассмотрения ее решение искалось методом имитационного моделирования.

Для проверки рассматриваемого алгоритма сепаратора было проведено исследование его эффективности по сравнению с алгоритмом Уидроу—Хоффа с ортогонализатором (УХО) [2]. Новый алгоритм, как описано выше, представлял из себя тот же ортогонализатор и тот же блок Уидроу—Хоффа для первого выхода сепаратора, но не имел второго блока Уидроу—Хоффа для вычисления весовых коэффициентов второго выхода, которые находились с помощью выражений (16). Поскольку этот алгоритм имел только один блок Уидроу—Хоффа, для краткости будем его называть алгоритмом ОУХО.

Для имитационного моделирования использовалась стандартная схема генерации входных смесей из двух сигналов типа ЧТ с одинаковой скоростью передачи символов [2], формирующихся по случайному закону, и разным частотным разнесом в полосе приема, что обеспечивало некоррелированность сигнала и сигналоподобной помехи.

К смеси сигнала и помехи в ветвях разнесения добавлялись некоррелированные тепловые шумы с гауссовским законом распределения комплексных амплитуд.

Обработка сигнально-помеховой смеси происходила следующим образом: сначала включался алгоритм УХО, в середине сеанса вместо него подключался алгоритм ОУХО. В корреляторах использовались фильтры низкой частоты с конечной полосой пропускания.

Результаты исследований приведены на рис. 4. На рис. 4а показана динамика ОСПШ для первого колебания (выход 1), которое можно считать полезным сигналом, и для второго колебания помехи, выделяющегося на выходе 2, при равной мощности двух этих колебаний. Исследованы два случая: $S = P = 20$ дБ и $S = P = 40$ дБ. Время приведено в условных единицах: 10 отсчетов соответствуют 64 битам. В середине сеанса (40-й отсчет) вместо базового алгоритма УХО включается алгоритм ОУХО. Полоса пропускания фильтров ФНЧ корреляторов составляла в данном случае 5 Гц. Из сравнения ОСПШ для этих алгоритмов следует, что эффективность компенсации колебания на втором выходе (жирная линия) в алгоритме УХО чуть ниже ОСПШ на первом выходе (на 2 дБ для входных уровней $S = P = 40$ дБ), т. е. второй выход несколько проигрывает первому.

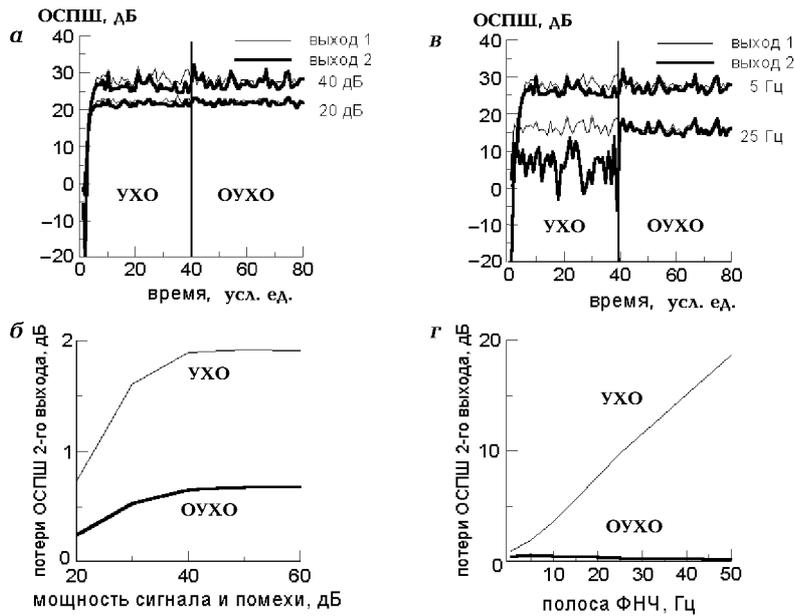


Рис. 4. а — поведение во времени ОСПШ на двух выходах алгоритмов УХО и ОУХО ($S = P = 20; 40$ дБ); б — зависимость потерь ОСПШ на втором выходе УХО и ОУХО от мощности входных колебаний; в — поведение во времени ОСПШ на двух выходах при полосе ФНЧ 5 Гц и 25 Гц, $S = P = 40$ дБ; г — зависимость от полосы ФНЧ потерь ОСПШ на втором выходе УХО и ОУХО, $S = P = 40$ дБ

Из графика видно, что алгоритм ОУХО работает лучше: неодинаковость двух выходов сепаратора значительно меньше (0,6 дБ), причем эта разница возрастает с увеличением мощности входных процессов. На рис. 4б построена зависимость потерь ОСПШ на втором выходе по сравнению с ОСПШ на первом выходе от мощности сигнала и помехи, из которой видно, что алгоритм ОУХО выигрывает у алгоритма УХО от 0,5 до 1,4 дБ.

Для установления причины этого эффекта были исследованы характеристики помехоустойчивости двух алгоритмов при изменении времени усреднения (или полосы ФНЧ) корреляторов. На рис. 4в аналогично графику на рис. 4а приведена временная зависимость ОСПШ для двух алгоритмов для полосы ФНЧ 5 Гц и 25 Гц. Из этого графика видно, что ухудшение помехоустойчивости второго выхода УХО возрастает с уменьшением времени усреднения, в то время как помехоустойчивость второго выхода ОУХО практически не меняется и соответствует помехоустойчивости первого выхода. Это хорошо видно из графика на рис. 4г, на котором приведены потери ОСПШ второго выхода (по отношению к ОСПШ первого выхода) для двух алгоритмов в зависимости от полосы ФНЧ. Анализируя эти результаты, можно сделать вывод о том, что единственной причиной ухудшения второго выхода УХО

является недоусредненность весового коэффициента в ортогонализаторе эталонных колебаний, в котором формируется второй эталон. В алгоритме ОУХО ортогонализатор отсутствует, что обеспечивает лучшую помехоустойчивость второго выходного канала ОУХО.

Таким образом, путем численных экспериментов показано, что алгоритм ОУХО эквивалентен УХО, но обладает лучшей помехоустойчивостью второго канала при конечном времени усреднения корреляторов.

4. ТЕСТИРОВАНИЕ АЛГОРИТМА ОУХО РЕАЛЬНЫМИ СИГНАЛАМИ

Имитационное моделирование алгоритмов в предыдущем разделе было проведено для стационарных условий сигнально-помеховой обстановки (СПО), что часто не соответствует реальным характеристикам каналов связи.

Для проверки эффективности нового алгоритма ОУХО в нестационарной СПО было проведено его тестирование на реальных сигналах, записанных в виде файлов данных в ходе проведенных нами ранее экспериментов.

В качестве тестового был выбран сеанс измерений, проведенный 17.09.96 в 11:37 MSK. В этом эксперименте прием сигналов осуществлялся на два приемника Р-160 («Вспышка») с общим опорным генератором при помощи двух разнесенных на 30 м вертикальных несимметричных вибраторов. Сигналы приемников на промежуточной частоте (с полосой 300 Гц) переносились на нулевую частоту с выделением квадратур и оцифровывались с частотой дискретизации 2 кГц 12-ти разрядным аналого-цифровым преобразователем.

Полезный сигнал представлял из себя сигнал типа ЧТ, который излучался локальным генератором с передающей антенны, расположенной на расстоянии 300 м от приемных антенн. Полезный сигнал излучался на частоте помехи с амплитудной модуляцией (радиовещательной станции) 15365 кГц, прошедшей через ионосферный канал и обладающей всеми свойствами нестационарного сигнала. Такая сигнально-помеховая смесь поступала на вход алгоритмов УХО и ОУХО. На одном выходе алгоритмов выделялся полезный сигнал, на другом — помеха. Узкие фильтры основной селекции приемников (с полосой 300 Гц) срезали боковые составляющие спектра амплитудно-модулированной (АМ) помехи, и она превращалась в почти монохроматическое колебание, частота которого несколько отличалась от частоты спектральных компонент полезного сигнала. Это позволяло вычислить ОСПШ полезного сигнала методом спектрального оценивания как сумму спектральных сигнальных компонент спектра мощности колебаний первого и второго выходов сепаратора к сумме всех оставшихся спектральных составляющих.

На рис. 5 приведена динамика ОСПШ для двух выходов сепаратора, работающего либо по алгоритму УХО (тонкие линии), либо по алгоритму ОУХО (жирные линии). На рис. 5 можно отметить следующие особенности:

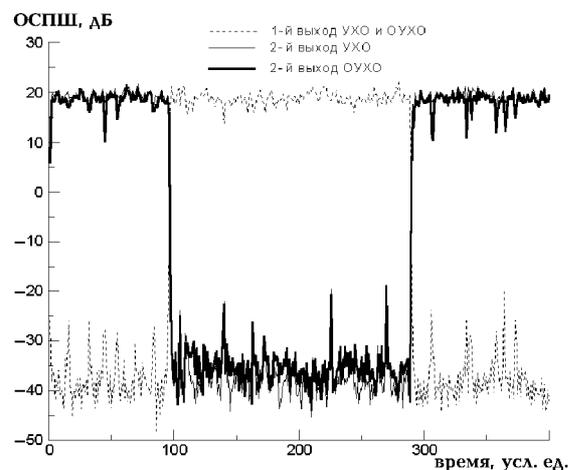


Рис. 5. Результаты обработки реальных сигналов (полезный сигнал ЧТ-200 и АМ помеха) алгоритмами УХО и ОУХО. Несущая частота $f = 15365$ кГц, измерения выполнены в 11:37 MSK, 17.09.96, общая длина реализации 5120 бит (128 секунд); 10 отсчетов по шкале времени соответствуют 128 битам

1) В процессе сеанса наблюдаются два перезахвата — перехода полезного сигнала из одного канала на другой, что является следствием резкой нестационарности СПО; перезахваты, естественно, происходят и в том, и в другом алгоритмах. В процессе перезахвата разность аргументов весовых коэффициентов первого блока Уидроу—Хоффа меняется на π , что, как следует из приведенного выше анализа, является следствием ортогональности входных процессов и противофазности сигнального и помехового колебаний. Этот факт можно использовать в качестве признака перезахвата для сохранения сигнала на одном и том же выходе путем коммутации выходных колебаний сепаратора.

2) Эффективность двух алгоритмов (УХО и ОУХО) близка: ОСПШ второго выхода ОУХО отличается от ОСПШ второго выхода УХО в среднем не более чем на 0,5 дБ.

Таким образом, результаты тестовых экспериментов доказали достаточно высокую эффективность алгоритма ОУХО (МПП) и в нестационарной СПО реальных каналов связи.

ВЫВОДЫ

Проведенное исследование показало, что при построении сепаратора с двумя ветвями разнесения предпроцессор ортонормировки обеспечивает возможность построения алгоритма разделения сигнала и помехи по упрощенной схеме. Функциональная связь между весовыми коэффициентами двух адаптивных сумматоров позволяет ограничить решение задачи поиском весовых коэффициентов для первой ветви разнесения.

При этом второй выход работает по критерию МПП, обладает свойствами, близкими к свойствам асимптотически оптимальных алгоритмов, построенных по критерию СКО, и требует вдвое меньшего числа корреляторов по сравнению с полным алгоритмом МСКО (УХО). Проведенные расчеты показали, что алгоритм МПП (ОУХО) проигрывает алгоритму МСКО при малых уровнях помехи в антенных колебаниях. Тестирование алгоритмов реальными сигналами также показало их близкие характеристики. Наблюдающиеся отличия находят свое объяснение в рамках проведенного в статье анализа.

Предложенный алгоритм может найти свое применение для компенсации помех в каналах радиосвязи, в частности, его целесообразно использовать в том случае, когда формирование эталонного сигнала для полезного сигнала затруднительно, а эталонное колебание для помехи выработать достаточно просто (например полезный сигнал — АМ колебание, а помеха — колебание с угловой модуляцией). В этом случае эталонное колебание для первого (помехового) выхода сепаратора формируется при помощи ограничения выходного колебания [2], а второй выход формируется по критерию МПП с компенсацией колебания, выделяемого на первом выходе.

Привлекательность алгоритма МПП обусловлена его вдвое меньшими вычислительными затратами по сравнению с алгоритмом МСКО при практически тех же характеристиках по помехоустойчивости. Это обеспечивает возможность пространственной обработки сигналов с более широкой частотной полосой при той же производительности вычислительного процессора в случае цифровой реализации алгоритма, либо позволяет вдвое сократить число радиоэлементов устройства сепарации сигнала и помех при построении сепаратора в аналоговом варианте.

ЛИТЕРАТУРА

1. Hackett C. M. // IEEE Trans. 1981. V. AES-17, № 2. P. 234.
2. Метелев С. А., Шишкин Ю. В. // Изв. вуз. Радиофизика. 1997. Т. 40, № 3. С. 378.
3. Ефименко В. С., Харисов В. Н. // Радиотехника. 1996. № 7. С. 87.

Государственное унитарное
предприятие НПП "Полет",
г. Н. Новгород, Россия

Поступила в редакцию
5 марта 1999 г.

**DESIGN SCHEME FOR A TWO-CHANNEL SPATIAL SEPARATOR OF SIGNAL AND INTERFERENCE USING
PRELIMINARY ORTHONORMALIZED INPUT PROCESSES***S. A. Metelev and Yu. V. Shishkin*

We propose and study an algorithm for the adaptive spatial separator of signal and interference in which the input processes are made orthogonal and normalized. It is shown that the weight coefficients of the second adaptive summator can be expressed in terms of the weight coefficients that form the first output of the separator. We analyze analytically and numerically the interference immunity of the device. The obtained conclusions are confirmed by the test using real radio signals.

УДК 538.574.6

ВЛИЯНИЕ ЛОКАЛЬНОЙ НЕОДНОРОДНОСТИ МАГНИТОСТАТИЧЕСКОГО ПОЛЯ НА СТАРТОВЫЕ И ВЫХОДНЫЕ ХАРАКТЕРИСТИКИ ГЕНЕРАТОРА ДИФРАКЦИОННОГО ИЗЛУЧЕНИЯ

А. И. Цвык, Г. С. Воробьев, А. В. Нестеренко, В. Н. Желтов

Теоретически и экспериментально изучаются возможности улучшения выходных характеристик нерелятивистского генератора дифракционного излучения (ГДИ) путём создания в однородном фокусирующем магнитном поле в области электронно-волнового взаимодействия локальной неоднородности магнитостатического поля (ЛНМП) при помощи ферромагнитного стержня. Получены соотношения для расчёта основных характеристик ЛНМП и показаны особенности изменения стартового тока ГДИ в зависимости от геометрических размеров ферромагнитного стержня и места его расположения относительно области взаимодействия. Результаты теории подтверждаются экспериментальными исследованиями, которые проводились в диапазоне частот 40–170 ГГц. Показано, что применение ЛНМП в ГДИ позволяет улучшить стартовые, энергетические и спектральные характеристики генератора. Делается вывод о перспективности создания на этом принципе низковольтных генераторов с улучшенными выходными характеристиками в коротковолновой области миллиметрового и в субмиллиметровом диапазонах длин волн.

ВВЕДЕНИЕ

Возможности применения пространственно неоднородных статических полей для повышения эффективности электронно-волнового взаимодействия и улучшения выходных характеристик СВЧ приборов исследовались во многих работах [1–4]. Наиболее полно в этом направлении изучены клинотроны, магнетроны и приборы на циклотронном резонансе. В частности, в клинотронах [1] улучшение выходных характеристик достигается применением наклонного относительно направления движения электронного потока (ЭП) фокусирующего магнитостатического поля, а в приборах М-типа и гиротронах — профилированием магнитостатического поля [2–4]. Предполагается, что таким же путем можно улучшить выходные характеристики генератора дифракционного излучения (ГДИ) [5]. Однако исследований в этом направлении (особенно экспериментальных) проведено недостаточно.

Как известно, ГДИ является источником высококогерентных электромагнитных колебаний миллиметрового и субмиллиметрового диапазонов длин волн. По сравнению с классическими генераторами эти приборы имеют ряд улучшенных выходных характеристик, что указывает на широкие возможности практического использования ГДИ в различных областях науки и техники. В то же время с укорочением длины волны генерации в нерелятивистских (низковольтных) ГДИ существенно уменьшается эффективная толщина ЭП и повышаются требования к технологии изготовления дифракционной решетки (в ГДИ эффективная толщина Δ ЭП оценивается соотношением $\Delta \approx \lambda\beta/\pi = l/\pi$, где λ — длина волны возбуждаемых колебаний, $\beta = v/c$ — отношение скорости электронов v к скорости света c , l — период дифракционной решетки). При генерации колебаний субмиллиметрового диапазона длин волн указанное обстоятельство вынуждает переходить к более высоким ускоряющим напряжениям и решать проблему формирования тонкого прямолинейного электронного потока [6]. В этой связи актуальными являются исследования возможности повышения эффективности процессов энергообмена в области взаимодействия с толстым (по сравнению с эффективной толщиной Δ) ЭП путем профилирования фокусирующего магнитостатического поля. Традиционно в ГДИ для фокусировки ЭП используется однородное фокусирующее магнитостатическое поле [7–9]. О возможности применения локальной неоднородности магнитостатического поля (ЛНМП) для улучшения выходных характеристик ГДИ

впервые сообщалось в экспериментальной работе [10], а перспективность такого подхода обосновывается теоретическими исследованиями [11, 12].

В данной работе исследуются характеристики ЛНМП, создаваемой в пространстве электронно-волнового взаимодействия ГДИ ферромагнитным стержнем, и влияние этой неоднородности на стартовый ток и другие характеристики генератора. Приводятся результаты экспериментального исследования выходных характеристик такого генератора в диапазоне частот $40 \div 170$ ГГц.

1. ТЕОРЕТИЧЕСКИЕ ИССЛЕДОВАНИЯ

Рассмотрим ГДИ, в котором локальная неоднородность магнитостатического поля в области взаимодействия создается ферромагнитным цилиндрическим стержнем, расположенным под дифракционной решеткой [10]. Теоретическая модель пространства взаимодействия такого генератора (условно назовем его ГДИ-МН) показана на рис. 1, где $2z_0$ — равновесная толщина электронного потока, a — прицельный параметр, r_e — амплитуда поперечных статических смещений (пульсаций) граничных электронов (в ГДИ обычно $a \leq r_e$), $a_z = z_0 + a + h_z$ — расстояние от оси ЭП до поверхности стержня, h_z — расстояние между дифракционной решеткой и поверхностью стержня, r — радиус стержня, μ — магнитная проницаемость стержня, y_m и z_m — координаты центра стержня, L — длина решетки, $0 < y_m < L$; $z_m = -(a_z + r)$.

Система координат выбирается таким образом, что ось y совпадает с осью ленточного электронного потока и с направлением индукции \mathbf{B}_0 однородного фокусирующего магнитостатического поля, ось z перпендикулярна поверхности отражательной дифракционной решетки, расположенной в плоскости $z = -(z_0 + a)$, ось x параллельна оси стержня и ламелям дифракционной решетки. Пунктирными вертикальными линиями здесь и на последующих рисунках условно показаны границы пятна электромагнитного поля на решетке, которое обычно образуется в открытом резонаторе ГДИ. Определим основные характеристики ЛНМП, создаваемой в области взаимодействия ферромагнитным стержнем, и проанализируем влияние этой неоднородности на возбуждение электромагнитных колебаний в ГДИ.

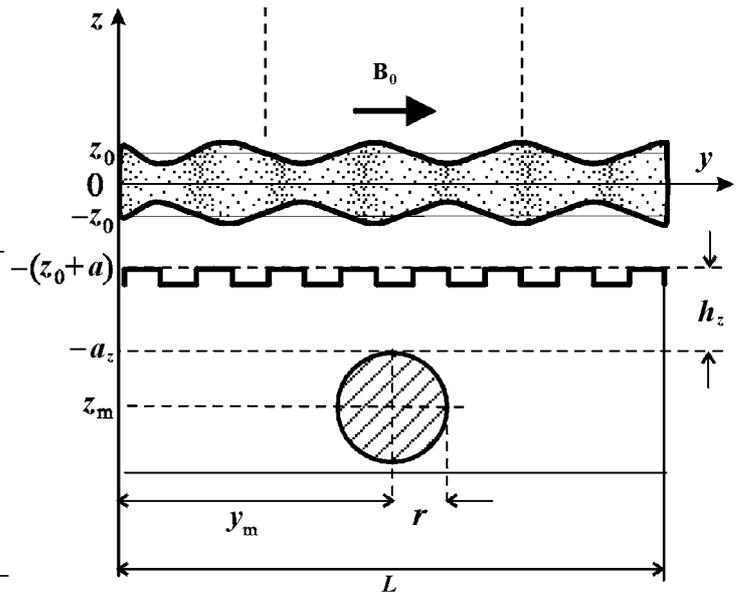


Рис. 1. Теоретическая модель пространства электронно-волнового взаимодействия ГДИ с локальной неоднородностью магнитостатического поля

1.1. Основные характеристики локальной неоднородности поля

Наличие ферромагнитного стержня в однородном статическом поле \mathbf{B}_0 приводит к “привисанию” силовых линий этого поля в сторону поверхности стержня и к образованию в области взаимодействия электронного потока с электромагнитным полем ЛНМП с продольной (B_y) и поперечной (B_z) составляющими:

$$B_y = 1 + \mu_r R^2 \frac{(Y - Y_m)^2 - (Z - Z_m)^2}{[(Y - Y_m)^2 + (Z - Z_m)^2]^2}, \quad (1)$$

$$B_z = 2\mu_r R^2 \frac{(Y - Y_m)(Z - Z_m)}{[(Y - Y_m)^2 + (Z - Z_m)^2]^2}. \quad (2)$$

В формулах (1), (2) значения B_y , B_z нормируются на B_0 , а радиус стержня и координаты — на длину решетки L ($R = r/L$, $Y = y/L$, $Z = z/L$, $Y_m = y_m/L$, $Z_m = z_m/L$, $\mu_r = (\mu - 1)/(\mu + 1)$).

Из соотношений (1), (2) следует, что ферромагнитный стержень создает в направлении движения ЭП характерную ЛНМП с минимальным значением продольной составляющей $(B_y)_{\min} = A_m$ в центре неоднородности и двумя симметричными максимумами $(B_y)_{\max} = A_1 = A_2$ на ее краях. В плоскости $Z = C_z$ в области взаимодействия эти значения B_y вычисляются по формулам

$$A_m = -\mu_r \frac{R^2}{(C_z - Z_m)^2}, \quad Y = Y_m, \quad (3)$$

$$A_{1,2} = |A_m|/8, \quad Y_{1,2} = Y_m \pm \sqrt{3}(C_z - Z_m), \quad (4)$$

где значения C_z лежат в интервале $-(Z_0 + a_L) \leq C_z \leq Z_0$. При этом на оси электронного потока $C_z = 0$, на верхней границе ЭП $C_z = z_0/L$, а на поверхности решетки $C_z = -(Z_0 + a_L) = -(z_0 + a)/L$. В области взаимодействия величина $D_y = 2R_y = Y_1 - Y_2$ определяет протяженность "пятна" ЛНМП, "радиус" которой

$$R_y = \sqrt{3}(C_z - Z_m). \quad (5)$$

В отличие от компоненты B_y поперечная составляющая B_z в направлении движения ЭП меняет знак, обращаясь в ноль в центре локальной неоднородности поля, и обладает двумя противоположными по знаку симметричными экстремумами $(B_z)_{\text{ext}} = B_{1,2}$, $|B_{1,2}| = B_m$, где

$$B_{1,2} = \pm 3\sqrt{3}|A_m|/8 \quad \text{при} \quad Y = Y_m \pm R_y/3. \quad (6)$$

На границах пятна ЛНМП $B_z = \sqrt{3}A_1$ (при $Y = Y_1$) и $B_z = -\sqrt{3}A_2$ (при $Y = Y_2$).

Воспользовавшись соотношениями (3), (5), получаем простую связь между амплитудой A_m , радиусом R_y пятна локальной неоднородности поля и радиусом R стержня:

$$A_m = -3\mu_r (R/R_y)^2, \quad R = R_y \sqrt{|A_m|/(3\mu_r)}. \quad (7)$$

В случае $C_z = -(Z_0 + a_L)$ из соотношений (3), (5) находим характеристики ЛНМП на поверхности дифракционной решетки:

$$A_m = -\frac{\mu_r}{(1 + \delta_z)^2}, \quad R_y = R \sqrt{3}(1 + \delta_z), \quad (8)$$

где $\delta_z = h_z/r$.

При условии $2Z_0 + a_L \ll R$ формулы (8) с достаточно высокой точностью определяют значения A_m и R_y в поперечном сечении электронного потока. Если это условие не выполняется (для малых значений δ_z и толстых электронных потоков, когда $2Z_0 \gg l/(L\pi)$), то параметры A_m и R_y

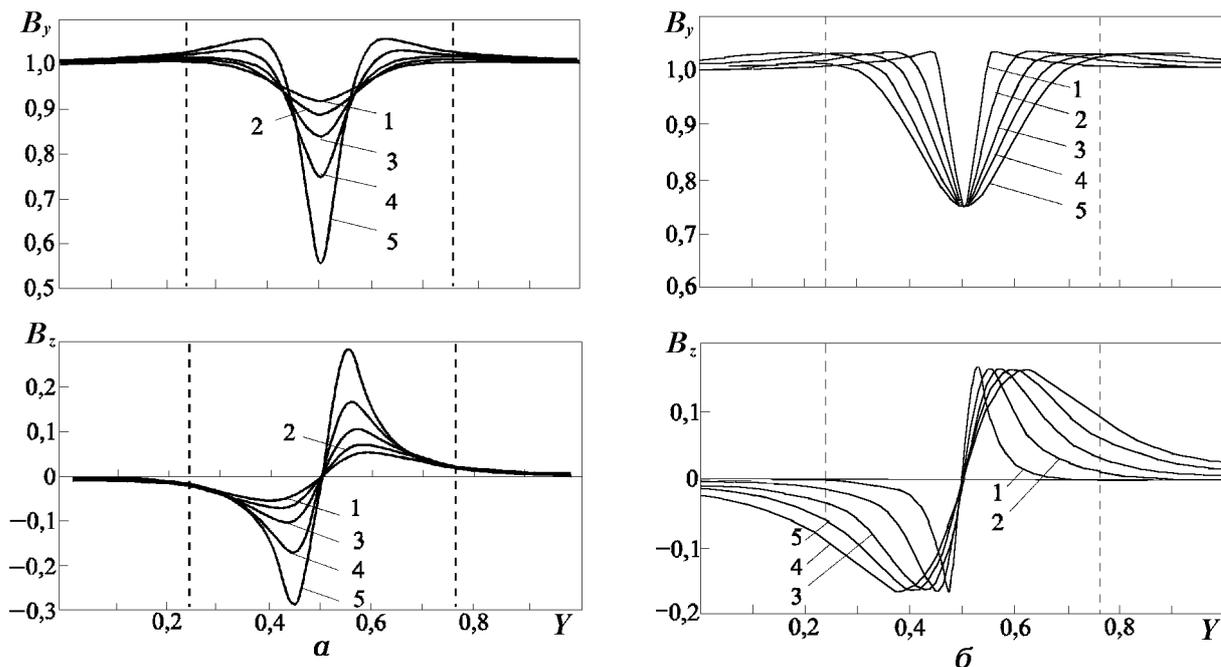


Рис. 2. Особенности изменения характеристик ЛНМП от параметров ферромагнитного стержня в пространстве взаимодействия ГДИ: а — $R = 0,05$; графики 1–5 соответствуют $\delta_z = 0,5; 1; 1,5; 2; 2,5$; б — $\delta_z = 1$; графики 1–5 соответствуют $R = 0,02; 0,04; 0,06; 0,08; 0,1$

необходимо вычислять по общим формулам (3)–(5), поскольку эти параметры в сечении ЭП, при $-z_0 < z < z_0$, могут существенно отличаться. Отметим, что для ферромагнетиков с проницаемостью $\mu \gg 1$ в соотношениях (1)–(8) можно положить $\mu_r \approx 1$. Полученные соотношения (3)–(8) позволяют при заданных геометрических параметрах $R, \delta_z, 2Z_0$ вычислить основные характеристики ЛНМП или по заданным характеристикам неоднородности магнитостатического поля (параметры A_m, B_m, R_y) находить радиус ферромагнитного стержня и место его расположения в пространстве взаимодействия ГДИ, что весьма важно при разработке ГДИ-МН.

На рис. 2 показаны рассчитанные по формулам (1), (2) зависимости составляющих B_y, B_z на поверхности решетки от координаты Y при $Y_m = 0,5$ и различных значениях параметров δ_z, R . Полученные графики отражают особенности формирования ЛНМП на поверхности решетки в двух случаях: при изменении расстояния δ_z между решеткой и поверхностью стержня заданного радиуса $R = 0,05$ (рис. 2а) и при изменении радиуса стержня R при постоянном параметре $\delta_z = 1$ (рис. 2б). Из этих зависимостей видно, что удаление стержня от поверхности решетки преимущественно уменьшает амплитуды A_m, B_m поля локальной неоднородности в области взаимодействия и в меньшей степени влияет на размеры пятна неоднородности. Во втором случае протяженность ЛНМП существенно возрастает с увеличением радиуса стержня. Если $Y_m \neq 0,5$, то рассмотренные зависимости сохраняются. Однако в этом случае нарушается симметрия расположения пятна ЛНМП относительно пятна СВЧ электромагнитного поля (пунктирные вертикальные линии на рис. 1–3). В результате при смещении стержня к концу решетки ($0,5 < Y < 1$) в область взаимодействия может попадать только спадающий участок распределения поля, а при смещении стержня к началу решетки ($0 < Y < 0,5$) — нарастающий.

Проведенный анализ позволяет выделить ряд свойств формирования ЛНМП ферромагнитным стержнем в области взаимодействия. Так, при наличии множества стержней с различными радиусами $R_i, i = 1, 2, 3, \dots$, при условии $C_z - (Z_m)_i = \text{const}$ каждый стержень создает в пространстве взаимодействия ЛНМП с одним и тем же радиусом $R_y = \text{const}$. Для двух произвольных стержней из этого множества выполняются равенства вида

$$\frac{R_1}{R_2} = \frac{(1 + \delta_{z2})}{(1 + \delta_{z1})}. \quad (9)$$

Из (9) следует, что для сохранения неизменного радиуса пятна ЛНМП при увеличении или уменьшении радиуса стержня необходимо соответственно уменьшить или увеличить расстояние между решеткой и поверхностью стержня на величину $\Delta h_z = |R_1 - R_2|$. Это свойство выполняется для стержней с концентрическими поверхностями, которые имеют одну и ту же поперечную координату центра Z_m и формируют в области взаимодействия локальные неоднородности поля с одинаковым радиусом R_y , но с различными параметрами A_m и B_m (согласно формулам (3), (8)). Смещением Z_m центра стержня и выбором его радиуса R определяются размеры локальной неоднородности (рис. 2б).

Следовательно, путем изменения геометрических параметров R , δ_z можно создавать в пространстве взаимодействия ГДИ различные профили ЛНМП и тем самым обеспечивать различные условия взаимодействия ЭП с электромагнитным полем резонатора.

1.2. Влияние ЛНМП на возбуждение ГДИ

Механизм и физические принципы преобразования энергии ЭП в электромагнитные колебания в ГДИ-МН впервые теоретически изучены в работах [11, 12], в которых при построении теории поле (1), (2) аппроксимируется экспонентами:

$$B_y = 1 + P_B \exp[-((Y - Y_m)/W_m)^2], \quad (10)$$

$$B_z = 2P_B \frac{(Y - Y_m)(Z - Z_m)}{W_m^2} \exp[-((Y - Y_m)/W_m)^2], \quad (11)$$

где P_B , W_m — значения амплитуды и размер пятна ЛНМП, которые являются неопределенными параметрами относительно геометрических размеров ферромагнитного стержня и расстояния от него до поверхности дифракционной решетки.

Стартовый ток I_s генератора определяется формулой

$$G_n = 1/S_1. \quad (12)$$

Здесь G_n — параметр эффективности взаимодействия, пропорциональный I_s ; S_1 — реальная часть крутизны колебательной характеристики генератора, которая находится из решения системы дифференциальных уравнений движения ЭП в СВЧ поле с гауссовой формой амплитудной огибающей в приближении слабого сигнала при наличии магнитостатических компонент (10), (11); в результате для S_1 получается достаточно сложное интегральное соотношение, которое можно проанализировать только численно.

Из теории следует, что повышение эффективности электронно-волнового взаимодействия в ГДИ-МН обусловлено двумя факторами: во-первых, изменением продольной амплитудной огибающей СВЧ поля на траектории электронов и, во-вторых, фазовой сортировкой частиц пучка с оседанием ускоренных электронов на поверхности дифракционной решетки. Первый фактор наиболее сильно проявляется в начальной (стартовой) стадии возбуждения колебания: существование ЛНМП приводит к тому, что электроны потока, приближаясь к дифракционной решетке, попадают в область с более сильным электромагнитным полем и увеличивают мощность взаимодействия, что и приводит к снижению стартового тока генератора. Второй фактор связан с удалением ускоренных электронов из области взаимодействия, что приводит к повышению электронного КПД генератора.

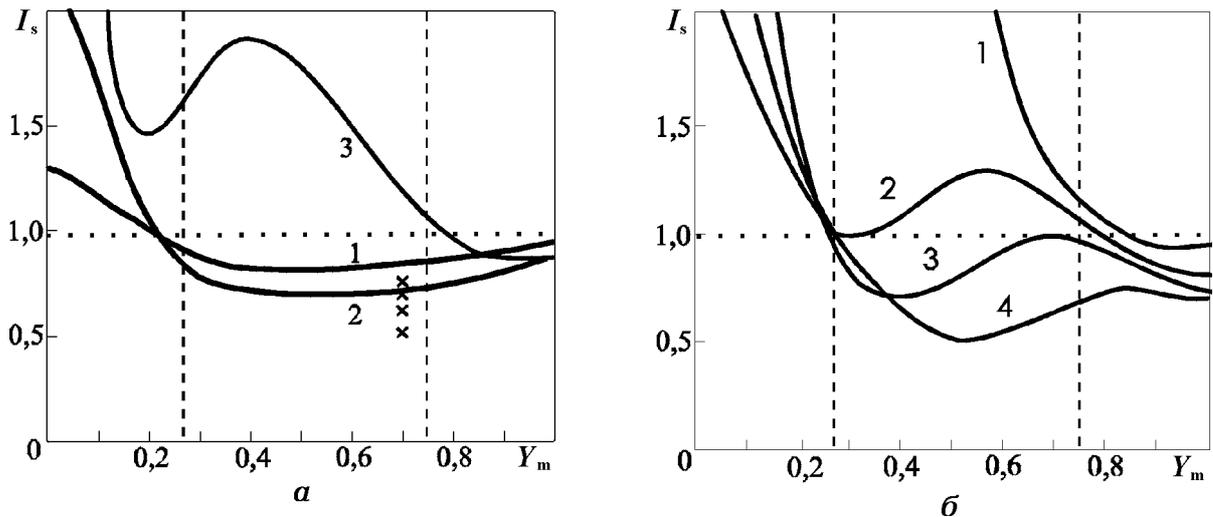


Рис. 3. Теоретические зависимости начального тока ГДИ от места расположения ферромагнитного стержня относительно области взаимодействия: *a* — стержни с концентрическими поверхностями (график 1 соответствует $R = 0,026$; $\delta_z = 3,487$; график 2 — $R = 0,044$; $\delta_z = 1,7$; график 3 — $R = 0,06$; $\delta_z = 1$); *б* — стержни с неконцентрическими поверхностями ($\delta_z = 1$, графики 1–4 соответствуют $R = 0,03$; $0,09$; $0,12$; $0,18$)

Однако неопределенность параметров P_B , W_m относительно характеристик ЛНМП, создаваемой ферромагнитным стержнем, не позволяет изучить особенности изменения начального тока и других характеристик генератора от геометрических размеров стержня и его расположения относительно поверхности решетки.

Такие исследования можно провести, если положить $P_B = A_m$, $W_m = R_y$ и произвести соответствующий перерасчет теоретических зависимостей. В частности, на рис. 3*a* показаны полученные таким путем зависимости минимального значения начального тока ГДИ-МН от места расположения на оси y ферромагнитного стержня при различных значениях параметров R и δ_z , соответствующих наличию в области взаимодействия ЛНМП с постоянной протяженностью $D_y = 0,4$ (случай стержней с концентрическими поверхностями; на рисунке по оси ординат отложены относительные значения начального тока $I_s = i_s/i_0$, где i_0 — минимальный пусковой ток ГДИ без локальной неоднородности поля).

Видно, что начальный ток ГДИ-МН можно существенно уменьшить, если ферромагнитный стержень размещать в области пятна СВЧ поля (вертикальные пунктирные линии) со стороны “коллекторного” конца решетки ($0,5 < Y < 1$). При этом степень уменьшения начального тока в значительной мере зависит от расстояния между решеткой и поверхностью стержня (т. е. от параметра δ_z). Так, с уменьшением δ_z (или с увеличением $|A_m|$, рис. 2*a*) оптимальное для начального тока место расположения стержня на оси y смещается в сторону коллектора. Согласно теории [11] это связано с тем, что с уменьшением δ_z возрастает интенсивность оседания на поверхность дифракционной решетки электронного потока, не взаимодействовавшего с СВЧ полем резонатора. В результате уменьшается эффективная длина области взаимодействия и увеличивается начальный ток генератора. Этими явлениями объясняются особенности изменения графика 3 на рис. 3*a* для значения $\delta_z = 1$, при котором уменьшить начальный ток в ГДИ удастся только при расположении стержня за пятном СВЧ поля (в конце решетки).

Если в рассматриваемом случае (при $R_y = 0,2$) параметр $\delta_z < 1$, то расположение стержня в области пятна СВЧ поля приведет к оседанию практически всего ЭП на решетку и к существенному увеличению начального тока ГДИ. Чтобы уменьшить начальный ток генератора, при значениях $\delta_z \leq 1$ необходимо увеличить в области взаимодействия протяженность пятна ЛНМП. Это обеспечивает-

ся путем увеличения радиуса ферромагнитного стержня (рис. 2б). В результате силовые линии поля в окрестности его локальной неоднородности сглаживаются, увеличивается эффективная длина дифракционной решетки и уменьшается I_s генератора. Это проиллюстрировано на рис. 3б, на котором графики отражают особенности изменения стартового тока I_s генератора для $\delta_z = 1$ при различных радиусах R (случай неконцентрических стержней — с увеличением радиуса R стержня для сохранения постоянного значения δ_z необходимо смещать координату Z_m центра стержня). В рассматриваемом случае уменьшение стартового тока наблюдается в генераторе с толстым стержнем, расположенным вблизи центра пятна СВЧ поля ($R = 0,18$; $Y_m = 0,5$; график 4 на рис. 3б). Такой стержень создает в пространстве взаимодействия генератора неоднородность поля с амплитудой $A_m = -0,25$ и с достаточно большой протяженностью пятна неоднородности ($D_y \approx 1,2$) по сравнению с D_y тонкого стержня (график 2 на рис. 3а).

Кроме того, согласно теории [12] с помощью ЛНМП также можно увеличить выходную мощность и КПД ГДИ, а также улучшить ряд других характеристик генератора. В частности, перемещение центра локальной неоднородности поля вдоль области взаимодействия приводит к изменению крутизны ν [МГц/В] электронной перестройки частоты ГДИ в пределах зоны генерации. Здесь возможны случаи, когда $\nu < \nu_0$ или $\nu > \nu_0$, где ν_0 — крутизна электронной перестройки частоты ГДИ без ЛНМП. Первый случай ($\nu < \nu_0$) указывает на возможность уменьшения степени влияния на спектр сигнала дестабилизирующих факторов, связанных с флуктуациями скорости электронов в потоке и обусловленных, например, нестабильностью источников питания ГДИ [13]. Второй случай ($\nu > \nu_0$) позволяет увеличить диапазон электронной перестройки частоты в зоне генерации. Следовательно, путем профилирования ЛНМП можно улучшить качество спектра выходного сигнала (уменьшить ширину спектральной линии), увеличить стабильность частоты и обеспечить возможности управления частотными характеристиками ГДИ. Перспективность применения ЛНМП в ГДИ подтверждается проведенными нами экспериментальными исследованиями.

2. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для экспериментальных исследований создан лабораторный макет ГДИ-МН разборной конструкции, в котором в процессе эксперимента можно менять электронную пушку, дифракционную решетку, открытый резонатор, ферромагнитный стержень и другие элементы генератора, а также при равных технических условиях корректно проводить сравнительный анализ выходных характеристик ГДИ-МН и ГДИ без локальной неоднородности поля (базового ГДИ). В исследуемом генераторе используются сфероцилиндрический открытый резонатор (ОР), образованный сферическим и цилиндрическим зеркалами с радиусами кривизны 110 мм. Волноводная согласованная щель, связывающая генератор с нагрузкой, располагается в центральной части сферического зеркала, а отражательная дифракционная решетка (шириной 10 мм) размещается в центральной части цилиндрического зеркала ОР. Длина решетки в направлении движения ЭП (катод—коллектор) определяется размерами апертуры зеркала ОР ($L = 36 \div 40$ мм). ГДИ располагается между полюсными наконечниками электромагнита, который обеспечивает в пространстве взаимодействия управляемое по величине фокусирующее магнитостатическое поле $B_0 = 0 \div 0,6$ Т. Ленточный электронный поток с размерами поперечного сечения $5 \times 0,2$ мм создается диодной электронной пушкой. Ферромагнитный стержень диаметром 3 мм и длиной 10 мм располагается на расстоянии 1,5 мм от поверхности решетки ($\delta_z = 1$; $A_m = -0,25$, см. рис. 1). Исследования проводились в диапазоне частот $40 \div 170$ ГГц, который перекрывается путем использования в приборе трех различных дифракционных решеток.

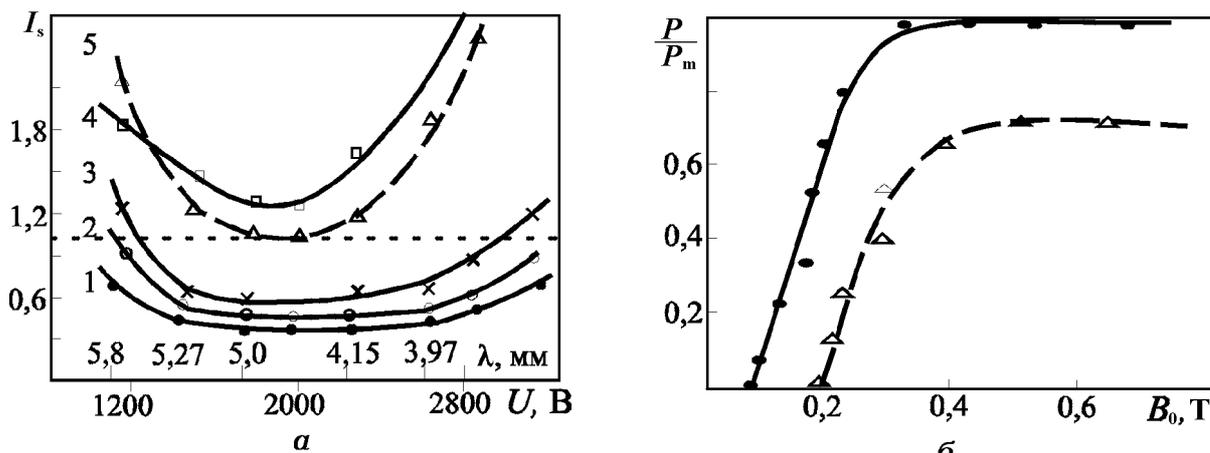


Рис. 4. Влияние величины фокусирующего магнитного поля на стартовый ток (панель *a*) и выходную мощность (панель *б*) электромагнитных колебаний в ГДИ-МН (сплошные линии) и в базовом ГДИ (пунктирные линии): *a* — графики 1–4 соответствуют $B_0 = 0,4; 0,25; 0,16; 0,1$ Т; график 5 построен для базового ГДИ при $B_0 = 0,4$ Т

На первом этапе эксперимента ферромагнитный стержень размещался в центральной части дифракционной решетки ($Y_m = 0,5$). В этом случае наблюдается значительное оседание ЭП в области центра ЛНМП и электромагнитные колебания при токах $I \leq 80$ мА не возбуждаются. Это согласуется с теорией, согласно которой для выбранных в эксперименте параметров R и δ_z должно наблюдаться значительное увеличение стартового тока (график 3, рис. 3*a*). Однако при смещении стержня к коллекторному концу решетки в ГДИ-МН эффективно возбуждаются электромагнитные колебания. Остановимся на результатах проведенных исследований.

Экспериментально установлено, что использование ЛНМП в ГДИ позволяет улучшить стартовые, энергетические, частотные и спектральные характеристики генератора. Это, в частности, показано на рис. 4–6 при размещении магнитного стержня вблизи границы пятна СВЧ поля ($Y_m \approx 0,7$, см. рис. 3).

Графики на рис. 4*a*, *б* отражают влияние фокусирующего магнитного поля на стартовый ток I_s и выходную мощность P генератора. Приведенные на рис. 4*a* зависимости стартового тока от длины волны λ электромагнитных колебаний, возбуждаемых при указанных здесь же значениях ускоряющего напряжения U , получены для различных значений B_0 (пунктирной линией (график 5) выделена аналогичная зависимость для ГДИ без локальной неоднородности поля при $B_0 = 0,4$ Т). Из сравнения графиков 1 и 5 следует, что при равных значениях B_0 минимальный стартовый ток в ГДИ-МН примерно в два раза меньше, чем в базовом генераторе. При этом с уменьшением B_0 до значения 0,16 Т стартовый ток в ГДИ-МН меняется незначительно (графики 1–3), в то время как в базовом ГДИ в этом случае наблюдается существенное увеличение стартового тока [5]. Следовательно, ГДИ-МН работает при меньших фокусирующих магнитных полях. Этот факт также иллюстрируют графики на рис. 4*б*, где показаны зависимости выходной мощности генераторов от величины B_0 при токе $I = 85$ мА электронного пучка (в процессе изменения B_0 генератор юстируется на максимальную выходную мощность P_m). Здесь минимальное стартовое значение фокусирующего поля в исследуемом генераторе составляет приблизительно 0,08 Т, а в базовом ГДИ — около 0,2 Т. Зависимость выходной мощности ГДИ-МН от B_0 , как и в базовом ГДИ имеет вид кривой насыщения, при этом максимальная выходная мощность в исследуемом генераторе примерно в 1,5 раза выше, чем в базовом ГДИ. Однако экспериментально установлено, что при $B_0 \geq 0,45$ Т и токе пучка $I \geq 130$ мА выходная мощность в базовом ГДИ может быть выше, чем в исследуемом генераторе, хотя при малых значениях фокусирующего поля ($B_0 < 0,45$ Т) преимущества ГДИ-МН сохраняются. Этот факт связан с особенностями изменения условий взаимодействия электронного потока с полем ОР в ГДИ-МН при больших токах

пучка и значениях B_0 . В этом случае для увеличения выходной мощности генератора необходимо менять характеристики ЛНМП в пространстве взаимодействия.

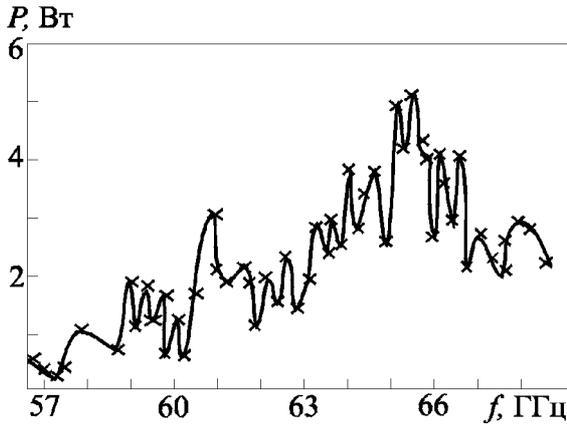


Рис. 5. Зависимость выходной мощности ГДИ-МН от частоты генерации в случае комбинированной перестройки частоты одного возбуждаемого в резонаторе типа колебаний

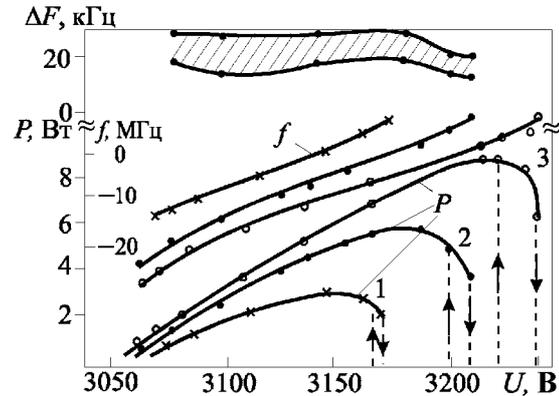


Рис. 6. Выходные характеристики ГДИ-МН в пределах одной зоны генерации. Графики 1–3 соответствуют значениям электронного тока $I = 70; 110; 160$ мА

На рис. 5 приведена зависимость выходной мощности исследуемого ГДИ-МН от частоты генерации при $I = 130$ мА и $B_0 = 0,32$ Т в случае непрерывной комбинированной перестройки частоты одного возбуждаемого в ОР типа колебаний (такая перестройка частоты в ГДИ обеспечивается путем плавного изменения расстояния между зеркалами ОР и подстройкой ускоряющего напряжения U). Анализ этих характеристик показывает, что введение в ГДИ неоднородности магнитостатического поля позволяет увеличить в несколько раз диапазон Δf комбинированной перестройки частоты. На рис. 5 в ГДИ-МН этот диапазон $\Delta f \approx 12,5$ ГГц, что составляет приблизительно 20 % относительно центральной частоты генерации. В случае использования двух типов колебаний, отличающихся по частоте, диапазон непрерывной комбинированной перестройки частоты возрастает до значения $\Delta f \approx 24$ ГГц.

Графики на рис. 6 иллюстрируют комплекс выходных характеристик ГДИ-МН в пределах отдельной зоны генерации в случае возбуждения одномодового колебания на частоте $f_0 = 65,5$ ГГц при постоянном значении $B_0 = 0,38$ Т и различных значениях электронного тока. Здесь P — выходная мощность генератора, f — электронная перестройка частоты, ΔF — ширина спектра выходного сигнала, наблюдаемого на спектроанализаторе в течение $0,3 \div 5$ с (на рис. 6 все измеренные значения ΔF попадают в заштрихованную область). Сравнительный анализ этих характеристик и аналогичных характеристик генератора без ЛНМП показывает, что особенности их в зависимости от различных параметров в обоих генераторах сохраняются. В частности, из зависимостей P следует, что электромагнитные колебания в ГДИ-МН в начале зоны генерации возбуждаются в мягком режиме, а в конце зоны срываются в жестком режиме с проявлениями характерных особенностей гистерезисных эффектов. Частота генерации f при изменении ускоряющего напряжения в пределах зоны генерации изменяется приблизительно по тангенсоидальному закону с линейной крутизной $\nu < \nu_0$ электронной перестройки частоты на уровне $0,5 P$. Однако в количественном отношении исследуемый генератор имеет ряд улучшенных выходных характеристик по сравнению с обычным ГДИ. Это следует из табл. 1, где выходные характеристики разных ГДИ получены при равных технических условиях ($B_0 = 0,38 \div 0,4$ Т, рабочий ток электронного потока $I = 160$ мА, используются одинаковые источники питания, электронная пушка, дифракционная решетка, открытый резонатор и равноценные условия измерения характеристик).

Уменьшение крутизны ν электронной перестройки частоты в ГДИ-МН приводит к уменьшению ширины ΔF спектра и к последующему уменьшению значения длительной нестабильности $\Delta F/f_0$ ча-

стоты выходного сигнала. Одной из интересных особенностей ГДИ-МН является высокая устойчивость его к разбюстировке электронного потока относительно области взаимодействия и фокусирующего магнитного поля, что повышает эксплуатационные характеристики этого генератора.

Полученные результаты исследований указывают на удовлетворительное качественное согласование теории с экспериментом. В то же время в количественном отношении эти

Таблица 1

Генератор	i_s , мА	P , Вт	ΔF , кГц	$\Delta F/f_0$	ν , МГц/В
ГДИ	36	5,2	25÷40	$\approx 5 \cdot 10^{-7}$	0,12÷0,15
ГДИ-МН	18	8,0	5÷25	$\approx 1 \cdot 10^{-7}$	0,08÷0,12

результаты могут существенно отличаться. В частности, это следует из рис. 3а, где крестиками отмечены полученные экспериментально значения пусковых токов исследуемого ГДИ-МН, который имеет в пространстве взаимодействия локальную неоднородность поля с характеристиками, соответствующими графику 3. Согласно приведенной выше теории создание в ГДИ такой неоднородности поля ($Y_m = 0,7$) должно незначительно влиять на величину пускового тока генератора. Однако в эксперименте в таком генераторе пусковой ток уменьшается в 1,5÷2 раза. Это различие объясняется, прежде всего, приближенностью теории по отношению к реальным условиям эксперимента. Важнейшим обстоятельством здесь, по-видимому, является изменение величины фокусирующего магнитостатического поля B_0 по параболическому закону относительно плоскости $Y_m = 0,5$, где достигается минимум поля, в направлении движения электронного потока, обусловленное наличием в эксперименте полюсных наконечников магнита и не учитывавшееся в теории.

Из проведенных исследований можно сформулировать ряд общих закономерностей, полезных для разработчиков ГДИ-МН: 1) чтобы уменьшить пусковой ток ГДИ необходимо располагать ферромагнитный стержень в области пятна СВЧ поля со стороны коллектора; 2) для заданного радиуса r стержня существует определенное расстояние между поверхностями решетки и стержня, при котором стартовый ток ГДИ существенно уменьшается ($h_z = 1,7r$, график 2, рис. 3а); 3) в интервале значений $r < h_z < 1,7r$ с уменьшением h_z минимальное значение пускового тока ГДИ обеспечивается смещением стержня от центра решетки в сторону коллектора; 4) при малых значениях $h_z \leq r$ для снижения пускового тока ГДИ необходимо использовать толстые ферромагнитные стержни (например, $0,09L \leq r < 0,18L$, рис. 3б); 5) путем профилирования ЛНМП можно управлять выходной мощностью, крутизной электронной перестройки частоты и спектром сигнала электромагнитного колебания ГДИ.

ЗАКЛЮЧЕНИЕ

Показана возможность повышения эффективности электронно-волнового взаимодействия в ГДИ путем применения в области взаимодействия локальной неоднородности магнитостатического поля, создаваемой в однородном фокусирующем магнитостатическом поле ферромагнитным стержнем. Найдены основные характеристики такой неоднородности. Установлено, что улучшить выходные параметры ГДИ предложенным методом можно только при определенных условиях расположения ферромагнитного стержня относительно пятна СВЧ поля на дифракционной решетке. Этим методом экспериментально в ГДИ миллиметрового диапазона уменьшен в два раза стартовый ток генератора, увеличена при малых фокусирующих магнитных полях мощность генерации, а также показаны возможности улучшения спектральных характеристик высокочастотного сигнала электромагнитного колебания. Результаты эксперимента качественно согласуются с теоретическими исследованиями и указывают на перспективу применения ЛНМП для улучшения выходных характеристик генераторов миллиметрового и субмиллиметрового диапазонов длин волн.

ЛИТЕРАТУРА

1. Левин Г. Я., Бородкин А. И., Кириченко А. Я. и др. Клиноотрон / Под ред. А. Я. Усикова. — Киев: Наук. думка, 1992. — 200 с.
2. Глущенко В. Н., Кошечая С. В., Прус В. А. // Изв. вуз. Радиоэлектроника. 1970. Т. 13, № 1. С. 12.
3. Макаров В. И., Нутович Л. М. // Изв. вуз. Радиоэлектроника. 1986. Т. 29, № 1. С. 86.
4. Васильев С. В., Васянович А. В. // Радиотехника и электроника. 1988. Т. 33, № 11. С. 2434.
5. Шестопалов В. П., Вертий А. А., Ермак Г. П., Скрынник Б. К., Хлопов Г. И., Цвык А. И. Генераторы дифракционного излучения / Под ред. В. П. Шестопалова. — Киев: Наук. думка, 1991. — 320 с.
6. Price E., Walsh J. // Proceedings SPTE. 1990. V. 1514. P. 135.
7. Исаев В. А. // Вопросы электроники сверхвысоких частот. Саратов, 1976. Вып. 9. С. 11.
8. Цвык А. И. // Изв. вуз. Радиофизика. 1978. Т. 21, № 8. С. 1216.
9. Русин Ф. С., Синенко Л. А. // Радиотехника и электроника. 1979. Т. 24, № 7. С. 1396.
10. Балаклицкий И. М., Воробьев Г. С., Цвык А. И., Шестопалов В. П. // Изв. вуз. Радиоэлектроника. 1977. Т. 20, № 10. С. 93.
11. Одаренко Е. Н., Шматько А. А. // Радиотехника и электроника. 1992. Т. 37, № 5. С. 901.
12. Одаренко Е. Н., Шматько А. А. // Радиотехника и электроника. 1994. Т. 39, № 9. С. 835.
13. Балаклицкий И. М., Воробьев Г. С., Годецкий А. Г. и др. // Изв. вуз. Радиоэлектроника. 1978. Т. 21, № 10. С. 105.

Институт радиофизики и
электроники им. А. Я. Усикова
НАН Украины, г. Харьков, Украина

Поступила в редакцию
20 ноября 1998 г.

INFLUENCE OF LOCAL INHOMOGENEITY OF THE MAGNETOSTATIC FIELD ON THE START AND OUTPUT CHARACTERISTICS OF THE DIFFRACTION RADIATION GENERATOR

A. I. Tsvyk, G. S. Vorob'yov, A. V. Nesterenko, and V. N. Zheltov

We study theoretically and experimentally the possibilities to improve the output characteristics of the non-relativistic generator of diffraction radiation (GDR) by creating local inhomogeneity of the magnetostatic field (LIMF) in the region of electron-wave interaction. The LIMF is created in the homogeneous focusing magnetic field by a ferromagnetic stem. We obtain the formulas for calculating the main characteristics of the LIMF and show the dependence of the start current on the size of the stem and its location relative to the interaction region. Theoretical results are confirmed by the experimental studies of GDR in the frequency range 40 to 170 GHz. Creation of the LIMF is shown to improve the start, power, and spectral characteristics of the generator. We conclude that the proposed technique can be useful for designing low-voltage generators with the improved output characteristics for the short-wavelength millimeter and submillimeter bands.

ДВИЖЕНИЕ КАПЛИ ЖИДКОСТИ В КАПИЛЛЯРЕ ПОД ДЕЙСТВИЕМ СТАТИЧЕСКОГО И АКУСТИЧЕСКОГО ПОЛЕЙ

В. С. Авербах, С. Н. Власов, Ю. М. Заславский

В работе рассматривается движения капли жидкости в капилляре пористой среды под воздействием статического и акустического полей. Выведено уравнение движения капли в капилляре и численно проанализировано его решение. Показано, что вибровоздействие приводит к уменьшению влияния сил поверхностного натяжения. В капилляре с неизменным вдоль оси сечением при наличии статического воздействия акустическая волна снижает порог начала движения капли и увеличивает скорость её поступательного движения.

При изучении эффективности фильтрации представляет интерес анализ воздействия акустической волны на частично заполненную жидкостью пористую среду. В [1, 2] исследуются различные физические механизмы такого воздействия и утверждается, что на низких частотах (до 1 кГц) защемленные в порах капли жидкости колеблются около положения равновесия. На высоких частотах (выше 10 кГц) существует механизм поступательного перемещения жидкости. Этот механизм известен под названием "акустический ветер". Согласно [1, 2] акустический ветер вызывает направленное смещение жидкости только в полностью насыщенной среде, при этом скорость перемещения пропорциональна декременту затухания и интенсивности звуковой волны. Однако в упомянутых работах не уделялось внимания особенностям движения одиночных капель в капиллярах, которые близки к особенностям движения материальных точек на поверхности с сухим трением [3] и имеют место вследствие своеобразной зависимости сил поверхностного натяжения от скорости перемещения капель. В данной работе этот пробел восполняется на примере рассмотрения капилляра со стенками, частично смачиваемыми жидкостью (капля воды в стеклянной капиллярной трубке, внутренний канал которой покрыт парафиновой пленкой).

Рассмотрим каплю в цилиндрическом капилляре (см. рис. 1), находящуюся под действием направленной вдоль оси капилляра постоянной силы \mathbf{G} , обусловленной внешним силовым полем (сила \mathbf{G} может быть, например, проекцией на ось z градиента давления или силы тяготения; в дальнейшем силой называется проекция последней на ось капилляра). Предположим, что:

1) линейные размеры капли $L = z_+ - z_-$, где z_{\pm} — координаты правого и левого менисков, существенно больше диаметра капилляра и все сечение последнего перекрывается каплей на некотором его участке;

2) движение капли сводится к перемещению собственного центра масс, тем самым пренебрегается деформацией капли вследствие изменения формы менисков;

3) температура капли и капилляра, а также максимальная скорость капли таковы, что силы поверхностного натяжения преобладают по сравнению со стоковыми силами трения, приложенными к поверхности контакта со стенкой; для справедливости последнего необходимо удовлетворить ограничению на скорость $u = dz_c/dt$ центра масс капли, где $z_c = (z_+ + z_-)/2$:

$$u \ll \frac{\sigma r}{L\mu}, \quad (1)$$

где μ — коэффициент вязкости жидкости, r — радиус капилляра, σ — константа поверхностного натяжения. Пусть вдоль оси капилляра по твердому материалу пористой среды распространяется продольная акустическая волна частоты ω . Будем считать, что материал каркаса достаточно жесткий, так

что удовлетворяется условие

$$\frac{cu}{2c_t^2} \ll 1, \tag{2}$$

где c и c_t — скорости продольных и поперечных волн в каркасе. В этом случае поперечным смещением стенки капилляра $\Delta r = -cur/(2c_t^2)$ в поле звуковой волны можно пренебречь.

В процессе распространения волны стенки капилляра совершают периодические продольные движения с амплитудой Y . При этом вследствие наличия капиллярных сил стенки действуют на каплю с силой, которую мы определим, используя формулу Лапласа для избыточного давления, действующего на каплю со стороны мениска. Для определенности сначала запишем выражение для капиллярной силы в случае мениска, ограничивающего столбик жидкости:

$$F = \pm 2\pi r \sigma \cos \theta, \tag{3}$$

где положительное значение силы F принято совпадающим с направлением оси z , а $\cos \theta$ — косинус краевого угла мениска, отсчитываемого от стенки капилляра со стороны жидкости (см. рис. 1). Знаки \pm отражают противоположную направленность силы поверхностного натяжения на столбик жидкости, находящийся относительно мениска либо слева, либо справа. Косинус краевого угла в статическом равновесии имеет положительное значение ($\cos \theta > 0$) в случае смачивания жидкостью стенки, а в случае несмачивания — отрицательное ($\cos \theta < 0$).

Вследствие относительного движения капли и стенок угол θ у каждого из менисков, ограничивающих каплю, будет меняться. Зависимость угла θ от относительной скорости движения мениска для пары вода—парафин была определена в [4]. Мы используем эту зависимость, отражающую типичные черты поведения мениска при движении жидкости, хотя известны и несколько другие ее виды [5–7]. Зависимость $\cos \theta$ от скорости движения мениска из [3] представлена на рис. 2 и имеет характерную черту: наличие неоднозначного (гистерезисного) участка при нулевой относительной скорости s . Угол θ при $s = 0$ может меняться от θ_- до θ_+ в зависимости от предыстории процесса. При $s \neq 0$ угол мениска однозначно связан со скоростью его движения и очень быстро, с характерной скоростью w , достигает предельных значений θ_{-0} и θ_{+0} .

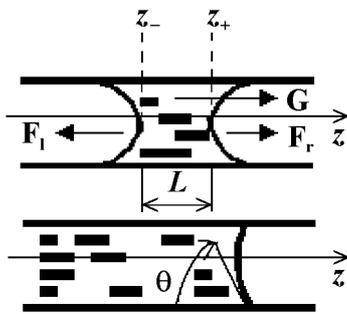


Рис. 1. Общий вид столбика жидкости и капли в капилляре при наличии сил поверхностного натяжения F_l, F_r и объемной силы G

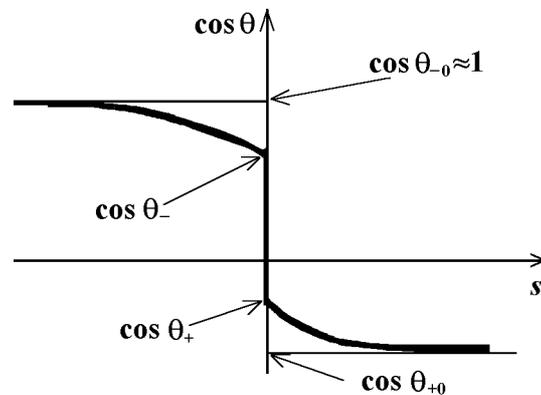


Рис. 2. Зависимость $\cos \theta$ от относительной скорости s движения мениска

В дальнейшем будем рассматривать малые скорости движения капли, при которых достигаются и предельные значения угла мениска, и выполняется условие (1). Описанную выше зависимость угла θ от скорости s можно представить в виде

$$\cos \theta_r = \begin{cases} \cos \theta_- - (\cos \theta_{-0} - \cos \theta_-) \operatorname{th} \frac{s}{w}, & s < 0, \\ \cos \theta_+ + (\cos \theta_{+0} - \cos \theta_+) \operatorname{th} \frac{s}{w}, & s > 0 \end{cases} \tag{4}$$

для правого мениска (столбик жидкости располагается слева от мениска), который при $s > 0$ является наступающим, а при $s < 0$ отступающим, и в виде

$$\cos \theta_1 = \begin{cases} \cos \theta_+ - (\cos \theta_{+0} - \cos \theta_+) \operatorname{th} \frac{s}{w}, & s < 0, \\ \cos \theta_- + (\cos \theta_{-0} - \cos \theta_-) \operatorname{th} \frac{s}{w}, & s > 0 \end{cases} \quad (5)$$

для левого мениска.

Уравнение движения капли в капилляре имеет вид

$$m \frac{du}{dt} = F_1(s) + F_r(s) + G = -2\pi r \sigma (\cos \theta_1 - \cos \theta_r) + G, \quad (6)$$

где m — масса капли, F_1 и F_r — силы, действующие на каплю со стороны соответственно левого и правого мениска, G — статическая сила, обусловленная, к примеру, действием гравитации.

Рассмотрим случай капилляра с постоянным сечением, для которого в пренебрежении изменением объема капли вследствие деформации менисков $m = \pi r^2 L \rho$, где ρ — плотность жидкости. Учитывая связь между u и s ($s = u - u_0$, где u_0 — заданная скорость движения стенки), получим уравнение движения в виде

$$m \frac{ds}{dt} + m \frac{du_0}{dt} = F_1(s) + F_r(s) + G. \quad (7)$$

Для случая внезапного включения вибрации, когда

$$u_0 = \begin{cases} 0, & t < 0, \\ \omega Y \sin \omega t, & t \geq 0, \end{cases} \quad (8)$$

справедливо следующее выражение для капиллярной силы, действующей на каплю со стороны обоих менисков:

$$F_1 + F_r = 2\pi r \sigma \left[-\cos \theta_+ + \cos \theta_- - (-\cos \theta_{+0} + \cos \theta_{-0} + \cos \theta_+ - \cos \theta_-) \operatorname{th} \frac{s}{w} \right] \quad (9)$$

при $s < 0$,

$$F_1 + F_r = 2\pi r \sigma \left[\cos \theta_+ - \cos \theta_- + (\cos \theta_{+0} - \cos \theta_{-0} - \cos \theta_+ + \cos \theta_-) \operatorname{th} \frac{s}{w} \right] \quad (10)$$

при $s > 0$.

Согласно [4] для контакта вода—парафин

$$\cos \theta_- - \cos \theta_+ = 1, \quad (11)$$

поэтому если обозначить

$$\cos \theta_{-0} - \cos \theta_- - \cos \theta_{+0} + \cos \theta_+ = 2\alpha, \quad (12)$$

где по данным [4] $\alpha \approx 0,25$, то уравнение движения для $t > 0$ примет вид

$$\frac{ds}{dt} = -\omega^2 Y \cos(\omega t) + \frac{G}{m} + \frac{2\pi r \sigma}{m} \left(\pm 1 - 2\alpha \operatorname{th} \frac{s}{w} \right), \quad (13)$$

где знак плюс следует брать при $s < 0$, знак минус — при $s > 0$.

Уравнение (13) описывает движение материальной точки под действием статической и переменной сил при наличии силы “сухого” трения, представленной членом $\pm 2\pi r \sigma / m$. В уравнение (13) входит

также выражение $\text{th}(s/w)$, близкое к обычному вязкому трению. Перейдем в (13) от переменных s, t к безразмерным переменным S, τ , для чего положим

$$g = \frac{Y\omega^2 m}{2\pi\sigma r}, \quad \tau = \frac{2\pi\sigma r}{m\omega} t, \quad \Omega = \frac{\omega m \omega}{2\pi\sigma r}, \quad f = \frac{G}{2\pi\sigma r}, \quad S = \frac{s}{w}. \tag{14}$$

В новых обозначениях уравнение (13) переписывается в форме

$$\frac{dS}{d\tau} + 2\alpha \text{th} S = -g \cos(\Omega\tau) + f \pm 1,$$

в котором выбор знаков \pm соответствует ранее указанному правилу. Если ввести координату $y = \int S d\tau$, описывающую смещение центра капли относительно стенок, то для y имеем уравнение

$$\frac{d^2 y}{d\tau^2} + 2\alpha \text{th} \frac{dy}{d\tau} = -g \cos(\Omega\tau) + f \pm 1. \tag{15}$$

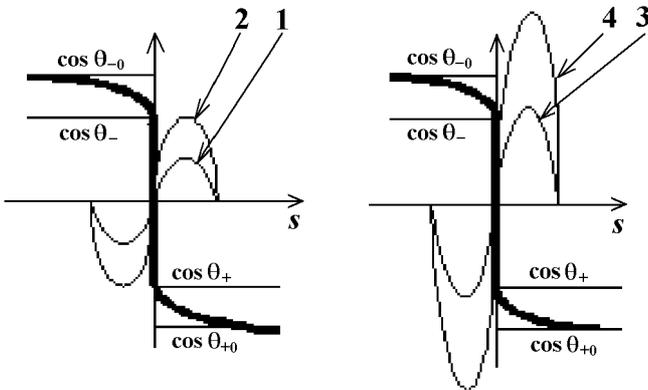


Рис. 3. Иллюстрация возможных соотношений уровня переменного воздействия f и порогов начала движения, обусловленных параметрами капиллярных сил: кривые 1, 2 — без смещения капли; кривые 3, 4 — с уходом капли от исходной точки

но меньше величины $G < G_{\text{пор}} + 4\pi\alpha\sigma r$, то на достаточно больших временах капля движется равномерно со скоростью $S = \text{arth}[(|f| - 1)/(2\alpha)]$, которая в размерных переменных имеет вид $u = w \text{arth}[(G - 2\pi r\sigma)/(4\pi r\alpha\sigma)]$. Если же постоянная сила превышает величину $2\pi r\sigma(1 + 2\alpha)$, то капля через некоторое время набирает скорость $u \gg w$, после чего движется с постоянным ускорением $du/dt = (G - 2\pi r\sigma(1 + 2\alpha))/m$. Отметим, что в [8, 9] использована упрощенная аппроксимация для избыточного давления, создаваемого мениском. В ней не учитывается гистерезисный участок, т. е. в соотношениях, аналогичных (4), (5), принимается, что $\cos \theta_+ = \cos \theta_- = \alpha$. Следствием этого является равенство нулю пороговой силы $G_{\text{пор}} = 0$. Действительно, если $\cos \theta_+ = \cos \theta_- = \alpha$, то капля движется равномерно при сколь угодно малой силе G . Последнее противоречит известным наблюдениям зависимости скорости движения капель в гладких капиллярах от градиента статического давления. Таким образом, для корректного описания движения капель необходимо введение гистерезисного участка.

При наличии переменной силы решения уравнения (15) исследовались численно. Заметим, что представленная зависимость координаты центра капли от времени дается в системе отсчета, связанной со стенкой капилляра; для перехода к лабораторной системе отсчета требуется учет периодического движения начала координат. Переменная сила существенно меняет характер движения капли. При

Пусть переменная сила $g = 0$. Капля остается на месте, если постоянная сила по модулю меньше силы “сухого” трения: $|f| < 1$. Имеется пороговая величина постоянной силы

$$G_{\text{пор}} = 2\pi\sigma r, \tag{16}$$

при превышении которой капля начинает двигаться (см. рис. 3). Пороговая сила может быть выражена через градиент давления $(dP/dz)_{\text{пор}} = 2\sigma/(rL)$, который необходим для смещения капли от исходного положения. В том случае, когда объемная сила G является силой тяжести в вертикально расположенном капилляре, f является так называемым числом Бонда [5]. Если выполняется условие $1 < |f| < 1 + 2\alpha$ (постоянная сила больше пороговой,

условии

$$f + |g| < 1 \quad \text{или} \quad G < G_{\text{пор}} - \omega^2 Y m$$

капля остается на месте (кривая 1 на рис. 4). Если

$$1 < f + |g| \quad \text{или} \quad G_{\text{пор}} < G + \omega^2 Y m, \quad (17)$$

капля перемещается. При небольших превышениях порога ($f + |g| - 1 \leq 1$) капля двигается в основном в те моменты времени, когда сила $|-g \cos(\Omega\tau) + f| > 1$.

Это иллюстрируется кривыми 3, 4 на рис. 3 и кривыми 2, 3 на рис. 4. При невыполнении последнего условия капля быстро останавливается. Перемещение капли носит, таким образом, ступенчатый характер. Средний уход центра масс капли за период осцилляций существенно меньше амплитуды смещения стенки в акустическом поле. Из условия (17), которое подтверждается кривыми на рис. 4, следует, что упругая волна в стенке капилляра снижает порог начала движения капли, находящейся в точке безразличного равновесия, на величину $\omega^2 Y m = G_{\text{пор}} \sqrt{I/I_0}$, где $I = Y^2 \omega^2 \rho_{\text{ск}} c / 2$ — интенсивность звуковой волны, $\rho_{\text{ск}}$ — плотность вещества стенки капилляра, $I_0 = \frac{2\pi\sigma^2 r^2 \rho_{\text{ск}} c}{\omega^2 m^2} = \frac{2\sigma^2 \rho_{\text{ск}} c}{\omega^2 r^2 L^2 \rho^2}$ — характерное значение интенсивности волны, при которой преодолевается статический силовой порог начала движения $G_{\text{пор}} = 2\pi\sigma r$. Снижение порога иллюстрируется кривыми 3, 4 на рис. 3.

Отметим, что характерная интенсивность I_0 весьма быстро растет с уменьшением радиуса капилляра и длины капли ($I_0 \propto r^{-2} L^{-2}$). Так, обращаясь к случаю рассматриваемого контакта воды и парафина, покрывающего пленкой стенки стеклянного капилляра, т. е. задаваясь параметрами $\sigma = 0,01$ Н/м, $\rho = 1\,000$ кг/м³, $\rho_{\text{ск}} = 2\,000$ кг/м³, $c = 4\,000$ м/с, для капилляра радиусом 10 мкм и капли в несколько раз большей длины на частоте 500 Гц получаем, что значение I_0 достигает 100 Вт/м². В то же время для миллиметровых диаметров капилляра интенсивность I_0 не превышает нескольких десятков милливольт на квадратный метр.

В области $1 < |f| < 1 + 2\alpha$, где под действием постоянной силы капля движется равномерно, появление переменного воздействия за счет акустической волны приводит к возрастанию поступательной скорости движения капли. Это иллюстрируется на рис. 5, где представлена зависимость средней скорости s движения капли, нормированной на ее скорость s_0 при отсутствии звуковой волны, от амплитуды акустического воздействия при трех разных значениях частоты в условиях постоянной по величине статической силы. С уменьшением частоты для достижения одинакового относительного приращения скорости требуется меньшая амплитуда поля. Очевидно, что в условиях уже существующих течений и при наличии статического градиента давления акустическое воздействие приводит к незначительному увеличению скорости протекания. Однако другая картина имеет место в случае, когда акустическое воздействие вызывает начало движения капель, которые под действием только статического давления двигаться не могли.

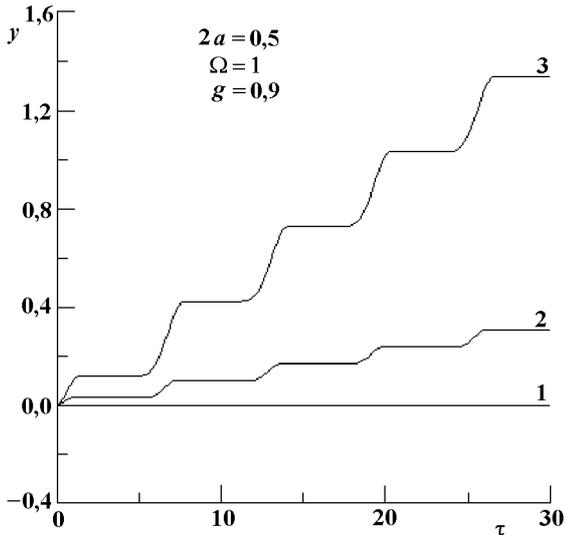


Рис. 4. Зависимость перемещения капли относительно стенок капилляра от времени. Кривая 1 соответствует $f + |g| < 1$, кривые 2 и 3 — соответственно $f + |g| = 1,3; 1,6$

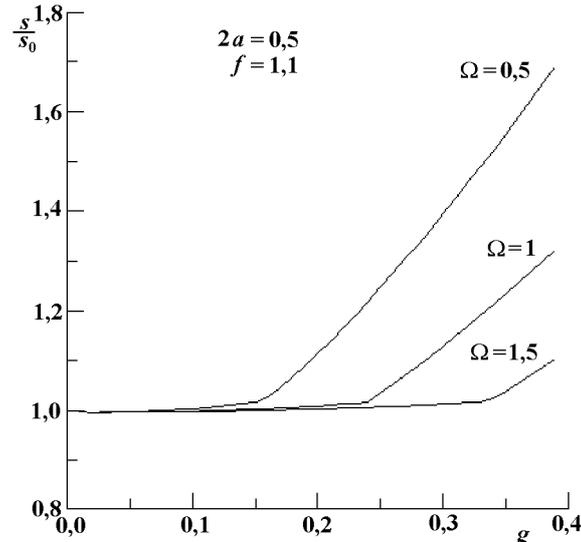


Рис. 5. Зависимость средней относительной скорости движения капли от уровня акустического поля при различной его частоте

В заключение отметим, что скорость движения капли при наличии капиллярных сил в зависимости от градиента статического давления и интенсивности акустической волны является существенно нелинейной. Очевидно, что нелинейной будет аналогичная зависимость ансамбля капель, составляющих поток жидкости, в связи с чем закон Дарси о пропорциональности скорости градиенту давления не выполняется. В зависимости скорости движения от градиента давления существует участок с нулевой скоростью до некоторого порогового значения градиента. При наличии акустической волны пороговое значение может падать, а крутизна самой зависимости нарастать. Такое поведение характерно для неньютоновых (бингамовских) жидкостей [10, 11].

Авторы признательны В. И. Таланову за внимание к работе. Работа выполнена при поддержке РФФИ (грант № 97-02-17537).

ЛИТЕРАТУРА

1. Сургучев М. Л., Кузнецов О. Л., Симкин Э. М. Гидравлическое, акустическое и тепловое воздействие на нефтяные пласты. — М.: Недра, 1975. — 184 с.
2. Вахитов Г. Г., Симкин Э. М. Использование физических полей для извлечения нефти из пластов. — М.: Недра, 1985. — 231 с.
3. Блехман И. И. Вибрационная механика. — М.: Наука, 1994. — 400 с.
4. Железный Б. В. // ДАН СССР. 1972. Т. 207, № 3. С. 647.
5. Самсонов В. М., Щербаков Л. М. // Коллоидный журнал. 1985. Т. 47, № 4. С. 729.
6. Самсонов В. М., Щербаков Л. М. // Коллоидный журнал. 1985. Т. 47, № 5. С. 907.
7. Rose W., Heins R. W. // Journal of colloid science. 1962. V. 17. P. 39.
8. Авербах В. С., Власов С. Н., Заславский Ю. М. // Сб. трудов VIII Сессии РАО Нелинейная акустика твердого тела. Н. Новгород, 1998. С. 160.
9. Авербах В. С., Власов С. Н., Заславский Ю. М. // Сб. трудов II РНКТ. М., 1998. Т. 5. С. 145.
10. Николаевский В. Н. Геомеханика и флюидодинамика. — М.: Недра, 1996. — 447 с.
11. Ампилов Ю. П. Поглощение и рассеяние сейсмических волн в неоднородных средах. — М.: Недра, 1992. — 157 с.

Институт прикладной физики РАН,
г. Н. Новгород, Россия

Поступила в редакцию
31 марта 1998 г.

MOTION OF THE LIQUID DROPLET IN THE CAPILLARY UNDER THE ACTION OF STATIC FORCE AND ACOUSTIC FIELD

V. S. Averbakh, S. N. Vlasov, and Yu. M. Zaslavsky

We consider motion of a liquid droplet inside a capillary of a porous medium under the action of static force and acoustic field. We derive the equation of motion of the droplet in the capillary and analyze numerically its solution. It is shown that the vibration attenuates the influence of the capillary force. Due to the effect by the acoustic field, the droplet motion in the capillary of the constant cross-section starts at a lower value of the static force, and the reciprocal velocity of the droplet increases.

УДК 537.868:536.3

ЭЛЕКТРОТЕПЛОВАЯ КОЛЕБАТЕЛЬНАЯ НЕУСТОЙЧИВОСТЬ В СЕГНЕТОЭЛЕКТРИЧЕСКИХ РЕЗОНАТОРАХ. ДВУХМОДОВЫЙ РЕЖИМ

Г. В. Белокопытов, М. В. Журавлев

Рассмотрены уравнения взаимодействия колебаний электромагнитного поля и температуры в нелинейном диэлектрическом резонаторе. Исследованы условия возникновения колебательной неустойчивости в системе. Рассчитаны пороговые условия электротеплового возбуждения (мощность, частота автомодуляции) в СВЧ резонаторах из виртуальных сегнетоэлектриков для простейшего, двухмодового варианта взаимодействия. Показано, что нелинейность диэлектрика накладывает принципиальные ограничения на осуществимость электротеплового возбуждения.

ВВЕДЕНИЕ

Твердотельные резонаторы СВЧ обладают кроме электромагнитных мод множеством других степеней свободы. В них, в частности, могут возбуждаться акустические колебания и происходить нестационарные процессы температурной релаксации. При этом специфической особенностью сегнетоэлектрических резонаторов является то, что между электромагнитными и акустическими модами и тепловыми степенями свободы имеется довольно сильная нелинейная связь. Вследствие этого в сегнетоэлектрических резонаторах оказываются возможными эффекты колебательной неустойчивости, когда при превышении некоторого порога электромагнитной накачки в системе возбуждаются упругие колебания или колебания температуры.

Первый эффект, получивший название стрикционного параметрического возбуждения (СПВ), был детально исследован экспериментально и теоретически [1]. По физическому механизму СПВ является аналогом вынужденного рассеяния Манделштама—Бриллюэна, отличаясь от него тем, что в диэлектрических резонаторах вызванная электрострикцией периодическая модуляция упругой жесткости ведет к параметрическому возбуждению не бегущей, а стоячей акустической волны. Существование второго эффекта, электротеплового возбуждения [2, 3], обусловлено наличием тепловой обратной связи: эффективная температура сегнетоэлектрического резонатора изменяется в зависимости от уровня рассеянной мощности вынужденных электромагнитных колебаний, амплитуда которых в свою очередь зависит от настройки системы и, соответственно, от эффективной температуры [3]. Параметрическое возбуждение температурных волн под действием интенсивной бегущей электромагнитной волны известно в нелинейной оптике как вынужденное температурное рассеяние света [4]. Электротепловое возбуждение (ЭТВ) является специфическим аналогом вынужденного температурного рассеяния в резонансных условиях.

Первоначальный анализ [3] показал, что наилучшие условия для наблюдения электротеплового возбуждения существуют в диэлектрических СВЧ резонаторах из виртуальных сегнетоэлектриков SrTiO_3 и KTaO_3 при температуре 4,2 К. Эти материалы обладают уникальным сочетанием чрезвычайно высокой диэлектрической проницаемости и малых потерь в диапазоне СВЧ [5, 6]. Кроме того, при гелиевых температурах их теплоемкость весьма мала, а температурный коэффициент диэлектрической проницаемости сравнительно велик. Совокупность этих обстоятельств позволяет ожидать, что пороговая мощность электротеплового возбуждения будет невысокой, составляя величину 10^{-3} Вт и ниже.

Для получения более точных оценок следует иметь в виду, что порог электротеплового возбуждения в большой степени зависит от пространственного распределения электромагнитного поля и температуры в объеме резонатора. Наиболее удобна для теоретического анализа и экспериментальных исследований сферическая конфигурация резонатора. Для нее распределения электромагнитного поля и температуры могут быть представлены через функции Риккати—Бесселя первого рода [7, 8]. Это упрощает вычисление интегральных коэффициентов нелинейного взаимодействия и определение собственных частот, что в свою очередь существенно облегчает идентификацию мод при сравнении с экспериментом. Изучение пороговых условий ЭТВ в сферических сегнетоэлектрических резонаторах составляет задачу настоящей работы.

1. УРАВНЕНИЯ ЭЛЕКТРОТЕПЛОГО ВЗАИМОДЕЙСТВИЯ

Сегнетоэлектрический резонатор представляет собой распределенную колебательную систему из материала с диэлектрической проницаемостью, имеющей выраженную зависимость от температуры. При описании электромагнитных колебаний в резонаторе будем исходить из уравнения

$$\frac{\partial^2 \mathbf{D}}{\partial t^2} + 4\pi \frac{\partial \mathbf{j}}{\partial t} + \frac{c^2}{\mu} \text{rot rot } \mathbf{E} = 0 \quad (1)$$

и полагать, что напряженность электрического поля \mathbf{E} и индукция \mathbf{D} связаны соотношением

$$\mathbf{E} = \chi(T)\mathbf{D} + \nu\mathbf{D}\mathbf{D}\mathbf{D}. \quad (2)$$

Здесь \mathbf{j} — плотность тока, μ — магнитная проницаемость диэлектрика, c — скорость света, ν — коэффициент кубической нелинейности, $\chi(T) = \chi_0(1 + a_\epsilon T)$ — зависящая от температуры величина обратной диэлектрической проницаемости, $\chi_0 = \epsilon^{-1}(0)$ (температура T отсчитывается от температуры окружающей среды). В общем случае величины χ и ν являются тензорами соответственно второго и четвертого ранга. Будем полагать, что распределение температуры в резонаторе удовлетворяет уравнению теплопроводности

$$\rho C_P \frac{\partial T}{\partial t} = \text{div}(k \text{ grad } T) + g(\mathbf{r}, t), \quad (3)$$

где ρ , C_P , k — соответственно плотность, удельная теплоемкость и теплопроводность среды, $g(\mathbf{r}, t) = \mathbf{j}\mathbf{E}$ — плотность тепловых источников. Будем считать джоулевы потери линейными, так что

$$\mathbf{j} = \sigma \mathbf{D}, \quad (4)$$

$$g(\mathbf{r}, t) = \sigma \mathbf{D} \chi_0 \mathbf{D},$$

где σ — удельная электрическая проводимость диэлектрика.

Возбуждение вынужденных электромагнитных колебаний происходит благодаря действию сторонних полей $\mathbf{E}^{\text{ст}}(\mathbf{r}, t)$, которые будем считать заданными на поверхности резонатора. Будем искать решение системы (1)–(4) в виде рядов по собственным функциям линейных краевых задач для уравнений

$$\frac{c^2}{\mu} \text{rot rot}(\chi_0 \mathbf{D}_a) = \omega_a^2 \mathbf{D}_a, \quad (5)$$

$$(\rho C_P)^{-1} \text{div}(k \text{ grad } T_s) + \lambda_s T_s = 0, \quad (6)$$

где λ_s и ω_a^2 — собственные значения соответствующей краевой задачи, с однородными граничными условиями (конкретный вид условий будет уточнен ниже). Собственные функции \mathbf{D}_a и T_s этих задач удовлетворяют соотношениям ортонормировки:

$$\int_V \chi_0 \mathbf{D}_a \sigma \mathbf{D}_b dV = \delta_{ab}, \quad \int_V \rho C_P T_s T_r dV = \delta_{sr}, \tag{7}$$

где δ_{ab} и δ_{sr} — символы Кронекера, V — объем резонатора. Таким образом, положим

$$\begin{aligned} \mathbf{D}(\mathbf{r}, t) &= \sum_a x_a(t) \mathbf{D}_a(\mathbf{r}), \\ T(\mathbf{r}, t) &= \sum_s \eta_s(t) T_s(\mathbf{r}). \end{aligned} \tag{8}$$

Подставим (2), (4) и (8) в уравнения (1) и (3), после чего умножим их соответственно на $\chi_0 \mathbf{D}_f(\mathbf{r})$ и $T_s(\mathbf{r})$ и проинтегрируем по объему V резонатора с учетом условий ортогональности (7). В результате получим цепочку уравнений, описывающую временные зависимости $x_a(t)$ и $\eta_a(t)$:

$$\ddot{x}_f + \omega_f^2 x_f = F_f(t) - R_{fa} \dot{x}_f - H_{faa} \eta_a x_a - N_{fabc} x_a x_b x_c, \tag{9}$$

$$\dot{\eta}_s = -\lambda_s \eta_s + K_{sab} x_a x_b. \tag{10}$$

Здесь точки обозначают дифференцирование по времени, по повторяющимся индексам подразумевается суммирование. В уравнениях (9), (10) обобщенные координаты x_f соответствуют электромагнитным, а η_s — температурным степеням свободы, R_{fa} — диссипативные коэффициенты:

$$R_{ab} = 4\pi \int_V \chi_0 \mathbf{D}_a \sigma \mathbf{D}_b dV. \tag{11}$$

В однородном диэлектрике они отличны от нуля только при $f = a$, причем $R_{ff} = \omega_f / Q_f$, где Q_f — добротность f -й моды резонатора. Эффективные сторонние силы определяются выражением

$$F_f(t) = -\frac{c^2}{\mu} \oint_{\Sigma} \left(\mathbf{E}^{ct}(\mathbf{r}, t) \times \text{rot}(\chi_0 \mathbf{D}_f) \right) \mathbf{n} dS, \tag{12}$$

где \mathbf{n} — единичный вектор внешней нормали к поверхности резонатора Σ . Интегральные коэффициенты взаимодействия имеют вид

$$N_{fabc} = \frac{c^2}{\mu} \int_V \chi_0 \mathbf{D}_f \text{rot rot}(\nu \mathbf{D}_a \mathbf{D}_b \mathbf{D}_c) dV, \tag{13}$$

$$K_{sab} = 4\pi \int_V T_s \sigma \mathbf{D}_a \chi_0 \mathbf{D}_b dV, \tag{14}$$

$$H_{sab} = \frac{c^2}{\mu} \int_V \chi_0 \mathbf{D}_b \text{rot rot}(\chi_0 a_\varepsilon T_s \mathbf{D}_a) dV. \tag{15}$$

2. ДИНАМИКА ЭЛЕКТРОТЕПЛОВОЙ НЕУСТОЙЧИВОСТИ В ДВУХМОДОВОМ РЕЖИМЕ

Переход от уравнений Максвелла и теплопроводности к цепочке уравнений (9), (10) позволяет раздельно исследовать динамику колебаний и зависимость интегральных коэффициентов взаимодействия от материальных параметров среды и распределений электромагнитного поля и температуры в резонаторе. В указанной цепочке совместные колебания электромагнитного поля и температуры описываются как взаимодействие совокупности нелинейно связанных осцилляторов, представляющих электромагнитные моды, и релаксационных звеньев, представляющих тепловые степени свободы. При этом, по крайней мере вблизи порога неустойчивости, имеют значительную амплитуду и эффективно участвуют в нелинейном взаимодействии лишь колебания, соответствующие небольшому числу степеней свободы (обычно двум или трем). Соответственно, можно рассматривать двух- или трехмодовый варианты электротеплового возбуждения.

В двухмодовом варианте в резонаторе эффективно взаимодействуют одна из электромагнитных мод и одна из мод температурной релаксации. Воздействие накачки на частоте, близкой к резонансной ($F(t) = F_0 \cos(\omega t)$, где $\omega \approx \omega_f$), приводит к нестационарному нагреву диэлектрика и, соответственно, к нестационарной расстройке резонатора. Как следствие, в спектре вынужденных электромагнитных колебаний появляются новые составляющие, которые влияют на временную зависимость температуры. Нестационарная тепловая расстройка может приобрести периодический характер с частотой Ω_T , если в результате взаимодействия с колебаниями накачки в системе возникнут спектральные составляющие электрического поля на разностной и суммарной частотах ($\omega \pm \Omega_T$) с такими амплитудами и фазами, что в результате их взаимодействия с накачкой произойдет усиление колебаний температуры на частоте Ω_T .

Чтобы исследовать двухмодовый режим ЭТВ, достаточно взять из полной цепочки (9), (10) одно уравнение вида (9) для электромагнитного осциллятора и одно уравнение вида (10) для температурного релаксатора. Анализ устойчивости колебаний в такой системе при одночастотной резонансной накачке был выполнен в [3]. Установлено, что в рассматриваемой системе существуют две области параметров, где происходит потеря устойчивости вынужденных колебаний. Параметрами задачи являются нормированная амплитуда A колебаний и нелинейная расстройка ξ_n . В [3] показано, что при условии

$$A^2 \geq \frac{2}{3\gamma \pm 4\theta_t^{-1}} \frac{\xi_n^2 + 4\theta_e^2}{\xi_n} \quad (16)$$

имеет место аperiodическая неустойчивость, а при условии

$$A^2 \geq \frac{2}{3\gamma \mp 2\theta_e^{-1}} \frac{\xi_n^2 + 4(\theta_e + \theta_t)^2}{\xi_n} \quad (17)$$

— колебательная неустойчивость. Здесь $\gamma = \omega N_{ffff}/(K_{sff}H_{ffs})$ — коэффициент диэлектрической нелинейности, $\theta_t = \lambda_s/\omega$ и $\theta_e = R_{ff}/(2\omega)$ — коэффициенты затухания тепловых и электромагнитных колебаний. Полная расстройка ξ_n содержит три слагаемых: обычную линейную расстройку $\xi_f = 1 - \omega_f^2/\omega^2$, нелинейную расстройку $-3\gamma A^2/4$, обусловленную диэлектрической нелинейностью, и тепловую расстройку A^2/θ_t . Таким образом,

$$\xi_n = \xi_f - A^2 (3\gamma/4 \pm \theta_t^{-1}). \quad (18)$$

Верхние знаки в (16)–(18) следует брать при $a_\varepsilon > 0$, а нижние — при $a_\varepsilon < 0$. Для неполярной фазы сегнетоэлектриков $a_\varepsilon > 0$.

Аperiodическая неустойчивость в области (16) имеет статический характер и обусловлена “опрокидыванием” резонансных кривых при совместном действии электрических и тепловых расстроек [9].

Возникновение колебательной неустойчивости является результатом тепловой обратной связи и происходит по описанному выше сценарию.

Формулу для пороговой мощности колебательной неустойчивости нетрудно получить, подставив в (17) выражение [3] для нормированной амплитуды A через мощность диэлектрических потерь P :

$$A^2 = 8\pi Q_f K_{sff} H_{ffs} P / \omega^4. \quad (19)$$

Запишем формулу пороговой мощности, разделив зависимости от интегральных коэффициентов и настройки колебательной системы:

$$P_{\text{пор}} = \frac{\omega^4}{8\pi Q_f^3 K_{sff} H_{ffs}} \frac{1 + \theta_t / \theta_e}{1 \mp 3\gamma / 4Q_f} \frac{1 + \bar{\xi}^2}{2\bar{\xi}}, \quad (20)$$

где $\bar{\xi} = \xi_n / \xi_0$, $\xi_0 = 2(\theta_t + \theta_e)$ — оптимальная настройка. При этом вблизи порога неустойчивости частота колебаний температуры (частота автомодуляции) равна

$$\Omega_T = \omega \left[\frac{\xi_n^2 / 4 + (\theta_t + \theta_e)^2}{1 \mp 3\gamma / 4Q_f} - \theta_t^2 \right]^{1/2}. \quad (21)$$

Примечательной особенностью электротеплового возбуждения является то, что пороговая мощность и частота автомодуляции существенно зависят не только от произведения интегральных коэффициентов $K_{sff} H_{ffs}$, ответственных за электротепловую связь, но и от интегрального коэффициента кубической диэлектрической нелинейности N_{fff} . При этом ориентировочные оценки, сделанные без учета неоднородности распределения электромагнитного поля и температуры, показывают, что в СВЧ резонаторах из SrTiO₃ и KTaO₃ величина γ / Q_f порядка единицы. В таких условиях точный расчет интегральных коэффициентов особенно необходим, поскольку от их отношения зависит сама возможность наблюдения колебательного ЭТВ. Так, если $\gamma \ll \theta_e^{-1}$, $\gamma \ll \theta_t^{-1}$, то колебательная и аperiodическая неустойчивости имеют место на разных склонах резонансных кривых. Возникновение автомодуляции в области (17) при $a_\varepsilon > 0$ происходит при отрицательной расстройке, когда частота накачки ω меньше резонансной частоты. Этим тепловая неустойчивость отличается от стрикционного параметрического возбуждения, наблюдаемого при $\xi > 0$. Если же $|\gamma|$ существенно превышает θ_e^{-1} и θ_t^{-1} , то области неустойчивости (16) и (17) находятся на одном склоне резонансной кривой, и в типичных для сегнетоэлектрических резонаторов условиях ($\gamma > 0$, $a_\varepsilon > 0$) область колебательной неустойчивости оказывается поглощенной областью аperiodической неустойчивости, которая и наблюдается в эксперименте.

Сравнение двухмодовых режимов ЭТВ и СПВ показывает, что в динамике этих эффектов имеется много общего. В обоих случаях нелинейно связанными оказываются колебания, частоты которых различаются на несколько порядков, причем нелинейность носит одинаковый (квадратичный) характер. Вместе с тем, между СПВ и ЭТВ имеется очевидное различие: акустические колебания при СПВ имеют выраженный резонансный характер, а температурные колебания при ЭТВ носят нерезонансный релаксационный характер. Как следствие, частота Ω_T колебаний температуры в зависимости от настройки системы может меняться в довольно широких пределах.

3. РАСЧЕТ ИНТЕГРАЛЬНЫХ КОЭФФИЦИЕНТОВ

Эффективность межмодового взаимодействия определяется интегральными коэффициентами K_{sab} , H_{sab} , N_{fabc} . Используя теорему Гаусса, величины H_{sab} и N_{fabc} можно представить в виде

$$N_{fabc} = N_{fabc}^{(v)} + N_{fabc}^{(s)}, \quad H_{sab} = H_{sab}^{(v)} + H_{sab}^{(s)}. \quad (22)$$

Первые слагаемые в равенствах (22) представляют вклад объемного взаимодействия:

$$N_{fabc}^{(v)} = \omega_f^2 \int_V \mathbf{D}_f \nu \mathbf{D}_a \mathbf{D}_b \mathbf{D}_c dV, \quad (23)$$

$$H_{sab}^{(v)} = \chi_0 a_\varepsilon \omega_f^2 \int_V T_f \mathbf{D}_a \mathbf{D}_c dV,$$

а вторые определяются значениями полей на поверхности резонатора

$$N_{fabc}^{(s)} = \frac{c^2}{\mu} \oint_S \left[(\chi_0 \mathbf{D}_f \times \text{rot } \nu \mathbf{D}_a \mathbf{D}_b \mathbf{D}_c) - (\nu \mathbf{D}_a \mathbf{D}_b \mathbf{D}_c \times \text{rot } \chi_0 \mathbf{D}_f) \right] \mathbf{n} dS, \quad (24)$$

$$H_{fac}^{(s)} = \frac{a_\varepsilon \chi_0^2 c^2}{\mu} \oint_S (\mathbf{D}_f \times (\text{grad } T_f \times \mathbf{D}_a)) \mathbf{n} dS.$$

Здесь S включает в себя поверхность резонатора, а также те поверхности, на которых диэлектрические характеристики χ_0 и ν терпят разрыв. Для нахождения интегральных коэффициентов требуется знать собственные функции колебательной системы, т. е. решения краевых задач для уравнений (5), (6) с соответствующими граничными условиями. Граничные условия для электромагнитных колебаний в общем случае состоят в непрерывности тангенциальных составляющих электрического и магнитного полей на поверхности диэлектрика. В сегнетоэлектрическом резонаторе ($\varepsilon \gg 1$) в волноводном тракте эффективно возбуждаются ТЕ-моды, для которых приближенное граничное условие имеет вид [10]

$$\mathbf{E}_\tau|_\Sigma = 0. \quad (25)$$

Будем полагать также, что для температуры имеет место граничное условие третьего рода

$$\frac{\partial T}{\partial r} + hT|_\Sigma = 0, \quad (26)$$

где h — коэффициент теплообмена.

Ограничиваясь здесь анализом двухмодового режима ЭТВ, рассмотрим только интегральные коэффициенты для H -мод, у которых индексы f , a , b и c одинаковы. С учетом условия (25) нетрудно установить, что в этом случае $N_{fabc}^{(s)} = 0$ и $H_{fac}^{(s)} = 0$, а вклад $H_{fac}^{(v)}$ пропорционален K_{sab} , так что

$$K_{sab} = \frac{\sigma}{\chi_0} S_{sab}, \quad H_{sab} = \frac{a_\varepsilon \omega_f^2}{\chi_0} S_{sab}, \quad N_{fabc}^{(v)} = \frac{\nu \omega_f^2}{\chi_0^4} C_{fabc}, \quad (27)$$

где

$$S_{sab} = \int_V T_s \mathbf{E}_a \mathbf{E}_b dV, \quad C_{fabc} = \int_V \mathbf{E}_f \mathbf{E}_a \mathbf{E}_b \mathbf{E}_c dV. \quad (28)$$

В случае сферического резонатора собственные функции линейных краевых задач для уравнений (5), (6) с граничными условиями (25), (26) хорошо известны [7, 8]. Выражения для компонент электрического поля и температуры представляют собой произведения функций Риккати—Бесселя 1-го рода (ФРБ1) от радиальной координаты на сферические гармоники $Y_n^{(m)}(\theta, \varphi)$ — функции угловых координат θ, φ . Таким образом, например, в качестве \mathbf{E}_a в (28) следует брать распределение электрического поля собственной моды сферического резонатора H_{mnl} или E_{mnl} , для которой $m = m_a, n = n_a, l = l_a$. Собственную функцию соответствующей тепловой задачи

$$T_{mnl}(r, \theta, \varphi) = \frac{1}{N_{mnl}\sqrt{r}} J_{n+1/2}(\mu_l^{n+1/2} r/a) Y_n^{(m)}(\theta, \varphi) \quad (29)$$

будем для краткости называть температурной модой и обозначать T_{mnl} . Здесь $J_{n+1/2}$ — функция Бесселя 1-го рода, N_{mnl} — нормировочный коэффициент. Для моды w , участвующей во взаимодействии, индексы n, m, l принимают значения n_w, m_w, l_w . Величину μ_l^n находят, решая трансцендентное уравнение

$$\mu_k^{n+1/2} J'_{n+1/2}(\mu_k^{n+1/2}) + (ah - 1/2) J_{n+1/2}(\mu_k^{n+1/2}) = 0, \quad (30)$$

которое получается при подстановке (29) в граничное условие (26).

Для сферического резонатора интегралы (28) сводятся к сумме произведений одномерных интегралов по радиальной и угловым координатам, аналогично тому, как это имеет место в случае электростриционного взаимодействия [11]. Наиболее простую структуру имеют интегралы по азимутальному углу φ , которые легко берутся аналитически. Рассматривая их, находим, что электротепловое взаимодействие в двухмодовом режиме может происходить либо с участием аксиально-симметричной температурной моды, для которой распределение температуры и электрического поля не зависит от φ ($m_w = 0$), либо между модами, для которых $m_w = 2m_a$.

Поскольку одномерные интегралы по θ имеют структуру, которая часто встречается в квантовой теории углового момента [12], можно, исследуя их, найти дополнительные ограничения на индексы m и n . В частности, из соображений симметрии очевидно, что индекс n_w температурной моды должен быть четным. При возбуждении сегнетоэлектрических резонаторов СВЧ представляют интерес электромагнитные моды типа H_{mnl} с небольшим индексом n_a ($n_a \leq 5$). Для таких мод подынтегральное выражение представляет собой тригонометрический полином невысокой степени относительно переменной θ , и соответствующий интеграл можно точно вычислить по квадратурной формуле [13].

Расчет одномерных интегралов по радиальной координате является наиболее трудоемким этапом вычисления интегральных коэффициентов. В настоящей работе для вычисления ФРБ1 и функций Бесселя, входящих в подынтегральные выражения, был использован алгоритм, предложенный в [14] и дополненный асимптотическими разложениями ФРБ1 в нуле [15]. Интегрирование производилось методом Симпсона.

Результаты расчета интегральных коэффициентов для ряда комбинаций электромагнитных и температурных мод представлены в табл. 1. Приведенные численные значения получены для сферических резонаторов, имеющих радиус $a = 0,5$ мм. В табл. 1 представлены также результаты вычисления интегральных коэффициентов C_{aaaa} , описывающих влияние нелинейных расстройек электрического происхождения.

В изотропной среде без дисперсии (жидкость, поликристалл) тензор кубической диэлектрической нелинейности имеет одну компоненту, при этом

$$N_{aaaa}^{(v)} = \omega_a^2 \chi_0^{-4} \nu C_{aaaa}. \quad (31)$$

Для монокристаллов KTaO_3 и SrTiO_3 , имеющих кубическую симметрию, тензор ν содержит две независимых компоненты, и соотношение (31) перестает быть точным. Однако более детальное рассмот-

Таблица 1

Параметры взаимодействия электромагнитных и температурных мод в сферических резонаторах.

Комбинации мод	Интегральный коэффициент C_{ffff}	Интегральный коэффициент S_{fss}	C_{ffff}/S_{fss}^2
$T_{mnl}-H_{mnl}$			
$T_{001}-H_{011}$	2400,01	34,76	1,98
$T_{021}-H_{011}$	2400,01	-18,75	6,82
$T_{221}-H_{111}$	900,05	97,47	0,09
$T_{222}-H_{111}$	900,05	-21,66	1,93
$T_{242}-H_{111}$	900,05	-0,34	$7,78 \cdot 10^3$
$T_{282}-H_{111}$	900,05	4,45	45,5
$T_{221}-H_{121}$	1619,01	97,26	0,17
$T_{441}-H_{221}$	1992,76	123,21	0,13
$T_{001}-H_{021}$	2988,99	30,74	3,16
$T_{442}-H_{221}$	1992,76	-21,10	4,43
$T_{221}-H_{131}$	2398,88	85,38	0,33
$T_{441}-H_{231}$	1706,74	101,52	0,17
$T_{442}-H_{231}$	1706,74	-31,36	1,74
$T_{441}-H_{241}$	2136,46	86,63	0,28
$T_{221}-H_{122}$	3128,6	107,76	0,27
$T_{442}-H_{222}$	3850,71	107,20	0,33
$T_{881}-H_{441}$	4029,88	160,86	0,16

рение показывает, что для Н-мод поправки, обусловленные анизотропией и вкладом $N_{aaaa}^{(s)}$, по крайней мере, на порядок меньше величины (31).

4. ОЦЕНКИ ПОРОГОВОЙ МОЩНОСТИ. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Используя введенные выше коэффициенты S_{waa} и C_{aaaa} , можно записать формулы (20) и (21) для основных динамических характеристик ЭТВ в виде, в котором разделены зависимости от пространственного распределения взаимодействующих полей и от материальных констант диэлектрика. Так, для пороговой мощности имеем

$$P_{\text{пор}} = \frac{1}{2} \frac{\rho C_P \chi_0}{a_\epsilon Q_f^2} \frac{1}{S_{sff}^2} \frac{1 + \theta_t/\theta_e}{1 \mp 3\gamma/4Q_f} \frac{1 + \bar{\xi}^2}{2\bar{\xi}}, \quad (32)$$

при этом

$$\frac{\gamma}{Q_f} = \frac{\epsilon^2 \nu \rho C_P}{a_\epsilon} \frac{C_{ffff}}{S_{sff}^2}. \quad (33)$$

Выясним вначале, при каких условиях наблюдение колебательной электротепловой неустойчивости принципиально осуществимо. Поскольку в монокристаллах KTaO_3 и SrTiO_3 $a_\epsilon > 0$ и $\nu > 0$, то в (32) следует брать знак минус, при этом в соответствии с результатами п. 2 колебательный режим электротепловой неустойчивости возможен только при

$$\gamma/Q_f \leq 4/3. \quad (34)$$

Если в (34) имеет место равенство, то $P_{\text{пор}} \rightarrow \infty$. Осуществимость этого условия зависит прежде всего от материального фактора $M(T) = \epsilon^2 \nu \rho C_P / a_\epsilon$. Оценим величину $M(T)$ при температуре

Таблица 2

Пороговая мощность и частота электротепловой автомодуляции при оптимальной расстройке для ДР из КТaO₃

Комбинации мод ($T_{mnl}-H_{mnl}$)	Собственные частоты ω_f , 10^{10} Гц	Оптимальная расстройка ξ_0 , 10^{-5}	Пороговая мощность $P_{\text{пор}}$, Вт	Частота автомодуляции Ω_T , МГц
$T_{221}-H_{111}$	2,86	8,35	$8,36 \cdot 10^{-3}$	1,984
$T_{221}-H_{121}$	4,10	7,98	$6,65 \cdot 10^{-3}$	3,286
$T_{441}-H_{221}$	4,10	8,84	$9,27 \cdot 10^{-4}$	3,256
$T_{221}-H_{131}$	5,26	7,80	$1,14 \cdot 10^{-1}$	10,721
$T_{441}-H_{231}$	5,26	8,46	$4,84 \cdot 10^{-4}$	4,388
$T_{441}-H_{241}$	6,37	8,24	$9,90 \cdot 10^{-4}$	9,288
$T_{221}-H_{112}$	5,73	7,75	$1,73 \cdot 10^{-1}$	6,000
$T_{221}-H_{122}$	7,04	5,79	$6,17 \cdot 10^{-2}$	8,409
$T_{222}-H_{112}$	5,73	8,64	$1,34 \cdot 10^{-1}$	5,896
$T_{442}-H_{222}$	7,04	9,17	$3,68 \cdot 10^{-1}$	4,478
$T_{881}-H_{441}$	6,37	9,61	$9,01 \cdot 10^{-7}$	5,956

жидкого гелия, используя известные значения параметров материалов. Для SrTiO₃ $\varepsilon = 2,5 \cdot 10^4$, $a_\varepsilon = 5 \cdot 10^{-3} \text{ K}^{-1}$ [5], $C_P/T^3 = 4,5 \cdot 10^{-6} \text{ Дж}/(\text{г} \cdot \text{K}^4)$ [16], $k = 6 \text{ мВт}/(\text{см}^2 \cdot \text{K})$ [17], $\rho = 5,13 \text{ г}/\text{см}^3$ [16], коэффициент нелинейности $\beta_1 = 6 \cdot 10^{-12} \text{ СГС}$, где $\beta_1 = (4\pi)^3 \nu$ [18]. Для КТaO₃ $\varepsilon = 4,3 \cdot 10^3$, $a_\varepsilon = 5 \cdot 10^{-3} \text{ K}^{-1}$ [5], $C_P/T^3 = 1,48 \cdot 10^{-6} \text{ Дж}/(\text{г} \cdot \text{K}^4)$ [16], $k = 10 \text{ мВт}/(\text{см}^2 \cdot \text{K})$ [19], $\rho = 6,97 \text{ г}/\text{см}^3$ [16], $\beta_1 = 1,1 \cdot 10^{-11} \text{ СГС}$ [18]. Отсюда находим для КТaO₃ $M(4,2\text{K}) \approx 0,4$ и для SrTiO₃ $M(4,2\text{K}) \approx 1,1$. Как видно из табл. 1, величины коэффициентов электротепловой связи и отношения C_{ffff}/S_{sff}^2 могут варьироваться в широких пределах. Вместе с тем, имеется довольно много сочетаний мод, для которых это отношение лежит в пределах $0,1 \div 0,3$. Для таких мод при гелиевой температуре возможен колебательный режим ЭТВ в диэлектрическом резонаторе (ДР) из КТaO₃.

В табл. 2 представлены результаты расчета пороговой мощности ЭТВ для резонаторов из КТaO₃. В расчетах использовались вычисленные ранее величины интегральных коэффициентов (табл. 1) и приведенные выше значения материальных констант. Кроме того, было принято, что $Q_f = 10^4$; это является удовлетворительной оценкой добротности ДР из КТaO₃ во всем диапазоне СВЧ [6]. Приведенные в таблице значения $P_{\text{пор}}$ и Ω_T соответствуют оптимальной расстройке ξ_0 .

В условиях эксперимента при фиксированной частоте накачки резонанс одновременно достигается для мод H_{mnl} , имеющих одинаковые индексы n и l , но различающихся по азимутальному индексу m . При повышении мощности накачки возбуждение температурных колебаний будет наблюдаться на той температурной моде, для которой имеется эффективная электротепловая связь с колебаниями H_{mnl} , а порог неустойчивости имеет наименьшее значение. Так, при возбуждении резонатора на основной частоте (моды $H_{011}-H_{111}$) наименьшая пороговая мощность достигается для комбинаций $T_{221}-T_{111}$. При возбуждении резонатора на более высокочастотных модах низкие пороги ЭТВ (менее 1 мВт) имеют место для моды T_{441} (в сочетаниях $T_{441}-H_{221}$, $T_{441}-H_{231}$).

Расчеты указывают на то, что при накачке ДР на модах более высокого порядка возможно еще более низкое значение порога колебательной неустойчивости, вплоть до 10^{-6} Вт. Примером этого может служить комбинация мод $T_{881}-H_{441}$. Малые значения порога достигаются для мод, поля которых сконцентрированы в поверхностном слое резонатора, так что эффективный объем моды оказывается существенно меньше, чем геометрический объем резонатора. Отметим, однако, что для мод высших порядков в сегнетоэлектрических резонаторах весьма трудно добиться достаточно большого коэффициента связи с СВЧ трактом.

Приведенные в табл. 2 результаты расчета носят ориентировочный характер, поскольку величины теплоемкости и температурного коэффициента диэлектрической проницаемости при гелиевых температурах сильно подвержены влиянию примесей [20–22]. В то же время точность, с которой вычислены интегральные коэффициенты, достаточно высока, чтобы установить, какие моды эффективно участвуют в электротепловом взаимодействии.

Колебательная электротепловая неустойчивость в сегнетоэлектрических резонаторах СВЧ — принципиально низкотемпературный эффект, что роднит его с другим низкотемпературным эффектом — возбуждением второго звука в виртуальных сегнетоэлектриках оптической волной [23, 24]. Действительно, для виртуальных сегнетоэлектриков эксперимент и теория дают $a_\epsilon \sim T$, $C_p \sim T^3$, ϵ при гелиевых температурах меняется незначительно, так что материальный параметр $M(T)$ растет пропорционально квадрату температуры. При превышении температуры над уровнем 4,2 К условие (34) довольно быстро перестает выполняться, что приводит к неограниченному росту $P_{\text{пор}}$, так что ЭТВ становится невозможным. Для КТaO₃ даже при наилучших сочетаниях мод наблюдение ЭТВ оказывается возможным лишь при $T \leq 10$ К. Еще более жесткие ограничения имеют место для SrTiO₃.

С другой стороны, в “классических” сегнетоэлектриках, где диэлектрическая аномалия имеет место при более высоких температурах, справедлив закон Кюри—Вейсса: $\epsilon = C_{\text{КВ}}/(T - T_e)$, где T_e — температура фазового перехода, $C_{\text{КВ}}$ — константа Кюри—Вейсса. Поскольку зависимость теплоемкости от температуры здесь незначительна, то

$$M(T) \propto \frac{\epsilon^2}{a_\epsilon} = C_{\text{КВ}}^2/(T - T_e).$$

Это значит, что в обычных сегнетоэлектриках по мере приближения к области диэлектрической аномалии возможности реализации колебательного режима неустойчивости ухудшаются, так что реально можно ожидать возникновения только аperiodической тепловой неустойчивости.

Своеобразие эффекта ЭТВ состоит в том, что хотя релаксация температуры обусловлена теплопроводностью кристаллов, однако частота релаксации в резонаторах СВЧ слабо зависит от величины коэффициента теплопроводности, а в гораздо большей степени зависит от характерного времени затухания электромагнитных колебаний. В резонаторах с принятыми для расчетов размерами не только частота автомодуляции, но и пороговая мощность самовозбуждения должны слабо зависеть от величины теплопроводности и граничных условий теплоотвода. На рис. 1 показана рассчитанная зависимость частоты автомодуляции от коэффициента теплообмена h : кривая 1 — для мод $T_{221}-H_{111}$, кривая 2 — для мод $T_{221}-H_{121}$. Соответствующая зависимость пороговой мощности ЭТВ показана на рис. 2. Можно было бы ожидать, что влияние граничных условий гораздо сильнее скажется на условиях возбуждения основной моды T_{001} температурной релаксации. Однако, как видно из табл. 1, отношение C_{ffff}/S_{sff}^2 для этой моды неблагоприятно для ее возбуждения.

Представляет интерес вопрос о том, существуют ли условия, при которых частота и пороговая мощность автомодуляции будут существенно зависеть от времени тепловой релаксации. Поскольку $\theta_t = 1/\tau_t \propto 1/a^2$, а $\theta_e \propto \omega \propto 1/a^2$, такие условия следует искать в резонаторах малых размеров $a \ll 0,1$ мм, что само по себе представляет существенное затруднение. Более того, резонансные частоты электромагнитных колебаний уходят при этом в диапазон субмиллиметровых волн. Электротепловое возбуждение в микрорезонаторах указанных размеров может иметь место в других физических системах — сферических частицах капельного аэрозоля [3]. Однако спецификой аэрозольных частиц является то, что высокочастотные электромагнитные колебания, для которых имеет смысл наблюдать ЭТВ, реализуются на высших типах мод (моды шепчущей галереи). Этот круг вопросов требует дополнительного исследования.

Примечательным обстоятельством является то, что для ДР из КТaO₃ частоты автомодуляции при ЭТВ должны быть одного порядка с частотами акустических колебаний, возбуждаемых при СПВ (это,

как правило, единицы МГц). Вместе с тем, имеется ряд отличий, которые помогают установить механизм возбуждения. Во-первых, СПВ имеет место при положительных расстройках ($\xi > 0$ [1]), а ЭТВ следует ожидать при $\xi < 0$ (см. (17)). Во-вторых, поскольку температурные колебания имеют релаксационный характер, для них зависимость частоты автомодуляции от расстройки выражена существенно более, чем для акустических колебаний. Наконец, пороговая мощность возбуждения акустических колебаний, как правило, ниже, чем температурных, и составляет $10^{-4} \div 10^{-6}$ Вт для большинства мод. Последнее обстоятельство объясняет, почему ЭТВ не было обнаружено наряду с СПВ в экспериментах. Наличие альтернативного нелинейного эффекта могло не только маскировать ЭТВ, но и препятствовать его проявлению. Выше указывалось, что в сферических СВЧ резонаторах возбуждение электромагнитных колебаний, эффективно взаимодействующих с колебаниями температуры при одночастотной накачке, возможно одновременно на нескольких модах. Это приводит к усложнению динамики колебаний и также требует дополнительного анализа, однако можно ожидать, что при этом будут найдены условия, при которых порог ЭТВ оказывается более низким и его экспериментальная реализация не составит труда.

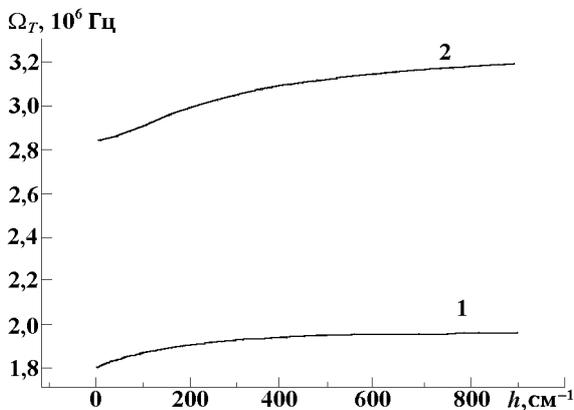


Рис. 1

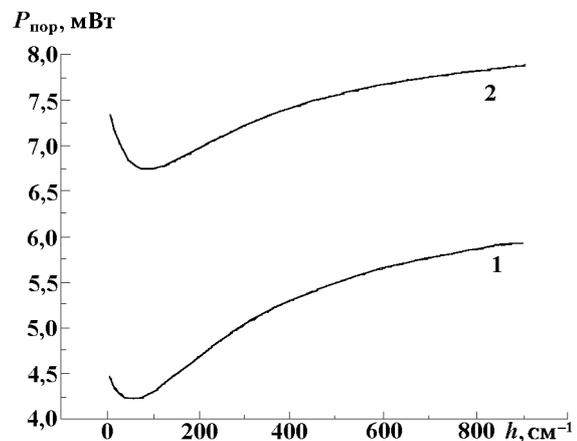


Рис. 2

ЛИТЕРАТУРА

1. Белокопытов Г. В., Иванов И. В., Семененко В. Н. // Изв. АН СССР. Сер. физ. 1990. Т. 54, № 4. С. 621.
2. Belokopytov G. V., Zhuravlev M. V. // Abstracts of the 9th International Meeting on Ferroelectricity. Seoul, Korea, 1997. P. 21.
3. Белокопытов Г. В. // Вест. Моск. ун-та. Сер.3: Физика, астрон. 1997. Т. 38, № 3. С. 11.
4. Старунов В. С., Фабелинский И. А. // УФН. 1969. Т. 98. С. 441.
5. Иванов И. В., Бузин И. М., Белокопытов Г. В., Сычев В. М., Чупраков В. Ф. // Изв. вуз. Физика. 1981. Т. 24, № 8. С. 6
6. Белокопытов Г. В., Иванов И. В., Катанов С. И., Моисеев Н. Н., Сырников П. П. // ФТТ. 1982. Т. 24, № 6. С. 1865.
7. Стрэттон Д. А. Теория электромагнетизма. — М.-Л.: ОГИЗ, 1948.
8. Свешников А. Г., Боголюбов А. Н., Кравцов В. В. Лекции по математической физике. — М.: МГУ, 1993.
9. Белокопытов Г. В. // Изв. вуз. Радиофизика. 1986. Т. 29, № 11. С. 1324.
10. Бурсиан Э. В. Нелинейный кристалл титанат бария. — М.: Наука, 1974.

11. Белокопытов Г. В., Пушечкин Н. П., Семенов В. Н. // Вест. Моск. ун-та. Сер.3: Физика, астрон. 1992. Т. 33, № 5. С. 18.
12. Варшалович Д. А., Москалев А. Н., Херсонский В. К. Квантовая теория углового момента. — Л., 1975.
13. Крылов В. И. Приближенное вычисление интегралов. — М., 1967.
14. Бабушкина Л. В., Керимов М. К., Никитин А. И. // ЖВМ и МФ. 1988. Т. 28, № 12. С. 1779.
15. Белов Н. Н., Журавлев М. В., Молодченко В. Л., Пьянков Е. Е. // Оптика и спектр. 1995. Т. 78, № 6. С. 885.
16. Landolt—Bornstein numerical data and functional relationships in science and technology. New Ser. Group III. V. 16a. Ferro and antiferroelectric substances. — N.Y.: Springer Verlag, 1981. — 584 p.
17. Suemune Y. J. // Phys. Soc. Japan. 1965. V. 20. P. 174.
18. Белокопытов Г. В., Моисеев Н. Н. // Изв. вуз. Радиофизика. 1982. Т. 25, № 10. С. 1210.
19. Salce B., Gravi J. L., Boatner L. A. // J. Phys. Condens. Matter. V. 6, № 22. P. 4077.
20. Lawless W. N. // Phys. Rev. B. 1976. V. 11, № 1. P. 134.
21. Steigmeir E. F. // Phys. Rev. 1968. V. 168, № 2. P. 532.
22. Belokopytov G. V. // Ferroelectrics. 1995. V. 168. P. 69.
23. Гуревич В. Л., Таганцев А. К. // ЖЭТФ. 1988. Т. 94, № 1. С. 370.
24. Nehlen B., Perou A. L., Courtens E., Vacher R. // Phys. Rev. Lett. 1995. V. 75, № 12. P. 2416.

Физический факультет
МГУ им. М.В. Ломоносова,
г. Москва, Россия

Поступила в редакцию
24 мая 1999 г.

ELECTROTHERMAL OSCILLATORY INSTABILITY IN FERROELECTRIC RESONATORS. TWO-MODE CASE

G. V. Belokopytov and M. V. Jouravlev

We consider the equations describing interaction of electromagnetic oscillations and temperature in a nonlinear dielectric resonator. The existence conditions for the oscillatory instability are analyzed. In the simplest two-mode case, we calculate the threshold values of power and automodulation frequency at which the electrothermal excitation of microwave virtual-segnetoelectric resonators takes place. Nonlinearity of the dielectric is shown to principally limit the appearance of the electrothermal instability.

УДК 537.862:530.182

ПРЕДЕЛЬНЫЕ ЦИКЛЫ НЕЛИНЕЙНОЙ РЕКУРСИВНОЙ ЦИФРОВОЙ СИСТЕМЫ ВТОРОГО ПОРЯДКА С УСЕЧЕНИЕМ ПО МОДУЛЮ ПОСЛЕ СЛОЖЕНИЯ

Ю. А. Брюханов, Е. А. Мясников

Исследованы свободные колебания в цифровой системе второго порядка с усечением по модулю после сложения и нелинейностью насыщения. Обнаружено наличие предельных циклов. Даны аналитические и экспериментальные оценки их периодов и условий существования.

Рассматриваются свободные колебания в цифровой рекурсивной системе второго порядка прямой формы реализации [1]. Данная система состоит из сумматора с нелинейной характеристикой, двух элементов задержки z^{-1} и умножителей a и b (см. рис. 1а). Вид нелинейности сумматора зависит от числа двоичных разрядов, метода представления чисел (в прямом, обратном или дополнительном коде), способа квантования результатов арифметических операций (усечение по модулю или по величине, округление) и типа глобальной нелинейности сумматора (насыщение, переполнение или обнуление) [2]. Под глобальной нелинейностью понимается нелинейность сумматора без учета эффектов квантования. В развитие работы [3], в которой рассматривается трехуровневое квантование, исследуется случай усечения результатов суммирования по модулю при многоуровневом квантовании. Полагается, что сумматор имеет характеристику с насыщением.

Исследуемое сочетание глобальной нелинейности, способа квантования и формы реализации обладает таким практически важным преимуществом по сравнению с другими вариантами, как наибольшая площадь области глобальной устойчивости системы. Кроме того, прямая форма реализации является канонической по количеству умножителей и сумматоров, а переход к иным формам приводит к изменению разрядности регистров на $1 \div 2$ бита [1].

Устойчивость систем с усечением по модулю после сложения, но без учета насыщения, исследовалась различными авторами при помощи разнообразных методов [4, 5]. Разработанные аналитические и алгоритмические методы позволили получить некоторые оценки границ области глобальной устойчивости вышеописанной системы. В литературе, посвященной теории колебаний в нелинейных системах дискретного времени, принято разделять предельные циклы большого и малого уровней. Под циклом большого уровня понимается периодическое колебание с максимальной амплитудой, обусловленное глобальной нелинейностью и нелинейностью квантования. Циклы малого уровня обусловлены только нелинейностью квантования. В работе [5] приводится доказательство отсутствия предельных циклов малого уровня при параметрах системы, лежащих внутри треугольника устойчивости.

Определим области устойчивости и существования предельных циклов. Свободные колебания в цифровой системе второго порядка с одним квантователем усечения по модулю после сложения описываются разностным уравнением

$$x_{n+2} = f(ax_{n+1} + bx_n),$$

где n — номер итерации, или соответствующим двумерным двухпараметрическим отображением $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ вида

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} y_n \\ f(bx_n + ay_n) \end{pmatrix}, \quad (1)$$

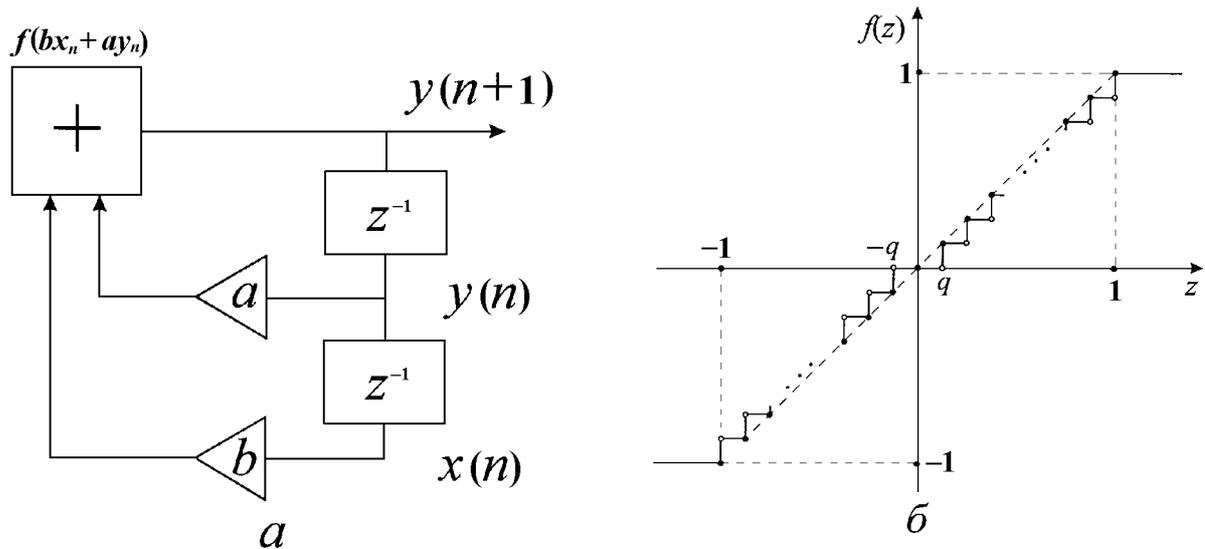


Рис. 1. Исследуемая система и вид нелинейности сумматора

где

$$f(z) = \begin{cases} \text{sign } z, & |z| \geq 1; \\ \varphi(z), & |z| < 1, \end{cases} \quad (2)$$

$$\varphi(z) = q \text{ floor}(|z|/q) \text{ sign } z.$$

Здесь функция $\text{floor}(\psi)$ задает ближайшее снизу к ψ целое число, q — шаг квантования. Вид функции (2) показан на рис. 1б. Величина q определяется числом шагов квантования L :

$$q = 2/(L - 1).$$

Все движения в системе вследствие характерного вида функции $f(z)$ суперустойчивы. Нечетность $f(z)$ и вид отображения (1) позволяют рассматривать одну полуплоскость плоскости параметров ab : либо $a \leq 0$, либо $a \geq 0$. Далее рассматривается только последний случай, а параметры системы выбираются внутри треугольника устойчивости

$$\Delta = \{(a, b) : |b| < 1 \cap b + |a| < 1\}.$$

Пересчет вида и периодов движений при смене знака параметра a производится по правилам, описанным в [6]. Кроме того, исследование предельных циклов малого уровня сводится к изучению предельных циклов большого уровня путем перехода от системы (1) к системе с шагом квантования q/A (с сохранением неизменных параметров a и b), где A — амплитуда предельного цикла малого уровня. При этом траектория данного цикла геометрически подобна траектории предельного цикла большого уровня, а коэффициент подобия траекторий есть отношение $1/A$. Доказательство данного утверждения следует из вида функции (2). Далее по вышеописанным причинам рассматриваются только предельные циклы большого уровня.

Вид функции (2) обуславливает особый интерес к области плоскости состояний xy , ограниченной квадратом

$$ABCD = \{(x, y) : |x| \leq 1 \cap |y| \leq 1\}$$

(считаем, что точки A, B, C и D имеют координаты $(1, 1), (1, -1), (-1, -1)$ и $(-1, 1)$ соответственно), и позволяет разбить плоскость состояний на несколько областей [7], определенных следующим образом:

$$D_{-1} = \{(x, y) : bx + ay \leq -1\}, \quad (3a)$$

$$D_0 = \{(x, y) : |bx + ay| < 1\}, \tag{3б}$$

$$D_{+1} = \{(x, y) : bx + ay \geq 1\}, \tag{3в}$$

$$D_{0,j} = \{(x, y) : (x, y) \in D_0, F[(x, y)] \in D_j\}, \tag{3г}$$

где $j \in \{-1; 0; +1\}$. Расположение зон (3а)–(3г) определяет динамику системы. Границей областей D_{-1} и D_0 является прямая

$$y(x) = -\frac{1 + bx}{a},$$

а граница областей $D_{0,0}$ и $D_{0,-1}$ описывается кусочно-линейной функцией, подобной функции $\varphi(z)$. Однако поскольку в рассматриваемых ниже случаях имеет место соотношение

$$(D_{0,-1} \cap ABCD) \subset (d_{0,-1} \cap ABCD),$$

где

$$d_{0,-1} = \{(x, y) : (x, y) \in D_0, (a^2 + b)y + abx \in D_{-1}\},$$

то вместо точной границы вышеописанных областей будем рассматривать прямую

$$y_1(x) = -\frac{1 + abx}{a^2 + b}. \tag{4}$$

Исследуем движения в системе (1) на плоскости состояний xy при параметрах $(a, b) \in \delta$ (см. рис. 2), где

$$\delta = \{(a, b) : |a| + |b| < 1\}.$$

При этом $ABCD$ целиком содержится в области D_0 и, если $(x_0, y_0) \in ABCD$, то движения описываются уравнениями

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} y_n \\ \varphi(bx_n + ay_n) \end{pmatrix}.$$

Здесь и ниже для доказательства устойчивости будем использовать векторную норму, определенную как

$$\|\mathbf{x}\| = \max[|x|, |y|],$$

где $\mathbf{x} = (x, y)$. Для данной нормы выполняются следующие неравенства

$$|\varphi(bx + ay)| \leq (|a| + |b|) \|\mathbf{x}\| < \|\mathbf{x}\|. \tag{5}$$

Рассматривая три последовательных состояния системы $\mathbf{x}_0 = (x_0, y_0)$, $\mathbf{x}_1 = F(\mathbf{x}_0)$, $\mathbf{x}_2 = F(\mathbf{x}_1)$ и используя неравенства (5), несложно показать, что имеют место два варианта соотношений между нормами данных векторов:

1. $\|\mathbf{x}_0\| > \|\mathbf{x}_1\|$, $\|\mathbf{x}_0\| = |x_0|$,
2. $\|\mathbf{x}_0\| = \|\mathbf{x}_1\| > \|\mathbf{x}_2\|$, $\|\mathbf{x}_0\| = |y_0|$,

что означает сходимости последовательности $\{\|\mathbf{x}_n\|\}$ к нулю. Следовательно, система (1) устойчива в нуле при параметрах a и b , принадлежащих области δ .

Исследуем динамику системы при параметрах a и b , выбранных за пределами области δ . При выборе параметров в области γ (см. рис. 2), определяемой из условия совпадения областей D_0 и $D_{0,0}$, верно, что

$$\begin{aligned} \forall \mathbf{x}_0 \in D_{-1} & \text{ имеем } F^{(2)}(\mathbf{x}_0) \in D_0, \\ \forall \mathbf{x}_0 \in D_0 & \text{ имеем } F^{(n)}(\mathbf{x}_0) \in D_0, \end{aligned}$$

где $n \geq 1$, $F^{(n)}$ — n -кратное отображение. Таким образом, в соответствии с [5] система (1) устойчива в нуле.

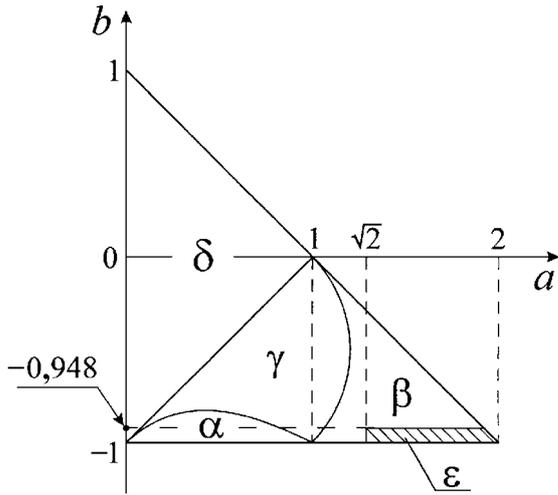


Рис. 2. Исследуемые области параметров системы

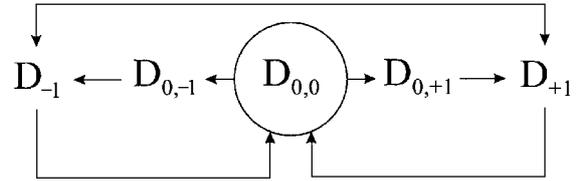


Рис. 3. Траектории движения изображающей точки

Рассмотрим случаи, когда параметры системы выбираются вне области $\gamma \cup \delta$. При этом изображающая точка перемещается согласно схеме, показанной на рис. 3. На данной схеме окружностью выделена зона, в которой возможно существование предельных циклов или m итераций. Из рис. 3 следует возможность возникновения процессов следующего вида:

$$D_{0,-1} \rightarrow D_{-1} \rightarrow \underbrace{D_{0,0} \rightarrow \dots \rightarrow D_{0,0}}_{m \text{ раз}} \rightarrow D_{0,+1} \rightarrow D_{+1} \rightarrow \underbrace{D_{0,0} \rightarrow \dots \rightarrow D_{0,0}}_{m \text{ раз}} \rightarrow D_{0,-1} \rightarrow \dots, \quad (6)$$

где m — количество итераций в области $D_{0,0}$. Развитие такого процесса может привести к появлению предельного цикла с периодом

$$T = 2m + 4.$$

Возможность существования процесса (6) определяется наличием внутри квадрата $ABCD$ зоны $D_{0,-1}$. Это условие удовлетворяется для областей

$$\alpha = \{(a, b) : 0 < a < 1 \cap -1 < b < -(1 + a^2)/(1 + a)\},$$

$$\beta = \{(a, b) : 1 < a < 2 \cap (a - 1/2)^2 + (b + 1/2)^2 > 1/2\},$$

которые показаны на рис. 2. На рис. 4а и рис. 4б показано расположение зон (3а)–(3г) соответственно при $(a, b) \in \alpha$ и $(a, b) \in \beta$.

Исследуем возможность возникновения периодических движений, обусловленных процессами (6). При этом можно свести исследование динамики таких процессов к изучению свойств одномерных нелинейных отображений. Опишем методику построения таких отображений. В случае $m \geq 1$ изображающая точка каждые m итераций оказывается на сторонах AB и CD квадрата $ABCD$, попадая в интервалы, образованные пересечениями $D_{0,-1} \cap AB$ и $D_{0,+1} \cap CD$ соответственно (рис. 4б). Данный процесс возможен при $(a, b) \in (\alpha \cup \beta)$ и условии, что изображающая точка попадает на вышеописанные интервалы. Особенности данного процесса позволяют выразить координату $y_{(m+2)n}$ изображающей точки на сторонах AB и CD посредством некоторого одномерного нелинейного разностного уравнения. При помощи преобразования $y_{(m+2)n} = (-1)^n u_n$ данное уравнение приводится к виду

$$u_n = s_m(u_{n-1}, a, b),$$

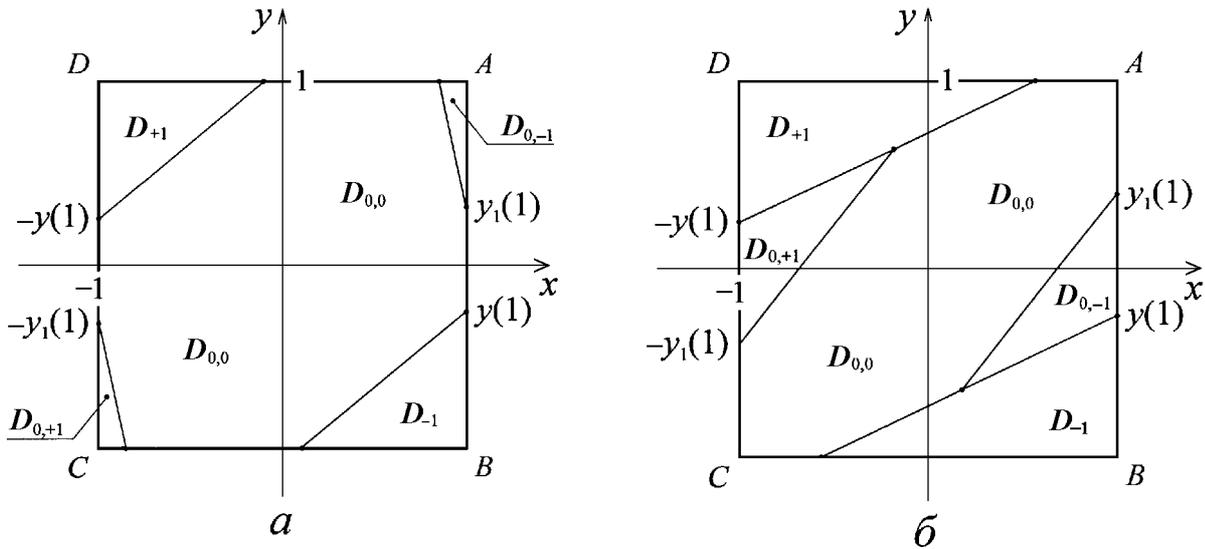


Рис. 4. Разбиения плоскости состояний системы

где $s_m(u)$ — функция последования. Условием существования периодического колебания является наличие у данной функции неподвижных точек (т. е. решений уравнения $u = s_m(u)$) на интервале $D_{0,-1} \cap AB$. Если такие точки отсутствуют, то процесс прекращается, не переходя в периодический.

Рассмотрим возможность существования предельных циклов при $m = 0$. В этом случае процесс (6) возможен при выборе параметров только в области α , а изображающая точка посещает интервалы, образованные пересечениями $D_{0,-1} \cap DA$ и $D_{0,+1} \cap BC$ (рис. 4а). Аналогично предыдущему случаю, рассматривая координату x_{2n} изображающей точки, осуществляем переход к одномерному нелинейному отображению

$$s_0(u) = -\varphi(bu + a),$$

исследуемому на интервале $(D_{0,-1} \cap DA) \subset (0, 1]$. При $(a, b) \in \alpha$ функция последования $s_0(u)$ неубывающая и не имеет неподвижных точек на интервале $(0, 1]$, где $s_0(u) < u$. Таким образом, данный процесс не переходит в предельный цикл.

В случае $m = 1$ соответствующее отображение имеет вид

$$s_1(u) = \varphi[a - b\varphi(b + au)].$$

Несложно показать, что при $a > 0$ и $b < 0$ функция $s_1(u)$ неубывающая. Поскольку для пересечения $D_{0,-1} \cap AB$ выполняются соотношения

$$D_{0,-1} \cap AB \subseteq \begin{cases} I_1 = [y_1(1), 1], & (a, b) \in \alpha; \\ I_2 = [y(1), y_1(1)], & (a, b) \in \beta, \end{cases}$$

исследуем функцию последования $s_1(u)$ на наличие неподвижных точек на отрезках I_1, I_2 . Пусть параметры системы принадлежат множеству $\beta_1 = \beta \cap (b > -1/a)$. Тогда выполняются следующие соотношения:

$$-1 < y(1) < y_1(1) < 0, \quad b + a > 0.$$

Вследствие этого $s_1(y(1)) = \varphi(a + b) \geq 0$, а поскольку $s_1(u)$ — неубывающая функция, то

$$\forall u \in I_2 \text{ имеем } s_1(u) \geq \varphi(a + b) \geq 0 > u.$$

Таким образом, на отрезке I_2 функция $s_1(u)$ не имеет неподвижных точек, а процесс не переходит в периодический. Далее, пусть параметры системы принадлежат оставшейся части множества β , т. е. подмножеству $\beta_2 = \beta \cap (b \leq -1/a)$. Тогда имеют место следующие неравенства:

$$y(1) < 0 \leq y_1(1) < -b/a < 1.$$

Сформируем последовательность

$$u_r = y(1) + rq/a, \quad (7)$$

где r — целое неотрицательное число. Для точек данной последовательности выполняются соотношения

$$\forall u < -b/a, \quad u \in (u_r, u_{r+1}] \quad \text{имеем} \quad \varphi(b + au) = b + au_r, \quad (8)$$

$$s_1(u_r) = \varphi(a - b^2 - abu_r). \quad (9)$$

В правой полуплоскости координатной плоскости us_1 сформируем сетку, определяемую по оси абсцисс последовательностью (7), а по оси ординат последовательностью $\{tq\}$, где t — целое число. Построим функцию $M(u)$, проходящую через узлы сетки, лежащие наиболее близко сверху к тождественной функции. Функция $M(u)$ образована отрезками семейства прямых

$$l_{n,k}(u) = au + m_n - kq, \quad (10)$$

где $m_n = 1 + b_n - nq$, $b_n \in (-1 + nq, -1 + nq + q]$, k — неотрицательное целое число. Несложно показать, что при вышеописанных условиях семейство прямых (10) (и функция $M(u)$) мажорируется прямой

$$l(u) = au + 1 + b.$$

С учетом соотношений (8) и (9) получим условие отсутствия неподвижных точек на интервале $[0, 1]$:

$$h(u) \geq l(u),$$

где $h(u) = a - b^2 - abu$. Данное неравенство выполняется для всех $(a, b) \in \beta_2$. Аналогично случаю $(a, b) \in \beta_1$ доказываем, что на интервале $[y(1), 0]$ справедливо соотношение $s_1(u) \geq 0$. Таким образом, на отрезке $[y(1), 1]$ выполняется условие $h(u) \geq l(u)$, а следовательно, и $s_1(u) > u$. Последнее неравенство равносильно отсутствию неподвижных точек на отрезке I_2 и прекращению процесса с течением времени. Доказательство прекращения процесса при $(a, b) \in \alpha$ аналогично приведенному выше.

Рассмотрим случай $m \geq 2$. Экспериментально установлено, что при данном условии и выборе параметров a, b в области

$$\varepsilon = \{(a, b): \sqrt{2} \leq a < 2, -1 \leq b \leq -0,948\} \subset \beta$$

(см. рис. 2) процессы (6) могут переходить в периодические. Точная форма области, в которой возникают предельные циклы меняется в зависимости от величины шага квантования, однако эта область всегда находится внутри области ε . Интерес к таким областям параметров, лежащим вблизи основания треугольника устойчивости, обусловлен возможностью получения высокодобротных цифровых резонаторов и узкополосных фильтров нижних и верхних частот [8]. Для данных типов фильтров особенно важен вопрос о наличии предельных циклов, поскольку последние могут серьезно нарушить их работу.

Обоснуем теоретически возникновение предельных циклов из процессов (6). Положим $m = 2$. Воспользуемся предложенной методикой и построим соответствующее одномерное нелинейное отображение

$$s_2(u) = -\varphi[b\varphi(bu + a\varphi(b + au)) - a].$$

При параметрах $(a, b) = (1,45; -0,99) \in \varepsilon$, $L = 19$ ($q = 1/9$) отображение $s_2(u)$ имеет неподвижную точку $u_0 = 8/9$, из чего следует существование предельного цикла с периодом $T = 8$. Данный предельный цикл возникает, например, при $(x_0, y_0) = (1, u_0)$. На рис. 5 с помощью диаграммы Кенигса—Ламерея показано графическое решение уравнения $s_2(u) = u$ на отрезке $u \in [0,4; 1]$. Аналогично, при $m = 3$ (период $T = 10$), $L = 21$ ($q = 0,1$) и параметрах $(a, b) = (1,64; -0,95) \in \varepsilon$ отображение

$$s_3(u) = -\varphi[b\varphi(b\Phi(u) + a\varphi(bu + a\Phi(u))) - a],$$

где $\Phi(u) = \varphi(b + au)$, имеет неподвижную точку $u_1 = 0,7$ и т. д.

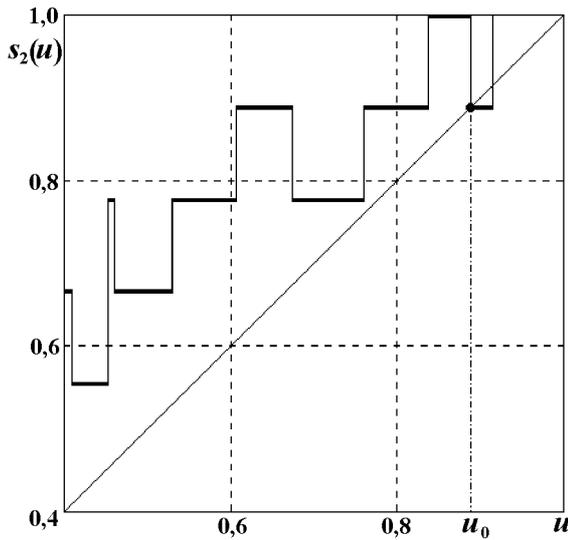


Рис. 5. Диаграмма Кенигса—Ламерея

Точные границы областей параметров, при которых реализуются предельные циклы могут быть получены путем перебора значений параметров a, b и численного решения уравнений $s_m(u, a, b) = u$ при $m \geq 2$. В качестве численного метода может использоваться проверка истинности решения данного уравнения на интервале $[y_1(1), 1]$ в тех точках, где значение переменной u кратно шагу квантования q . В принципе, точные границы указанных областей могут быть определены из отображений (1), (2). Однако это потребует знания области притяжения предельных циклов.

Таким образом, аналитически доказано, что в системе (1) при $a \geq 0$ не существует предельных циклов с периодами $T = 1; 2; 4; 6$. Экспериментально установлено и теоретически обосновано, что при выборе параметров в области ε возникают предельные циклы с периодами $T = 2s$, где $s = m + 2$, $m \geq 2$. Из свойств процессов (6) и симметрии функции φ следует, что траектория таких предельных циклов обладает следующим свойством симметрии:

$$\forall n \geq 0 \text{ справедливо } \mathbf{x}_n = -\mathbf{x}_{n+s}.$$

Исходя из этого свойства и используя выкладки об изменении структуры предельного цикла при смене знака параметра a [6], заключаем, что при $a < 0$ не существует предельных циклов с периодами $T = 1; 2; 3; 4$. Далее, если при $a \geq 0$ существовал предельный цикл с периодом $T = 2s$, то при $a < 0$ и нечетном s существуют два предельных цикла с периодом $T = s$, а при четном s — один цикл с периодом $T = 2s$.

ЗАКЛЮЧЕНИЕ

Исследованы свободные колебания в нелинейной цифровой системе второго порядка с усечением по модулю после сложения и насыщением. Экспериментально установлено и теоретически обосновано, что при выборе параметров в области ε возникают предельные циклы большого и малого уровней. Предложена численно-аналитическая методика их исследования, основанная на переходе к одномерному нелинейному отображению и поиске его неподвижных точек. При помощи этой методики аналитически найдены условия существования циклов. Вычислены периоды допустимых предельных циклов и установлены особенности расположения областей их существования на плоскости параметров и плоскости состояний.

Результаты работы могут использоваться при анализе устойчивости нелинейных цифровых фильтров второго порядка и синтезе систем цифровой обработки информации.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 99–02–17939).

ЛИТЕРАТУРА

1. Коча В. М., Торяник Н. Н. // Зарубежная радиоэлектроника. 1985. № 11. С. 16.
2. Ogorzalek M. J. // Int. J. of Bifurcation and Chaos. 1992. V. 2, № 1. P. 11.
3. Bryukhanov Y. A., Myasnikov E. A. // Proceedings of 5th International Specialist Workshop "Nonlinear Dynamics of Electronic Systems"(NDES'97). — Moscow, 1997. — P. 87.
4. Claasen T., Mecklenbrauker W. F. G., Peek J. B. H. // IEEE Trans. on Circuits and Systems. 1975. V. CAS-22, № 3. P. 232.
5. Круковский-Синевиц К. Б., Азаров В. С. // Вестник Киевского Политехнического Института. Радиотехника. 1989. № 26. С. 46.
6. Брюханов Ю. А., Глызин С. Д., Рахманова Н. К. // Динамика цифрового фильтра с кусочно-линейной характеристикой типа насыщения. Аналитическое исследование: Препринт № 17 ИПМ РАН им. М. В. Келдыша. — М., 1994.
7. Брюханов Ю. А. // Изв. вуз. Радиофизика. 1998. Т. 41, № 4. С. 534.
8. Брюханов Ю. А. // Радиотехника и электроника. 1997. № 10. С. 53.

Ярославский государственный
университет, г. Ярославль, Россия

Поступила в редакцию
20 ноября 1998 г.

LIMITING CYCLES OF A NONLINEAR RECURSIVE DIGITAL SECOND-ORDER SYSTEM WITH TRUNCATION BY MODULUS AFTER SUMMATION

Yu. A. Bryukhanov and E. N. Myasnikov

We analyze free oscillations in a digital second-order system with truncation by modulus after summation and nonlinear saturation. The limiting cycles are found to exist in the system. Analytical and experimental estimates for the periods and existence conditions of the limiting cycles are given.

**ОПЕЧАТКИ К СТАТЬЕ Т. М. Заборонковой, А. В. Кудрина, Е. Ю. Петрова
“О РАСПРЕДЕЛЕНИИ ТОКА ВДОЛЬ ЦИЛИНДРИЧЕСКОЙ АНТЕННЫ
В МАГНИТОАКТИВНОЙ ПЛАЗМЕ
В ДИАПАЗОНЕ ОЧЕНЬ НИЗКИХ ЧАСТОТ”
(Известия ВУЗов. Радиофизика. 1999. Т. 42, №. 8, С. 750–764)**

В статье Т. М. Заборонковой, А. В. Кудрина, Е. Ю. Петрова “О распределении тока вдоль цилиндрической антенны в магнитоактивной плазме в диапазоне очень низких частот” (Известия ВУЗов. Радиофизика. 1999. Т. 42, №. 8, С. 750–764) допущены опечатки в номерах ряда формул.

Формулу (20) следует читать как формулу без номера.

Формулы с номерами (21)–(38) следует читать как формулы (20)–(37).