МИНИСТЕРСТВО ОБЩЕГО И ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Известия высших учебных заведений

РАДИОФИЗИКА

ежемесячный научно–технический журнал

Издается с апреля 1958 г.

Том XLI N 9 **Нижний Новгород** **1998**

Содержание

УДК 551.510.535

ИССЛЕДОВАНИЕ НИЖНЕЙ ИОНОСФЕРЫ С ПОМОЩЬЮ ИСКУССТВЕННЫХ ПЕРИОДИЧЕСКИХ НЕОДНОРОДНОСТЕЙ

В. В. Беликович, Н. В. Бахметьева, В. Н. Бубукина, А. А. Караштин, А. В. Толмачева

Приведены результаты исследования нижней ионосферы с помощью искусственных периодических неоднородностей плазмы с увеличенным высотным разрешением. Обращается внимание на возможность изучения неоднородной структуры ионосферы таким методом.

ВВЕДЕНИЕ

К настоящему времени проведено значительное количество исследований нижней ионосферы с помощью искусственных периодических неоднородностей плазмы (ИПН), созданных полем мощной стоячей радиоволны. В работах [1, 2] изложены основы этого метода и показана перспективность его применения для определения профилей электронной концентрации, измерения температуры и плотности атмосферы, скоростей вертикальных и турбулентных движений. Однако, подавляющее большинство этих исследований проведено со сравнительно низким техническим высотным разрешением (шаг по высоте 3–5 км) [1–7]. Только в исследованиях авроральной ионосферы в Тромсе [8] был использован шаг 1,5 км. В связи с этим представляют интерес первые исследования с помощью ИПН, проведённые на стенде "СУРА"с цифровой регистрацией амплитуд рассеянного сигнала с высотным шагом 1,2 км. Как будет видно из дальнейшего, такое высотное разрешение кроме обычных параметров позволяет получить интересные сведения о неоднородной структуре ионосферы и, в частности, о спорадических слоях.

1. МЕТОДИКА ЭКСПЕРИМЕНТА

В августе 1996 года на стенде "СУРА" (координаты 56,13° с. ш., 46,1° в. д.) проведена серия экспериментов по исследованию ионосферы при помощи ИПН с применением цифровой регистрации данных. Было проведено 12 сеансов измерений длительностью от 5 до 20 минут каждый. В каждом сеансе периодически проводилось воздействие на ионосферу необыкновенной радиоволной на частоте 4,7 МГц с эффективной мощностью 60–100 МВт в течение трёх секунд и паузой продолжительностью 12 с. После каждого цикла нагрева в течение двух секунд излучались импульсы длительностью 20 мкс при частоте повторения 40 Гц. Рассеянные сигналы принимались антенной, имеющей 12 синфазных диполей на каждой линейной поляризации и выделяющей необыкновенную компоненту. Усиление сигналов осуществлялось приёмником с полосой пропускания 50 кГц. Сигналы, рассеянные ИПН, регистрировались с высотным шагом 1,2 км в виде двух квадратурных компонент, кодировались 10-ти разрядным АЦП и подвергались последующей компьютерной обработке.

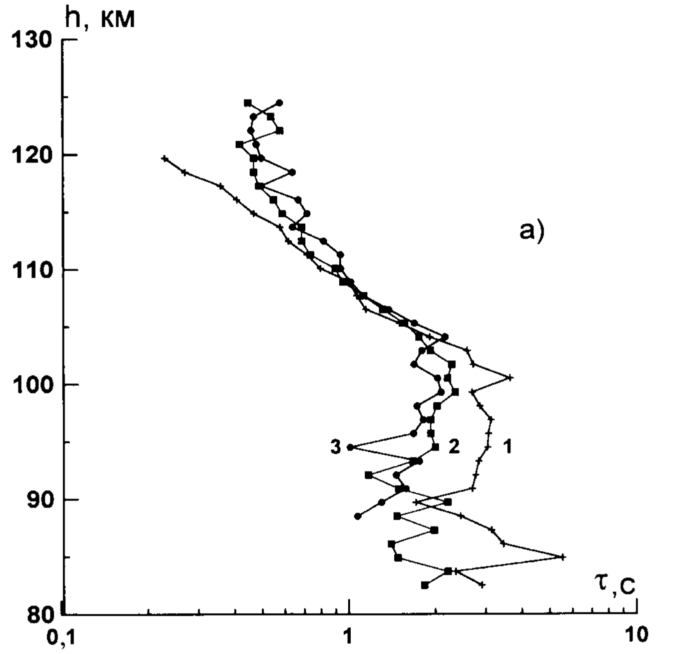
Дальнейшая обработка данных состояла в вычислении амплитуды и фазы сигнала на каждой высоте и аппроксимации их временной зависимости линейными функциями вида

$$\ln A(t) = \ln A_0 - \frac{t}{\tau} \quad \text{и} \quad \phi(t) = \phi_0 + bt\,.$$

Напомним, что τ характеризует время жизни неоднородностей после выключения мощного передатчика, а b — скорость их вертикального движения $V = \dfrac{\lambda}{4\pi}b$ [9]. Нахождение параметров аппроксимации проводилось по методу наименьших квадратов с весовой функцией $\exp(-2t/\tau_0)$, которая, согласно [10], позволяет минимизировать дисперсию находимых параметров (здесь τ_0 — первое приближённое значение параметра). При этом вычислялись дисперсии параметров τ и b, которые позволяли судить о степени надёжности полученных данных.

2. АНАЛИЗ ВЫСОТНЫХ ПРОФИЛЕЙ $\tau(h)$

Для каждого сеанса измерений анализировались средние высотные профили $\tau(h)$ и $V(h)$. Примеры таких профилей можно видеть на рис. 1. При нахождении средних отбрасывались регистрации с начальным отношением сигнал/шум менее 10 дБ и с высокой дисперсией, а также выходящие за рамки $\pm1{,}6\sigma$, что соответствует учёту $\sim$ 90% данных при их нормальном распределении. На рис. 1а представлены профили $\tau(h)$ для трёх моментов времени 15.08.96. Приведённые примеры характеризуются участком экспоненциального убывания τ с ростом высоты в интервале $\sim$ 100−120 км, что свидетельствует о диффузионном механизме релаксации неоднородностей на этих высотах. На высоте $\sim$ 100 км в высотном профиле τ можно видеть излом. Уменьшение τ ниже этого уровня обусловлено турбулентностью, разрушающей периодическую структуру быстрее, чем амбиполярная диффузия. При этом очевидно, что высота излома соответствует турбопаузе — высоте, на которой турбулентное перемешивание атмосферы сменяется диффузионным разделением газов. Отметим, что высота излома τ несколько увеличивается в течение дня (от 10 к 12 часам) и одновременно происходит уменьшение $\tau(h)$ ниже 100 км. Оба эти факта свидетельствуют о некотором увеличении интенсивности турбулентности при приближении к полудню.



Некоторые высотные профили τ немонотонны на высотах более 100 км и имеют локальные максимумы. Пример такого профиля для сеанса 12 час. 7 мин. 15.08.96 можно видеть на рис. 2. Здесь прямой линией показана высотная зависимость τ, которая наблюдалась при отсутствии максимума. На фоне этой линии отчётливо виден локальный максимум τ на высоте 104 км. Анализ показал, что

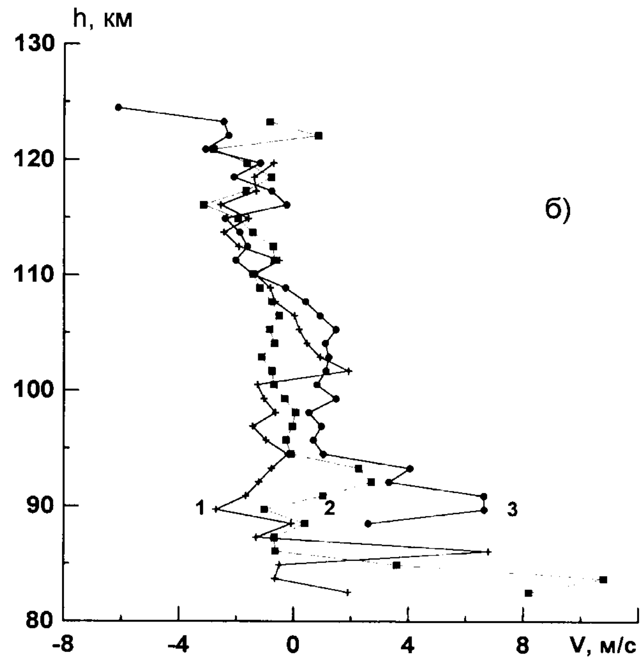


Рис. 1. Высотные профили времени релаксации ИПН (а) и вертикальной скорости (б) для трёх сеансов 15.08.1996: кривые 1 — 9 час. 55 мин., 2 — 11 час. 42 мин., 3 — 12 час. 07 мин.

такие локальные максимумы сопровождаются увеличением дисперсии амплитуды рассеянного сигнала.

По нашему мнению, эти максимумы могут быть обусловлены спорадическими слоями, которые очень часто наблюдаются в летнее время. Действительно, в области E $\tau(h)$ определяется амбиполярной диффузией [1]

$$\tau = (K^2 D_{\text{a}})^{-1}, \quad \text{где} \quad D_{\text{a}} = \frac{\kappa(T_e + T_i)}{M_i \nu_m},$$

D_{a} — коэффициент амбиполярной диффузии, $K = K_0 n$ — волновой вектор в среде, $n = \sqrt{\varepsilon}$ — коэффициент преломления среды. Превышение электронной концентрации в E_{s} над фоновой приводит к уменьшению показателя преломления, увеличению длины волны в слое и увеличению $\tau(h)$, причём

$$\Delta\tau/\tau = \Delta\varepsilon/\varepsilon. \tag{1}$$

Одновременно с этим увеличивается коэффициент отражения от неоднородностей, поскольку он пропорционален фоновой электронной концентрации [1], которой в данном случае является концентрация в E_{s}.

Другой причиной изменения τ в E_{s} могут быть металлические ионы. Наибольшее влияние на $\tau(h)$ могли бы оказать ионы железа с атомной массой 56, но их относительная доля в E_{s} невелика. Атомные массы других наиболее распространённых ионов магния и кремния (24 и 28, соответственно) близки к средней молекулярной массе атмосферных ионов (~ 29), поэтому их присутствие в E_{s} должно слабо влиять на $\tau(h)$.

Для примера рассмотрим сеанс в 9 час. 55 мин., в котором на высоте 104 км наблюдается локальный максимум. Значение τ в нём составило 2,3 с, что больше фонового для этой высоты на 35 процентов. Используя равенство (1) и выражение для ε в квазипродольном приближении, легко получить

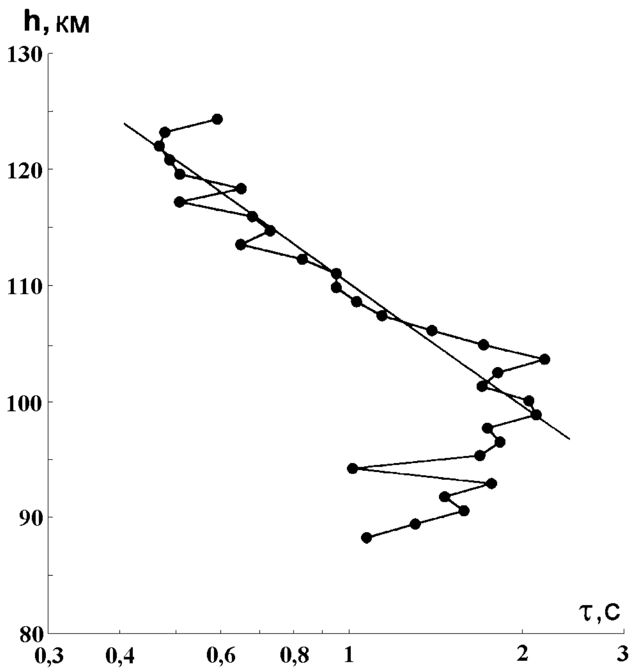


Рис. 2. Высотный профиль времени релаксации ИПН для сеанса 12 час. 07 мин. 15.08.1996. Прямая линия соответствует высотному профилю τ при чисто диффузионном механизме релаксации.

выражение

$$\frac{f_{0\mathrm{s}} - f_0}{f_0} = \frac{\Delta\tau}{2\tau}\left[\frac{f(f - f_\mathrm{L})}{f_0^2} - 1\right],$$

где f, f_0 — рабочая и плазменная частоты, f_L — продольная компонента гирочастоты, $f_{0\mathrm{s}}$ — плазменная частота в E_s. Подстановка в эту формулу значений $f = 4{,}7$, $f_0 = 3,,3$, $f_\mathrm{L} = 1{,}3$ МГц и $\Delta\tau/\tau = 0{,}35$ приводит к значению $\Delta f_{0\mathrm{s}} = 0{,}08$. То есть плазменная частота в E_s выше фоновой на 8%. Таким образом, регистрация высотной зависимости $\tau(h)$ является хорошим инструментом для изучения спорадических слоев.

Рассматривая высотные зависимости для приведённых сеансов, можно отметить, что выше 100 км $\tau(h)$ несколько возрастает в течение дня. Этот эффект, очевидно, обусловлен суточным ростом электронной концентрации в E-слое.

3. СКОРОСТИ ВЕРТИКАЛЬНЫХ ДВИЖЕНИЙ

По измерениям фазы обратно рассеянных сигналов с использованием методик [3, 9] определены скорости вертикальных движений и построены их высотные профили. Положительная скорость соответствует направлению вниз. Усреднённые за сеанс значения скорости вертикального движения находились в пределах от 1 до 5 м/с, при этом на высотах 90—100 км, как правило, наблюдалась смена направления движения. На рис. 1б приведены высотные профили V для тех же сеансов 15.08.96, что и времени релаксации. Для высот, на которых на профилях τ видны локальные максимумы, в профилях V также имеются некоторые особенности. Так, например, для сеанса 12 час. 7 мин. на высотах 104—100 км имеется положительный градиент направленной вниз скорости порядка $5 \cdot 10^{-4}$ с$^{-1}$. По оценкам, такой градиент может обеспечить сгонку плазмы (на высотах 100—104 км) и быть причиной возникновения E_s. В целом, величина dV/dh на высотах, где обнаружены локальные максимумы в высотном ходе τ, составила $5 \cdot 10^{-4} - 4 \cdot 10^{-3}$ с$^{-1}$.

Усреднённые за минуту значения τ и V для того же сеанса приведены на рис. 3а и 3б соответственно для высот 100, 104, 110 и 120 км. Как и в измерениях, проведённых этим методом ранее [7], они испытывают быстрые колебания с характерным временем порядка и меньше минуты, при этом амплитуда колебаний на высотах ниже 100 км существенно выше, чем на больших высотах. Причиной этого, по мнению авторов, является наличие двух источников колебаний. Один из них — атмосферная турбулентность, которая проявляется на высотах ниже турбопаузы, а другой, действующий на всех высотах, нам пока не известен.

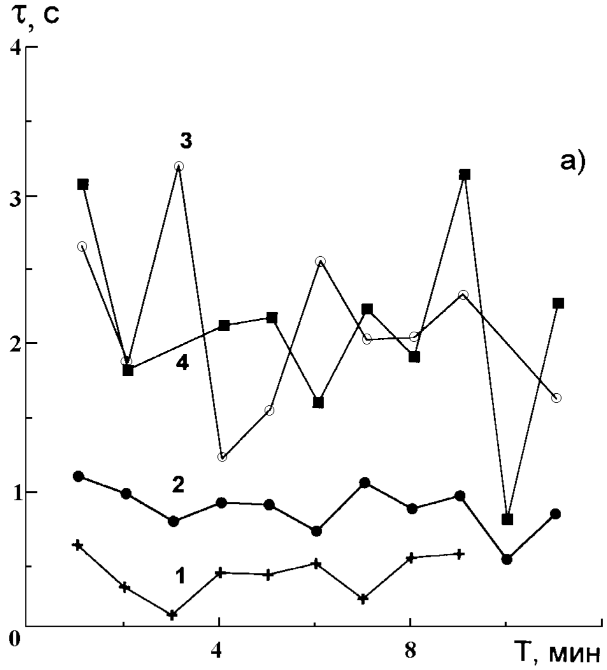


Рис. 3а. Временные вариации минутных значений τ в сеансе 12 час. 07 мин. 15.08.1996: кривые 1, 2, 3 и 4 для действующих высот 120, 110, 104 и 100 км, соответственно.

По разработанной ранее методике [4, 5] рассчитаны значения вертикальной компоненты скорости турбулентных движений v_{t} на высотах ниже 100 км для того же сеанса. Средние значения v_{t} для высот 90 и 95 км составили 1,8 м/с и 2,3 м/с, соответственно, что близко к величинам, полученным в [5]. Высота 100 км в данном сеансе соответствует высоте турбопаузы.

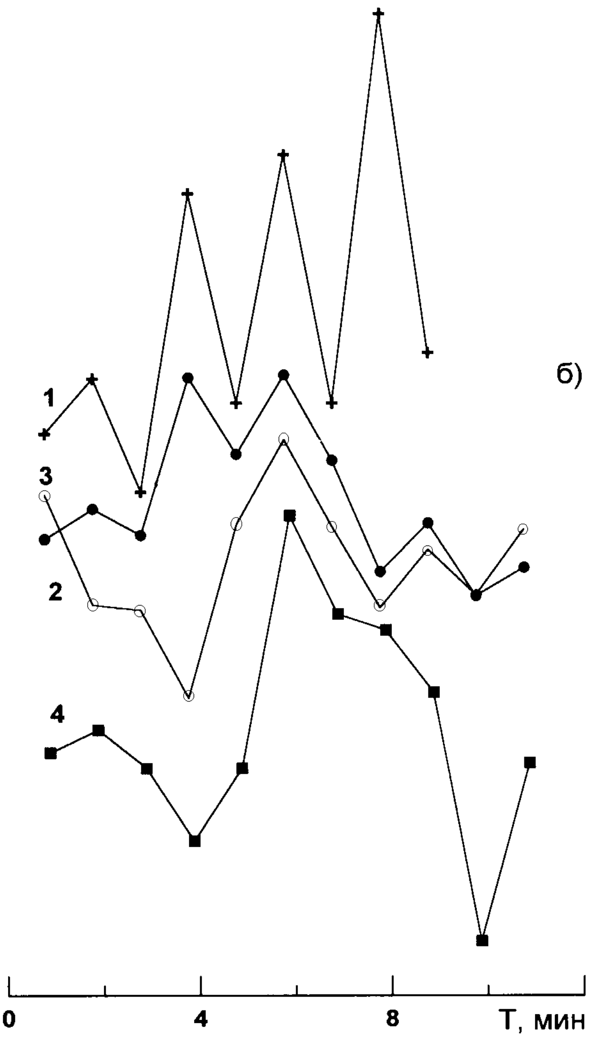


Рис. 3б. Временные вариации минутных значений вертикальной скорости в сеансе 12 час. 07 мин. 15.08.1996: кривые 1, 2, 3 и 4 для действующих высот 120, 110, 104 и 100 км, соответственно.

Необходимо отметить, что выше приведены зависимости величин τ и V от действующей высоты рассеянного сигнала. Обычно на высотах E-слоя ионосферы она близка к истинной высоте. Однако, при относительно низкой рабочей частоте f = 4,7 МГц с необыкновенной поляризацией и критической частоте E-слоя $\sim$ 3,3 МГц групповое запаздывание становится существенным и его необходимо учитывать. В таблице приведены групповое запаздывание L и действующие высоты $h_{\text{д}} = h + L$, рассчитанные в квазипродольном приближении для модельного ионосферного профиля $N(h)$. Из таблицы видно, что на высоте $h = 120$ км $L = 10{,}4$ км и соответственно $h_{\text{д}} = 130{,}4$ км.

Таблица

Групповое запаздывание на частоте 4,9 МГц

h, км	88	92	96	100	104	108	112	116	120	124
L, км	0	0,18	0,6	1,42	2,74	4,49	6,49	8,49	10,4	12,1
$h_{\text{д}}$, км	88	92,2	96,6	101,4	106,7	112,5	118,5	124,5	130,4	136,1

ВЫВОДЫ

Проведены исследования ионосферы с помощью ИПН с высоким пространственно–временным разрешением и при полной автоматизации процесса измерения амплитуд рассеянных сигналов и первичной обработки данных. Показано, что использование высотного шага съёма данных ~ 1 км открывает новые возможности метода. В частности, в высотной зависимости времени релаксации ИПН обнаружены локальные максимумы, которые можно отождествить со спорадическими слоями. Это обстоятельство позволяет предложить новый способ исследования спорадических слоев.

Авторы выражают благодарность Ю. В. Шлюгаеву за помощь в проведении эксперимента и В. В. Беликовичу младшему за составление программ обработки экспериментальных данных.

Работа выполнена при поддержке Российского Фонда фундаментальных исследований (гранты № 96–05–65130 и № 97–05–64397).

ЛИТЕРАТУРА

1. Беликович В. В., Бенедиктов Е. А., Гончаров Н. П., Толмачева А. В. //Геомагнетизм и аэрономия, 1995. Т. 35. № 4. С. 64.
2. Belikovich V. V., Benediktov E. A., Goncharov N. P., and Tolmacheva A. V. //J. Atmos. Sol. Terr. Phys., 1997. V. 59. № 18. P. 2447.
3. Бахметьева Н. В., Беликович В. В., Бенедиктов Е. А., Бубукина В. Н., Гончаров Н. П., Игнатьев Ю. А. //Геомагнетизм и аэрономия, 1996. Т. 36. № 5. С. 120.
4. Беликович В. В., Бенедиктов Е. А. //Геомагнетизм и аэрономия, 1995. Т. 35. № 2. С. 91.
5. Бахметьева Н. В., Беликович В. В., Коротина Г. С. //Геомагнетизм и аэрономия, 1996. Т. 36. № 5. С. 180.
6. Bakhmet'eva N. V., Belikovich V. V., Benediktov E. A., Bubukina V. N., Goncharov N. P., and Ignat'ev Yu. A. // Изв. вузов. Радиофизика, 1996. Т. 39. № 3. С. 329.
7. Бахметьева Н. В., Беликович В. В., Бенедиктов Е. А., Бубукина В. Н., Игнатьев Ю. А. //Изв. вузов. Радиофизика, 1997. Т. 40. № 3. С. 308.
8. Rietveld M. T., Turunen E., Matveinen H., Goncharov N. P., and Pollari P. //Ann. Geophys., 1996. V. 14. P. 1437.
9. Беликович В. В., Бенедиктов Е. А., Гончаров Н. П. //Геомагнетизм и аэрономия, 1991. Т. 31. № 2. С. 381.
10. Беликович В. В. //Изв. вузов. Радиофизика, 1993. Т. 36. № 12. С. 1105.

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

Поступила в редакцию
26 февраля 1998 г.

THE STUDY OF THE LOWER IONOSPHERE BY ARTIFICIAL PERIODIC INHOMOGENEITIES

V. V. Belikovich, N. V. Bakhmet'eva, V. N. Bubukina, A. A. Karashtin, A. V. Tolmacheva

The results have been given of the lower ionosphere studies with improved height resolution using artificial periodic inhomogeneities. A special attention has been paid to a possible investigation of the ionosphere inhomogeneous structure by this method.

УДК 550.388

ЭВОЛЮЦИЯ ВОЛНОВЫХ ВОЗМУЩЕНИЙ В ВЕРХНЕЙ ИОНОСФЕРЕ. Часть II

В. Б. Иванов, В. М. Поляков

Представлен теоретический анализ вертикального распространения крупномасштабных возмущений волновой природы в области F и внешней ионосфере. Анализ проведён с учётом всех основных факторов, влияющих на динамику ионосферной плазмы в условиях средних широт. Показано, что возмущения, распространяющиеся сверху вниз в сильно неоднородной среде, нарастают по интенсивности до высот максимума слоя F_2, а затем затухают в нижележащих слоях ионосферы. Исследуемый механизм может рассматриваться как один из источников формирования неоднородной структуры верхней ионосферы.

Настоящая работа является продолжением исследований, представленных в публикации авторов [1], где рассматривались возможности распространения в ионосферной плазме наряду с возмущениями диффузионной природы волновых возмущений, генерируемых источниками в самых верхних слоях ионосферы. Для условий верхней части области F ионосферы в представлениях о диффузионном равновесии фоновой плазмы было получено аналитическое решение уравнения, описывающего распространение по вертикали малых возмущений концентрации и скорости электронно—ионного газа. Было обнаружено, что в таких условиях могут существовать волновые возмущения, распространяющиеся сверху вниз с увеличением амплитуды.

В данной статье мы представляем результаты решения более полной задачи — анализа возможностей распространения волновых возмущений с учётом вертикальных потоков частиц и ионизационно—рекомбинационных процессов. Таким образом, эти результаты соответствуют условиям реальной среднеширотной и умеренно высокоширотной ионосферной области F.

Учёт потоков плазмы и процессов образования и исчезновения заряженных частиц делает невозможным получение строгого аналитического описания распределения характеристик фоновой ионосферы и, тем более, возмущений. В этой связи мы были вынуждены обратиться к численным методам анализа.

В приближении амбиполярного движения зарядов система гидродинамических уравнений столкновительной плазмы [2], описывающая вертикальное распределение концентрации N и скорости V, включает в себя уравнение непрерывности

$$\frac{\partial N}{\partial t} + \frac{\partial (NV)}{\partial z} = q(z) - \beta(z)N \qquad (1)$$

и уравнение движения (без учёта нелинейного ускорения)

$$N\frac{\partial V}{\partial t} = Ng - N\nu(V - V_{\rm n}) - c^2\frac{\partial N}{\partial z}\,. \qquad (2)$$

Здесь t — время, z — вертикальная координата, отсчитываемая вниз от некоторой верхней границы $z = 0$, $q(z)$ — скорость ионообразования, $\beta(z)$ — коэффициент линейной рекомбинации для области F, g — ускорение свободного падения (в выбранной системе координат g — величина положительная), ν — частота столкновений ионов с нейтральными частицами, $V_{\rm n}$ — продольная к геомагнитному полю составляющая скорости нейтральной атмосферы, $c = \sqrt{k(T_e + T_i)/m_i}$ — скорость ионного звука.

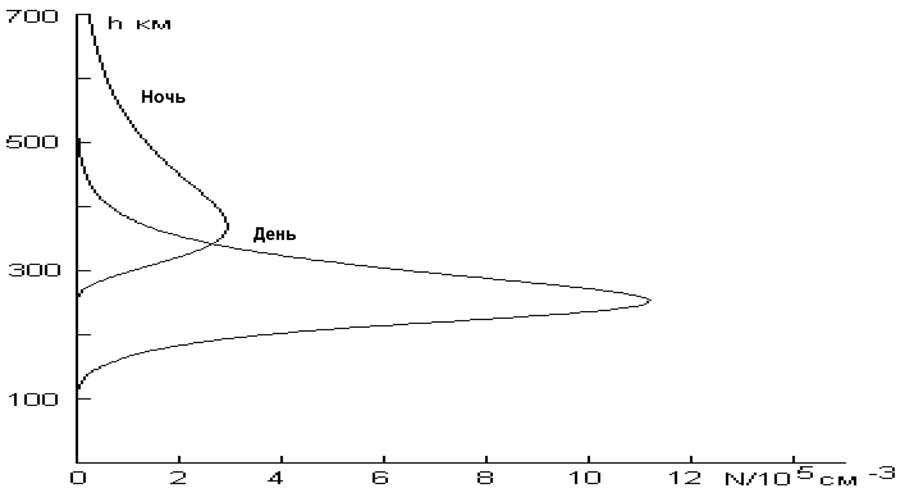


Рис. 1.

Считается, что в верхней атмосфере концентрация нейтрального газа — атомного кислорода — распределена по барометрическому закону со шкалой высот H. Тогда и частота столкновений ионов с нейтральными частицами распределена по формуле $\nu = \nu_m \exp{(z/H)}$, где ν_m — частота столкновений на верхней границе. Вертикальное распределение коэффициента рекомбинации $\beta(z)$ также имеет экспоненциальный характер, но с вдвое меньшей шкалой высот. Скорость ионообразования в уравнении (1) задаётся известной чепменовской формулой.

Система уравнений (1), (2) широко используется для модельного описания среднеширотной ионосферы, и методы её численного решения хорошо известны [3]. Мы также реализовали численный алгоритм решения системы в стационарном варианте $\dfrac{\partial}{\partial t} = 0$ для получения данных по фоновой ионосфере для описания возмущений. Отметим, что согласование модели ночной области F с наблюдаемыми параметрами в части высоты расположения максимума слоя достигается введением так называемой ветровой коррекции — учётом продольной к геомагнитному полю составляющей нейтрального ветра в уравнении движения [3]. Величина концентрации плазмы в максимуме слоя определяется потоком плазмы, втекающим в ионосферу из плазмосферного резервуара.

Результаты расчётов вертикального распределения концентрации плазмы для ночных и дневных условий, принятого в качестве модели фоновой ионосферы, представлены на рис. 1.

Концентрация плазмы представляется в виде суммы стационарной “фоновой"величины N_0 и малого возмущения n. Малая скорость зарядов в возмущениях обозначается как v при невозмущённой постоянной во времени скорости V_0. После линеаризации системы уравнений (1) и (2) по малым возмущениям концентрации и скорости эта система принимает вид

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial z}(nV_0 + N_0 v) + \beta n = 0, \tag{3}$$

$$N_0 \frac{\partial v}{\partial t} = ng - N_0 \nu v - \nu n(V_0 - V_{\mathrm{n}}) - c^2 \frac{\partial n}{\partial z}\,. \tag{4}$$

Далее будут рассматриваться возмущения концентрации и скорости волновой природы, гармонически изменяющиеся во времени с частотой ω — решение ищется в виде $\sim e^{i\omega t}$. При этом временная производная $\dfrac{\partial}{\partial t}$ преобразуется в оператор умножения на $i\omega$. После этого система уравнений (3), (4) может быть приведена к одному уравнению второго порядка для возмущений концентрации:

$$\frac{d^2 n}{dz^2} - \frac{dn}{dz}\left(\frac{1}{H_\mathrm{p}} + \frac{1}{H}\frac{\nu}{\nu + i\omega} + \frac{\nu V_\mathrm{n} + i\omega V}{c^2}\right) +$$

$$+ n\left[\left(\frac{1}{HH_\mathrm{p}} + \frac{i\omega(V - V_\mathrm{n})}{c^2 H}\right)\frac{\nu}{\nu + i\omega} - \frac{i\omega V' + (\beta + i\omega)(\nu + i\omega)}{c^2}\right] = 0. \qquad (5)$$

Используется понятие плазменной шкалы высот — масштаба, определяемого выражением $H_\mathrm{p} = c^2/g = k(T_e + T_i)/m_i g$. Величина V' есть производная скорости фоновой плазмы по вертикальной координате.

Для уравнения (5) необходимо сформулировать краевые условия. Прежде всего заметим, что решение уравнения представляет собой комплексную функцию от координаты z. Вещественная и мнимая части n являются двумя линейно независимыми решениями. В однородной среде без диссипации они соответствовали бы просто синусу и косинусу. Рассматривая возмущения, распространяющиеся сверху вниз, разумно задать нижнее (в области сильной рекомбинации) краевое условие нулевым и для вещественной, и для мнимой частей. В физической постановке задачи предполагается, что на верхней границе задаётся гармонически изменяющееся во времени возмущение. При этом верхнее краевое условие на n можно задать в виде 1 или i, что совершенно эквивалентно, поскольку при этом мнимая и вещественная части решения только “меняются местами". Построенная из вещественной и мнимой частей решения уравнения (5) пространственно—временная функция

$$\Delta N(z,t) = \mathrm{Re}\, n(z)\sin(\omega t) + \mathrm{Im}\, n(z)\cos(\omega t) \qquad (6)$$

удовлетворяет системе (3), (4) и необходимым краевым условиям, и, стало быть, является искомым решением.

В отличие от работы [1] уравнение (5) не может быть решено аналитически и решалось численно. Сразу укажем на то, что включение в рассмотрение рекомбинационно—ионизационных процессов и наличие отличной от нулевой и изменяющейся с высотой вертикальной скорости фоновой плазмы существенно изменило картину распространения возмущений. Напомним, что в условиях диффузионного равновесия, рассматриваемых в [1], мы констатировали то, что при распространении волновых возмущений сверху вниз их длина волны уменьшалась, а амплитуда увеличивалась. Учёт перечисленных факторов приводит к существенной и, в определённой степени, неожиданной модификации картины.

Рис. 2 представляет решение (6) для пяти последовательных моментов времени, охватывающих половину временного периода, для частоты колебаний ω, равной 0,05 с$^{-1}$. Расчёт проведён для условий ночной ионосферы.

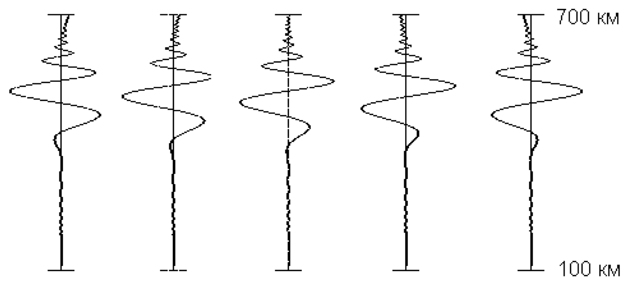


Рис. 2.

Во-первых, как и следовало ожидать, по мере распространения сверху вниз амплитуда возмущений сначала нарастает, а затем, в силу рекомбинации электронно—ионного газа уменьшается. Во-вторых, дисперсионные свойства возмущений оказались такими, что длина волны сначала растёт, а затем уменьшается. Численный анализ показал, что это обусловлено наличием неоднородной ненулевой вертикальной скорости фоновой плазмы.

Интересным является поведение решения вблизи верхней границы. Здесь возмущение можно представить суперпозицией двух частей. Первая описывает диффузионно—затухающую в пространстве составляющую, которая гармонически осциллирует во времени. Вторая представляет собой волновую составляющую. Реально, если источник возмущений находится значительно выше рассматриваемой области, первую составляющую не следует принимать во внимание. Волновая часть возмущения при этом будет выглядеть так, как это представлено средним из графиков рис. 2. Следует также отметить, что область максимальной возмущённости лежит выше максимума слоя F_2.

Достаточно нетривиальный результат был получен из анализа вертикального распределения огибающей амплитуды возмущений. Поведение огибающих для частот 0,07, 0,05 и 0,03 с^{-1} представлено на рис. 3. Можно видеть, что на низких частотах имеется “модуляция" огибающей. Видимо единственным объяснением этому может быть то, что при распространении волн сверху вниз имеет место их частичное отражение. В такой интерпретации частотная зависимость эффекта вполне понятна. Для больших длин волн регулярная неоднородность среды распространения относительно больше и отражение более эффективно.

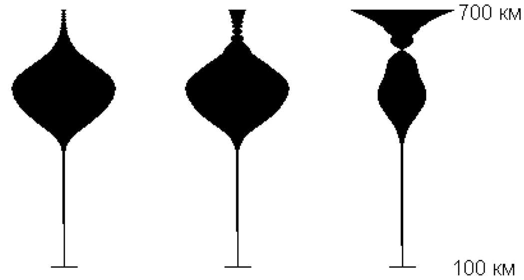


Рис. 3.

Представляемая методика численного моделирования позволяет рассмотреть распространение волновых возмущений внешней ионосферы, вызванных источниками, не только находящимися в вышележащих слоях, но и проанализировать реакцию ионосферы на источники колебательных возмущений в “схеме СНИЗУ ВВЕРХ". Такая постановка задачи может быть применена к анализу возмущений, обусловленных землетрясениями, радионагревом ионосферы мощными радиоволнами и активным воздействием на ионосферу с помощью инжекции газов или плазмы с космических аппаратов. В математической постановке задачи в этом случае следует задавать источник на нижней границе, а на верхней границе задавать отсутствие возмущений.

Численные расчёты показывают, что если источник расположен на достаточно малых высотах — менее $\sim$ 280 километров, где ещё сильны рекомбинационные процессы, то вверх распространяется сильно затухающее возмущение диффузионной природы. Если-же источник находится на больших высотах, то возникает интересная картина, представленная на рис. 4.

Здесь источник задавался на высоте 300 км. Частота колебаний выбиралась равной 0,05 с^{-1}. Представлен мгновенный “снимок" вертикального профиля возмущений и огибающая амплитуды. Можно видеть, что волновая составляющая, распространяясь снизу вверх, сначала затухает, но затем усиливается во внешней ионосфере и вновь начинает затухать. Аналогичным образом изменяется и длина волны. Таким образом, если возмущениям “удается"проникнуть в область выше максимума слоя F_2, то здесь они создают зону неоднородностей.

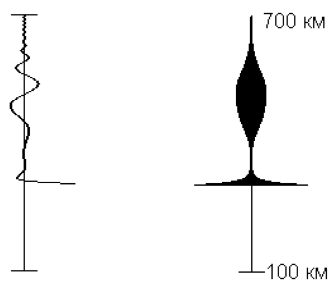


Рис. 4.

Напомним, что описана ситуация, характерная для условий ночной ионосферы, когда она формируется под действием нисходящего потока плазмы, запасённой в плазмосферном резервуаре. Аналогичные расчёты, проведённые для условий дневной ионосферы, показали, что, в силу наличия общего потока плазмы снизу вверх, волновая компонента возмущений в таких условиях практически отсутствует, по крайней мере, при задании источника на верхней границе, как это и предполагалось в работе [1].

Рассмотренные здесь процессы представляют наибольший интерес с точки зрения поиска механизмов формирования случайных неоднородностей верхней ионосферы. Очевидно, что описанные волновые возмущения могут образовывать такие случайные структуры в области F и выше. Первичными источниками здесь могут быть волновые процессы в верхней нейтральной атмосфере — акустические и внутренние гравитационные волны, которые практически всегда существуют в самых верхних слоях атмосферы. Поскольку описанный механизм возбуждения плазменных возмущений нейтральными волнами во внешней ионосфере, с дальнейшим их переносом вниз, наиболее эффективен в ночных условиях, это может являться одним из объяснений того, что неоднородности области F наиболее сильно проявляются ночью [4].

ЛИТЕРАТУРА

1. Иванов В. Б., Поляков В. М. //Изв. вузов. Радиофизика, 1998. Т. 41. № 4. С. 432.
2. Ионосферные процессы /В. М. Поляков, Л. А. Щепкин, Э. С. Казимировский, В. Д. Кокоуров. — Новосибирск: Наука, 1968. С. 536.
3. Коен М. А. Моделирование ионосферы в прикладных задачах геофизики. — Иркутск: Изд-во Иркут. ун-та, 1983. С. 278.
4. Явление F-рассеяния в ионосфере /Б. Н. Гершман, Э. С. Казимировский, В. Д. Кокоуров и др. — М.: Наука, 1984. С. 141.

Иркутский государственный
университет, Россия

Поступила в редакцию
10 ноября 1997 г.

THE EVOLUTION OF WAVE DISTURBANCES IN THE UPPER IONOSPHERE. Part II

V. B. Ivanov, V. M. Polyakov

The theoretical analysis of vertical propagation of wave large-scale disturbances in the F-region and outer ionosphere is presented. The analysis has included all major factors influencing the ionospheric plasma dynamics under conditions of mid latitudes. It is shown that the disturbances propagating down in a strongly non-uniform medium grow in the intensity up to the heights of F_2-layer maximum and then decay in lower layers of the ionosphere. The mechanism studied can be considered as one of the sources of outer ionosphere non-uniform structure formation.

УДК 520.16

ИСПОЛЬЗОВАНИЕ СИГНАЛОВ НАВИГАЦИОННЫХ ИСЗ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ АТМОСФЕРЫ

А. А. Азизов, К. П. Гайкович, С. С. Кашкаров, М. Б. Черняева

В работе рассматривается задача определения высотного профиля показателя преломления по измерениям фазы сигнала навигационных ИСЗ из решения обратной задачи. Соответствующее интегральное уравнение 1-го рода с нелинейным ядром решалось на основе итерационного алгоритма, в котором на каждом шаге решалось линеаризованное уравнение Фредгольма 1-го рода с применением методов статистической регуляризации и Тихонова. По вкладу атмосферного водяного пара, выделенному из восстановленного высотного профиля индекса рефракции, определялся профиль концентрации водяного пара и интегральное влагосодержание в вертикальном столбе атмосферы. С помощью численного эксперимента получены статистические оценки точности возможного определения параметров атмосферы в зависимости от точности измерений для ансамблей метеоданных, соответствующих различным климатическим условиям. Приводятся первые результаты восстановления высотных профилей показателя преломления по данным измерений сигналов ИСЗ навигационной системы “Транзит".

ВВЕДЕНИЕ

Рефрактометрические методы дистанционного зондирования атмосферы находят всё более широкое применение по мере развития космических исследований и их приложений к решению проблем метеорологии, связи и навигации. Восстановление высотного профиля показателя преломления по вкладу атмосферы в допплеровский сдвиг сигнала, излучаемого с космического аппарата, широко использовалось при радиопросвечивании атмосфер планет Солнечной системы [1—9] и Земли [10—18]. В ситуации, когда и источник, и приёмник расположены вне исследуемой атмосферы, задача решается путём применения обратного преобразования Абеля к наблюдаемой зависимости измеряемой величины от прицельного параметра. К аналогичному способу решения сводится задача и в случае внутриатмосферных лимбовых измерений рефракции для восстановления профиля показателя преломления ниже уровня наблюдателя [18—20].

Для случая измерений, когда приёмник сигнала ИСЗ находится на поверхности Земли, задача становится некорректной и её решение существенно усложняется. Задача впервые была сформулирована в [22]. Основные вопросы решения этой некорректной задачи рассматривались в [22—30]. В [19] задача была сведена к линейному уравнению Фредгольма 1-го рода; в [25] было получено последовательное решение этой задачи как некорректной, и этот подход был применён к результатам измерений рефракции звёзд в оптическом диапазоне [26—28].

В радиодиапазоне необходимо учитывать дополнительный вклад влажности в показатель преломления, что требует отдельного анализа, который проводился в [24, 29—30]. В [29—30] использовалась связь атмосферного вклада в допплеровский сдвиг частоты сигнала ИСЗ с величиной астрономической рефракции ε эквивалентного, бесконечно удалённого вдоль луча источника, установленная в [23]. Это позволило применить в исследованиях разработанные ранее алгоритмы решения обратной задачи оптической рефракции. Был определён диапазон углов места, информативных для измерения. Этот диапазон ограничен сверху углами, на которых вариации рефракции становятся сравнимыми с погрешностями измерений, что является следствием “теоремы Лапласа-[31]. При точности измерений рефракции $\delta\varepsilon = 1''$ диапазон информативных углов находится при $\theta_0 \leq 5°$, а при $\delta\varepsilon = 10''$ — соответственно при $\theta_0 \leq 2{-}3°$. Далее, в [29—30] с помощью численного эксперимента были получены статистические оценки точности восстановления показателя преломления по измерениям допплеровского сдвига частоты сигнала ИСЗ. Кроме того, исходя из заданного уровня погрешности $\delta\varepsilon$, в [29—30] определены требования к точности определения параметров задачи, таких как допплеровская частота $\nu_{\rm D}$, радиальная и поперечная компоненты вектора скорости $V_{\rm R}$, $V_{\perp}$, высота H спутника, приземное значение индекса рефракции N_0, высота h_0 пункта наблюдения. Оказалось, что характеристики действовавших в то время навигационных систем типа “Транзит"этим требованиям не удовлетворяли в полной мере. На первом этапе исследований удалось лишь использовать результаты амплитудных измерений сигнала “Транзит"для индикации наличия и определения высоты отражающих слоёв над морем [32—33].

В последнее время, в связи с введением в действие новых навигационных систем, таких как GPS и GLONASS, которые обладают существенно лучшей частотной стабильностью сигнала и траекторные параметры которых известны с большей точностью, чем в прежних системах, появилась реальная возможность применить развитый в [29—30] метод для дистанционного зондирования параметров атмосферы. Первые шаги на этом пути уже сделаны. Решена задача определения интегрального содержания водяного пара по величине влажной части электрической длины пути радиоволн (ЭДПР), которая получается по данным фазовых измерений сигнала GPS на больших углах места [34—36]. Характеристики российской системы GLONASS [37] также позволяют рассчитывать на успех в решении аналогичных задач. Разрабатываются приёмники, способные работать с обеими упомянутыми системами, что может увеличить точность определения атмосферных параметров.

Вместе с тем, имея в виду перспективы решения проблемы на базе более совершенных систем, продолжались попытки извлечь информацию об атмосферных параметрах из полученных данных измерений сигналов “Транзит"на основе более совершенных методик и алгоритмов. Опыт такой работы показал, что использование атмосферного вклада в допплеровскую частоту в качестве исходной величины для анализа, возможно, не является оптимальным. Во-первых, реально измеряемой величиной является изменение фазы сигнала, т. е. для определения допплеровского сдвига необходимо выполнить операцию дифференцирования экспериментальных данных, что, как хорошо известно [38], является некорректной задачей и может приводить к серьёзным погрешностям. Особенно это становится существенным на малых (наиболее информативных) углах места, где в эксперименте наблюдались сильные вариации фазы сигнала, существенно превосходящие величины, ожидаемые из расчётов по метеозондовым данным. Во-вторых, процедура пересчёта допплеровской частоты в рефракцию связана с использованием нелинейного соотношения между величинами, что также может увеличивать погрешность.

В настоящей работе рассматривается новый подход к решению задачи восстановления профиля показателя преломления (индекса рефракции), основанный на решении интегрального уравнения непосредственно для атмосферного вклада в фазу сигнала. Это уравнение, в отличие от уравнения для рефракции, нелинейно, но, как оказалось, отсутствие упомянутых выше погрешностей приводит к более точному решению и позволило применить развитый подход к имеющимся данным измерений сигналов спутников системы “Транзит".

Оценки точности метода были получены методом численного моделирования с использованием ансамблей метеозондовых данных. Получены также оценки возможной точности определения высотного профиля концентрации водяного пара $q(h)$ и интегрального содержания водяного пара Q по высотному профилю индекса рефракции $N(h)$, восстановленному из решения обратной задачи измерений фазы (ЭДПР) на малых углах места. В результате численных экспериментов определены требования к точности информативных измерений.

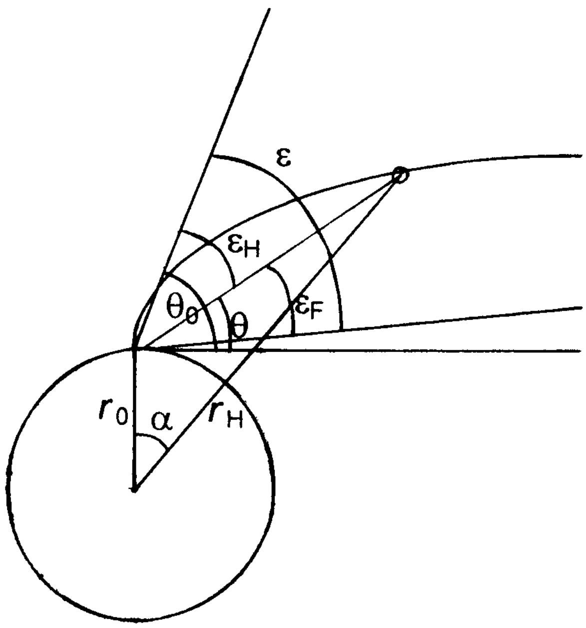


Рис. 1.

1. МЕТОДИКА ВОССТАНОВЛЕНИЯ ПРОФИЛЯ ПОКАЗАТЕЛЯ ПРЕЛОМЛЕНИЯ ПО ХАРАКТЕРИСТИКАМ СИГНАЛА ИСЗ

Геометрия задачи представлена на рис. 1. Для исключения вклада ионосферы во всех навигационных системах используется излучение на двух разнесённых частотах, поэтому к этому вопросу мы далее не будем возвращаться. Формулировка обратной задачи относительно фазы сигнала ИСЗ является достаточно сложной проблемой. Реально измеряемой величиной является разность фазы $\Delta\varphi$ сигнала ИСЗ, измеренной в двух точках его орбиты, которым соответствуют истинные углы места ϑ_1 и ϑ_2. Кроме того, известными данными можно считать координаты и компоненты вектора скорости ИСЗ, которые передаются в сигнале в закодированном виде вместе с соответствующими данными о времени.

Разность фаз сигнала источника при углах места ϑ_1 и ϑ_2 может быть представлена как

$$\begin{aligned}\Delta\varphi &= \varphi(\theta_1) - \varphi(\theta_2) = \frac{\omega}{c}\int\limits_{r_0}^{r_{H_1}} n(r)\frac{dr}{\sqrt{1-\left(\dfrac{n_0 r_0\cos\theta_{0_1}}{nr}\right)^2}} - \\ &- \frac{\omega}{c}\int\limits_{r_0}^{r_{H_2}} n(r)\frac{dr}{\sqrt{1-\left(\dfrac{n_0 r_0\cos\theta_{0_2}}{nr}\right)^2}},\end{aligned} \tag{1}$$

где $\omega = 2\pi\nu$ (ν — частота сигнала), c — скорость света, $n(r)$ — зависимость показателя преломления атмосферного воздуха от геоцентрического расстояния, r_H — геоцентрическое расстояние до спутника в двух точках его траектории с высотами над земной поверхностью H_1 и H_2, r_0, n_0 — радиус Земли и приземное значение показателя преломления, а θ_0 — углы места прихода луча в точку приёма, которые отличаются от истинных углов места на величину атмосферной рефракции ε_H:

$$\theta_0 = \theta + \varepsilon_H\,. \tag{2}$$

В свою очередь, величина атмосферной рефракции ε_H, зависящая от координат передатчика, отличается от величины астрономической рефракции (эквивалентного бесконечно удалённого вдоль луча источника) ε на величину фотограмметрической рефракции ε_{F}:

$$\varepsilon_H = \varepsilon - \varepsilon_{\mathrm{F}}\,, \tag{3}$$

откуда

$$\varepsilon = \varepsilon_H + \varepsilon_{\mathrm{F}}\,, \tag{4}$$

где

$$\varepsilon = -\int\limits_{\kappa_0}^{\infty} \frac{d\ln n}{dr}\,\frac{n_0 r_0 \cos\theta_0}{\sqrt{(nr)^2 - (n_0 r_0 \cos\theta_0)^2}}dr\,. \tag{5}$$

Здесь следует отметить следующее. Наиболее последовательным, в математическом смысле, способом решения задачи представляется изложенный в [29–30] метод, основанный на определении угла места прихода луча θ_0 из выражения для допплеровского сдвига частоты сигнала в атмосфере [23]:

$$\nu_{\mathrm{D}}(\theta) = \frac{\nu}{c r_H}\left(p_0 \cos\theta_0\, V_\perp + \sqrt{r_H^2 - p_\theta^2}\, V_{\mathrm{R}}\right)\,, \tag{6}$$

где $p_0 = r_0 n_0$, $p_\theta = r_0 n_0 \cos\theta_0$ с последующим решением интегрального уравнения (5), записанного относительно индекса рефракции $N = 10^6(n-1)$ в виде линейного интегрального уравнения типа Фредгольма 1-го рода:

$$\begin{aligned}\widetilde{\varepsilon}(p_\theta) &= \int\limits_{p_0}^{\infty} N(p)\,\frac{p p_\theta}{(p^2 - p_\theta^2)^{3/2}}\,dp\,,\\ \widetilde{\varepsilon} &= -10^6\varepsilon + p_\theta \frac{N(p_0)}{\sqrt{p^2 - p_\theta^2}}\end{aligned} \tag{7}$$

и вычислением масштаба высоты из

$$h = \frac{p}{n} - r_0 = \frac{p}{1 + 10^{-6}N(p)}\,. \tag{8}$$

Однако, как уже отмечалось, такой подход оказался неприменим при интерпретации данных системы “Транзит". В данной работе разработан и применён альтернативный подход, основанный на использовании измерений разности фаз. Соотношение (1) можно записать через разность электрической длины пути радиоволн (ЭДПР), связанную с измеряемой разностью фаз $\Delta\varphi$ и длин пути Δl (вдоль луча) как

$$\begin{aligned}\Delta L &= L(\theta_1) - L(\theta_2) = \int\limits_0^{\infty} N(h)\frac{10^{-6}dh}{\sqrt{1 - \left(\dfrac{n_0 r_0 \cos\theta_{0_1}}{nr}\right)^2}} - \\ &- \int\limits_0^{\infty} N(h)\frac{10^{-6}dh}{\sqrt{1 - \left(\dfrac{n_0 r_0 \cos\theta_{0_2}}{nr}\right)^2}}\,, \\ \Delta L &= \frac{\lambda}{2\pi}\Delta\varphi - \Delta l\,,\end{aligned} \tag{9}$$

где длина пути от точки приёма до передатчика вдоль луча l определяется из

$$l = \int\limits_{r_0}^{r_H} \frac{dr}{\sqrt{1 - \left(\frac{n_0 r_0}{n_0 r_0 \cos\theta_0}\right)^2}}\,. \tag{10}$$

Дальность вдоль луча из-за его кривизны в атмосфере отличается от прямой дальности d на некоторую величину Δd, которую следует учитывать при анализе, т. е.

$$l = d + \Delta d\,. \tag{11}$$

Для измерений разности фаз, начиная от некоторого фиксированного значения угла места θ_2, соотношение (9) переходит в

$$\Delta L(\theta) = L(\theta) - L(\theta_2) = \tag{12}$$

$$= \int\limits_0^\infty N(h)\cdot 10^{-6} \left(\frac{1}{\sqrt{1 - \left(\frac{n_0 r_0 \cos\theta_{0_1}}{nr}\right)^2}} - \frac{1}{\sqrt{1 - \left(\frac{n_0 r_0 \cos\theta_{0_2}}{nr}\right)^2}} \right) dh\,.$$

Разность ЭДПР (12) может также определяться по времени распространения сигнала от спутника до приёмника из кодов времени, передаваемых в сигнале ИСЗ. В частности, в системе GPS такую информацию содержат коды P и C/A. Передаваемая в кодах сигнала траекторная информация имеет некоторую избыточность, что также может использоваться для выбора наиболее надёжных данных. Для исключения ошибок в ходе часов применяется метод разностных измерений для двух разнесённых пунктов с точно известными координатами и метод двойных разностей, когда в этих двух пунктах сравниваются разностные сигналы двух спутников [34–36]. В первом случае исключается погрешность хода часов спутника, во втором — также и хода часов в точке приёма.

Итак, задача состоит в решении интегрального уравнения 1-го рода (12) относительно высотного профиля индекса рефракции. Уравнение имеет нелинейное ядро, в которое входит искомый профиль, поэтому представляется естественным применить для его решения итерационный алгоритм. Были применены два варианта. В первом из них на первом шаге итерационного алгоритма в ядро (12) подставлялся средний профиль $\langle N(h)\rangle$, вычисленный с использованием ансамбля профилей метеорологического зондирования, а во втором — профиль индекса рефракции, полученный методом статистической экстраполяции от его приземного значения:

$$N^{\text{э}}(h) = \langle N(h)\rangle + \frac{\mathbf{B}_{NN}(0,h)}{\sigma_N^2(0,0)}\,(N_0 - \langle N_0\rangle), \tag{13}$$

где $\mathbf{B}_{NN}$ — межуровневая (по высоте) ковариационная матрица индекса рефракции. Линеаризованное одним из указанных способов уравнение (12) представляет собой интегральное уравнение Фредгольма 1-го рода. Итерационный алгоритм в операторном виде можно записать как

$$\mathbf{K}(N_{i-1})\vec{N}_i^* = \Delta\vec{L}^\delta\,, \tag{14}$$

где $\Delta\vec{L}^\delta$ — вектор данных, заданных с некоторой погрешностью δ, звёздочка соответствует вектор-строке. Для решения уравнений рассматриваемого типа, которые являются некорректными [38], должны применяться методы регуляризации, вносящие дополнительную априорную информацию о точном решении. В данном случае так же, как и в [24—30], применялись методы статистической регуляризации и Тихонова. После того, как решение линеаризованного уравнения оказывается получено в первом приближении с использованием любого из методов регуляризации, на следующем шаге итерационного алгоритма в ядро $\mathbf{K}$ подставляется восстановленный профиль $N(h)$. Двух—трёх итераций оказывается вполне достаточно для достижения сходимости к приближённому решению.

Существенным моментом постановки задачи является определение вектора данных и угла прихода луча θ_0 в ядре интегрального уравнения в (12), (14). И угол прихода, и расстояние до передатчика вдоль луча, входящее в вектор данных, зависят от искомого распределения показателя преломления, что привносит в задачу дополнительную нелинейность. Если при учёте нелинейности ядра мы исходим в первом приближении из среднего $\langle N(h)\rangle$ или статистически экстраполированного профиля $N^{э}(h)$, то, естественно, этот же профиль должен быть использован при определении других параметров. Будем исходить из координат передатчика (θ, H) — истинного угла места и высоты передатчика над земной поверхностью. Тогда длина трассы по прямой определяется как

$$d = \sqrt{r_0^2 - 2r_0 r_H \cos\theta + r_H^2}\,, \tag{15}$$

а центральный угол α, отсчитываемый от положения приёмника (см. рис. 1), можно вычислить из

$$\alpha = \arcsin\left[\frac{d}{r_H}\cos(\theta)\right]. \tag{16}$$

Этот же угол определяется интегрированием до высоты передатчика вдоль пути луча:

$$\alpha = \int\limits_{r_0}^{r_H} \frac{n_0 r_0 \cos\theta_0}{nr^2\sqrt{1-\left(\dfrac{n_0 r_0\cos\theta_0}{nr}\right)}}\,dr\,. \tag{17}$$

Условие равенства (16) и (17) определяет угол прихода луча θ_0, соответствующий заданным профилю показателя преломления и положению передатчика. Это значение используется как в ядре интегрального уравнения (12), (14), так и в (10) для определения дальности l вдоль пути луча, которая также используется в (12), (14) при определении ΔL. На последующих шагах итерационного процесса для аналогичных целей используется уже профиль показателя преломления, восстановленный из решения уравнения (12), (14). Можно также использовать на первом шаге итерационного процесса значение угла θ_0, определённое из (6) по допплеровскому сдвигу частоты, если соответствующая ошибка меньше величины естественных вариаций рефракции.

Кратко опишем методы регуляризации, использованные для решения задачи на каждом шаге итерационного алгоритма. В методе статистической регуляризации (или максимальной энтропии) [39] решение ищется на априорном ансамбле, заданном межуровневой (по высоте) ковариационной матрицей индекса рефракции $\mathbf{B}_{NN}$, где плотность вероятности определяется из условия максимума энтропии (минимума дополнительной информации) об искомом векторе $\vec{N}$. Мера сглаживания приближённого решения зависит от ковариационной матрицы погрешностей данных измерений $\mathbf{W}$. В данной работе полагалось, что ошибки измерений некоррелированы, поэтому в матрице $\mathbf{W}$ ненулевыми являются лишь диагональные элементы, равные дисперсиям правой части (14) на соответствующих углах места. Апостериорное распределение находится по формуле Бейеса, а средний по этому распределению вектор $\vec{N}$ при заданной правой части $\Delta\vec{L}$ определяет решение

$$\vec{N} = \langle\vec{N}\rangle + \left(\mathbf{K}^*\mathbf{W}^{-1}\mathbf{K} + \mathbf{B}_{NN}^{-1}\right)\mathbf{K}^*\mathbf{W}^{-1}\left(\Delta\vec{L}^{\delta} - \langle\Delta\vec{L}\rangle\right). \tag{18}$$

В методе статистической регуляризации при стремлении к нулю погрешностей измерений решение сходится к точному в среднеквадратичном смысле.

Метод Тихонова [38] в форме принципа обобщённой невязки использует весьма общую информацию о квадратичной суммируемости искомой функции и её производной, а также увязывает меру сглаживания решения с уровнем интегральной погрешности данных измерений δL, определяемая как

$$\delta L^2 = \left\|\mathbf{K}\vec{N} - \Delta\vec{L}^\delta\right\|^2_{L_2} = \frac{1}{\theta_2-\theta_1}\int\limits_{\theta_1}^{\theta_2}\left[\Delta L(\theta) - \Delta L^\delta(\theta)\right]^2 d\theta\,, \tag{19}$$

где ΔL — правая часть (14), которая соответствует точному решению $N(h)$, ΔL^δ — данные, измеренные с некоторой ошибкой, $\left\|x\right\|^2_{L_2}$ — норма функции x в пространстве L_2 [38].

В методе Тихонова [38] приближённое решение минимизирует сглаживающий функционал

$$M^\alpha(N) = \left\|\mathbf{K}_h\vec{N} - \Delta\vec{L}\right\|^2_{L_2} + \alpha\left\|\vec{N}\right\|^2_{W_2^1}\,, \tag{20}$$

т. е. может быть найдено путём его минимизации. В приведённых выше соотношениях $\left\|N\right\|^2_{W_2^1} = \dfrac{1}{h^{\max}}\displaystyle\int\limits_0^{h^{\max}}\left[N^2(h) + \left(h^{\max}\frac{dN(h)}{dh}\right)^2\right]dh$ обозначает норму функции $N(h)$ как элемента функционального пространства W_2^1 (пространство квадратично суммируемых функций, имеющих квадратично суммируемые производные) [38], $\mathbf{K}_h$ — приближённое ядро уравнения. Задача минимизации выпуклого функционала, каковым является (20), после соответствующей дискретизации сводится к своему конечномерному аналогу, представляющему собой хорошо изученную, с вычислительной точки зрения, задачу квадратичного программирования, и решается стандартными градиентными методами. В данной работе применялся метод сопряжённых градиентов (в [38] приводится его алгоритм на языке Fortran). Численная реализация метода, в данной работе написанного на языке Borland Pascal 7.0, решает задачу за время около секунды на персональном компьютере типа IBM Pentium-200.

Принципиальным элементом уравнения (20) является параметр регуляризации α, который определяет степень сглаженности приближённого решения. Именно второе (стабилизирующее) слагаемое в (20) обеспечивает выпуклость, а следовательно, и саму возможность минимизации функционала и решения задачи. Полученное решение выделяет среди множества функций, удовлетворяющих исходному некорректному уравнению, такую, которая является минимальной в смысле используемой нормы стабилизирующего слагаемого W_2^1, содержащей и саму функцию, и её производную, т. е. реализует условие некоторого компромисса минимальности по абсолютной величине и гладкости для искомой функции. Если искать решение как отклонение от разумного среднего и модельного распределения, то минимизироваться будет именно отклонение от этого распределения, что может использоваться для разработки оптимальных алгоритмов.

Как показано в [38], параметр регуляризации (и это является фундаментальным достоинством метода обобщённой невязки) оказывается однозначно связанным с интегральной мерой погрешности данных (числом), убывая по мере убывания уровня погрешности, но более медленно. При этом по мере увеличения точности роль второго стабилизирующего слагаемого в (20) постепенно уменьшается. Параметр α находится как корень одномерного нелинейного уравнения обобщённой невязки:

$$\rho(\alpha) = \left\|\mathbf{K}_h\vec{N}^\alpha - \Delta\vec{L}^\delta\right\|^2_{l_2} - \delta^2 = 0\,, \tag{21}$$

где N^α — функция, минимизирующая функционал (20), т. е. алгебраическое уравнение (21) решается совместно с функциональным уравнением (20). Смысл (21) состоит в том, что норма невязки полученного решения должна быть в точности равна норме погрешности, поскольку нет оснований минимизировать отклонение от данных измерений за пределами уровня ошибок. В уравнение (21) входит параметр эффективной погрешности δ, который должен быть определён априори, исходя из конкретных условий решения задачи. Этот параметр должен включать в себя все составляющие ошибок измерения и интерпретации. В частности, в δ должна быть включена ошибка измерений δL, как случайная, так и систематическая. В рамках метода может быть учтена также погрешность ядра $\mathbf{K}_h$, которая включает в себя погрешность линеаризации и дискретизации при численном решении

$$\delta_h^2 = \left\| \mathbf{K}_h \vec{N} - \mathbf{K} \vec{N} \right\|_{L_2}^2 , \tag{22}$$

где $\mathbf{K}_h$ — задаваемое при решении (14) приближённое ядро. Указанные погрешности могут приводить также к несовместности вектора данных с решаемым уравнением, поскольку сглаживающее действие ядра ограничивает класс возможных реализаций $\Delta L(\theta)$; и при наличии случайной погрешности функция $\Delta L^\delta(\theta)$ может выйти из допустимого класса, т. е. ни при каком распределении $N(h)$ невозможно получить измеренное распределение $\Delta L^\delta(\theta)$. Мера несовместности δ_μ не может, естественно, превосходить суммарной погрешности ядра и измерений:

$$\delta_\mu^2 = \left\| \mathbf{K}_h \vec{N} - \Delta \vec{L}^\delta \right\|_{L_2}^2 \leq (\delta L + \delta_h)^2 . \tag{23}$$

Таким образом, в качестве параметра эффективной погрешности принимается

$$\delta^2 = (\delta L + \delta_h)^2 + \delta_\mu^2 , \tag{24}$$

где учтены ошибки измерений, дискретизации, линеаризации и других неточностей описания ядра, а также зависящая от этих факторов мера возможной несовместности уравнения со своей правой частью. В методе Тихонова значения входящих в (24) параметров должны представлять собой соответствующие оценки по максимуму на классе возможных реализаций для искомой функции. При этом величина параметра регуляризации α, а следовательно, и степень сглаженности решения определяются одним числом — значением параметра эффективной погрешности δ. Последнее обстоятельство является весьма существенным достоинством метода, поскольку теперь субъективизм исследователя переносится с области интерпретации экспериментальных данных на оценку погрешности своих измерений. Поскольку оценка погрешности всегда содержит некоторую неопределённость, существует возможность выбора стратегии решения. Так, если ставится задача заведомого исключения несуществующих деталей в решении, то лучше принять оценку погрешности с некоторым избытком, что, конечно, может привести к заглаживанию некоторых реальных деталей тонкой структуры. Если при решении задачи более важным представляется не пропустить эти детали, то следует принять наименьшую ошибку из области её реально возможных значений. В этом случае, однако, становится возможным появление в решении и не существующих реально (ложных) деталей. Правильная оценка ошибки даёт оптимальное, в смысле метода Тихонова, решение. После определения величины δ процедура получения конечного результата становиться формальной.

Очень важное преимущество метода обобщённой невязки по сравнению с другими известными методами состоит в том, что при стремлении δ к нулю в интегральной метрике приближённое решение сходится к точному равномерно, т. е. в метрике, где нормой является максимум модуля, хотя, как правило, в отличие от корректных задач, скорость сходимости не пропорциональна уменьшению δ, а более медленная. Равномерная сходимость позволяет использовать максимум модуля отклонения решения от точного для типичных распределений индекса рефракции в качестве меры точности восстановления и не привлекать для анализа статистические ансамбли.

2. МЕТОДИКА НАХОЖДЕНИЯ ПАРАМЕТРОВ АТМОСФЕРЫ ПО ПРОФИЛЮ ИНДЕКСА РЕФРАКЦИИ

В радиодиапазоне для безоблачной атмосферы индекс рефракции (приведённый показатель преломления) определяется соотношением [31]

$$N = K_1 \frac{P}{T} + K_2 \frac{e}{T} + K_3 \frac{e}{T^2}\,, \tag{25}$$

где $K_1 = 77{,}607 \pm 0{,}013$ К/мбар, $K_2 = -6{,}007 \pm 8{,}5$ К/мбар, $K_3 = = (3{,}747 \pm 0{,}031) \cdot 10^5$ К^2/мбар, P — давление воздуха, e — парциальное давление водяного пара, Т — абсолютная температура.

Зная высотный профиль индекса рефракции $N(h)$, из (25) можно найти высотное распределение парциального давления водяного пара, считая известными температуру и давление:

$$e(h) = \frac{N(h) - K_1 \dfrac{P}{T}}{\dfrac{K_2}{T} + \dfrac{K_3}{T^2}}\,. \tag{26}$$

И, далее, можно определить профиль плотности

$$\rho(h) \equiv 2{,}16\, \frac{e(h)}{T}\,, \tag{27}$$

интегральное содержание водяного пара в атмосфере

$$Q = \int\limits_0^\infty \rho(h)\, dh\,, \tag{28}$$

и высотный профиль концентрации (удельной влажности) водяного пара

$$q(h) = \frac{0{,}6\, e(h)}{P - 0{,}37\, e(h)}\,. \tag{29}$$

Очевидно, что при определении характеристик влагосодержания из восстановленного профиля индекса рефракции профили температуры и давления, входящие в (26)–(29), должны задаваться, исходя из тех или иных модельных представлений, либо из результатов их независимых измерений. Естественные вариации вычитаемого в числителе (26) вклада сухой части индекса рефракции слишком велики, чтобы можно было обойтись в его оценке без использования хотя бы приземного значения давления. При этом естественно считать известной и приземную температуру. Тогда в (26)–(29) используется модельный линейный профиль температуры со стандартным градиентом температуры 6,5 К/км

$$T(h) = T_0 - 6{,}5\, h\,, \tag{30}$$

и соответствующий профиль давления вычисляется из профиля температуры (30) по барометрической формуле с использованием приземного значения давления P_0.

3. СТАТИСТИЧЕСКИЕ ОЦЕНКИ ТОЧНОСТИ ОПРЕДЕЛЕНИЯ АТМОСФЕРНЫХ ПАРАМЕТРОВ

В данной работе проводится численный эксперимент с использованием тех же статистических ансамблей радиозондовых данных (по 80 профилей для одного времени суток в каждом для климатических условий ЕТ РФ и тропического океана), что и в [25–29]. Схема численного эксперимента включала расчёт разности ЭДПР $\Delta L(\theta_0)$ и интегрального содержания водяного пара Q по зондовым профилям метеопараметров; набрасывание случайной нормально распределённой погрешности с заданной дисперсией δL^2 на правую часть (14); решение (14) с применением одного из методов регуляризации; вычисление по восстановленному профилю индекса рефракции профилей концентрации водяного пара и его интегрального содержания; статистический анализ погрешностей определения атмосферных параметров с целью определения зависимости точности восстановления от уровня моделируемой погрешности. Полагалось, что задаваемая при моделировании ошибка включает в себя все виды ошибок, кроме погрешности линеаризации.

Как обычно (см. [25–29]), решение обратной задачи включает расчёты средних значений и среднеквадратичных естественных вариаций измеряемых и восстанавливаемых величин с целью получения исходных критериев для качества измерений и восстановления. Результаты такого анализа для решения обратной задачи, а именно, расчёты вариаций рефракции, а также её вариаций при фиксированном приземном значении индекса рефракции, определяющие информативность измерений, представлены в [29]. Здесь мы приведём данные о связанных с погодной изменчивостью вариациях параметров для рассматриваемой обратной задачи фазовых измерений. В таблице представлены результаты для ансамбля данных ЕТ РФ, соответствующие летним континентальным условиям.

Таблица

θ_0	0,5°	1°	2°	5°	10°	20°	60°	90°
$\langle L\rangle$, см	8102	6680	4851	2528	1359	704,6	280,1	242,5
σ_L, см	267	185	112	50	26	13	5,3	4,6
σ_L/N_0, см	49	48	42	23	12	6,3	2,5	2,2
$\langle\Delta d\rangle$, см	248	147	60	9,3	1,4	-	-	-
σ_d, см	38	18	5,6	0,7	0,1	-	-	-
$\langle\varepsilon\rangle$	0,58°	0,48°	0,35°	0,18°	-	-	-	-
σ_ε	$2'35''$	$1'55''$	$1'13''$	$33''$	-	-	-	-
σ_ε/N_0	$41''$	$21''$	$7''$	$1''$	-	-	-	-

Для тропических условий все вариации и погрешности возрастают примерно в 1,5–2 раза. Помимо погодных вариаций, приведённых в таблице, в естественную изменчивость приведённых величин добавляются суточная и сезонная динамики. Суточная динамика имеет приблизительно тот же порядок величины, а сезонная может быть существенно больше. Из таблицы видно, что для информативных измерений уровень суммарной погрешности правой части δL в (14) не должен превышать величины естественных погодных вариаций σ_L, а более строго — величины вариаций при известном приземном значении показателя преломления σ_L/N_0. Погрешность определения L в системе GPS составляет в настоящее время не хуже 1,5 см для зенитного направления [34–36], что позволяет считать, что уровень информативности измерений достигнут.

Точность определения ЭДПР на низких углах места для навигационных систем, включая GPS, точно неизвестна, поскольку все приложения систем и исследования относились к достаточно высоким углам места. Поэтому в численном моделировании исследовалась зависимость точности восстановления от уровня ошибок в широком интервале их значений. Оказалось, что измерения в этом интервале являются информативными, причём даже без использования приземного значения индекса рефракции, т. е. в ядро уравнения (12), (14) на первом шаге может подставляться среднее, а не экстраполированное от приземного значения распределение (13). Исходя из этого, удалось разработать метод в варианте, где не используются никакие другие измерения, кроме параметров спутникового сигнала и в котором приземное значение индекса рефракции также получается из решения обратной задачи.

Результаты показали, что так же, как и при решении задачи рефракции из измерений доплеровской частоты в [29], практически вся информация о высотном профиле индекса рефракции содержится в данных о фазе сигнала на углах места ниже 5°, поэтому в численных экспериментах были использованы значения ЭДПР на углах места от $0{,}5^\circ$ до $5{,}5^\circ$ (11 углов). В рассмотренном интервале моделируемых ошибок дальнейшее увеличение числа и интервала углов практически не влияло на точность восстановления.

На рис. 2а представлена зависимость погрешности восстановления профиля индекса рефракции для различного уровня моделируемой ошибки (постоянной по углу места) для случая применения метода статистической регуляризации в итерационном алгоритме. На рис. 2б можно видеть соответствующие результаты для метода Тихонова. Уровень стандартного отклонения моделируемой ошибки в представленных данных перекрывает большой диапазон ($\sigma_L == 1{,}5$, 100 и 500 см).

Статистический анализ в аспекте сравнения двух методов требует установления связи интегральной ошибки (19) метода Тихонова со стандартным отклонением σ_L и средним значением (систематической ошибкой) Δ_L моделируемой случайной погрешности. Это непростая задача, поскольку оценка (19) по максимуму приводит к излишнему заглаживанию большей части данных. Если исходить из минимальной среднеквадратичной погрешности восстановления, то наилучшие результаты получались, когда использовалось не максимальное, а среднее значение (19), для которого можно получить простое соотношение

$$\delta L^2 = \frac{1}{\Delta\theta}\int\limits_{\theta_1}^{\theta_2}\left\langle\left[\Delta L(\theta)-\Delta L_m^\delta(\theta)\right]^2\right\rangle d\theta = \frac{1}{\Delta\theta}\int\limits_{\theta_1}^{\theta_2}\left[\sigma_L^2(\theta)+\Delta_L^2(\theta)\right]d\theta\,. \tag{31}$$

При постоянных параметрах Δ_L и σ_L

$$\delta L = \sqrt{\sigma_L^2(\theta)+\Delta_L^2(\theta)}\,. \tag{32}$$

Можно видеть, что при нулевой систематической ошибке $\delta L = \sigma_L$. Наоборот, при нулевой случайной ошибке — $\delta L = \Delta_L$.

Видно, как точность восстановления профилей индекса рефракции зависит от уровня моделируемой погрешности. Даже при уровне погрешности разности ЭДПР $\delta L = 5$ м измерения оказываются информативными, несмотря на то, что этот уровень превышает величину естественных вариаций. Этот, на первый взгляд, странный результат связан с коррелированностью фазы по углу места, которая при некоторой статистической избыточности дискретизации по углу места приводит к уменьшению эффективной среднеквадратичной погрешности данных, что свидетельствует о целесообразности увеличения числа точек по углу места, если есть основания считать погрешность некоррелированной.

В методе статистической регуляризации при малой моделируемой ошибке данных проявляется роль неучтённой погрешности, связанной с нелинейностью ядра уравнения (14), которая в этом случае заведомо превышает моделируемый уровень ошибки. Для $\delta L = 1{,}5$ см эффект проявляется в результатах восстановления на высотах больше 1 км (см. рис. 2а). Метод Тихонова позволяет гибко учитывать погрешность ядра, связанную с линеаризацией, включая её в параметр ошибки δ (24). Так, при $\delta L = 1{,}5$ см наилучшие результаты в методе Тихонова получались при величине $\delta = 5$ см на первом шаге итерационного процесса. Надо отметить, что некоторое преимущество метода статистической регуляризации в среднеквадратичной точности восстановления не свидетельствует о том,

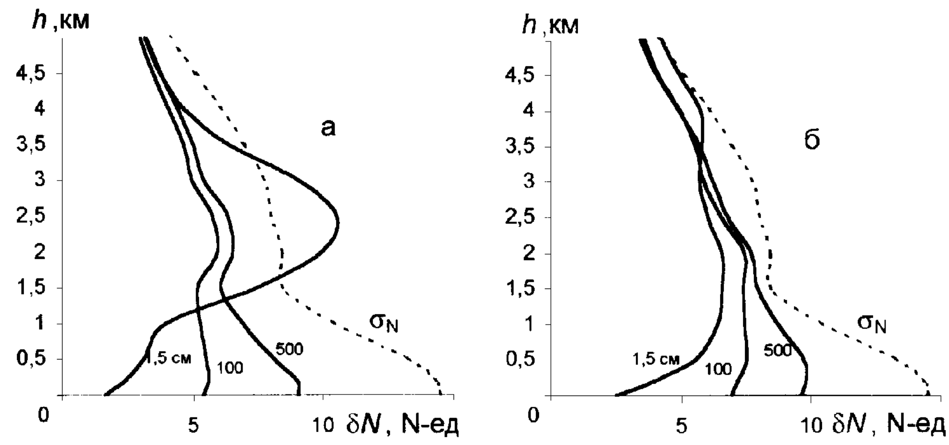


Рис. 2.

что метод Тихонова хуже, поскольку метод статистической регуляризации по определению минимизирует эту ошибку для данного статистического ансамбля. Ясно, что при наличии вариаций, связанных с суточной и сезонной динамикой, а также для других климатических условий, потери в точности для метода статистической регуляризации будут неизбежны, тогда как метод Тихонова свободен от таких ограничений. Более того, анализ восстановления конкретных профилей показывает, что метод Тихонова является более чувствительным и даёт лучшее качество восстановления, но при неизбежном наличии в статистическом ансамбле редких больших погрешностей чувствительность метода приводит к большим отклонениям и потерям в среднеквадратичной точности. Наконец, стоит напомнить, что для метода Тихонова естественной мерой погрешности является не среднеквадратичная ошибка, а отклонение по максимуму модуля.

На рис. 3 представлены среднеквадратичные погрешности восстановления профилей концентрации по восстановленному двумя методами индексу рефракции (рис. 3а — метод статистической регуляризации, рис. 3б — метод Тихонова) на основе описанной выше методики, где используются дополнительные данные о значениях приземной температуры и давления.

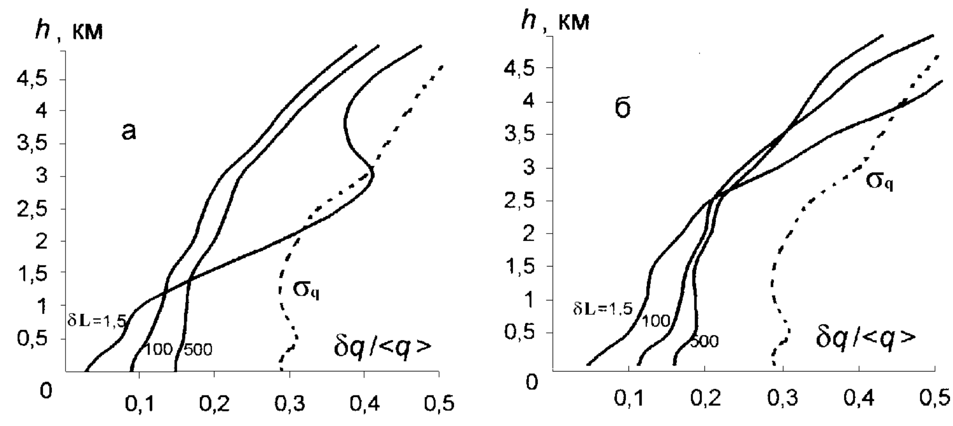


Рис. 3.

Можно видеть, что предложенный метод весьма эффективно уменьшает априорную неопределённость в профиле концентрации водяного пара, представленную величиной стандартного отклонения естественных вариаций концентрации водяного пара (отнесённых к средним значениям на соответствующих высотах). Отметим, что, если вместо линейного профиля (30) использовать точный профиль температуры, результаты улучшатся незначительно, причём, главным образом, в интервале высот выше 2 км.

Величина интегрального влагосодержания Q в [34–36] определяется по значению ЭДПР в зенитном направлении, которое легко вычисляется из разности фаз сигнала на высоких углах места. Между этими величинами существует связь, близкая к детерминированной [40]:

$$L(\pi/2) = 0{,}22717\, P_0 + \left(0{,}107 + \frac{1729{,}1}{\bar{T}}\right) Q\,, \tag{33}$$

куда входит единственный, вносящий небольшую неопределённость, параметр — средняя (с некоторым весом по высоте) температура. Для ансамбля данных, использованных в данной работе, справедливо регрессионное соотношение для статистической оценки L по величине Q

$$L(\pi/2) = 2{,}09 + 0{,}22717\, P_0 + 5{,}9\, Q\,. \tag{34}$$

Из (33)–(34) легко получаются полученные в [34–36] оценки точности определения величины Q. При достигнутой точности измерений $\delta L(90^\circ) = 1{,}5$ см погрешность определения интегрального влагосодержания составляет $\delta Q = 0{,}2$ г/см^2 или 2 мм осаждённой воды (при условии, что известно приземное значение давления воздуха, входящее в соответствующие формулы). Погрешность определения Q из (33)–(34), связанная с неточностью определения приземного давления, составляет 2 мм/мбар.

Показанная в данной работе возможность использования измерений фазы только на низких углах места ($\theta_0 \le 5^\circ$), где измерения информативны даже при погрешностях, превышающих 1 м, позволяет рассчитывать как на определение профиля концентрации водяного пара, так и его интегрального содержания путём интегрирования этого профиля из (28) при существенном снижении требований к точности определения орбитальных параметров, что, возможно, позволит отказаться от методов многопозиционных измерений, применяемых с целью уменьшения траекторных ошибок до требуемого уровня. Следует также отметить, что многие спутники не поднимаются достаточно высоко над горизонтом, чтобы применять методику [34–36], связанную с использованием измерений на высоких углах места, так что метод, использующий измерения на малых углах позволяет уменьшить среднее время между измерениями.

На рис. 4 представлены погрешности определения δQ из восстановленных профилей индекса рефракции по измерениям фазы на углах места $\theta_0 \le 5^\circ$ в зависимости от величины погрешности измерений. Сплошными линиями представлены результаты для метода Тихонова, а пунктиром — для метода статистической регуляризации. Показаны погрешности определения полной массы водяного пара $\delta Q/T_0$ при учёте линейного профиля (30) и $\delta Q/T(h)$ при использовании точного профиля температуры, по сравнению с величиной среднеквадратичных погодных вариаций σ_Q. В данном случае способ учёта температуры в (26)–(29) оказался существенным. Можно видеть, что если имеется возможность учесть текущий профиль температуры из результатов его независимых измерений, то в интервале ошибок $\delta L < 50$ см точность определения Q значительно возрастает.

Можно также видеть, что результаты применения метода Тихонова для задачи определения полной массы водяного пара оказались в целом более точными по сравнению с методом статистической регуляризации, что является объективным количественным подтверждением приведённых выше аргументов в пользу лучшего качества восстановления с использованием этого метода. Приведённые результаты показывают, что метод, основанный на измерениях при малых углах места, даёт сравнимые с [34–36] точности определения полной массы водяного пара при существенном (на два порядка) снижении требований к точности фазовых измерений и наличии преимуществ, связанных с небольшим интервалом времени измерений и независимостью от траектории спутника.

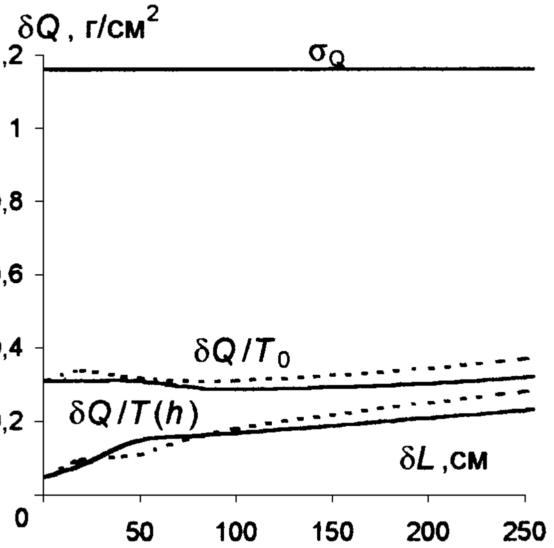


Рис. 4.

4. РЕЗУЛЬТАТЫ ВОССТАНОВЛЕНИЯ ПО ДАННЫМ ИЗМЕРЕНИЙ ПАРАМЕТРОВ СИГНАЛОВ СИСТЕМЫ "ТРАНЗИТ"

В 1993 г. в Звенигородской лаборатории ИФА РАН в пункте с точно известными координатами был выполнен цикл измерений сигналов навигационной системы "Транзит"(частоты 400 и 150 МГц). При измерениях использовался приёмник "Челн–1"с двумя антеннами типа волновой канал на каждый из каналов с шириной диаграммы направленности около 40°. Действительное значение частоты опорного генератора не выходило за пределы $\pm 1 \cdot 10^{-11}$ и непрерывно контролировалось по сигналам радиостанции РБУ с помощью фазового компаратора. Точность измерений фазы составляла $1/160$ часть фазового цикла. Номинальная точность определения допплеровской частоты при реализованных параметрах системы составляла около 0,005 Гц, что по замыслу должно было позволить реализовать точность определения рефракции $5''$, при которой измерения были бы информативны для восстановления профиля индекса рефракции, согласно [29]. В реальности из-за сильных вариаций фазы на малых углах места $\theta_0 \leq 5°$ среднеквадратичная ошибка в определении допплеровской частоты составила 0,28 Гц, что привело к ошибкам определения угла прихода луча и, соответственно, к ошибкам в определении рефракции $\delta\varepsilon = 4{,}5'$, которые, как можно видеть из таблицы, превышают величину погодных вариаций. Такие измерения, следовательно, являются неинформативными для решения обратной задачи рефракции по методике [29].

Помимо погрешностей, связанных с возможным возрастанием влияния турбулентности, влияния помех сигнала, отражённого от земной поверхности, а также возрастания роли горизонтальных неоднородностей атмосферы, следует отметить, что весьма вероятной причиной наблюдаемой в эксперименте сильной фазовой изменчивости на малых углах места может быть тонкая структура вертикальной стратификации приземного слоя атмосферы (толщиной порядка 100 м). Используемые для анализа метеозондовые данные не обладают необходимой для исследования этой тонкой структуры разрешающей способностью измерений по высоте, поэтому соответствующие вариации фазы не могут быть определены на основе расчётов. Это свидетельствует о необходимости специальных исследований с привлечением необходимых методов измерений параметров приземного слоя. Тогда, можно надеяться, что разработанный метод, может оказаться эффективен и для восстановления тонкой структуры профиля индекса рефракции в приземном слое.

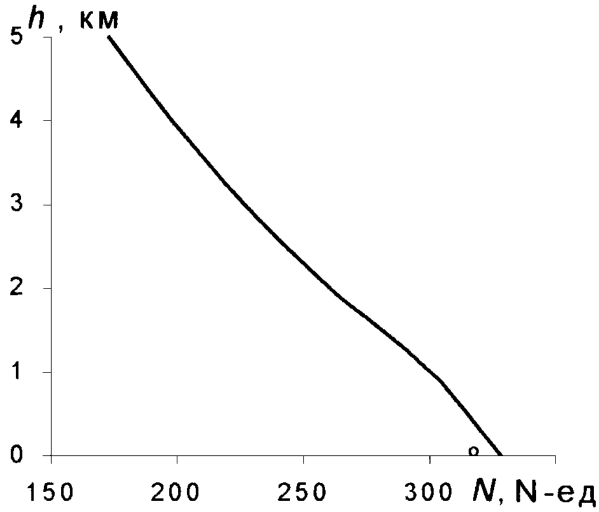


Рис. 5.

В описываемом эксперименте ошибка определения разности ЭДПР, связанная с погрешностью измерения фазы, составляла около 1 см. Определяющую роль в ошибку правой части (12), (14) вносила, очевидно, неопределённость разности расстояний вдоль луча, связанная с погрешностью определения орбитальных параметров. Номинальная точность передаваемых координат в системе "Транзит"составляет 10–20 м, однако, как показывают оценки вклада разности расстояний вдоль луча в погрешность разности ЭДПР на короткой дуге в интервале углов места $\theta_0 \le 5^\circ$, который спутник проходит за 1–2 мин, соответствующая ошибка должна быть не более 1 м. Имеет место сильная зависимость траекторной ошибки от конкретной орбиты спутника. Встречающиеся иногда потери фазового цикла, связанные с сильными вариациями фазы на низких углах места, также не превосходят указанной погрешности. Такой уровень точности, как показано в приведённых результатах численных экспериментов, является информативным для восстановления профиля индекса рефракции.

С учётом всех факторов, влияющих на точность измерений, из полученных данных для дальнейшего анализа было отобрано около 30 случаев. В ряде случаев в течение короткого времени удавалось измерить фазовые зависимости для нескольких спутников, что позволяло путём усреднения получить более надёжные результаты. К сожалению, одновременные независимые метеозондовые измерения не проводились, поэтому определялись лишь приземные метеопараметры в точке приёма.

Поскольку условия реального эксперимента не слишком хорошо соответствовали применённым в статистическом анализе ансамблям метеоданных и соответствующие ковариационные матрицы могли бы быть неадекватными реальности, для восстановления по данным реальных измерений был использован метод Тихонова. Результаты решения обратной задачи (14) без использования информации о независимо измеренных приземных значениях индекса рефракции позволили оценить их среднеквадратичное отклонение от восстановленных из решения обратной задачи значений, которое составило 7,3 N-ед , что близко к оценке ошибки для $\delta L = 100$ см (см. рис. 2). На рис. 5 приведён пример восстановления высотного профиля индекса рефракции для 15.11.93. Кружком представлено приземное значение индекса рефракции в точке приёма, определённое из прямых измерений давления, температуры и влажности.

Параметры влагосодержания не восстанавливались, поскольку измерения проводились, в основном, при отрицательных или близких к ним температурах, когда влагосодержание столь мало, что трудно рассчитывать информативность соответствующих оценок при реализованном уровне точности.

ЗАКЛЮЧЕНИЕ

Разработан итерационный метод восстановления по наблюдаемой при низких углах места фазовой зависимости сигналов навигационных ИСЗ высотных профилей индекса рефракции и связанных с ними атмосферных параметров: концентрации водяного пара его интегрального содержания. Определены среднеквадратичные указанных параметров в зависимости от величины моделируемой погрешности измерений рефракции.

Метод применён для восстановления профилей показателя преломления по данным измерений характеристик сигналов спутниковой навигационной системы “Транзит". Результаты показывают наличие перспективы применения развитой методики для дистанционного зондирования параметров атмосферы с использованием более совершенных навигационных систем GPS и GLONASS.

Авторы признательны проф. А. С. Гурвичу за внимание к работе и полезные обсуждения.

Работа выполнена при поддержке CRDF, грант № RG2-357, и Минобразования России, грант № 97–0–8.1–27.

ЛИТЕРАТУРА

1. Kliore A. J., Gain D. L., Levy G. S., Eshelman V. R. //Astron. Aeron., 1965. № T-7. P. 72.
2. Lions J. R., Sweetnam D. L., Eshleman V. R. et al. //J. Geophys. Res., 1987. V. 92. № 13. P. 14.987.
3. Fjedlbo G., Eshleman V. R. //Radio Sci., 1969. V. 4. № 10. P. 879.
4. Kliore A. J., Fjedlbo G., Seidel B. //Radio Sci., 1970. V. 5. № 2. P. 373.
5. Kliore A. J., Patel J. R., Seidel B. et al. //J. Geophys. Res., 1980. V. 85. № A-11. P. 5857.
6. Lindal G. F., Sweetnam D. L., Eshleman V. R. //Astron. J., 1985. V. 90. № 6. P. 1136.
7. Lindal G. F., Lions J. R., Sweetnam D. L. et al. //J. Geophys. Res., 1987. V. 92. № A-13. P. 14.987.
8. Lindal G. F., Lions J. R., Sweetnam D. L. et al. //J. Geophys. Res., 1990. V. 17. № 10. P. 1733.
9. Колосов М. А., Яковлев О. И., Круглов Ю. М. и др. //Радиотехника и электроника, 1972. Т. 17. № 12. С. 2483.
10. Гречко Г. М., Гурвич А. С., Романенко Ю. В., Савченко С. А., Соколовский С. В. //ДАН СССР, 1979. № 4. С. 828.
11. Соколовский С. В. //Изв. АН СССР. Сер. ФАО, 1981. Т. 17. № 6. С. 574.
12. Гурвич А. С., Кан В., Попов Л. И., Рюмин В. В., Савченко С. А. //Изв. АН СССР. Сер. ФАО, 1982. Т. 18. № 1. С. 3.
13. Гайкович К. П., Наумов А. П. //Исследование Земли из космоса, 1983. № 4. С. 25.
14. Гайкович К. П. //Изв. АН СССР. Сер. ФАО, 1984. Т. 20. № 8. С. 675.
15. Соколовский С. В. //Изв. АН СССР. Сер. ФАО, 1986. Т. 22. № 8. С. 890.
16. Волков А. А., Гречко Г. М., Гурвич А. С. и др. //Изв. АН СССР. Сер. ФАО, 1987. Т. 23. № 11. С. 1228.
17. Гречко Г. М., Гурвич А. С., Казбанов В. А. и др. //Труды ГОИ, 1989. Т. 71. Вып. 205. С. 121.
18. Бесчастнов С. П., Гречко Г. М., Гурвич А. С. и др. // Изв. АН СССР. Сер. ФАО, 1984. Т. 20. № 4. С. 231.
19. Гайкович К. П., Гурвич А. С., Наумов А. П. //Изв. АН СССР. Сер. ФАО, 1983. Т. 19. № 7. С. 675.
20. Бесчастнов С. П., Гречко Г. М., Гурвич А. С. и др. //Изв. АН СССР. Сер. ФАО, 1984. Т. 20. № 2. С. 231.
21. Загоруйко С. В., Кан В. //Радиотехника и электроника, 1984. Т. 29. № 5. С. 95.
22. Павельев А. Г. //Радиотехника и электроника, 1980. Т. 25. № 12. С. 2504.
23. Колосов М. А., Павельев А. Г. //Радиотехника и электроника, 1982. Т. 27. № 12. С. 2310.

24. Арманд Н. А., Андрианов В. А., Смирнов В. М. //Радиотехника и электроника, 1987. Т. 32. № 4. С. 673.
25. Гайкович К. П., Сумин М. И. //Изв. АН СССР. Сер. ФАО, 1986. Т. 22. № 9. С. 917.
26. Василенко Н. А., Гайкович К. П., Сумин М. И. //ДАН СССР, 1986. Т. 290. № 6. С. 1332.
27. Василенко Н. А., Гайкович К. П., Сумин М. И. //Изв. АН СССР. Сер. ФАО, 1986. Т. 22. № 10. С. 1026.
28. Василенко Н. А., Гайкович К. П., Черняева М. Б. //Изв. ВУЗов. Радиофизика, 1997. Т. 40. № 6. С. 682.
29. Гайкович К. П. //Изв. ВУЗов. Радиофизика, 1992. Т. 35. № 3–4. С. 211.
30. Gaikovich K. P. — In: Digest of IGARSS'94: Pasadena, USA, August 8–12, 1994. V. 1. P. 7.
31. Колосов М. А., Шабельников А. В. Рефракция электромагнитных волн в атмосферах Земли, Венеры, Марса. — М: Сов. радио, 1976, – 219 с.
32. Богатуров А. Н., Гайкович К. П., Гурвич А. С. и др. //ДАН СССР, 1990. Т. 315. № 4. С. 830.
33. Gaikovich K. P., Bogaturov A. N., Gurvich A. S. et al. — In: Digest of IGARSS'96: Lincoln, Nebraska, USA. May 27–31, 1996. V. 1. P.369.
34. Bevis M., Businger S., Herring T. A. et al. //J. Geophys. Res., 1992. V. 97. № D14. P. 15.787.
35. Rocken Ch., Ware R., Van Hove T. et al. //Geophys. Res. Lett., 1993. V. 20. № 23. P. 2631.
36. Tralli D. M., Lichten S. M. //Bull. Geod., 1990. V. 64. P. 127.
37. Russian's Global Navigation Satellite System. /Ed. and integr. by ANSER. USA, 1994.
38. Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. Регуляризирующие алгоритмы и априорная информация. — М.: Наука, 1983. – 200 с.
39. Турчин В. Ф., Козлов В. И., Малкевич М. С. //УФН, Т. 102. Вып. 3. С. 345.
40. Алексеев В. А., Гайкович К. П., Наумов А. П. — В сб.:Тез. докл. 13-й Всесоюзн. конф. по распространению радиоволн. Ч.2. — М: Наука, 1981. С. 80.

Институт физики атмосферы РАН,
г. Москва,
Научно-исследовательский
радиофизический институт,
г. Нижний Новгород,
Нижегородский государственный
университет
им. Н. И. Лобачевского,
Россия

Поступила в редакцию
4 декабря 1997 г.

DETERMINATION OF ATMOSPHERE PARAMETERS BY NAVIGATION SATELLITES SIGNALS

A. A. Azizov, K. P. Gaikovich, S. S. Kashkarov, M. B. Chernyaeva

The refractive index height profile is determined from the inverse problem solution using data on the elevation angle dependence of the navigation satellites signal phase. The corresponding integral equation of the 1-st kind with the nonlinear kernel is solved on the basis of iterative algorithm, on each step of which the linear Fredholm equation of the 1-st kind is solved with the application of statistical or Tikhonov's regularization methods. Using the humidity part of refraction index derived from its retrieved profile, the water vapor concentration profile and water vapor total content are determined. On the basis of numerical simulation, the statistical estimations of accuracy of the possible determination of atmosphere parameters as dependent on the measurement accuracy are obtained for meteorological data ensembles under different climate conditions. The first results of retrieval of the refraction index height profiles by satellite navigation system "Transit"are presented.

УДК 621.396, 537.87

ОСОБЕННОСТИ ПОВЕДЕНИЯ ОГИБАЮЩЕЙ СОЛИТОННОГО ИМПУЛЬСА В НЕЛИНЕЙНОЙ СРЕДЕ В СУБПИКОСЕКУНДНОМ ДИАПАЗОНЕ

И. А. Молотков, Й. Х. Повлсен, Н. И. Манаенкова

Исследуется воздействие дополнительных высших дисперсионных и нелинейных членов в модельном уравнении на свойства коротких импульсов. Для анализа использованы два взаимно дополняющих асимптотических подхода. Найдены явные формулы, описывающие искажения амплитуды и фазы огибающей сигнала. Исследованы условия, при которых дополнительные члены компенсируют друг друга.

ВВЕДЕНИЕ

В работе речь идёт об использовании обобщённого НУШ (нелинейного уравнения Шредингера)

$$i\Psi_x + \Psi_{tt} + 2|\Psi|^2\Psi - i\beta\Psi_{ttt} + i\gamma(|\Psi|^2\Psi)_t = 0, \quad \beta > 0, \ \ \gamma > 0, \tag{1}$$

для комплексной амплитуды $\Psi(x,t)$ огибающей световодного импульса. Применение субпикосекундных и фемтосекундных импульсов обусловлено многочисленными дополнительными возможностями для световодных устройств и, в частности, возможностью увеличить передаваемые мощности. Однако в указанных диапазонах необходим гораздо более аккуратный учёт нелинейных и дисперсионных эффектов. Имеются серьёзные теоретические доводы [1–4] в пользу того, что дополнительные члены с β и γ в (1) позволяют описать переход в субпикосекундный диапазон. Следует отметить, что такие члены естественно появляются при последовательном асимптотическом выводе [5] уравнения (1). Член $(|\Psi|^2\Psi)_t$ учитывался в [1, 6–11], а член Ψ_{ttt} — в [10–12]. Класс возмущений НУШ был изучен также в [13]. Полезную информацию в этом вопросе можно получить, исследуя частный световод с квадратичной зависимостью показателя преломления от поперечной координаты [14].

Однако полученные до сих пор результаты не полны и не дают возможности сравнить влияние указанных дополнительных членов и их воздействие на свойства импульсов. Наша цель здесь — проанализировать влияние членов с β и γ на уединённые решения. В последующих разделах 1 и 2 мы используем два асимптотических подхода (см. [15]). Эти подходы новые и взаимно дополняющие. Один из этих подходов связан с тем, что уравнение (1) при $\beta = \gamma = 0$ имеет двухпараметрическое солитонное решение

$$\Psi_{\mathrm{s}} = a\frac{\exp\left[i\dfrac{b}{2}t - i\left(\dfrac{b^2}{4} - a^2\right)x\right]}{\mathrm{ch}[a(t-bx)]}\,. \tag{2}$$

Здесь параметр a описывает солитонную амплитуду и его обратную ширину, b — солитонная скорость (по отношению к общему движению с групповой скоростью). Использование этих параметров — важный элемент дальнейших построений. В результате оказывается, что влияние отмеченных дополнительных членов в уравнении (1) весьма велико и ведёт к значительным изменениям структуры сигнала.

1. ДЕФОРМАЦИЯ КОРОТКИХ ИМПУЛЬСОВ ПРИ МАЛЫХ β И γ

Чтобы выяснить роль дополнительных членов в уравнении (1) введём степенные по β и γ ряды в фазу и амплитуду анзатца:

$$\left.\begin{aligned} &\Psi = a\frac{\mathrm{e}^{i\Phi}}{\operatorname{ch} Q} \\ &\Phi = \Phi_0(x,t) + \beta\varphi(x,t) + \gamma\xi(x,t) + \ldots, \\ &Q = Q_0(x,t) + \beta p(x,t) + \gamma q(x,t) + \ldots \end{aligned}\right\} \tag{3}$$

Предполагаем, что β и γ — вещественные постоянные, а φ, ξ, p, q — вещественные функции. Естественно положить

$$\Phi_0 = \frac{b}{2}t - \left(\frac{b^2}{4} - a^2\right)x\,, \quad Q_0 = a(t - bx)$$

в соответствии с формулой (2).

Подставляя анзатц (3) в уравнение (1), получаем две системы уравнений

$$\left.\begin{aligned} &\varphi_{tt} + (\operatorname{th} Q_0)p_{tt} - L\varphi + 2a\left(\frac{2}{\operatorname{ch}^2 Q_0} - 1\right)p_t = \\ &\qquad = \frac{3}{2}a^2 b\left(1 - \frac{2}{\operatorname{ch}^2 Q_0}\right) + \frac{b^3}{8}, \\ &Lp - 2a\varphi_t = -\frac{3}{4}ab^2 + a^3\left(1 - \frac{6}{\operatorname{ch}^2 Q_0}\right), \end{aligned}\right\} \tag{4}$$

$$\left.\begin{aligned} &\xi_{tt} + (\operatorname{th} Q_0)q_{tt} - L\xi + 2a\left(\frac{2}{\operatorname{ch}^2 Q_0} - 1\right)q_{tt} = \frac{a^2 b}{2\operatorname{ch}^2 Q_0}, \\ &Lq - 2a\xi_t = \frac{3a^3}{\operatorname{ch}^2 Q_0}\,. \end{aligned}\right\} \tag{5}$$

Здесь $L \equiv \dfrac{\partial}{\partial x} + b\dfrac{\partial}{\partial t}$. Конечно, в уравнениях старшего порядка влияния членов с β и с γ разделяются. Система (4) описывает воздействие члена с β, система (5) — аналогичное воздействие члена с γ.

Наиболее интересный случай решения выписанных систем соответствует условию

$$a \gg 1, \tag{6}$$

когда начальная амплитуда солитона велика.

Начнём с системы (5). В случае (6) эта система имеет простое решение

$$\xi = -\frac{3a}{2}\operatorname{th} Q_0 + O(a^0),$$

$$q = -\frac{a}{5}\left(\operatorname{ch} Q_0 + 3\ln\operatorname{ch} Q_0\right) + O(a^0),$$

которое ведёт к следующей формуле для вещественной амплитуды:

$$A = \frac{1}{\operatorname{ch}[Q_0 + \gamma q(Q_0)]}\,, \quad Q_0 = a(t - bx), \tag{7}$$

нормированной на единицу в точке её максимума.

Будем рассматривать поведение функции $A(Q_0)$. Предположим, что произведение $a\gamma = \alpha$ мало, несмотря на неравенство (6). Даже для малых значений параметра α формула (7) даёт значительное изменение формы импульса. Амплитудный максимум сдвигается вправо, в точку $Q_0 = 0{,}2\alpha$. График амплитуды делается несимметричным. Правая точка перегиба (на заднем фронте импульса) смещается влево на величину $-0{,}1\alpha$ и вверх на $0{,}4\alpha$. Левая точка перегиба (на переднем фронте) уходит заметно вправо на величину $1{,}4\alpha$ и снова вверх на $0{,}4\alpha$. Общая ширина графика уменьшается. Крутизна графика возрастает на переднем фронте и уменьшается на заднем, что вообще характерно для действия нелинейных членов. Общее изменение нормированной амплитуды $A(Q_0)$ под действием члена $i\gamma(|\Psi|^2\Psi)_t$ весьма близко к аналогичному изменению амплитуды под влиянием некоторых типов слабых продольных неоднородностей при $\beta = \gamma = 0$ [16]. По-видимому, можно указать конкретные виды продольных неоднородностей, эквивалентные по вызываемой амплитудной модуляции действию обсуждаемого дополнительного нелинейного члена.

Перейдём к случаю системы (4), описывающей влияние высшего дисперсионного члена Ψ_{ttt}. Снова предположим, что условие (6) выполнено. Получаем решение

$$\varphi = a\left(3\,\mathrm{th}\,Q_0 - \frac{1}{2}Q_0\right) + O(a^0),$$

$$p = \frac{2}{5}a\left(\mathrm{ch}^2 Q_0 + 3\ln\mathrm{ch}\,Q_0\right) + O(a^0)$$

системы (4) и амплитудную формулу

$$A = \frac{1}{\mathrm{ch}[Q_0 + \beta\, p(Q_0)]}\,. \tag{8}$$

Здесь по-прежнему $Q_0 = a(t - bx)$.

Другой знак функции $p(Q_0)$, по сравнению с $q(Q_0)$, говорит о том, что член с β вызывает противоположные по знаку деформации амплитудного графика, по сравнению с описанными выше в связи с формулой (7).

Из формул (7), (8) вытекает (см. разложение (3) для Q) общая формула для нормированной амплитуды

$$A + \frac{1}{\mathrm{ch}[Q_0 + \beta\, p(Q_0) + \gamma\, q(Q_0)]}\,,$$

описывающая амплитудную модуляцию сигнала. После подстановки $p(Q_0)$ и $q(Q_0)$ эта формула принимает вид

$$A = \frac{1}{\mathrm{ch}\left[Q_0 + \dfrac{(2\beta - \gamma)a}{5}(\mathrm{ch}^2 Q_0 + 3\ln\mathrm{ch}\,Q_0)\right]}\,. \tag{9}$$

Видно, что при

$$2\beta = \gamma \tag{10}$$

форма огибающей сигнала, по крайней мере, в старшем порядке не изменяется. Таким образом, при условии (10) дополнительные члены в уравнении (1) компенсируют друг друга.

2. ДЕФОРМАЦИЯ КОРОТКИХ ИМПУЛЬСОВ НА МАЛЫХ ДИСТАНЦИЯХ ПРИ КОНЕЧНЫХ β И γ

В этом пункте рассмотрим модельную ситуацию, когда отличные от нуля коэффициенты β и γ появляются лишь при $x \geq 0$. Зона $x < 0$ вообще не будет рассматриваться. Предполагая коэффициенты β и γ конечными, будем решать задачу Коши для уравнения (1) при однопараметрическом начальном условии

$$\Psi\Big|_{x=0} = a\frac{\mathrm{e}^{it}}{\mathrm{ch}(at)}\,. \tag{11}$$

Условие (11) согласуется с формулой (2). Из-за конечности β и γ изучение сформулированной задачи соответствует более глубокому проникновению в фемтосекундный диапазон.

Уравнение (1) содержит комплекснозначную огибающую

$$\Psi = T(x,t)\exp[i\,S(s,t)].$$

Для положительных функций T и S рассматриваемая задача Коши имеет вид

$$T_{tt} - (S_x + S_t^2)T + 2T^3 + \beta(S_{ttt} + 3S_{tt}T_t + 3S_tT_{tt} - S_t^2T) - \gamma S_tT^3 = 0,$$

$$TS_{tt} + T_x + 2S_tT_t + \beta(3S_tS_{tt}T + 3S_t^2T_t - T_{ttt}) + 3\gamma T^2T_t = 0,$$

$$T\Big|_{x=0} = \frac{a}{\mathrm{ch}(at)}\,, \quad S\Big|_{x=0} = t\,.$$

Будем искать сосредоточенные решения в виде степенных рядов по расстоянию x:

$$\left.\begin{aligned} T &= \frac{a}{\mathrm{ch}\,Q_0} + xT_1(t) + \ldots \\ S &= t + x[(a^2-1) + S_1(t)] + \ldots \end{aligned}\right\} \tag{12}$$

Здесь $Q_0 = a(t-2x)$, мы имеем дело с частным случаем $b = 2$ по сравнению с формулами предыдущего пункта.

На этом пути находим $T_1(t)$ и $S_1(t)$. Объединяя члены $a\,\mathrm{ch}^{-1}\,Q_0$ и $xT_1(t)$, находим приближённую формулу для нормированной амплитуды

$$A = \frac{T(x,t)}{a} + \frac{1}{\mathrm{ch}[f(Q_0,x)]}\,, \tag{13}$$

$$f(Q_0,x) = Q_0 + a(a^2-3)\beta x - \frac{3a^3(2\beta+\gamma)}{\mathrm{ch}\,Q_0}\,,$$

внешне подобную формулам (7)–(9). Однако здесь β и γ конечны, а дистанция x вдоль волновода не может быть протяжённой.

При $Q_0 > 0$ член с $\mathrm{ch}^{-1}\,Q_0$ в выражении $f(Q_0,x)$ становится несущественным.

Отсюда важный вывод из сравнения формулы (13) с формулой (9): при конечных β и γ исчезает эффект компенсации дополнительных членов в уравнении (1). Амплитудная модуляция при конечных β и γ вызывается, в старшем порядке, лишь членом с β.

3. ФАЗОВЫЕ ИСКАЖЕНИЯ

А. Случай малых β и γ

На основании формул разд. 1 при условии (6) получаем

$$\Phi = \Phi_0 + a\left[\left(3\beta - \frac{3}{2}\gamma\right)\operatorname{th} Q_0 - \frac{\beta}{2}Q_0\right] + O(\beta) + O(\gamma)\,. \tag{14}$$

Нарастание или убывание возмущённой фазы Φ по сравнению с невозмущённой Φ_0 зависит от соотношения констант β и γ, а также от $Q_0 = a(t - bx)$. При малых Q_0 (центральная часть импульса) и $5\beta > 3\gamma$ имеет место дополнительное нарастание фазы. При выполнении обратного неравенства происходит убывание фазы.

При больших $|Q_0|$ (на крыльях импульса) член с γ делается несущественным, при этом положительным Q_0 отвечает уменьшение фазы по сравнению с невозмущёнными значениями.

Б. Случай конечных β, γ и малых дистанций x

Используем формулу (12) для фазы S, в которой главная возмущающая добавка обозначена через S_1. Для S_1 может быть выведена асимптотическая формула

$$S_1(Q_0) = a^2\left[3\beta - \frac{6\beta + \gamma}{\operatorname{ch}^2 Q_0}\right] + O(a^0). \tag{15}$$

При положительных Q_0 в выписанной в (15) скобке основным является первое слагаемое, а фаза S практически всегда нарастает за счёт возмущения. В отличие от случая малых β и γ конкуренция между этими коэффициентами отсутствует.

Сопоставление формул, относящихся к случаям А и Б, позволяет сделать следующий вывод. В случае малых коэффициентов β и γ, таких, что

$$\beta \approx 0{,}6\gamma\,, \tag{16}$$

имеет место эффект компенсации рассматриваемых дополнительных членов. При увеличении этих коэффициентов явление компенсации исчезает. Дополнительный нелинейный член с γ перестаёт конкурировать с высшим дисперсионным членом (с коэффициентом β). В свою очередь, этот дисперсионный член даёт дополнительное нарастание фазы. Следует отметить, что главные (невозмущённые) члены фазы Φ_0 и $t + x(a^2 - 1)$ при двух рассмотренных подходах совпадают.

ЗАКЛЮЧЕНИЕ

В работе исследовано влияние добавочных, по сравнению со стандартным НУШ, членов $-i\beta\Psi_{ttt}$, $i\gamma(|\Psi|^2\Psi)_t$ на поведение огибающей сверхкороткого импульса. Главные результаты состоят в следующем.

1. Получены явные формулы (7)–(9), (13), определяющие деформацию огибающей импульса. Эти формулы описывают появление асимметрии, изменение крутизны и другие конкретные характеристики изменения формы импульса.
2. Получены явные формулы (14) и (15), характеризующие изменения фазы под воздействием указанных высших дисперсионных и нелинейных членов.

3. Установлено, что при малых значениях коэффициентов β и γ имеет место конкуренция указанных дополнительных членов уравнения (1). При определённых соотношениях между β и γ (см. (10) и (16)) возможна взаимная компенсация эффектов, порождаемых этими членами. При бóльших значениях коэффициентов β, γ явление компенсации исчезает, а из двух дополнительных членов член с β оказывается превалирующим.

Работа выполнена при поддержке РФФИ (проект 96–02–18050).

ЛИТЕРАТУРА

1. Anderson D. and Lisak M. //Phys. Rev. A., 1983. V. 27. P. 1393.
2. Christodoulides D. N. and Joseph R. I. //Electron Lett., 1984. V. 20. P. 659.
3. Hasegava A. and Kodama Y. //Proc. IEEE, 1981. V. 69. P. 1145.
4. Грудинин А. Б., Меньшов В. Н, Фурса Т. Н. //ЖЭТФ, 1990. Т. 97. № 2. С. 449.
5. Bisyarin M. A. and Molotkov I. A. //Opt. and Quant. Electr., 1992. V. 24. № 3. P. 303.
6. Камчатнов А. М. //ЖЭТФ, 1990. Т. 97. № 1. С. 144.
7. Земсков В. Р. //ЖТФ, 1992. Т. 62. № 11. С. 167.
8. Hayata K. and Koshiba M. //J. Opt. Soc. Amer., ser. B, 1994. V. 11. № 1. P. 61.
9. Маймистов А. И. //ЖЭТФ, 1993. Т. 104. № 5(11). С. 3620.
10. Громов Е. М., Таланов В. И. //ЖЭТФ, 1996. Т. 110. № 1(7). С. 137.
11. Abdullaev F. K., Darmanyan S. A., Bischoff S., and Soerensen M. P. //J. Opt. Soc. Amer., ser. B, 1997. V. 14. № 1. P. 27.
12. Elgin J. N. //Opt. Lett., 1992. V. 17. № 20. P. 1409.
13. Rottwitt K., Hermann B., Povlsen J. H., and Elgin J. N. //J. Opt. Soc. Amer., ser. B, 1995. V. 12. № 7. P. 1307.
14. Bisyarin M. A. and Molotkov I. A. //Proc. of the SPIE, 1995. V. 2943. P. 24.
15. Молотков И. А., Вакуленко С. А. Сосредоточенные нелинейные волны. — Л.: Изд-во Ленингр. универ., 1988. – 240 с.
16. Бисярин М. А., Молотков И. А. //Оптика и спектроскопия, 1989. Т. 67. № 2. С. 442.

Институт земного магнетизма,
ионосферы и распространения
радиоволн РАН, г. Троицк, Россия

Поступила в редакцию
17 марта 1998 г.

PECULIARITIES OF THE BEHAVIOUR OF SOLITON PULSE ENVELOPE IN THE NONLINEAR MEDIUM IN THE SUBPICOSECOND RANGE

I. A. Molotkov, J. H. Povlsen, N. I. Manaenkova

The impact of additional higher dispersion and nonlinear terms in the model equation on properties of the short pulses has been investigated. The analysis is based on two mutually complementary asymptotic approaches. We found explicit formulae describing distortions of the amplitude and phase of the signal envelope. The conditions under which the additional terms compensate each other have been studied.

УДК 535.2+535.854

НЕГОЛОНОМНОСТЬ СВЯЗИ СОСТОЯНИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В ОДНОМОДОВОМ СВЕТОВОДЕ С ЛИНЕЙНЫМ ДВУЛУЧЕПРЕЛОМЛЕНИЕМ И УГЛА ЕГО КРУЧЕНИЯ

Г. Б. Малыкин, Ю. И. Неймарк

Рассмотрена связь между состоянием электромагнитного поля в скрученном одномодовом волоконном световоде с линейным двулучепреломлением и углом его кручения, для случая, когда световод вытянут в прямую линию. Показано, что эта связь носит неголономный характер, что следует из некоммутативности для двух бесконечно малых, но различных изменений длины световода и его кручения. В частности, отсюда следует, что связь между состоянием поляризации излучения в скрученном световоде и углом его кручения также носит неголономный характер. Показано также, что связь между геометрооптической фазой Панчаратнама, которая набегает дополнительно, вследствие эволюции состояния поляризации света при его распространении в анизотропной оптической среде, которой является скрученный ОВС, и углом кручения ОВС также является неголономной.

Около 60 лет назад в работах Рытова [1, 2] было показано, что в случае распространения светового луча по неплоской траектории происходит вращение плоскости его поляризации относительно естественного трёхгранника Дарбу, образуемого единичными векторами — касательной, нормалью и бинормалью к искривлённой траектории луча. В вышедшей вскоре работе Владимирского [3] было показано, что если в некоторой точке траектории луча касательная возвращается в исходное состояние, то плоскость поляризации света, в общем случае, будет отличаться от исходной, однако если траектория луча является плоской кривой, то этого явления не будет. В [3] была также рассмотрена связь между величиной эффекта Рытова и пространственной эволюцией трёхгранника Дарбу. Как показано в [3], для случая циклической эволюции трёхгранника Дарбу, когда в некоторой точке луча трёхгранник возвращается в исходное положение, величина угла поворота плоскости поляризации света относительно исходного численно равна площади на поверхности сферы единичного радиуса, которая заключена внутри замкнутой кривой, описанной на сфере одним из векторов трёхганника Дарбу (касательной) в процессе его пространственной эволюции.

В работах [4, 5] было показано, что в одномодовом волоконном световоде (ОВС) с неплоской конфигурацией также имеет место эффект Рытова, причём, поскольку в этом случае траектория луча совпадает со световодом, то трёхгранник Дарбу можно ориентировать относительно последнего. Изменение плоскости поляризации света, обусловленное неплоской конфигурацией световода, в [5, 6] было предложено называть, по аналогии с классической механикой, параллельным переносом, который можно интерпретировать как неголономность — неинтегрируемую (неголономную) связь направления вектора электрического поля с пространственной эволюцией трёхгранника Дарбу, который ориентируется непосредственно световодом.

Отметим здесь, что величина геометрооптической фазы Рытова—Владимирского [7, 8], которая набегает дополнительно вследствие неплоской траектории луча, численно равна величине эффекта Рытова и также неголономно связана с пространственной эволюцией трёхгранника Дарбу [9, 10].

Цель настоящей работы — показать, что в ОВС с собственным линейным двулучепреломлением при наличии кручения световода (торсионной скрутки) будет иметь место неголономная связь между направлением вектора электрического поля в ОВС и углом кручения световода даже в том случае, если световод вытянут в прямую линию. (На рис. 1 изображён скрученный на угол α ОВС типа PANDA,

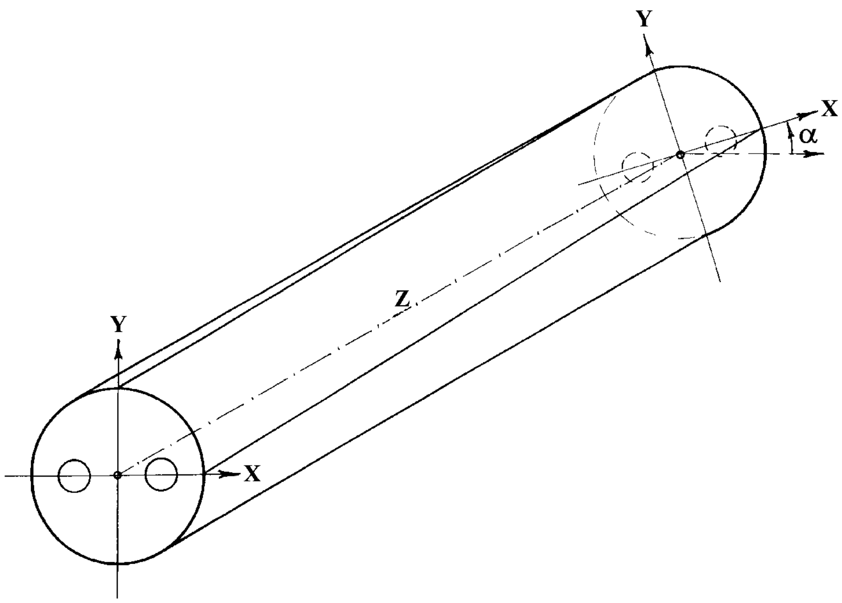


Рис. 1. Скрученный на угол α отрезок ОВС типа PANDA, с собственным линейным двулучепреломлением, длина отрезка — z. X, Y — медленная и быстрая оси линейного двулучепреломления ОВС.

линейное двулучепреломление которого создаётся за счёт двух, расположенных параллельно светонесущей жиле, цилиндрических вставок, коэффициент термического расширения которых отличается от коэффициента термического расширения световода. После вытяжки и остывания световода вставки создают механическое напряжение во всём световоде, в том числе и в светонесущей жиле, что и вызывает линейное двулучепреломление, одна из осей которого совпадает с линией, соединяющей центры вставок, другая — перпендикулярна этой линии.) В частности, если известны состояние поляризации света на входе отрезка ОВС, ориентация осей его линейного двулучепреломления на входе (α_1) и на выходе (α_2) отрезка, то, в общем случае, невозможно получить функциональное выражение для состояния поляризации на выходе отрезка ОВС через его значение на входе отрезка и значения углов α_1 и α_2. Другая цель работы — показать, что связь между геометрооптической фазой Панчаратнама [11] (см также обзоры [7–10]) и углом кручения ОВС также является неголономной.

Отметим здесь, что фаза Панчаратнама, которая в отличие от фазы Рытова–Владимирского связана не с пространственной эволюцией луча, а с эволюцией состояния поляризации света при его распространении в анизотропной оптической среде (в данном случае — в скрученном ОВС), набегает дополнительно динамической фазе, связанной с распространением света в ОВС и равной произведению волнового числа быстрой или медленной моды на длину световода. Фаза Панчаратнама определена для случая циклической эволюции состояния поляризации света и численно равна площади на сфере Пуанкаре единичного радиуса, которая ограничена замкнутой кривой на сфере Пуанкаре, соответствующей эволюции состояния поляризации света при его распространении в анизотропной среде.

Запишем дифференциальное уравнение для нормированного вектора Джонса [12]

$$\vec{E} = \begin{vmatrix} E_x e^{i\psi_x} \\ E_y e^{i\psi_y} \end{vmatrix}$$

(где E_x и E_y — амплитуды компонент электрического поля ($E_x^2 + E_y^2 = 1$), а ψ_x и ψ_y — их фазы) в ОВС без потерь света и дихроизма в декартовой, сопровождающей кручение световода, системе координат

$$\frac{d\vec{E}}{dz} = N(z)\vec{E}(z). \tag{1}$$

Здесь z — длина, отсчитываемая от начала отрезка ОВС,

$$N(z) = \begin{vmatrix} -i\beta_x & (1-g)\dfrac{d\alpha}{dz} \\ -(1-g)\dfrac{d\alpha}{dz} & -i\beta_y \end{vmatrix} \quad —$$

дифференциальная матрица Джонса ОВС [13, 14] с собственным линейным двулучепреломлением $\beta = \beta_x - \beta_y$, $\beta_{x,y} = \dfrac{2\pi}{\lambda} n_{x,y}$ (λ — длина волны света, n_x, n_y — показатели преломления в медленной и быстрой оси ОВС, здесь полагаем $n_x,\ n_y = \mathrm{const}$), и углом кручения α, $g = \mathrm{const}$ — фотоупругий коэффициент материала, из которого изготовлен световод.

Векторное соотношение (1) определяет дифференциальную связь между амплитудами и фазами ортогональных компонент волны, распространяющейся вдоль волокна при её перемещении (dz) вдоль него и соответствующем изменении угла кручения $d\alpha$. Имеется в виду, что на отрезке длины dz световода величина α изменяется на $d\alpha$. В каждом конкретном волокне изменение dz определяет $d\alpha$, но эта зависимость может быть нам не известна или известна. Но в абстрактной математической постановке задачи о голономности или неголономности связи состояния электромагнитного поля с кручением осей ОВС мы можем рассмотреть ансамбль независимых реализаций волокон, считая дифференциалы dz и $d\alpha$ независимыми. Действительно, из комплексной векторной связи (1) следуют четыре скалярных вещественных дифференциальных соотношения:

$$\begin{aligned} &\cos\psi_x\, dE_x - E_x \sin\psi_x\, d\psi_x = \beta_x E_x \sin\psi_x\, dz + (1-g)E_y\cos\psi_y\, d\alpha, \\ &\sin\psi_x\, dE_x + E_x \cos\psi_x\, d\psi_x = -\beta_x E_x \cos\psi_x\, dz + (1-g)E_y\sin\psi_y\, d\alpha, \\ &\cos\psi_y\, dE_y - E_y \sin\psi_y\, d\psi_y = \beta_y E_y \sin\psi_y\, dz - (1-g)E_x\cos\psi_x\, d\alpha, \\ &\sin\psi_y\, dE_y + E_y \cos\psi_y\, d\psi_y = -\beta_y E_y \cos\psi_y\, dz - (1-g)E_x\sin\psi_x\, d\alpha, \end{aligned} \tag{2}$$

или после разрешения их относительно dE_x, dE_y, $d\psi_x$ и $d\psi_y$:

$$\begin{aligned} &dE_x = (1-g)E_y\cos(\psi_x - \psi_y)\, d\alpha, \\ &dE_y = -(1-g)E_x\cos(\psi_x - \psi_y)\, d\alpha, \\ &d\psi_x = -\beta_x dz + (1-g)\frac{E_y}{E_x}\sin(\psi_y - \psi_x)\, d\alpha, \\ &d\psi_y = -\beta_y dz + (1-g)\frac{E_x}{E_y}\sin(\psi_y - \psi_x)\, d\alpha. \end{aligned} \tag{3}$$

Казалось бы, из (3) путём интегрирования можно найти E_x, E_y, ψ_x и ψ_y как функции z и α. Однако это не так. Не так — в общем случае, и так — в частном случае, при выполнении некоторых условий. Невозможность интегрирования носит принципиальный характер, а не обусловлена, например, технической причиной невычислимости интегралов. Дело в том, что изменения E_x, E_y, ψ_x и ψ_y не определяются только начальными и конечными значениями z и α. Они зависят от того, как менялись z и α от начальных значений к конечным, по какому пути точка с координатами z и α движется от начального положения к конечному.

В своё время, к удивлению механиков, принципиальная неинтегрируемость была обнаружена для связей качения без проскальзывания одного тела по другому и повлекла неприменимость уравнений Лагранжа, необходимость их соответствующего изменения [15]. Такие неинтегрируемые связи были названы неголономными, так же были названы и системы с такими связями. Неголономность была позднее обнаружена в электромеханических системах со скользящими электрическими контактами, в частности, в коллекторных электрических машинах [16], в гироскопических системах [17], в явлениях распространения электромагнитных волн в неоднородных средах и волокнах [5, 6]. Неголономность влечёт различие в числе степеней свободы и фазовых координат: фазовых координат оказывается больше, чем степеней свободы. В данном случае имеется пять независимых фазовых координат — E_y/E_x, ψ_x, ψ_y, z, α (одна фазовая координата утрачена вследствие того, что при отсутствии потерь в ОВС выполняется соотношение $E_x^2 + E_y^2 = 1$) — и две степени свободы, соответствующие независимым изменениям z и α.

Интегрируемость системы уравнений (2) означает принципиальную возможность получения из них конечных соотношений, определяющих E_x, E_y, ψ_x и ψ_y как однозначные функции z и α.

Математические критерии неголономности хорошо известны. Они были получены Фробениусом и носят его имя [15, 18]. Для их применения необходимо ввести в рассмотрение однородные Пфаффовы дифференциальные формы

$$\begin{aligned}
d\omega_1 &= dE_x - (1-g)E_y\cos(\psi_x - \psi_y)\,d\alpha,\\
d\omega_2 &= dE_y + (1-g)E_x\cos(\psi_x - \psi_y)\,d\alpha,\\
d\omega_3 &= d\psi_x + \beta_x dz - (1-g)\frac{E_y}{E_x}\sin(\psi_y - \psi_x)\,d\alpha,\\
d\omega_4 &= d\psi_y + \beta_y dz - (1-g)\frac{E_x}{E_y}\sin(\psi_y - \psi_x)\,d\alpha,
\end{aligned} \tag{4}$$

которые обращаются в нуль при соблюдении сотношений (3) или (2). Затем следует взять два произвольных различных изменения переменных z, α, E_x, E_y, ψ_x, ψ_y: dz, $d\alpha$, dE_x, dE_y, $d\psi_x$, $d\psi_y$ и δz, $\delta\alpha$, δE_x, δE_y, $\delta\psi_x$, $\delta\psi_y$, — каждое из которых удовлетворяет (2) или, что то же, обращает в нуль Пфаффовы формы (4). (Как было отмечено выше, изменения переменных z и α можно считать независимыми.) При этом, согласно (3), dz, $d\alpha$ и соответственно δz, $\delta\alpha$ могут быть выбраны произвольно, а dE_x, dE_y, $d\psi_x$, $d\psi_y$ и соответственно δE_x, δE_y, $\delta\psi_x$, $\delta\psi_y$ ими определяются. Связи (2) неголономны, если хотя бы для одной из дифференциальной форм (4) не имеет места обращение в нуль так называемых билинейных дифференциальных форм, т. е. тождественного равенства

$$\delta d\omega_s - d\delta\omega_s \equiv 0 \tag{5}$$

для $s = 1,\ 2,\ 3,\ 4.$

В рассматриваемом случае для первой Пфаффовой формы (4) имеем

$$
\begin{gathered}
(\delta d - d\delta)\omega_1 = \delta\{dE_x - (1-g)E_y\cos(\psi_x-\psi_y)\,d\alpha\} - \\
- d\{\delta E_x - (1-g)E_y\cos(\psi_x-\psi_y)\,\delta\alpha\} = \\
= (1-g)E_y\sin(\psi_x-\psi_y)(\delta\psi_x-\delta\psi_y)d\alpha - (1-g)\cos(\psi_x-\psi_y)\,\delta E_y\,d\alpha - \\
- (1-g)E_y\sin(\psi_x-\psi_y)(d\psi_x-d\psi_y)\,d\alpha + (1-g)\cos(\psi_x-\psi_y)\,dE_y\,\delta\alpha
\end{gathered} \tag{6}
$$

или, выражая, согласно (3), дифференциалы d и δ от E_x, E_y, ψ_x, ψ_y через дифференциалы d и δ от z и α в виде

$$
\begin{gathered}
(1-g)E_y\sin(\psi_x-\psi_y)\left\{\left[-\beta_x z + (1-g)\frac{E_y}{E_x}\sin(\psi_y-\psi_x)\delta\alpha\right]d\alpha - \right. \\
\left. - \left[-\beta_x dz + (1-g)\frac{E_x}{E_y}\sin(\psi_y-\psi_x)d\alpha\right]\delta\alpha\right\} + (1-g)\cos(\psi_x-\psi_y)\times \\
\times\{-(1-g)E_x\cos(\psi_x-\psi_y)\delta\alpha\cdot d\alpha + (1-g)E_x\cos(\psi_x-\psi_y)d\alpha\cdot\delta\alpha\} = \\
= (1-g)E_y\sin(\psi_x-\psi_y)\left\{\beta_x(\delta z\cdot d\alpha - d\alpha\cdot\delta z) + \right. \\
\left. +(1-g)\sin(\psi_y-\psi_x)\left(\frac{E_y}{E_x}-\frac{E_x}{E_y}\right)\delta\alpha\cdot d\alpha\right\}.
\end{gathered} \tag{7}
$$

Это последнее выражение не равно тождественно нулю при произвольных dz, δz и $d\alpha$, $\delta\alpha$, и поэтому имеет место неголономность (кроме случая $g = 1$). В реальных оптических материалах $0 < g < 1$. Так, в кварцевых световодах $g = 0{,}08 - 0{,}065$ [19–21].

Смысл условий Фробениуса прост: изменения величин E_x, E_y, ψ_x, ψ_y должны быть одинаковыми при любых изменениях dz, $d\alpha$ и затем δz, $\delta\alpha$, или наоборот: сначала δz, $\delta\alpha$ и затем dz, $d\alpha$. Рассмотрим бесконечно малые перемещения на плоскости z, α (рис. 2). Переход из точки A в бесконечно близкую ей точку B можно осуществить двумя способами: 1. сначала по прямой AC (этому перемещению соответствует изменение dz, $d\alpha$) и затем по прямой CB (этому перемещению соответствует изменение δz, $\delta\alpha$); 2. сначала по прямой AD (этому перемещению соответствует изменение δz, $\delta\alpha$) и затем по прямой DB (этому перемещению соответствует изменение dz, $d\alpha$).

Билинейной форме $\delta d\omega_s$ соответствует разность между изменением Пфаффовой формы ω_s при перемещении из точки A в точку C и её изменением при перемещении из точки D в точку B. Аналогично, билинейной форме $d\delta\omega_s$ соответствует разность между изменением Пфаффовой формы ω_s при перемещении из точки A в точку D и её изменением при перемещении из точки C в точку B. Билинейные формы $\delta d\omega_s$ и $d\delta\omega_s$ являются бесконечно малыми второго порядка, разность которых равна разности изменений при перемещении вдоль путей ACB и ADB. Для интегрируемости необходимо обращение её в нуль, что соответствует выполнению условий Фробениуса. Таким образом, невыполнение условий Фробениуса можно трактовать как некоммутативность рассмотренных выше переходов. Напротив, голономность означает их коммутативность. В конкретных случаях она может быть выявлена как некоммутативность некоторых конечных изменений.

Неголономность связи состояния электроманитного поля, распространяющегося вдоль ОВС, и кручения осей его двулучепреломления можно проиллюстрировать на простом примере: когда неголономность является следствием некоммутативности матриц Джонса для отрезков ОВС с различным,

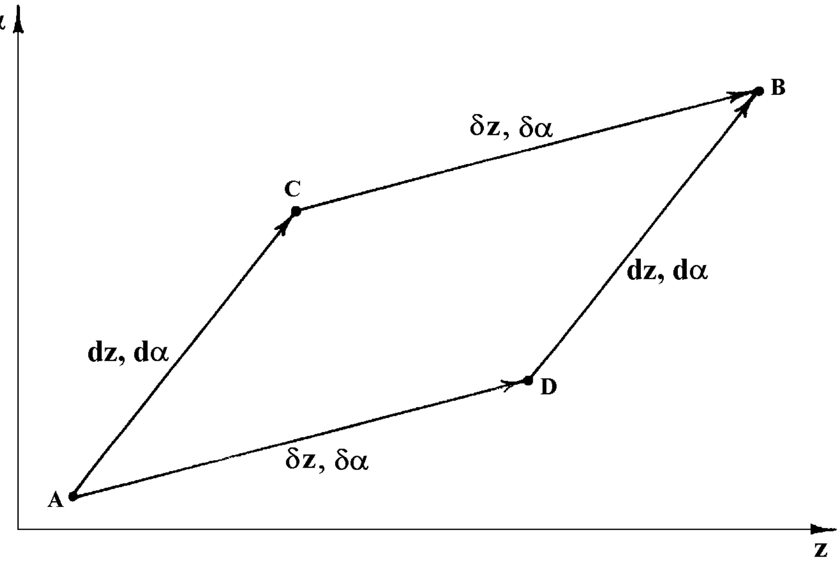


Рис. 2. Бесконечно малые перемещения на плоскости $z,\ \alpha$.

обусловленным кручением, эллиптическим двулучепреломлением, например, для двух отрезков ОВС с постоянным, но различным кручением осей. Другим примером некоммутативности может служить различие матриц Джонса для отрезков ОВС с положительной и отрицательной экспоненциальными зависимостями $\alpha(z)$ в случае, когда значения величины α для начал и концов отрезков совпадают [22]. Отметим, что матрицы Джонса, которые являются квадратными матрицами, в общем случае некоммутативны [23].

В некоторых, специально подобранных, частных случаях матрицы Джонса отрезков ОВС с различным кручением коммутируют, например, для двух отрезков ОВС одинаковой длины и с одинаковым собственным линейным двулучепреломлением, которые имеют одинаковое по величине и противоположное по знаку кручение, при условии, что оптическая толщина отрезков кратна длине волны света, т. е. разность фаз собственных волн, прошедших по медленной и быстрой осям эллиптического двулучепреломления, кратна 2π рад.

Таким образом, при произвольных соотношениях между z и α, зная состояние поляризации света на входе отрезка ОВС и азимуты осей анизотропии на входе и выходе отрезка, невозможно вычислить значения E_x, E_y, ψ_x и ψ_y, которые определяют состояние поляризации на выходе отрезка. В случае же, когда зависимость $\alpha(z)$ задана в явном виде, то, как показано в [13, 14], из уравнения (1) можно получить частное уравнение Рикатти для состояния поляризации света в ОВС $\dfrac{d\chi}{dz} = -n_{12}\chi^2 + (n_{22} - n_{11})\chi + n_{21}$ (где $\chi = \dfrac{|E_y|}{|E_x|}\mathrm{e}^{i(\psi_y-\psi_x)}$, n_{ij} — элементы дифференциальной матрицы Джонса), которое, как известно, для произвольного вида $\alpha(z)$ не решается в квадратурах.

Отметим, что ранее аналогичная задача рассматривалась для случая распространения электромагнитных волн в анизотропной плазме при наличии обусловленной эффектом Фарадея гиротропии: уравнения Бадена—Кравцова [24], описывающие это явление, так же сводятся к уравнению Рикатти [25, 26]. В работе [22] для ряда частных случаев зависимостей $\alpha(z)$ получены точные или приближённые решения. Отметим также, что в частном случае $\dfrac{d\alpha}{dz} = \mathrm{const}$ выражение для состояния поляризации может быть получено непосредственно решением дифференциального уравнения (1) [13, 14] или методом полиномов Чебышева [27]. В этом случае, в сопровождающей кручение системе координат собственные поляризационные моды равномерно скрученного ОВС — эллиптические, причём двулучепреломление составляет $\beta_э = \sqrt{\beta^2 + \left[2(1-g)\dfrac{d\alpha}{dz}\right]^2}$, а эллиптичность $\varepsilon = \mathrm{tg}\left\{\arcsin\left[2(1-g)\dfrac{d\alpha}{dz}\Big/\beta_э\right]\right\}$. Такие моды именуются винтовыми, поскольку они распространяются вдоль ОВС с вращением главных осей эллипсов [22, 28–30].

Следует отметить, что в отсутствие потерь ($E_x^2 + E_y^2 = 1$) состояние поляризации излучения определяется двумя параметрами: отношением E_y/E_x и разностью фаз $\psi_y - \psi_x$, — а состояние электромагнитного поля определяется этими двумя параметрами и величиной ψ_x. Таким образом, если для вычисления состояния поляризации излучения достаточно решить соответствующее уравнение Рикатти, то для вычисления состояния электромагнитного поля необходимо решить систему дифференциальных уравнений (3). Для вычисления результата интерференции излучения, прошедшего отрезок произвольно скрученного ОВС с линейным двулучепреломлением, с некоторым опорным излучением, также необходимо знать три параметра: E_y/E_x, ψ_x и ψ_y.

Рассмотрим теперь связь геометрооптической фазы Панчаратнама [11] с углом кручения ОВС. Как было отмечено выше, для её вычисления необходимо знать не только состояние поляризации излучения на входе и выходе скрученного отрезка ОВС, но и как оно изменяется при распространении излучения в ОВС. Но, как было нами показано, зная состояние поляризации на входе ОВС и ориентацию осей его анизотропии на входе (α_1) и выходе (α_2), невозможно не только вычислить, как меняется состояние поляризации излучения при распространении в ОВС, но и даже его состояние на выходе отрезка. Таким образом, связь между фазой Панчаратнама [11] и углом кручения ОВС носит заведомо неголономный характер. Отметим здесь, что и фаза Рытова—Владимирского, и фаза Панчаратнама являются частным случаем проявления в оптике геометрической (топологической) фазы Берри [31] (см. также обзоры [6–10]).

Кручение осей линейного двулучепреломления ОВС возникает как в процессе его вытяжки из заготовки, так и в процессе укладки в волоконную линию связи или на катушку волоконного датчика тех или иных физических параметров. Отметим, что виток световода на катушке является уже неплоской кривой и в нём будет иметь место изменение состояния поляризации света, связанное не только с кручениями осей линейного двулучепреломления, но и с эффектом Рытова [1, 2].

В реальном случае зависимость $\alpha(z)$ в волоконной линии связи или катушке установить практически невозможно, по причине отсутствия доступа к волокну по всей его длине; известными являются только значения α_1 и α_2 в начале и конце отрезка ОВС. Как было показано выше, в этом случае состояние поляризации света не может быть найдено теоретически. Однако в случае, если изменение величины кручения осей происходит достаточно медленно на длине поляризационных биений, равной в ОВС $2\pi/\beta_э$, то возможно получение решения в адиабатическом приближении [22, 27, 32].

Поскольку матрицы Джонса являются оператором (пропагатором) для связи апмлитуд и фаз компонент электрического поля на входе и выходе двулучепреломляющей среды, частным случаем которой является отрезок ОВС, то можно сделать заключение, что уравнения Максвелла для среды с линейным двулучепреломлением и неоднородным кручением также являются неголономными. Другой случай, когда уравнения Максвелла также являются неголономными, — это среда без двулучепреломления, но с неоднородным коэффициентом преломления, где имеет место эффект Рытова. Отметим, что Рытов [1, 2] и Берри [6] проводили вычисления непосредственно с помощью уравнений Максвелла.

В работе [33] высказывается утверждение, что выбором некоторой криволинейной системы координат всегда можно добиться голономности уравнений Максвелла. Это будет локально декартовая система координат, жёстко связанная с вектором электрического поля распространяющегося в световоде. Но для нахождения связи этой криволинейной, сопровождающей вектор $\vec{E}$, системы координат с лабораторной или сопровождающей кручение осей двулучепреломления ОВС системой координат, необходимо найти угловое положение вектора $\vec{E}$ в каждой точке ОВС, т. е. решить уравнение (1), которое само, как было показано выше, является неголономным. Таким образом, предложенная в [33] криволинейная система координат не даёт никакого преимущества в решении рассматриваемой задачи. (По поводу неголономных систем координат см. также [34, 35]).

В заключение авторы выражают благодарность Н. К. Вдовичевой и И. А. Шерешевскому (ИФМ РАН) за полезные консультации.

Работа частично поддержана грантами РФФИ № 96–02–18568 и № 96–15–96742.

ЛИТЕРАТУРА

1. Рытов С. М. //ДАН СССР, 1938. Т. 18. № 4–5. С. 263.
2. Рытов С. М. //Труды ФИАН, 1940. Т. 2. № 1. С. 41.
3. Владимирский В. В. //ДАН СССР, 1941. Т. 31. № 3. С. 222.
4. Ross J. N. //Opt. and Quant. Electr., 1984. V. 16. № 5. P. 455.
5. Haldane F. D. //Optics Lett., 1986. V. 11. № 11. P. 730.
6. Berry M. V. //Nature, 1987. V. 326. № 6110. P. 277.
7. Виницкий С. И., Дербов В. Л., Дубовик В. Н., Марковски Б. Л., Степановский Ю. П. //УФН, 1990. Т. 160. № 6. С. 1.
8. Малыкин Г. Б. //Изв. ВУЗов. Радиофизика, 1997. Т. 40. № 3. С. 265.
9. Berry M. //Phys. Today, 1990. V. 43. № 12. P. 34.
10. Клышко Д. Н. //УФН, 1993. Т. 163. № 11. С. 1.
11. Pancharatnam S. //Proc. Ind. Acad. Sci., 1956. V. A44. № 5. P. 247.
12. Monerie M., Jeunhomme L. //Opt. and Quant. Electr., 1980. V. 12. № 6. P. 449.
13. Azzam R. M., Bashara N. M. //JOSA, 1972. V. 62. № 11. P. 1252.
14. Аззам Р., Башара Н. Эллипсометрия и поляризованный свет. — М.: Мир, 1981. – 583 с.
15. Неймарк Ю. И., Фуфаев Н. А. Динамика неголономных систем. — М.: Наука 1967, – 519 с.
16. Гапонов А. В. //ДАН СССР, 1952. Т. 87. № 3. С. 401.
17. Ишлинский А. Ю. Механика гироскопических систем. — М.: Изд-во АН СССР, 1963. – 480 с.
18. Рашевский П. К. Геометрическая теория уравнений с частными производными. — М.-Л.: ГИТТЛ, 1947. – 354 с.
19. Rashleigh S. C., Ulrich R. //Optics Lett., 1978. V. 3. № 2. P. 60.
20. Ulrich R., Simon A. //Appl. Optics, 1979. V. 18. № 13. P. 2241.
21. Barlow A. J., Ramskov–Hansen J. J., Payne D. N. //Appl. Optic, 1981. V. 20. № 17. P. 2962.
22. Железняков В. В., Кочаровский В. В., Кочаровский Вл. В. //УФН, 1983. Т. 141. № 2. С. 257.
23. Гантмахер Ф. Р. Теория матриц. — М.: Наука, 1966. – 576 с.
24. Кравцов Ю. А. //ДАН СССР, 1968. Т. 183. № 1. С. 74.
25. Найда О. Н. //Изв. ВУЗов. Радиофизика, 1970. Т. 13. № 10. С. 1496.
26. Кравцов Ю. А., Орлов Ю. И. Теоретическая оптика неоднородных сред. — М.: Наука, 1980. – 304 с.
27. Ярив А., Юх П. Оптические волны в кристаллах. — М.: Мир, 1987. – 616 с.
28. Drucker D. C., Mindlin R. D. //J. of Appl. Phys., 1940. V. 11. № 11. P. 724.
29. Гинзбург В. Л. //ЖТФ, 1944. Т. 14. № 3. С. 181.
30. Суворов Е. В. //Изв. ВУЗов. Радиофизика, 1972. Т. 14. № 9. С. 1320.
31. Berry M. V. //Proc. Royal Soc. London, 1984. V. A392. № 1802. P. 45.
32. Даршт М. Я., Зельдович Б. Я., Кундикова Н. Д. //Оптика и спектроскопия, 1997. Т. 82. № 4. С. 660.

33. Post E. J. Formal structure of electromagnetics. — Amsterdam: North Holl. Publ., 1962. – 204 p.
34. Schouten J. A. Ricci-calculus an introduction in tensor analysis. — Berlin: Springer-Verlag, 1954. – 516 p.
35. Схоутен А. Я. Тензорный анализ для физиков. — М.: Наука, 1965. – 455 с.

Институт прикладной физики РАН,
Научно–исследовательский
институт прикладной математики и
кибернетики,
г. Н. Новгород, Россия

Поступила в редакцию
5 декабря 1997 г.

NONHOLONOMIC LINK OF UNIMODE LINEAR BIREFRINGENCE LIGHTGUIDE ELECTROMAGNETIC FIELD AND ITS TORSION ANGLE

G. B. Malykin, Yu. I. Neimark

A link between the electromagnetic field state in a stranded unimode fiber lightguide (UFL) with linear birefringence and the torsion angle has been considered for the case when the lightguide is stretched into a straight line. This link has been shown to be nonholonomic that comes from the noncommutativity of two infinitesimal but different changes of the lightguide length and its torsion. In particular, it follows that the interrelation between the radiation polarization in the stranded lightguide and its torsion angle is also nonholonomic. It has been also shown, that the relation between geometric Pancharatnam's phase and UFL torsion angle is nonholonomic as well.

УДК 621.385.69.001.573:[530.132+530.182]

К ВОПРОСУ О МЕХАНИЗМЕ ВОЗНИКНОВЕНИЯ ХАОТИЧЕСКОЙ ДИНАМИКИ В ВАКУУМНОМ СВЧ ГЕНЕРАТОРЕ НА ВИРТУАЛЬНОМ КАТОДЕ

В. Г. Анфиногентов, А. Е. Храмов

Представлены результаты численного исследования нелинейной динамики электронного потока с виртуальным катодом в виркаторной системе на пролётном токе. Показано существование сложных режимов, включая хаотические. Выявлена динамическая природа хаоса в рассматриваемой системе. Исследованы физические процессы в потоке и обнаружено, что возникновение хаотической динамики электронного потока с виртуальным катодом связано с нелинейным взаимодействием образующихся в системе структур.

1. Хорошо известно, что СВЧ генераторы, использующие в качестве активной среды электронный поток с виртуальным катодом (ВК), или виркаторы [1], обладают сложной нерегулярной динамикой (см., например, [2–4]).

Исследование простейшей модели диода Пирса в режиме с ВК [5], описывающей механизм генерации сверхмощного излучения в виркаторах, показало, что формирование ВК в потоке, полностью нейтрализованном ионным фоном, сопровождается возникновением двух электронных структур, взаимодействие между которыми приводит к хаотической динамике в потоке [6, 7].

Однако, традиционные виркаторы работают обычно без нейтрализации электронного потока, как это имеет место в диоде Пирса. Поэтому особый интерес вызывает выяснение процессов, сопровождающих нерегулярное поведение электронного потока с ВК в вакуумной дрейфовой камере. Понимание физических процессов в потоке с ВК позволит выработать достаточно эффективные методы воздействия на виркаторную систему как для повышения эффективности энергообмена между потоком и электромагнитным полем, так и для управления режимами колебаний ВК.

В представленной работе мы исследуем систему, представляющую собой замкнутый отрезок цилиндрического волновода, в который инжектируется моноскоростной сильноточный электронный релятивистский пучок. Вдоль продольной оси волновода приложено достаточно сильное магнитное поле, так что электроны пучка полностью замагничены.

При превышении током пучка предельного вакуумного тока для данной геометрии в потоке формируется ВК [8], т. е. область с потенциалом, примерно равным ускоряющему потенциалу, от которой отражается часть потока обратно к плоскости инжекции.

Исследование нелинейных нестационарных процессов в исследуемой системе осуществлялось с помощью модели, основанной на решении самосогласованной системы кинетического уравнения Власова и уравнений Максвелла. Построение соответствующей разностной схемы и её конкретная численная реализация основывается на работах [9, 10].

2. При заданной геометрии основным управляющим параметром, от которого зависит поведение исследуемой системы, является отношение тока пучка к предельному вакуумному току, обозначаемое α.

Анализ и диагностика динамических режимов в исследуемой системе проводились по временным реализациям тока пучка $I(t)$ из области ВК, по которым строились спектры мощности и восстанавливались фазовые портреты по методу Такенса [11] (рис. 1). Было выделено три качественно отличающихся режима колебаний ВК с увеличением тока пучка.

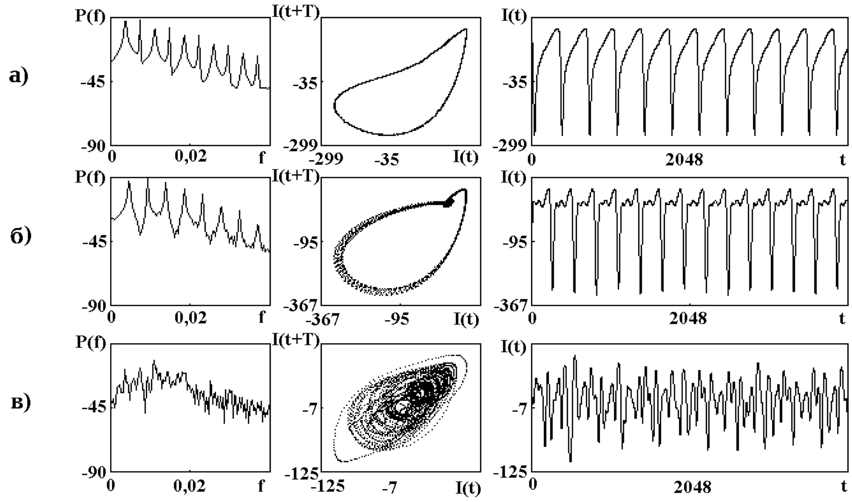


Рис. 1.

При малых значениях α ($\alpha < 1{,}7$) в системе устанавливаются регулярные колебания релаксационного типа (рис. 1а). Спектр мощности содержит узкие пики, которые являются гармониками основной частоты $\omega_0 \approx 2{,}6\omega_\mathrm{p}$, где ω_p — плазменная частота электронного потока. Отметим, что значение частоты генерации в численном эксперименте близко к наблюдаемой в физических экспериментах (см., например, [1] (Гл. 14—15), а также библиографию к ним).

В диапазоне $1{,}7 < \alpha < 3{,}0$ в системе устанавливаются слаборегулярные колебания (рис. 1б); хаотический аттрактор появляется на базе одного неустойчивого предельного цикла, соответствующего притягивающему множеству периодических движений.

При $\alpha > 3{,}0$ (рис. 1в) колебания усложняются; фазовый портрет более однороден; структура аттрактора сложна и состоит из множества неустойчивых орбит; спектр мощности содержит высокий шумовой пьедестал, на фоне которого выделяется несколько спектральных составляющих.

Количественный анализ сложности динамики системы, проведённый с помощью вычисления размерности аттракторов по алгоритму Гроссбергера—Прокаччио [12], свидетельствует о детерминированной природе сложных колебаний ВК. Во всех режимах размерность испытывает насыщение с ростом размерности пространства вложения, что характерно для динамического хаоса [13]. Число активных степеней свободы, вовлекаемых в колебательное движение в системе, определённое по величине размерности пространства вложения, при которой размерность аттрактора испытывает насыщение, растёт с увеличением тока пучка и насыщается при $\alpha > 5{,}0$. Тем не менее, во всех режимах их число остаётся небольшим (порядка $3 \div 4$), что позволяет описать динамику системы малым числом собственных мод.

Это подтверждается выделением когерентных структур в потоке с помощью алгоритма Карунена—Лоэва [14, 15]. В таблице представлены энергии первых десяти мод Карунена—Лоэва, отнесённые к средней энергии потока. Видно, что число мод, энергия которых превышает 1 % от полной энергии потока, невелико (порядка $6 \div 7$), причём энергии мод с номерами выше $6 \div 7$ практически не меняются с ростом α. С другой стороны, с ростом надкритичности наблюдается уменьшение энергии в высшей моде, с одновременным увеличением энергии $3 \div 6$ мод, что свидетельствует об увеличении числа влияющих на поведение потока внутренних структур.

Таблица

Распределение энергии по модам Карунена–Лоэва
для различных режимов колебаний ВК

РЕЖИМ	ЭНЕРГИЯ МОД (в %)									
	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}
Регулярные колебания ($\alpha = 1{,}4$)	64,08	23,67	7,33	2,73	1,54	0,44	0,15	0,05	0,01	0,00
Слаборазвитый хаос($\alpha = 2{,}0$)	53,00	24,41	13,27	5,73	2,43	0,74	0,29	0,08	0,02	0,00
Развитый хаос ($\alpha = 4{,}0$)	38,73	21,34	19,50	12,27	5,00	1,84	0,67	0,15	0,04	0,00

3. В работах [6, 16, 17] типичные моды, полученные с помощью КЛ разложения, связываются с характерными структурами (сгустками заряженных частиц) нейтрализованного потока в пространстве взаимодействия. Такой подход является весьма плодотворным и применительно к электронному потоку с ВК в вакуумной трубе дрейфа. Однако, если в потоке с полной нейтрализацией формирование второго сгустка связано либо с развитием кинематической неустойчивочти электронного потока в меняющемся тормозящем поле “открывающегося” ВК [6], либо с группировкой предварительно промодулированного потока сигналом обратной связи [17], то в нашем случае можно говорить о формировании в пространстве дрейфа нескольких отражающих плоскостей (нескольких ВК), взаимодействием между которыми может быть объяснена хаотическая динамика в электронном потоке без нейтрализации его пространственного заряда.

Эту точку зрения хорошо иллюстрирует функция распределения $\Phi(\tau)$ заряженных частиц по временам жизни в пространстве взаимодействия (рис. 2). В регулярном режиме (рис. 2а, сплошная линия) $\Phi(\tau)$ имеет двугорбый вид. Площадь под максимумами пропорциональна числу пролётных и отражённых к плоскости инжекции частиц в потоке. Характерные траектории заряженных частиц в координатах (z, t) и (v, z) с временами жизни, соответствующими максимумам $\Phi(\tau)$, приведены на рис. 3а. Видно, что в потоке имеется одна структура — ВК — и, соответственно, два характерных типа частиц — пролетные частицы и отражённые от ВК частицы.

С увеличением α отражённые частицы начинают доминировать в общем числе инжектируемых частиц (рис. 2а, пунктирная линия), спектр времен жизни отражённых частиц уширяется. За счёт появления долгоживущих частиц в потоке возникает внутренняя распределённая обратная связь, обеспечивающая взаимосвязь между основной структурой (ВК) и возникающей вторичной структурой, которой соответствует третий максимум на кривой $\Phi(\tau)$ (отмечен стрелкой на рис. 2а). Однако, при небольшой надкритичности эффективность этой связи мала (общая масса частиц, отражённых от вторичного ВК, мала) и движение слабонерегулярно (в фазовом пространстве наблюдается размытый предельный цикл).

Для развитого хаоса (рис. 2б) характерна сильно изрезанная форма $\Phi(\tau)$, которая позволяет выделить несколько примерно равных по массе групп заряженных частиц с различным временем жизни. В этом случае в потоке нет хорошо сформированного ВК (ср. пространственно–временные диаграммы электронного потока в пространстве дрейфа для регулярных (рис. 4а) и хаотических (рис. 4б) движений. На основании рис. 3б, на котором изображены траектории частиц, времена жизни которых соответствуют максимумам $\Phi(\tau)$ (отмечены цифрами на рис. 2б и 3б), процессы в потоке могут быть интерпретированы как формирование нескольких ВК (нескольких колебательных структур) на различном расстоянии от плоскости инжекции. Это также подтверждается приведёнными на рис. 5 кривыми распределения плоскостей, в которых наблюдается отражение заряженных частиц, по продольной

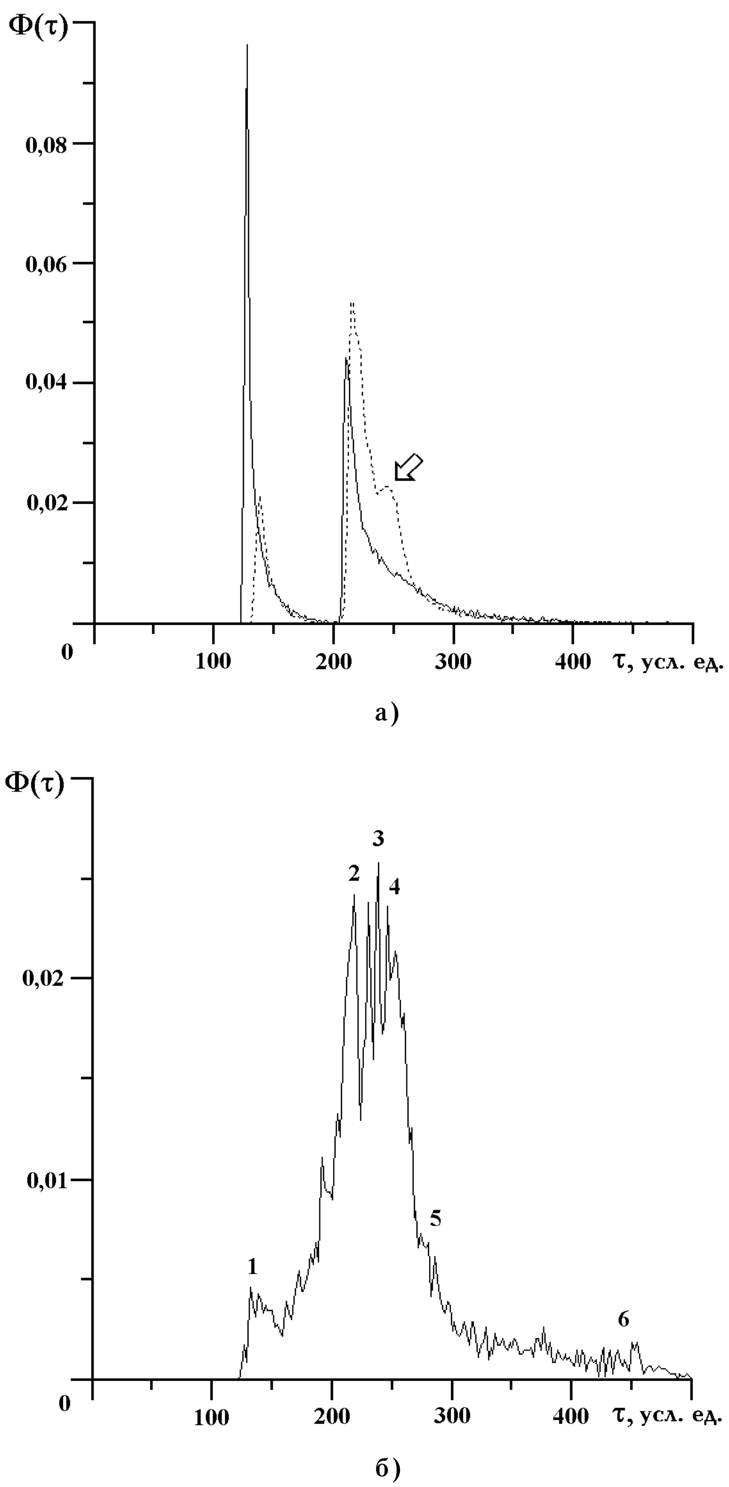

Φ(τ)
0,08
0,06
0,04
0,02
0
100
200
300
400
τ, усл. ед.
а)
Φ(τ)
3
2
4
0,02
0,01
1
5
6
0
100
200
300
400
τ, усл. ед.
б)

Рис. 2.

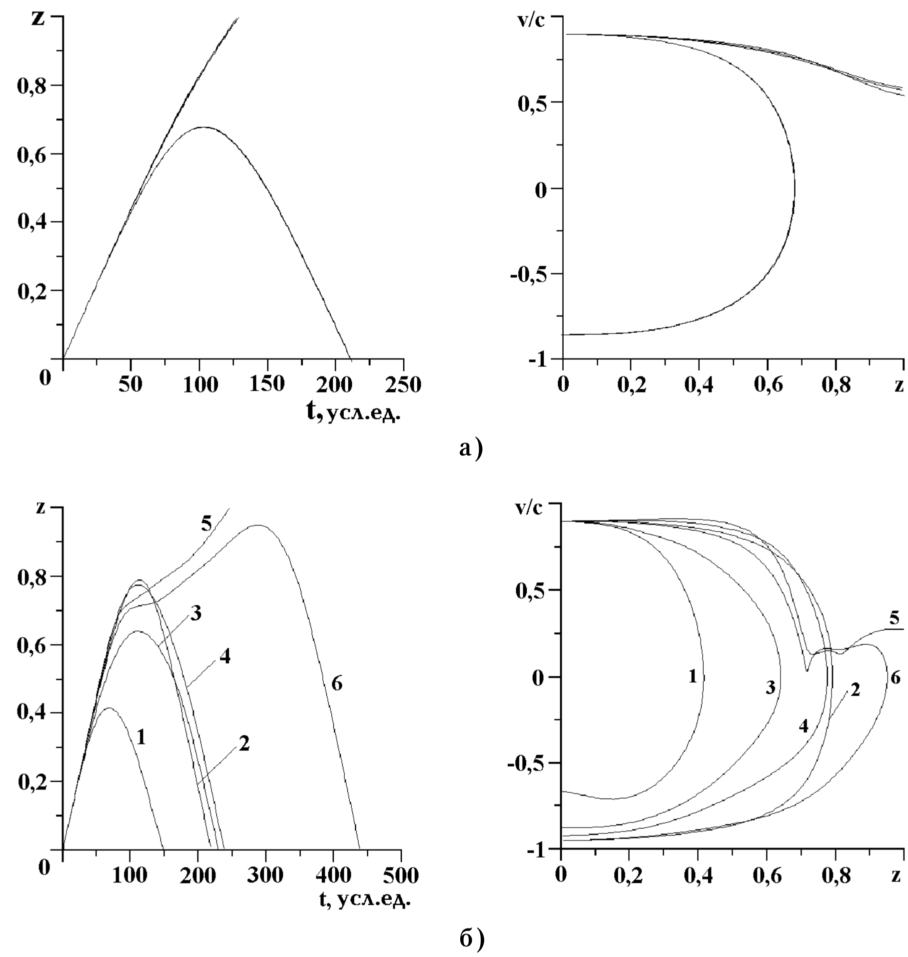

z
0,8
0,6
0,4
0,2
0
50
100
150
200
250
t, усл.ед.
v/c
0,5
0
-0,5
-1
0
0,2
0,4
0,6
0,8
z
а)
z
5
3
4
6
1
2
0,8
0,6
0,4
0,2
0
100
200
300
400
500
t, усл.ед.
v/c
0,5
0
-0,5
-1
1
3
5
2
6
4
0
0,2
0,4
0,6
0,8
z
б)

Рис. 3.

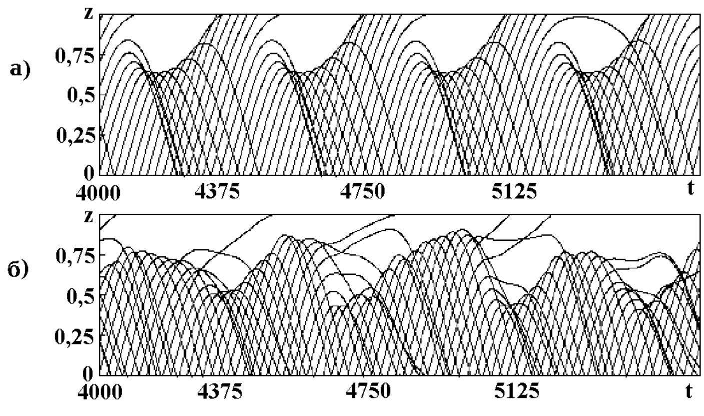

z
а)
0,75
0,5
0,25
0
4000
4375
4750
5125
t
z
б)
0,75
0,5
0,25
0
4000
4375
4750
5125
t

Рис. 4.

координате. Видно, что в случае регулярных движений функция распределения имеет ярко выраженный максимум, соответствующий тому, что в потоке имеет место единственная отражающая плоскость (ВК), которая чётко локализована в пространстве. Рост тока приводит к уширению функции распределения, причём в режиме слаборазвитого хаоса на кривой выделяются два глобальных максимума, соответствующих двум структурам в потоке, а в режиме развитого хаоса форма распределения сильно изрезана, при этом каждый максимум соответствует своему ВК. Отметим, что максимумы на кривых покоятся на высоком основании, связанном с тем, что каждый ВК колеблется как во времени, так и в пространстве, однако наиболее вероятные места локализации отражающих плоскостей соответствуют именно максимумам распределения.

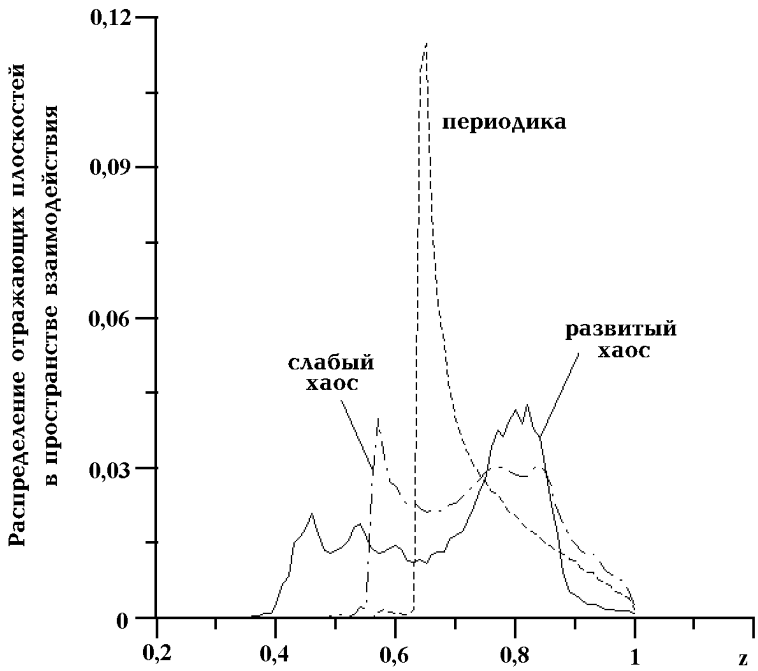


Рис. 5.

Отражение части потока от каждого из ВК оказывает влияние на условия формирования других структур в потоке, что обеспечивает несколько петель внутренней обратной связи с различными временами запаздывания. Такая распределённая связь между всеми структурами в потоке приводит к сильно нерегулярной динамике системы при большой надкритичности α.

Отметим, что в системе появляются и метастабильные частицы (траектория № 6 на рис. 3б) [2], однако их влияние на динамику потока не велико. Учёт расслоения потока по продольной координате (поток представляется набором колец различного диаметра) показал, что развитый метастабильный сгусток, живущий в пространстве взаимодействия примерно $1.5 \div 2$ периода колебаний ВК, формируется в слое, максимально удалённом от оси волновода, и его воздействие на динамику системы подобно описанному в [7].

В заключение заметим, что описанное выше явление взаимодействия и конкуренции возбуждаемых в потоке мод является одной из причин понижения эффективности генерации при токах пучка, значительно превышающих стартовый, как видно из рис. 6, на котором представлена зависимость электронного к.п.д. в системе от безразмерного тока пучка α.

4. Таким образом, электронный поток с ВК без нейтрализации ионным фоном в ограниченной трубе

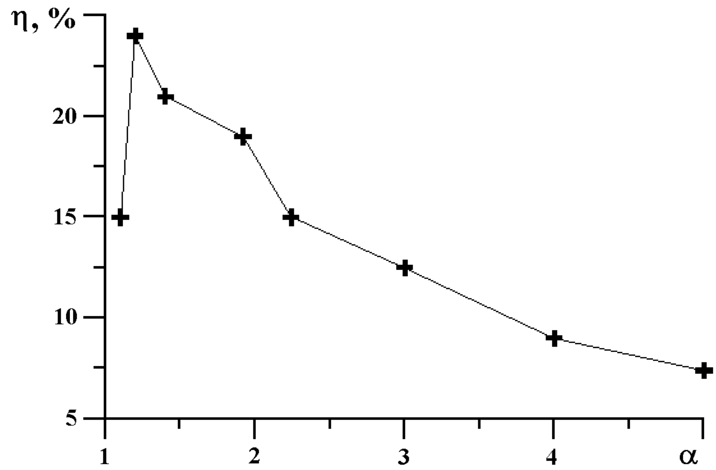


Рис. 6.

дрейфа демонстрирует различные типы нелинейных колебаний. При малой надкритичности ($\alpha < 3$) в системе устанавливается малоразмерный хаос. С увеличением α система демонстрирует развитый хаос. Хаотическая динамика объясняется формированием и взаимодействием нескольких структур, которые связываются с расслоением потока на несколько групп заряженных частиц с различными временами жизни в пространстве дрейфа и характерными траекториями в фазовом пространстве. Такая ситуация интерпретируется как формирование нескольких ВК, сильно связанных друг с другом через поток.

Понимание физических процессов в исследуемой системе с ВК позволяет выработать эффективные методы управления системой за счёт воздействия на функцию распределения электронов по временам жизни. К таким механизмам, осуществляющим дополнительную селекцию электронов, видимо может быть отнесена либо предварительная модуляция электронного потока, поступающего в пространство взаимодействия [18, 19], либо введение в систему с ВК внешней или внутренней обратной связи [17, 20–22]. Такой подход позволяет подавить образование вторичных структур в потоке и, в итоге, повысить качество генерируемого спектра и эффективность преобразования энергии потока в энергию электромагнитного поля в приборе с ВК.

Работа частично поддержана Российским фондом фундаментальных исследований (грант № 96–02–16753).

ЛИТЕРАТУРА

1. High Power Microwave Sources /Eds. by Granatstein V.L. and Alexeff I. — Boston: Artech House, 1987. Ch. 14–15.
2. Селемир В. Д., Алехин Б. В., Ватрунин В. Е. и др. //Физика плазмы, 1994. Т. 20. № 7, 8. С. 689.
3. Афонин А. М., Диденко А. М., Пауткин А. Ф., Рошаль А. С. //Радиотехника и электроника, 1992. Т. 37. № 10. С. 1889.
4. Привезенцев А. П., Саблин Н. И., Фоменко Г. П. //Радиотехника и электроника, 1990. Т. 35. № 64. С. 832.
5. Pierce J. //J. Appl. Phys., 1944. V. 15. P. 721.
6. Anfinogentov V. G. — In: Proc. of the III Intern. Spec. Workshop on Nonlinear Dynamics of Electronic Systems. University College Dublin. 28–29 July 1995. P. 79.

7. Анфиногентов В. Г. //Изв. вузов. Радиофизика, 1995. Т. 38. № 3-4. С. 268.
8. Богданкевич Л. С., Кузелев М. Б., Рухадзе А. А. //УФН, 1983. Т. 133. № 1, 2. С. 3.
9. Свешников А. Г., Якунин С. А. //Математическое моделирование, 1989. Т. 1. № 4. С. 1.
10. Храмов А. Е. — В сб.: Труды 7 Межвуз. конф. "Математическое моделирование и краевые задачи", Самара. 28—30 мая 1997. Т. 2. С. 94.
11. Takens F. //Lect. Notes in Math. — Warwick: Springler-Verlag. 1980. V. 898. P. 366.
12. Grassberger P., Procaccia J. //Physica D., 1983. V. 19. № 1, 2. P. 189.
13. Берже П., Помо И., Видаль К. Порядок в хаосе. — М.: Мир, 1991.
14. Ватанабе С. — В сб.: Автоматический анализ сложных изображений /Под ред. Э. М. Бравермана. — М.: Мир, 1969. С. 254.
15. Lumley J. L. — In: Atmospheric Turbulence and Radio Wave Propagation: Proc. of the Intern. Coll. /Eds. by Yaglom A. M. and Tutarsky V. I. — Moscow: Nauka, 1967. P. 166.
16. Анфиногентов В. Г. //Письма в ЖТФ, 1995. Т. 21. Вып. 8. Р. 70.
17. Anfinogentov V. G. — In: Proc. of 11th Intern. Conf. on High Power Particle Beams, BEAMS'96, Prague, 10—14 June 1996. P. 1-29.
18. Hramov A. E. — In: Proc. of the 5th Intern. Spec. Workshop on Nonlinear Dynamics of Electronic Systems, NDES'97, Moscow, Russia, 26—27 June 1997. P. 443.
19. Jiang W., Masugata K., Yatsui K. //Physics of Plasmas, 1995. V. 2. № 12. P. 4635.
20. Гадецкий Н. Н., Магда И. И., Найстетер С. И. и др. //Физика плазмы, 1993. Т. 19. Вып. 4. С. 530.
21. Анфиногентов В. Г., Храмов А. Е. — В сб.: Труды VI Всеросс. школы-семинара "Физика и применение микроволн", май 1997 г. Красновидово Моск. обл. С. 7.
22. Анфиногентов В. Г., Храмов А. Е. — В сб.: Труды Всеросс. межвуз. конф. "Современные проблемы электроники и радиофизики СВЧ", 4—8 сентября 1997. Саратов, С. 6.

Государственный учебно—научный
центр "Колледж"при Саратовском
госуниверситете, Россия

Поступила в редакцию
27 октября 1997 г.

ON MECHANISM OF CHAOTIC DYNAMICS ORIGIN IN MICROWAVE VACUUM GENERATOR WITH VIRTUAL CATHODE

V. G. Anfinogentov, A. E. Khramov

The nonlinear dynamics of an electron beam with a virtual cathode in the vircator on transit current is studied numerically. The existence of complex regimes including the chaotic ones the is shown. Physical processes in the beam are investigated and it is shown that chaotic dynamics in the electron beam with the virtual cathode is associated with the nonlinear interaction between structures in the beam.

УДК 523.164+519.216.2

АВТОКОРРЕЛЯЦИОННАЯ ФУНКЦИЯ СУММЫ НОРМАЛЬНОГО ШУМА И ДВУХ ГАРМОНИЧЕСКИХ КОЛЕБАНИЙ ПРИ ОГРАНИЧЕНИИ С ПРЕОБРАЗОВАНИЕМ ВАН ФЛЕКА

Л. Г. Содин

В статье выполнен строгий анализ статистических характеристик суммы гауссового процесса и двух гармонических колебаний, прошедших безынерционный ограничитель — типичный элемент радиоастрономических спектрометров. Приведены расчёты для двух наиболее интересных случаев: преобладания интенсивности компоненты с непрерывным спектром и, наоборот, преобладания интенсивности компонент с дискретным спектром. Получены расчётные формулы для автокорреляционной функции, позволяющие оценить спектр полностью ограниченного процесса, а также его спектр после преобразования Ван Флека. Показано, что в любом случае такое преобразование повышает точность определения автокорреляционной функции и спектра, снижает уровень нелинейных продуктов, в том числе, компонент с комбинационными частотами.

ВВЕДЕНИЕ

В работе исследована задача о воздействии на идеальный ограничитель (нелинейное устройство с характеристикой $y = \mathrm{sign}[e]$) суммы нормального случайного процесса и двух квазигармонических колебаний. Интерес к этой задаче связан с двумя обстоятельствами. Во-первых, при наблюдении излучений космических объектов в спектральных линиях атомов и молекул практически во всех обсерваториях применяются спектрометры со знаковыми коррелометрами (с “полным ограничением") (см. [1], с.244). Во-вторых, аналогичные устройства используются при приёме и обработке радиолокационных, связных и других сигналов.

Рассматриваемая задача решалась Джонсом [2] и кратко изложена в [3]. Однако упомянутые решения недостаточны. Это связано с тем, что в реальных приборах для компенсации нелинейности, обязанной ограничителю, используется преобразование, обратное преобразованию Ван Флека (ОПВФ) [4], требующее неискажённой формы автокорреляционной функции (АКФ) процесса на выходе ограничителя. А в работе [2] после ограничителя введена фильтрация, изменяющая форму АКФ, что делает применение ОПВФ бессмысленным. Поясним это. Если на ограничитель воздействует нормальный процесс с АКФ $\rho(\tau)$, то на выходе его “знаковая” АКФ $R(\tau)$ связана с $\rho(\tau)$ нелинейным преобразованием Ван Флека:

$$R(\tau) = (2/\pi)\arcsin\rho(\tau) = (2/\pi)(\rho + \rho^3/6 + 3\rho^5/40 + \ldots).$$

Обратное преобразование Ван Флека $\widehat{\rho}(\tau) = \sin\left[\frac{\pi}{2}R(\tau)\right]$ обеспечивает точное восстановление АКФ процесса. А полосовая фильтрация резко снижает нелинейные члены в зависимости $R(\rho)$ и устраняет возможность точного восстановления АКФ $\rho(\tau)$ по знаковой АКФ $R(\tau)$.

Введение в исходное колебание двух гармонических сигналов позволяет учесть влияние помех радиоастрономическим наблюдениям, всегда присутствующих в эфире из-за множества радиостанций различного назначения. При этом возникают колебания с комбинационными частотами, а оценку их мешающего действия можно сделать, если в рассматриваемый процесс ввести как минимум два монохроматических сигнала.

В соответствии с указанными двумя применениями исследуемого устройства мы будем вести расчёты для двух случаев:

1) радиоастрономия, интенсивность гауссовой компоненты выше интенсивности помех;

2) радиотехника, интенсивность гармонических составляющих превышает интенсивность шума.

Входной сигнал представим в следующем виде:

$$e(t) = n(t) + A\cos(\omega_1 t + \phi_1) + Ax\cos(\omega_2 t + \phi_2),$$

где $n(t)$ — нормальный процесс с $\langle n(t)n(t+\tau)\rangle = \sigma^2\rho(\tau)$, x — отношение амплитуд гармоник (далее без потери общности полагаем $x \le 1$). Все три компоненты $e(t)$ полагаем статистически независимыми.

1. ПОСТАНОВКА ЗАДАЧИ. ИСХОДНЫЕ СООТНОШЕНИЯ

Характеристику ограничителя представим интегралом Дирихле, связывающим выход $y(t)$ со входом $e(t)$:

$$y(t) = \mathrm{sign}[e(t)] = \frac{2}{\pi}\int\limits_0^\infty \frac{\sin[ue(t)]}{u}\,\mathrm{d}u\,. \tag{1}$$

При этом АКФ процесса $y(t)$ равна

$$R(\tau) = \langle y(t)y(t+\tau)\rangle = \frac{2}{\pi^2}\iint\limits_0^\infty \mathrm{Re}\Big[\theta(u,v) - \theta(u,-v)\Big]\frac{\mathrm{d}u\mathrm{d}v}{uv}\,, \tag{2}$$

где $\theta(u,v)$ — характеристическая функция входного процесса.

Для косинусоидальных компонент $e(t)$ возможны два представления:

1) гармонические колебания со случайными взаимнонезависимыми фазами, равномерно распределёнными на интервале $-\pi$, π;

2) неслучайные гармонические колебания.

Во втором случае R зависит от t и τ, но если усреднение $R(t,\tau)$ по текущему времени проводится на большом сравнительно с $2\pi/\omega$ интервале, то, как нетрудно показать, результат идентичен первому случаю. В первом случае характеристическая функция равна [2]

$$\begin{aligned}\theta_1(u,v) &= \exp\left[-\sigma^2\frac{u^2+v^2+2\rho uv}{2}\right] J_0(A\sqrt{u^2+v^2+2\rho_1 uv}\,)\times\\ &\times J_0(Ax\sqrt{u^2+v^2+2\rho_2 uv}\,),\end{aligned} \tag{3}$$

где $\rho_1 = \cos\gamma_1$, $\rho_2 = \cos\gamma_2$, $\gamma_1 = \omega_1\tau$, $\gamma_2 = \omega_2\tau$.

Представим $R(\tau)$ из выражения (2), с учётом (3), в следующем виде:

$$\begin{aligned}R(\tau) &= \frac{2}{\pi^2}\iint\limits_0^\infty \mathrm{e}^{-\sigma^2(u^2+v^2)/2}\Big[\mathrm{e}^{\sigma^2\rho uv}J_0(A\sqrt{u^2+v^2-2\rho_1 uv}\,)\times\\ &\times J_0(Ax\sqrt{u^2+v^2-2\rho_2 uv}\,) - \mathrm{e}^{-\sigma^2\rho uv}J_0(A\sqrt{u^2+v^2+2\rho_1 uv}\,)\times\\ &\times J_0(Ax\sqrt{u^2+v^2+2\rho_2 uv})\Big]\frac{\mathrm{d}u\mathrm{d}v}{uv}\,.\end{aligned} \tag{4}$$

Используем разложения

$$J_0(A\sqrt{u^2+v^2\mp 2\rho uv}) = \sum_{k=0}^{\infty}\epsilon_k(\pm 1)^k J_k(u)J_k(v)\cos k\psi\,,$$

$$\cos\psi=\rho\,,\qquad \epsilon_0=1\,,\qquad \epsilon_k=2\ \ (k\geq 1),$$

и, подставив их в (4), представим $R(\tau)$ в следующем виде:

$$R(\tau)=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\epsilon_k\epsilon_l H_{kl}\cos k\gamma_1\cos l\gamma_2\,, \tag{5}$$

где

$$\begin{aligned} H_{kl} &= \frac{4}{\pi^2}\iint\limits_0^{\infty} \mathrm{e}^{-\sigma^2(u^2+v^2)/2} J_k(Au)J_k(Av)J_l(Axu)\times \\ &\times\ J_l(Axv)\begin{bmatrix}\mathrm{sh}\,\sigma^2\rho uv\\ \mathrm{ch}\,\sigma^2\rho uv\end{bmatrix}\frac{rmdu\mathrm{d}v}{uv} \end{aligned} \tag{6}$$

(sh — при $k+l$ чётном, ch — при $k+l$ нечётном).
При $A=0$ из (5) получаем

$$R(\tau)=\frac{4}{\pi^2}\iint\limits_0^{\infty}\mathrm{e}^{-\sigma^2(u^2+v^2)/2}\,\frac{\mathrm{sh}\,\sigma^2\rho uv}{uv}\,\mathrm{d}u\mathrm{d}v=\frac{2}{\pi}\arcsin\rho(\tau).$$

Очевидно, $\rho(\tau)=\sin\left[\frac{\pi}{2}R(\tau)\right]$, что и определяет использование ОПВФ.

2. РАДИОАСТРОНОМИЧЕСКИЙ СЛУЧАЙ, $\sigma > A$

Поскольку A мало, разложим произведение функций Бесселя в ряд Тейлора:

$$\begin{aligned} J_k(Au)J_k(Av)J_l(Axu)J_l(Axv) &= \frac{x^{2l}(A^2uv)^{k+l}}{(l!)^2 4^{k+l}}\sum_{s=0}^{\infty}\sum_{t=0}^{\infty}\left(-\frac{A^2u^2v^2}{4}\right)^{s+t}\times \\ &\times\frac{F(-s,-s-k;l+1;x^2)}{s!(s+k)!}\frac{F(-t,-t-k;l+1;x^2)}{t!(t+k)!}\,, \end{aligned}$$

где $F(a,b;c;z)$ — гипергеометрическая функция Гаусса. Подставив это разложение в интеграл (6), получим

$$\begin{aligned} H_{kl} &= \frac{x^{2l}\beta^{k+l}}{(l!)^2 2^{k+l}}\sum_{s,t=0}^{\infty}\left(-\frac{\beta}{2}\right)^{s+t}\times \\ &\times\ \frac{F(-s,-s-k;l+1;x^2)F(-t,-t-k;l+1;x^2)}{s!\,(s+k)!\,t!\,(t+k)!}I_{st}^{kl}, \end{aligned} \tag{7}$$

где $\beta = A^2/(2\sigma^2)$,

$$I_{st}^{kl} = \frac{4}{\pi^2}\iint\limits_0^\infty \mathrm{e}^{-(u^2+v^2)/2}\left\{\begin{matrix}\mathrm{sh}\,\rho uv\\ \mathrm{ch}\,\rho uv\end{matrix}\right\}u^{k+l+2s-1}v^{k+l+2t-1}\mathrm{d}u\mathrm{d}v =$$

$$= \frac{2}{\pi}(-1)^{s+t}\frac{\partial^{k+l+s+t}}{\partial\rho^{k+l}\partial a^s\partial b^t}\arcsin(\rho/\sqrt{4ab}), \quad a = b = 1/2. \tag{8}$$

Выражения (7) и (8) позволяют рассчитать $R(\tau)$ в зависимости от исходных параметров $x,\ \rho,\ \rho_1,\ \rho_2$ и интенсивности гармонических помех.

Формулы для H_{kl} с точностью до β^3 приведены в Приложении 1. Здесь мы выпишем $R(\tau)$ с точностью до β^2: $R(\tau) = \dfrac{2}{\pi}(\arcsin\rho + \Delta)$, где

$$\Delta = -\beta\frac{\rho-\rho_1+x^2(\rho-\rho_2)}{(1-\rho^2)^{1/2}} + \frac{\beta^2}{4(1-\rho^2)^{3/2}}\Big[(\rho-\rho_1)(2-\rho^2+\rho\rho_1) +$$

$$+ x^4(\rho-\rho_2)(2-\rho^2+\rho\rho_2) + 4x^2(2\rho-\rho_1-\rho_2-\rho^3+\rho\rho_1\rho_2)\Big]. \tag{9}$$

Применим теперь к выражению (9) ОПВФ, представив

$$\widehat{\rho}(\tau) = \sin[\pi R(\tau)/2] = \rho\cos\Delta + \sqrt{1-\rho^2}\sin\Delta. \tag{10}$$

Для $\cos\Delta$ и $\sin\Delta$ учтём члены порядка β^2:

$$\cos\Delta \approx 1 - \Delta^2/2 = 1 - \frac{\beta^2}{2(1-\rho^2)}\Big[\rho-\rho_1+x^2(\rho-\rho_2\Big]^2 + O(\beta^3),$$

$$\sin\Delta \approx \Delta = -\frac{\beta}{\sqrt{1-\rho^2}}\Big[\rho-\rho_1+x^2(\rho-\rho_2)\Big] + \frac{\beta^2}{(1-\rho^2)^{3/2}}Q + O(\beta^3),$$

где Q — выражение в квадратных скобках в формуле (9). В результате после ОПВФ получим

$$\widehat{\rho} = \rho[1-\beta(1+x^2)] + \beta\rho_1 + x^2\beta\rho_2 + \frac{\beta^2}{4(1-\rho^2)} \times$$

$$\times \Big\{(\rho-\rho_1)(2-3\rho^2+\rho\rho_1) + x^4(\rho-\rho_2)(2-3\rho^2+\rho\rho_2) +$$

$$+ x^2[3\rho(\rho^2-\rho_1\rho_2) - (4\rho^2-1)(2\rho-\rho_1-\rho_2)]\Big\} + O(\beta^3). \tag{11}$$

Отметим, что при $\tau = 0,\ \rho = \rho_1 = \rho_2 = 1$ в (11) нули в знаменателях сокращаются с нулями числителей.

Как следует из соотношения (11), нелинейные искажения АКФ в знаковом коррелометре при использовании преобразования Ван Флека имеют второй порядок малости по отношению интенсивностей гармонических помех и исследуемого нормального процесса. Изменяются лишь относительные уровни шумовой и гармонических компонент, что является следствием эффекта подавления слабого сигнала сильным в ограничителе.

АКФ входного процесса $e(t)$ равна $\rho_e = \rho + \beta\rho_1 + x^2\beta\rho_2$, выходного — $R(\tau) = \rho - \dfrac{\rho^3}{6} + \dfrac{3\rho^5}{40} + \ldots - \beta(1+x^2)\rho + \beta\rho_1 + \beta x^2\rho_2 + \ldots$, а после ОПВФ —

$$\begin{aligned}\widehat{\rho}(\tau) &= \rho[1-\beta(1+x^2)+\beta^2(1+x^2+x^4)/2] + \beta\rho_1[1-\beta(2+x^2)/4] + \\ &+ \beta x^2\rho_2[1-\beta x^2(1+2x^2)/4] + \beta^2\,[\text{нелинейные члены}].\end{aligned}$$

Видно, что в $R(\tau)$ в отличие от $\widehat{\rho}(\tau)$ нелинейность проявляется даже при $\beta \to 0$.

Таким образом, показано, что ОПВФ полезно использовать и при измерениях АКФ гауссового процесса в смеси с гармоническими помехами.

3. ИНТЕНСИВНОСТИ ГАРМОНИЧЕСКИХ КОМПОНЕНТ ВЕЛИКИ, $A > \sigma$

Этот случай сложнее для анализа, чем предыдущий. Большие трудности представляет также учёт ОПВФ, поскольку в выражении (10) $\Delta > \arcsin\rho$. Так как теперь $\sigma < A$, разложим в выражении (6) гиперболические функции в ряды Тейлора, сохранив лишь члены, пропорциональные $\sigma^2\rho$:

$$\begin{aligned}R(\tau) &= \frac{4}{\pi^2}\Bigg[\sigma^2\rho\sum\sum_{k+l=2m}\epsilon_k\epsilon_l g_{kl}^2(\alpha,x)\cos k\gamma_1\cos l\gamma_2 + \\ &+ \sum\sum_{k+l=2m+1}\epsilon_k\epsilon_l h_{kl}^2(\alpha,x)\cos k\gamma_1\cos l\gamma_2\Bigg] + O(\sigma^4\rho^2),\end{aligned} \tag{12}$$

где

$$g_{kl}(\alpha,x) = \frac{1}{\alpha}\int\limits_0^\infty \mathrm{e}^{-u^2/\alpha^2} J_k(u)J_l(xu)\,\mathrm{d}u\,, \tag{13}$$

$$h_{kl}(\alpha,x) = \int\limits_0^\infty \mathrm{e}^{-u^2/\alpha^2} J_k(u)J_l(xu)\frac{\mathrm{d}u}{u} \tag{14}$$

и введён амплитудный параметр $\alpha = \sqrt{2}\,A/\sigma = 2\sqrt{\beta} \gg 1$. Первая сумма в (12) даёт непрерывную (“шумовую") часть спектра, вторая — дискретную. Для коэффициентов g_{kl} и h_{kl} необходимо получить асимптотики при $\alpha \gg 1$. Для g главный член асимптотики из (13) получается лишь при $x < 1$:

$$\begin{aligned}g_{kl}(\infty,x) &\sim \frac{x^l\sin[(l-k+1)\pi/2]}{\pi\alpha\, l!}\Gamma\left(\frac{l{+}k{+}1}{2}\right)\Gamma\left(\frac{l{-}k{+}1}{2}\right)\times \\ &\times F\left(\frac{l{+}k{+}1}{2}, \frac{l{-}k{+}1}{2}; l{+}1; x^2\right),\end{aligned} \tag{15}$$

При $x = 1$ гипергеометрическая функция в (15) обращается в бесконечность:

$$g_{00}(\alpha,x) = \frac{F(1/2,1/2;1;x^2)}{\alpha} = \frac{2\mathbf{K}(x)}{\alpha\pi}\,,$$

где $\mathbf{K}(x)$ — полный эллиптический интеграл первого рода, $\mathbf{K}(1) = \infty$.

Коэффициенты h_{kl} при $\alpha \to \infty$ остаются конечными, при любых значениях x они равны

$$h_{kl}(\infty,x) \sim \frac{x^l\sin[(l{-}k)\pi/2]}{2\pi\, l!}\Gamma\left(\frac{l{+}k}{2}\right)\Gamma\left(\frac{l{-}k}{2}\right)F\left(\frac{l{+}k}{2}, \frac{l{-}k}{2}; l{+}1; x^2\right), \tag{16}$$

аналогичное выражение приведено в [2] (формула (61)). Для расчёта коэффициентов $h_{kl}(\infty, x)$ достаточно определить два первых $h_{10},\ h_{01}$, остальные вычисляются реккурентно:

$$h_{21} = (2xh_{10} - h_{01})/3, h_{12} = (2h_{01} - xh_{10})/3x,$$

$$h_{30} = [h_{12} + 2(1 - 2x^2)h_{21}/x]/3, h_{03} = [h_{21} + 2(x^2 - 2)h_{12}/x]/3 \text{ и т.д.}$$

Для h_{10} и h_{01}, используя связь гипергеометрических функций с полными эллиптическими интегралами первого и второго рода, получим

$$h_{10}(\infty, x) = F(1/2,\, -1/2;\, 1;\, x^2) = (2/\pi)\mathbf{E}(x),$$

$$h_{01}(\infty, x) = (x/2)F(1/2, 1/2; 2; x^2) = (2/\pi x)[\mathbf{E}(x) - (1 - x^2)\mathbf{K}(x)].$$

Для коэффициентов g_{kl} из (13) получаются следующие реккурентные соотношения:

$$g_{kl}(\alpha, x) = 2(k-1)h_{k-1,l}(\alpha, x) - g_{k-2,l}(\alpha, x) = \\ = 2(l-1)h_{k,l-1}(\alpha, x)/x - g_{k,l-2}(\alpha, x),$$

поэтому, если известны h_{kl}, достаточно определить только g_{00} и g_{11}. В Приложениях 2 и 3 дан вывод соответствующих асимптотик. Ниже даны результаты и условия их применимости, проверенные на ЭВМ.

$$h_{kl}(\alpha, x) \sim h_{kl}(\infty, x)\left[1 + \frac{l^2 - k^2}{\alpha^2(1 - x^2)}\right], \quad \alpha^2(1 - x^2) \geq 10,$$

$$g_{kl}(\alpha, x) \sim g_{kl}(\infty, x)\left[1 + \frac{(l+1)^2 - k^2}{\alpha^2(1 - x^2)^2}\frac{A}{B}\right], \quad \alpha(1 - x^2) \geq 10,$$

$$A = F\left(\frac{l+k-1}{2}, \frac{l-k-1}{2}; l+1; x^2\right),$$

$$B = F\left(\frac{l+k+1}{2}, \frac{l-k+1}{2}; l+1; x^2\right),$$

$$h_{10}(\alpha, 1) = h_{01}(\alpha, 1) \sim \frac{2}{\pi}\left(1 - \frac{\ln\alpha + 1,886}{4\alpha^2}\right), \quad \alpha \geq 2,$$

$$g_{00}(\alpha, 1) \sim \frac{1}{\pi\alpha}\left(\ln\alpha + 2,368 + \frac{\ln\alpha - 1.132}{8\alpha^2}\right), \quad \alpha \geq 2,$$

$$g_{11}(\alpha, 1) \sim \frac{1}{\pi\alpha}\left(\ln\alpha + 0,368 + 3\frac{\ln\alpha - 1,045}{8\alpha^2}\right), \quad \alpha \geq 2.$$

Эти соотношения позволяют рассчитать все составляющие спектра ограниченного процесса при любых параметрах его компонент.

Отметим следующий факт. При $x \to 1$ шумовая составляющая АКФ с ростом входного отношения сигнала к шуму ($\alpha > 1$) убывает медленнее первой степени α из-за логарифмического члена в выражении для g_{00}. Этим случай двух близких по интенсивности косинусоид существенно отличается от случая одного монохроматического сигнала. При $x \ll 1$ $g_{00} \sim \alpha^{-1}$, как и следовало бы ожидать из физических соображений. В работе [2] отмеченное свойство ограничителя было обнаружено при численных расчётах (см. первый абзац на с. 38 и последний абзац на с. 39). Однако Джонсу не удалось получить асимптотики при $\alpha \to \infty$ в аналитическом виде, и он ограничился утверждением, что $\alpha g_{00} \to \infty$ при $\alpha \to \infty$.

4. КОМБИНАЦИОННЫЕ ПОМЕХИ С УЧЁТОМ ОПВФ

Рассмотрим наиболее неблагоприятный случай $\sigma = 0$, поскольку при $\sigma > 0$ интенсивности гармонических и комбинационных помех будут меньше.

$$\begin{aligned} R(\tau) &= \frac{4}{\pi^2}\sum\sum \epsilon_k\epsilon_l h_{kl}^2 \cos k\gamma_1 \cos l\gamma_2 = b_1\cos\gamma_1 + b_2\cos\gamma_2 + \\ &+ b_3\cos(2\gamma_1+\gamma_2) + b_4\cos(2\gamma_1-\gamma_2) + b_5\cos(2\gamma_2+\gamma_1) + \\ &+ b_6\cos(2\gamma_2-\gamma_1) + b_7\cos 3\gamma_1 + \\ &+ b_8\cos 3\gamma_2 + (\text{члены высших порядков}), \end{aligned} \qquad (17)$$

где $b_1 = (4/\pi^2)h_{10}^2$, $b_2 = (4/\pi^2)h_{01}^2$ и т.д..

Оценку результатов ОПВФ суммы (17) можно сделать двумя способами.

1. Используя представление $\sin(z\cos\theta) = \mathrm{Im}\sum\limits_{k=-\infty}^{\infty} J_k(z)\exp\left[\mathrm{i}k\left(\theta+\dfrac{\pi}{2}\right)\right]$,

2. Используя ряд Тейлора $\sin\left(\dfrac{\pi R}{2}\right) = \dfrac{\pi R}{2} - \dfrac{(\pi R/2)^3}{6} + \dfrac{(\pi R/2)^5}{120} - \ldots$

Члены этих разложений быстро убывают, что позволяет заменить их конечными суммами. Оба способа дали практически совпадающие коэффициенты разложения $\widehat{\rho}$:

$$\widehat{\rho} = \rho_{10}\cos\omega_1\tau + \rho_{01}\cos\omega_2\tau + \rho_{21}\cos(2\omega_1\tau - \omega_2\tau) + \rho_{12}\cos(2\omega_2\tau - \omega_1\tau). \qquad (18)$$

Сравним относительные уровни комбинационных частот АКФ до ОПВФ со спектром после ОПВФ. Результаты расчётов приведены в таблице ниже, где b_i/b_1 — в соответствии с (17), $\rho ik/\rho_{10}$ — по (18).

Таблица

x^2	0,2	0,4	0,6	0,8	1,0
b_2/b_1	0,059	0,141	0,266	0,476	1,00
ρ_{01}/ρ_{10}	0,031	0,101	0,221	0,437	1,00
b_3/b_1	0,047	0,088	0,119	0,134	0,111
ρ_{21}/ρ_{10}	0,027	0,059	0,082	0,091	0,061
b_4/b_1	0,001	0,004	0,012	0,033	0,111
ρ_{12}/ρ_{10}	0	0	0,005	0,023	0,061

Как видно из таблицы, уровни комбинационных помех заметно снизились сравнительно с их уровнями до ОПВФ.

Отметим, что при $x = 0$, т. е. в присутствии только одной гармоники во входном сигнале, знаковый коррелометр с преобразованием Ван Флека даёт точный результат. Покажем это простым способом. Если $e(t) = A\cos\omega t$, то после ограничителя $y(t) = \mathrm{sign}[e(t)]$ — “прямоугольная"волна единичной амплитуды, а её АКФ — волна “треугольной"формы:

$$R(\tau) = \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{\cos(2k+1)\omega\tau}{(2k+1)^2} = 1 - \frac{2}{\pi}\omega|\tau| \quad (\omega|\tau| \le \pi/2).$$

При этом $\widehat{\rho}(\tau) = \sin(\pi/2 - \omega|\tau|) = \cos\omega\tau$, $\rho_{10} = 1$, остальные равны нулю, т. е. после ОПВФ получается точный вид АКФ без нелинейных искажений.

На самом деле, как показано в [5], преобразование Ван Флека после знакового коррелометра даёт точное восстановление АКФ случайного процесса, если его двумерная плотность вероятностей симметрична, т. е. зависит от $x(t)$ и $x(t+\tau)$ через комбинацию $x^2(t) + x^2(t+\tau) + Qx(t)x(t+\tau)$ (Q — постоянная). Этому требованию удовлетворяют многие случайные функции, в частности, центрированный нормальный процесс и косинусоида с абсолютно случайной фазой. К сожалению , сумма нормального процесса и гармонического колебания этому условию не удовлетворяет. Но тем не менее и для таких сигналов ОПВФ оказывается полезным, что видно из выражения (11) и из приведённых выше данных. Если до преобразования относительный уровень комбинационной компоненты при $x = 1$ равен $h_{21}/h_{10} = 1/9$, то после ОПВФ он снижается до 0,061. В худшем случае при $x^2 = 0{,}8$ соответствующие амплитуды равны 0,134 и 0,091.

Отметим, что в монографии [8] на с. 183 ошибочно утверждается, что вывод о применимости преобразования $R = (2/\pi)\arcsin\rho$ к любым случайным процессам с симметричными плотностями распределения неверен. Вероятно, это некорректное утверждение вызвано тем, что автор [8] не обратил внимание на тот факт, что в соответствии с [5] требуется эллиптическая симметрия двумерной плотности, а не симметрия одномерной плотности.

ЗАКЛЮЧЕНИЕ

1. Приведённые в работе соотношения позволяют рассчитать автокорреляционную функцию ограниченного по амплитуде случайного процесса в виде суммы нормального шума и двух косинусоид с абсолютно случайными взаимнонезависимыми фазами при любых соотношениях интенсивностей компонент.

2. Показано, что синусное преобразование Ван Флека, обеспечивающее в отсутствие гармоник точное восстановление АКФ, полезно и при измерении АКФ составного процесса.

ПРИЛОЖЕНИЕ 1

Выражения для коэффициентов $H_{kl}(\beta,\rho,x)$, рассчитанные по формулам (8) и (9), имеют следующий вид:

$$H_{00} = \frac{2}{\pi}\left\{\arcsin\rho - \frac{\beta\rho(1+x^2)}{(1-\rho^2)^{1/2}}\left[1 - \frac{\beta(1+x^2)}{4(1-\rho^2)} - \frac{\beta(1+4x^2+x^4)}{8(1+x^2)}\frac{3-2\rho^2}{1-\rho^2} + \right.\right.$$

$$\left.\left. + \; 3\beta^2\frac{1+4x^2+x^4}{16(1-\rho^2)} + \beta^2\frac{1+8x^2+x^4}{144}\,\frac{15-20\rho^2+8\rho^4}{(1-\rho^2)^2}\right] + O(\beta^4)\right\},$$

$$H_{10} = \frac{\beta}{\pi(1-\rho^2)^{1/2}}\left[1 - \beta\frac{1+2x^2}{2(1-\rho^2)} + \frac{\beta^2}{16}(1+2x^2)^2\frac{1+2\rho^2}{(1-\rho^2)^2} + \right.$$

$$\left. + \; \frac{\beta^2}{8}\,\frac{1+6x^2+3x^4}{(1-\rho^2)^2}\right] + O(\beta^4),$$

$$H_{01} = \frac{x^2\beta}{\pi(1-\rho^2)^{1/2}}\left[1 - \beta\frac{2+x^2}{2(1-\rho^2)} + \frac{\beta^2}{16}(2+x^2)^2\frac{1+2\rho^2}{(1-\rho^2)^2} + \right.$$

$$\left. + \; \frac{\beta^2}{8}\,\frac{3+6x^2+x^4}{(1-\rho^2)^2}\right] + O(\beta^4),$$

$$H_{11} = \frac{x^2\beta^2\rho}{2\pi(1-\rho^2)^{3/2}}\left(1 - \frac{3}{2}\beta\frac{1+x^2}{1-\rho^2}\right) + O(\beta^4),$$

$$H_{20} = \frac{\beta^2\rho}{8\pi(1-\rho^2)^{3/2}}\left(1 - \beta\frac{1+3x^2}{1-\rho^2}\right) + O(\beta^4),$$

$$H_{02} = \frac{x^4\beta^2\rho}{8\pi(1-\rho^2)^{3/2}}\left(1 - \beta\frac{3+x^2}{1-\rho^2}\right) + O(\beta^4),$$

$$H_{21} = \frac{x^2\beta^3(1+2\rho^2)}{16\pi(1-\rho^2)^{5/2}} + O(\beta^4), \quad H_{12} = \frac{x^4\beta^3(1+2\rho^2)}{16\pi(1-\rho^2)^{5/2}} + O(\beta^4),$$

$$H_{30} = \frac{\beta^3(1+2\rho^2)}{144\pi(1-\rho^2)^{5/2}} + O(\beta^4), \quad H_{03} = \frac{x^6\beta^3(1+2\rho^2)}{144\pi(1-\rho^2)^{5/2}} + O(\beta^4).$$

ПРИЛОЖЕНИЕ 2

Расчёт асимптотики для $g_{kl} = \dfrac{1}{\alpha}\int\limits_0^\infty \exp\left(-\dfrac{u^2}{\alpha^2}\right) J_k(u)\, J_l(xu)\, \mathrm{d}u$ при $\alpha \to \infty$. Разложив $J_l(xu)$ в ряд по степеням xu: $J_l(xu) == \left(\dfrac{xu}{2}\right)^l \sum\limits_{m=0}^{\infty} \dfrac{(-1)^m}{m!(m+l)!}\left(\dfrac{xu}{2}\right)^{2m}$, получим

$$g_{kl}(\alpha, x) = \frac{x^l}{2^l\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m}{(m+l)!}\frac{x^{2m}}{2^{2m}}I_m, \quad \text{где } I_m = \int\limits_0^\infty \mathrm{e}^{-u^2/\alpha^2} J_k(u) u^{2m+l}\mathrm{d}u.$$

Интегралы I_m (см. [6], с. 186) равны

$$I_m = \frac{\alpha^{2m+l+k+1}}{2^{k+1}}\,\frac{\Gamma[m+(l+k+1)/2]}{k!}\,{}_1F_1\left(m + \frac{l+k+1}{2}; k+1; -\frac{\alpha^2}{4}\right).$$

При $\alpha \gg 1$ для вырожденной гипергеометрической функции (см. [7], с. 310) используем асимптотику ${}_1F_1(a;b;-z) \sim \Gamma(a)\Gamma^{-1}(b)z^{-a}{}_2F_0(a, a-b+1; z^{-1})$. Тогда

$$\begin{aligned} g_{kl}(\alpha, x) &= \frac{x^l}{\alpha}\sum_{m=0}^{\infty}\frac{(-1)^m x^{2m}\Gamma[m+(l+k+1)/2]}{m!(m+l)!\Gamma[1-m-(l-k+1)/2]} \times \\ &\times\ {}_2F_0\left(m + \frac{l+k+1}{2}, m + \frac{l-k+1}{2}; \frac{4}{\alpha^2}\right), \end{aligned}$$

$$\begin{aligned} &{}_2F_0\left(m + \frac{l+k+1}{2}, m + \frac{l-k+1}{2}; \frac{4}{\alpha^2}\right) = \\ &= \sum_{p=0}^{\infty}\frac{(4/\alpha)^p}{p!}\,\frac{\Gamma[p+m+(l+k+1)/2]}{\Gamma[m+(l+k+1)/2]}\,\frac{\Gamma[p+m+(l-k+1)/2]}{\Gamma[m+(l-k+1)/2]}\,. \end{aligned}$$

Учтём, что $\Gamma[1-m-(l-k+1)/2]\,\Gamma[m+(l-k+1)/2] = \Gamma(1-y)\,\Gamma(y) == \pi/\sin y\pi = (-1)^m\pi/\sin(l-k+1)\pi/2$. В результате получим для g_{kl} двойную сумму

$$g_{kl}(\alpha,x) = \frac{x^l \sin(l-k+1)\pi/2}{\alpha\pi}\sum_{p=0}^{\infty}\frac{(4/\alpha^2)^p}{p!}\sum_{m=0}^{\infty}\frac{x^{2m}}{m!(m+l)!}\times$$

$$\times \quad \Gamma[p+m+(l+k+1)/2]\Gamma[p+m+(l+k-1)/2].$$

Так как сумма по m представляется гипергеометрической функцией Гаусса, то

$$g_{kl} \sim \frac{\sin(l-k+1)\pi/2}{\alpha\pi}\frac{x^l}{l!}\sum_{p=0}^{\infty}\frac{(4/\alpha^2)^p}{p!}\Gamma\left(p+\frac{l+k+1}{2}\right)\times$$

$$\times \quad \Gamma\left(p+\frac{l-k+1}{2}\right)F\left(p+\frac{l+k+1}{2}, p+\frac{l-k+1}{2}; l+1; x^2\right).$$

При $\alpha\to\infty$ сохраним лишь слагаемое с $p=0$:

$$g_{kl}(\infty,x) = \frac{\sin(l-k+1)\pi/2}{\alpha\pi}\frac{x^l}{l!}\Gamma\left(\frac{l+k+1}{2}\right)\Gamma\left(\frac{l-k+1}{2}\right)\times$$

$$\times \quad F\left(\frac{l+k+1}{2},\frac{l-k+1}{2}; l+1; x^2\right).$$

В частности,

$$g_{00}(\infty,x) = F(1/2,1/2;1;x^2)/\alpha = 2\mathbf{K}(x)/\alpha\pi,$$
$$g_{11}(\infty,x) = xF(3/2,1/2;2;x^2)/2\alpha = 2[\mathbf{K}(x)-\mathbf{E}(x)]/\alpha\pi x.$$

При $x=1$ все члены асимптотического ряда обращаются в бесконечность. Для подтверждения используем тождество $F(a,b;c;z) == (1-z)^{c-a-b}F(c-b,c-a;c;z)$. В нашем случае $a=p+(l+k+1)/2$, $b=p+(l-k+1)/2$, $c=l+1$, $z=x^2$:

$$F\left(p+\frac{l+k+1}{2}, p+\frac{l-k+1}{2}; l+1; x^2\right) =$$

$$= (1-x^2)^{-2p}F\left(\frac{l+k+1}{2}-p, \frac{l-k+1}{2}-p; l+1; x^2\right).$$

Окончательно асимптотический ряд для g_{kl} при $\alpha\to\infty$ представим в следующем виде:

$$g_{kl}(\alpha,x) \sim \frac{\sin(l-k+1)\pi/2}{\alpha\pi}\sum_{p=0}^{\infty}\frac{1}{p!}\left[\frac{2}{\alpha(1-x^2)}\right]^{2p}\Gamma\left(p+\frac{l+k+1}{2}\right)\times$$

$$\times \quad \Gamma\left(p+\frac{l-k+1}{2}\right)F\left(\frac{l+k+1}{2}-p, \frac{l-k+1}{2}-p; l+1; x^2\right).$$

В последнем выражении гипергеометрическая функция остаётся конечной при $x=1$, а расходимости связаны со знаменателями $(1-x^2)^{2p}$. Полученное выражение пригодно при условии, что $\alpha(1-x^2)\gg 1$. При $x=1$ асимптотики для g_{kl} будут иметь вид рядов по степеням $\ln\alpha/\alpha$, и этот случай требует особого рассмотрения. Поскольку для g_{kl} были получены рекуррентные соотношения, ограничимся расчётом $g_{00}(\alpha,1)$ и $g_{11}(\alpha,1)$:

$$g_{00}(\alpha,1) = \frac{1}{\alpha}\int_0^{\infty}\exp(-u^2/\alpha^2)J_0^2(u)\,\mathrm{d}u.$$

Используя представление $J_0^2(u) = \frac{2}{\pi}\int\limits_0^{\pi/2} J_0(2u\sin\phi)\,\mathrm{d}\phi$ и интеграл $\int\limits_0^{\infty}\exp\left(-\frac{u^2}{\alpha^2}\right) J_0(2u\sin\phi)\,\mathrm{d}u = \frac{\alpha\sqrt{\pi}}{2}\exp\left(-\frac{\alpha^2}{2}\sin^2\phi\right) I_0\left(\frac{\alpha^2}{2}\sin^2\phi\right)$, где I_0 — функция Бесселя от мнимого аргумента, получим

$$g_{00}(\alpha,1) = \frac{1}{\sqrt{\pi}}\int\limits_0^{\pi/2}\exp\left(-\frac{\alpha^2}{2}\sin^2\phi\right) I_0\left(\frac{\alpha^2}{2}\sin^2\phi\right)\,\mathrm{d}\phi.$$

Разобьём интеграл на две части: от 0 до ψ и от ψ до $\pi/2$.

Первый интеграл

$$\begin{aligned} L_1 &= \int\limits_0^{\psi}\exp\left(-\frac{\alpha^2}{2}\sin^2\phi\right) I_0\left(\frac{\alpha^2}{2}\sin^2\phi\right)\,\mathrm{d}\phi = \\ &= \frac{\sqrt{2}}{\alpha}\int\limits_0^{q}\exp(-z^2)I_0(z^2)\,\mathrm{d}z + O(\alpha^{-3}), \end{aligned}$$

где $q = (\alpha/\sqrt{2})\sin\psi$.

Второй интеграл $L_2 = \int\limits_{\psi}^{\pi/2}(\cdot)\mathrm{d}\phi$ вычислим, заменив I_0 асимптотикой:

$$I_0(z) = \frac{\exp(z)}{\sqrt{2\pi z}}\,{}_2F_0(1/2,\,1/2;\,1/2z) = \frac{\exp(z)}{\sqrt{2z}\pi^{3/2}}\sum_{m=0}^{\infty}\frac{\Gamma^2(m+1/2)}{m!(2z)^m}\,.$$

Тогда $L_2 = \frac{1}{\sqrt{\pi}\alpha}\sum\limits_{m=0}^{\infty}a_m K_m$, где $K_m = \frac{1}{\alpha^{2m}}\int\limits_{\psi}^{\pi/2}\sin^{-2m-1}\phi\,\mathrm{d}\phi$, $a_0 = 1, a_1 = 1/4, a_2 = 9/32,\ \ldots,$

$$K_0 = -\ln\mathrm{tg}(\psi/2) \sim \ln\alpha - \ln(q/\sqrt{2}) - q^2/2\alpha^2 + O(\alpha^{-4}),$$
$$K_1 = 1/4q^2 + [\ln\alpha - \ln(q/\sqrt{2})]/2\alpha^2 - 1/4\alpha^2 + O(\alpha^{-4}),$$
$$K_2 = 1/16q^4 + 1/8q^2\alpha^2 + O(\alpha^{-4})\,,\ \ldots$$

В результате $g_{00} = [\ln\alpha + C + (\ln\alpha + C_1)/8\alpha^2 + O(\ln\alpha/\alpha^4)]/\pi\alpha$, где

$$C = \sqrt{2\pi}\int\limits_0^{q}\exp(-z^2)I_0(z^2)dz - \ln(q/\sqrt{2}) - \frac{1}{16}q^2 + \frac{9}{512}q^4 + \frac{25}{2048}q^6 + \ldots,$$

$$C_1 = 8\sqrt{2\pi}\int\limits_0^{q}z^2\exp(-z^2)I_0(z^2)dz - \ln(q/\sqrt{2}) - 4q^2 - 2 + \frac{9}{32}q^2 + \frac{75}{512}q^4 + \ldots$$

Константы C и C_1 вычислены на ЭВМ. Как и следовало ожидать, они не зависят от q и равны: $C = 2{,}368$, $C_1 = -1{,}132$. Таким образом,

$$g_{00}(\alpha,1) = [\ln\alpha + 2,368 + (\ln\alpha - 1,132)/8\alpha^2 + O(\ln\alpha/\alpha^4)]/\pi\alpha.$$

При расчёте коэффициента $g_{11}(\alpha,1)$ выкладки аналогичны предыдущим. В результате

$$g_{11} = [\ln\alpha + 0,368 - 3(\ln\alpha + 1,041)/8\alpha^2 + O(\ln\alpha/\alpha^{-4})]/\pi\alpha.$$

ПРИЛОЖЕНИЕ 3

Расчёт асимптотики для $h_{kl} = \int\limits_0^\infty \exp\left(-\dfrac{u^2}{\alpha^2}\right) J_k(u) J_l(xu)\,\dfrac{\mathrm{d}u}{u}$ при $\alpha\to\infty$. Разложив $J_l(xu)$ в ряд по степеням xu и использовав интеграл (см. [6], с. 186)

$$\int\limits_0^\infty u^{2m+l-1}\mathrm{e}^{-u^2/\alpha^2} J_k(u)\,\mathrm{d}u = \\ = \frac{\alpha^{2m+l+k}}{2^{k+1}}\,\frac{\Gamma[m+(l+k)/2]}{k!}\,{}_1F_1[m+(l+k)/2;k+1;-\alpha^2/4],$$

а также асимптотику для вырожденной гипергеометрической функции (см. [7], с. 310) ${}_1F_1(a;b;-z)\sim \Gamma(a)\Gamma^{-1}(b-a)\,{}_2F_0(a;a-b;1/z)$, получим

$$h_{kl} \sim \frac{x^l\sin(l-k)\pi/2}{2\pi\,l!}\sum_{s=0}^{\infty}\frac{(4/\alpha^2)^s}{s!}\Gamma\left(s+\frac{l+k}{2}\right)\Gamma\left(s+\frac{l-k}{2}\right)\times \\ \times F\left(s+\frac{l+k}{2},s+\frac{l-k}{2};l+1;x^2\right).$$

Представим гипергеометрическую функцию при $s\geq 1$ в соответствии с формулой $F(a,b;c;z)=(1-z)^{c-a-b}F(c-b,c-a;c;z)$. Тогда

$$h_{kl}\sim h_{kl}(\infty,x)\left\{1+\frac{l^2-k^2}{\alpha^2(1-x^2)}+O\Big[\alpha^{-4}(1-x^2)^{-2}\Big]\right\},$$

где

$$h_{kl}(\infty,x)=\frac{x^l\sin(l-k)\pi/2}{2\pi\,l!}\Gamma\Big(\frac{l+k}{2}\Big)\Gamma\Big(\frac{l-k}{2}\Big)F\Big(\frac{l+k}{2},\frac{l-k}{2};l+1;x^2\Big).$$

Полученное асимптотическое разложение пригодно, если $\alpha^2(1-x^2)\gg 1$. При $x=1$ требуется иное выражение. В этом случае

$$h_{kl}(\alpha,1)=\int\limits_0^\infty \mathrm{e}^{-u^2/\alpha^2}J_k(u)\,J_l(u)\frac{\mathrm{d}u}{u}=h_{lk}(\alpha,1).$$

Приведём соответствующие расчёты лишь для $h_{10}(\alpha,1)=h_{01}(\alpha,1)$:

$$h_{10}(\alpha,1)=h_{10}(\infty,1)-\int\limits_0^\infty\frac{1-\exp(-u^2/\alpha^2)}{u}J_1(u)J_0(u)\,\mathrm{d}u=h_{10}(\infty,1)-\Delta\,,$$

где $h_{10}(\infty,1)=h_{01}(\infty,1)=2/\pi$. Учитывая, что $2J_1(u)J_0(u)=-\dfrac{\mathrm{d}}{\mathrm{d}u}J_0^2(u)$, возьмём интеграл для Δ по частям. Тогда

$$\Delta=\frac{1}{\alpha}g_{00}(\alpha,1)-\frac{1}{2}T,\quad \text{где }\ T=\int\limits_0^\infty\frac{1-\exp(-u^2/\alpha^2)}{u^2}J_0^2(u)\,\mathrm{d}u.$$

Для асимптотической оценки T разобьём интеграл на две части: от 0 до q и от q до ∞.

$$T_1 = \frac{1}{\alpha^2}\int\limits_0^q J_0^2(u)\,\mathrm{d}u - \frac{1}{2\alpha^4}\int\limits_0^q u^2 J_0^2(u)\,\mathrm{d}u + O(\alpha^{-6}),$$

$$T_2 = \int\limits_q^\infty \frac{1-\exp(-u^2/\alpha^2)}{u^2} J_0^2(u)\,\mathrm{d}u.$$

В интеграле от q до ∞ заменим $J_0^2(u)$ асимптотическим разложением при $u \to \infty$. Подставив это разложение в интеграл T_2, получим

$$T_2 \sim \frac{1}{\pi\alpha^2}\left(\ln\alpha + C_2\right) + O\left(\frac{\ln\alpha}{\alpha^4}\right), \quad T \sim \frac{1}{\pi\alpha^2}\left(\ln\alpha + C_3\right) + O\left(\frac{\ln\alpha}{\alpha^4}\right),$$

где C_3 — постоянная, не зависящая от α и от q, $C_3 = 1{,}886$.

В результате

$$h_{10}(\alpha,1) = h_{01}(\alpha,1) = \frac{2}{\pi}\left[1 - \frac{\ln\alpha + 1{,}886}{4\alpha^2} + O\left(\frac{\ln\alpha}{\alpha^4}\right)\right].$$

ЛИТЕРАТУРА

1. Томсон Р., Моран Дж., Свенсон Дж. Интерферометрия и синтез в радиоастрономии. — М.: Мир, 1989. – 567 с.
2. Jones J. J. //IEEE Trans., 1963. V. IT-9. P. 34.
3. Давенпорт В. Б., Рут В. А. Введение в теорию случайных процессов и шумов. — М.: ИЛ, 1960. – 468 с.
4. Weinreb S. //Proc. IEEE, 1961. V. 49. № 6. P. 1099.
5. McGraw D. K., Wagner J. F. //IEEE Trans., 1968. V. IT-14. P. 110.
6. Прудников А. П., Брычков Ю. А., Марычев О. И. Интегралы и ряды, специальные функции. — М.: Мир, 1983. – 750 с.
7. Люк Ю. Специальные математические функции и их аппроксимации. — М.: Мир, 1980. – 608 с.
8. Мирский Г. Я. Характеристики статистической взаимосвязи и их измерения. — М.: Энергоиздат, 1982. – 319 с.

Радиоастрономический институт
НАН Украины, г. Харьков, Украина

Поступила в редакцию
17 ноября 1998 г.

THE AUTOCORRELATION FUNCTION OF A SUM OF NORMALLY DISTRIBUTED NOISE AND A TWO HARMONIC VOLTAGES AT THE HARD-LIMITING WITH VAN VLECK TRANSFORMATION

L. G. Sodin

The paper presents a rigorous analysis of statistical characteristics for a sum of a Gaussian random process and two harmonic oscillations passed through an inertialess limiter, a typical unit of radio astronomical spectrometers. The analysis has been done for two most interesting cases, when the dominant intensity has the continious-spectrum component and when the dominant is the discrete-spectrum component. The formulas obtained for the autocorrelation function allow one to estimate the power spectrum of the clipped process as well as that of its Van Vleck transform. In all cases the transformation improves the estimates for the autocorrelation function and the spectrum, reducing the level of nonlinear products including components with combination frequences.

УДК 519.2+534.(2+6)

ДИСТАНЦИОННАЯ ДИАГНОСТИКА ШИРОКОПОЛОСНЫХ ДВИЖУЩИХСЯ ИСТОЧНИКОВ

Н. К. Вдовичева, В. И. Турчин, И. Ш. Фикс

Для диагностики протяжённого шумоподобного излучателя, движущегося вблизи приёмной антенной решётки, разработаны методы реконструкции матрицы ковариации сторонних элементарных источников, распределённых вдоль излучателя. Получены линеаризованные уравнения правдоподобия для искомой матрицы ковариации в случаях априорно известной и неизвестной матрицы ковариации фоновой помехи, исследована дисперсия оценок элементов ковариационной матрицы, характеризующей диагностируемый излучатель.

1. ВВЕДЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

Дистанционная диагностика протяжённых широкополосных источников различной природы с использованием многоэлементных приёмных систем (построение изображения, радиовидение и т.п.) рассматривалась ранее для сценариев, обычно включающих удалённость источника, априорные данные о его функции когерентности и внешних помехах и т.п. (см., например, [1–4]). В настоящей работе рассматривается задача дистанционной диагностики протяжённого шумоподобного источника, движущегося на небольшом расстоянии вдоль линейной решётки приёмных элементов*. Часто хорошей моделью такого источника можно считать одномерный ("нитевидный") излучатель со случайным распределением вдоль его оси x спектральных компонент элементарных источников $c_\omega(x)$ в некотором диапазоне частот ω. Диагностика при этом заключается в реконструкции функции когерентности на источнике $G_\omega(x', x'') = \mathrm{E}\{c_\omega(x') c_\omega^*(x'')\}$ по результатам наблюдения широкополосных сигналов $p_m(t)$ на m-х приёмных элементах линейной решётки $(m = 1, \ldots, M)$ в процессе движения источника с постоянной скоростью v; $\mathrm{E}\{\cdot\}$ здесь и далее означает математическое ожидание.

Особенностями рассматриваемого сценария можно считать:

– отсутствие каких-либо априорных предположений относительно вида функции когерентности G_ω;

– минимальная дистанция между приёмной решёткой и источником существенно меньше границы дальней зоны источника во всём диапазоне частот*; при движении источника функция Грина, описывающая распространение поля от элементарного источника в точке x до m-го приёмного элемента, меняется существенно;

– сигнал от источника наблюдается на фоне стационарной помехи с известными (измеренными заранее) или неизвестными статистическими характеристиками;

– скорость движения источника полагается достаточно малой, так что можно пренебречь трансформацией из-за эффекта Допплера спектральной плотности мощности внутри узких частотных полос анализа.

В дальнейшем будем предполагать, что принимаемые сигналы $p_m(t)$ пропускаются через узкополосные фильтры с центральными частотами ω с шириной полосы $\Delta\omega/2\pi < c/L_\mathrm{a}$, где c — скорость

*Практическим приложением этой задачи может являться диагностика протяжённых акустических источников шума: кораблей, поездов и т.п. [5–7].

*Малость дистанции позволяет, во-первых, достаточно хорошо прогнозировать функцию Грина при движении излучателя в неоднородной среде, во-вторых, получать высокое пространственное разрешение при разумных размерах приёмной антенны и, наконец, проводить измерения при максимальном отношении сигнал/помеха.

звука, L_a — длина приёмной решётки. Полагая, что интервал между отсчётами на выходе фильтров $p_{m,\omega}(t_s)$ в моменты t_s, $s = 1,\ldots,S$, составляет $2\pi/\Delta\omega$, будем считать эти отсчёты независимыми по времени и рассматривать их как $M \times 1$ векторы $\boldsymbol{p}_s$ (индекс ω в дальнейшем будем опускать). Всего при длине траектории движения источника L число независимых временных отсчётов составляет $S = L(\Delta\omega/2\pi)/v$.

Далее для каждой узкой полосы будем рассматривать конечномерную аппроксимацию излучателя совокупностью точечных элементарных источников с комплексными амплитудами c_n, $n = 1,\ldots,N$, расположенных в узлах эквидистантной сетки на оси излучателя с шагом $\lambda/2$, где λ — длина волны, соответствующая центральной частоте данной полосы. Будем полагать c_n стационарными случайными гауссовыми величинами, независимые отсчёты которых образуют $N\times 1$ вектора $\boldsymbol{c}_s$ (s — номер отсчёта) с нулевым средним и матрицей ковариации $\boldsymbol{G} = \mathrm{E}\left\{\boldsymbol{c}_s\boldsymbol{c}_s^\dagger\right\}$, где † означает эрмитовое сопряжение.

При сделанных выше предположениях связь между $\boldsymbol{c}_s$ и $\boldsymbol{p}_s$ определяется соотношением

$$\boldsymbol{p}_s = \boldsymbol{A}_s\boldsymbol{c}_s + \boldsymbol{\xi}_s, \qquad s = 1,\,\ldots,\,S, \tag{1}$$

где $\boldsymbol{A}_s$ — $M \times N$ матрица, элементами которой являются значения функции Грина для пар "n-й элементарный источник — m-й приёмный элемент" в момент t_s, характеризующий текущее положение источника относительно приёмной решётки. В случае свободного пространства и монопольных элементарных источников $\boldsymbol{A}_s = \left\| R_{mn,s}^{-1}\,\exp\left(i\dfrac{\omega}{c}R_{mn,s}\right)\right\|$, где $R_{mn,s}$ — расстояние между n-м источником и m-м приёмником в момент t_s. Вектор $\boldsymbol{\xi}_s$ в (1) характеризует помеху, которую будем считать распределённой по нормальному закону с нулевым средним и матрицей ковариации $\boldsymbol{K} = \mathrm{E}\left\{\boldsymbol{\xi}_s\boldsymbol{\xi}_s^\dagger\right\}$, не зависящей от s при сделанном выше предположении о стационарности помехи.

С учётом вышесказанного, наблюдаемые на приёмной решётке отсчёты сигнальных векторов $\boldsymbol{p}_s$ являются независимыми по s нормально распределёнными векторами с нулевым средним и матрицей ковариации $\boldsymbol{T}_s$:

$$\boldsymbol{T}_s = \mathrm{E}\left\{\boldsymbol{p}_s\boldsymbol{p}_s^\dagger\right\} = \boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger + \boldsymbol{K}\,. \tag{2}$$

Задача диагностики тем самым сводится к нахождению оценок $\hat{\boldsymbol{G}}$ матрицы ковариации $\boldsymbol{G}$ в каждой узкой полосе частот по совокупности результатов наблюдений $\boldsymbol{p}_s$, $s = 1,\ldots,S$ при известной либо неизвестной матрице ковариации помехи $\boldsymbol{K}$, а также к исследованию свойств полученных оценок, в частности, дисперсий элементов $\hat{\boldsymbol{G}}$

$$\mathrm{disp}\,\left\{\hat{G}_{nk}\right\} = \mathrm{E}\left\{\hat{G}_{nk}\hat{G}_{nk}^*\right\} - \mathrm{E}\left\{\hat{G}_{nk}\right\}\mathrm{E}\left\{\hat{G}_{nk}^*\right\} \tag{3}$$

в зависимости от параметров задачи.

2. РЕКОНСТРУКЦИЯ ИЗЛУЧАТЕЛЯ ПРИ ИЗВЕСТНОЙ КОВАРИАЦИОННОЙ МАТРИЦЕ ПОМЕХИ

Построение оценки $\hat{\boldsymbol{G}}$ для описанных выше моделей принимаемых антенной сигналов и помех (1), (2) проведём на основе метода максимального правдоподобия. Известно (см., например, [12, 13]), что для такой модели логарифм функции правдоподобия w для выборки принимаемого сигнала $\boldsymbol{p}_s$ с точностью до несущественных констант есть

$$\ln w = -\sum_s(\ln\det\boldsymbol{T}_s + \boldsymbol{p}_s^\dagger\boldsymbol{T}_s^{-1}\boldsymbol{p}_s)\,. \tag{4}$$

Заметим, что

$$\ln\det\boldsymbol{T}_s = \ln\det\boldsymbol{K} + \ln\det(\mathbf{I} + \boldsymbol{R}_s), \qquad \boldsymbol{T}_s^{-1} = (\mathbf{I} + \boldsymbol{R}_s)^{-1}\boldsymbol{K}^{-1}, \tag{5}$$

где $\mathbf{I}$ — единичная матрица, а $\boldsymbol{R}_s = \boldsymbol{K}^{-1}\boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger$. Полагая норму $\boldsymbol{R}_s$ меньше 1, разложим (5) в ряды по степеням $\boldsymbol{R}_s$ и найдём уравнение правдоподобия, приравнивая нулю первую производную $\ln w$ по $\boldsymbol{G}$:

$$\left(\frac{\partial \ln w}{\partial \boldsymbol{G}}\right)^T = -\sum_s [\boldsymbol{B}_s - \boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s + \boldsymbol{B}_s(\boldsymbol{G}\boldsymbol{B}_s)^2 - \ldots$$
$$\ldots - \boldsymbol{u}_s\boldsymbol{u}_s^\dagger + \boldsymbol{B}_s\boldsymbol{G}\boldsymbol{u}_s\boldsymbol{u}_s^\dagger + \boldsymbol{u}_s\boldsymbol{u}_s^\dagger\boldsymbol{G}\boldsymbol{B}_s - \ldots] = 0, \tag{6}$$

где $\boldsymbol{B}_s$ и $\boldsymbol{u}_s$ определяются как

$$\boldsymbol{B}_s = \boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s, \qquad \boldsymbol{u}_s = \boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{p}_s\,. \tag{7}$$

Линеаризуем (6), отбрасывая слагаемые порядка малости $O(P_{\text{с}}^2/P_{\text{п}}^2)$, где $P_{\text{с}}$ и $P_{\text{п}}$— мощности сигнала от излучателя и помехи, соответственно, на приёмных элементах [8]. При этом мы должны учесть, что слагаемые вида $\boldsymbol{B}_s(\boldsymbol{G}\boldsymbol{B}_s)^n$, $n = 0, 1, \ldots$ имеют порядок малости $O(P_{\text{с}}^n/P_{\text{п}}^n)$. Далее, слагаемые, включающие произведения сигнальных векторов $\boldsymbol{u}_s\boldsymbol{u}_s^\dagger$, в силу очевидного соотношения

$$\mathrm{E}\left\{\boldsymbol{u}_s\boldsymbol{u}_s^\dagger\right\} = \boldsymbol{B}_s + \boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s \tag{8}$$

содержат члены двух порядков малости отношения $P_{\text{с}}/P_{\text{п}}$. Поскольку эти слагаемые нельзя разделить, заменим два последних выписанных в (6) слагаемых на их математическое ожидание, полагая

$$\sum_s(\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{u}_s\boldsymbol{u}_s^\dagger + \boldsymbol{u}_s\boldsymbol{u}_s^\dagger\boldsymbol{G}\boldsymbol{B}_s) = 2\sum_s(\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s + \boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s) + \sum_s \boldsymbol{\Delta}_s\,, \tag{9}$$

где $\sum_s \boldsymbol{\Delta}_s$ — случайная матрица с нулевым математическим ожиданием и среднеквадратичным отклонением, имеющим порядок малости $O(1/\sqrt{S})$ относительно суммы математических ожиданий. Окончательно линеаризованное уравнение правдоподобия приобретает вид

$$\sum_s \boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s = \sum_s \boldsymbol{u}_s\boldsymbol{u}_s^\dagger - \sum_s \boldsymbol{B}_s\,. \tag{10}$$

Используя (8), нетрудно убедиться, что математические ожидания левой и правой частей (10) тождественно совпадают, т. е. оценка $\hat{\boldsymbol{G}}$, удовлетворяющая (10), является несмещённой при любых отношениях $P_{\text{с}}/P_{\text{п}}$,* если решение (10) существует.

Для нахождения элементов $\hat{\boldsymbol{G}}$ мы должны решить систему линейных уравнений (10). Свойства решения, как в случае любой обратной задачи, будут существенно зависеть от спектральных свойств матрицы этой системы. Для исследования этого вопроса рассмотрим вначале случай неподвижного источника; матрицы $\boldsymbol{A}_s$ и $\boldsymbol{B}_s$ при этом от s не зависят и уравнение (10), с учётом (7), приобретает вид**

$$\boldsymbol{B}\hat{\boldsymbol{G}}\boldsymbol{B} = \frac{1}{S_0}\sum_s \boldsymbol{A}^\dagger\boldsymbol{K}^{-1}\boldsymbol{p}_s\boldsymbol{p}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A} - \boldsymbol{B}\,.$$

Решение этого уравнения находится в явном виде

$$\hat{\boldsymbol{G}} = \boldsymbol{B}^{-1}\boldsymbol{A}^\dagger\boldsymbol{K}^{-1}\hat{\boldsymbol{T}}_0\boldsymbol{K}^{-1}\boldsymbol{A}\boldsymbol{B}^{-1} - \boldsymbol{B}^{-1}, \tag{11}$$

*Величина этого отношения влияет лишь на дисперсию оценки. При малых $P_{\text{с}}/P_{\text{п}}$ полученная оценка обладает свойством асимптотической эффективности, как любая максимально правдоподобная оценка. При больших $P_{\text{с}}/P_{\text{п}}$ дисперсия оценки может ухудшиться по сравнению с минимальной дисперсионной границей, однако её увеличение не столь существенно из—за общего снижения дисперсии оценки с ростом входного отношения сигнал/шум.

**Предполагается, что неподвижный источник наблюдается в течение некоторого времени t_0; соответствующее число независимых отсчётов обозначим S_0.

где $\hat{\boldsymbol{T}}_{\mathbf{0}} = \dfrac{1}{S_0}\sum\limits_s \boldsymbol{p}_s\boldsymbol{p}_s^\dagger$.

Характеризовать погрешность оценивания удобно величиной средней дисперсии σ^2:

$$\sigma^2 = \frac{1}{N^2}\sum_{n,k=1}^{N} \mathrm{disp}\left\{\hat{G}_{nk}\right\}, \tag{12}$$

которая находится непосредственно из (11):

$$\sigma^2 = S_0{}^{-1}N^{-2}\left[\mathrm{tr}\,\boldsymbol{G} + \mathrm{tr}\,(\boldsymbol{A}^\dagger\boldsymbol{K}^{-1}\boldsymbol{A})^{-1}\right]^2; \tag{13}$$

первое и второе слагаемые в (13) представляют собой составляющие дисперсии оценки, вносимые "шумовым" характером источника и внешней помехой, соответственно. Для изотропной помехи $\boldsymbol{K} = \varepsilon_0^2\mathbf{I}$ составляющая дисперсии σ_1^2, связанная только с помехой, может быть представлена в виде

$$\sigma_1^2 = S_0{}^{-1}N^{-2}\left[\varepsilon_0^2\mathrm{tr}\,(\boldsymbol{A}^\dagger\boldsymbol{A})^{-1}\right]^2 = S_0{}^{-1}N^{-2}\varepsilon_0^4\left(\sum_{i=1}^{N}\mu_i^{-1}\right)^2, \tag{14}$$

где μ_i — собственные значения $\boldsymbol{A}^\dagger\boldsymbol{A}$; плохая обусловленность $\boldsymbol{A}^\dagger\boldsymbol{A}$ (большое отношение $\mu_{\max}/\mu_{\min}$, где $\mu_{\max} \ge \mu_i \ge \mu_{\min}$) приводит к резкому увеличению дисперсии оценки.

При численных расчётах будем полагать, во-первых, что приёмные элементы расположены эквидистантно с шагом $\lambda/2$, во-вторых, что источник и приёмная решётка расположены в свободном пространстве параллельно на расстоянии R друг от друга, и центр источника смещён вдоль оси на величину x_0 относительно центра приёмной решётки. Очевидно, что минимальная величина дисперсии достигается при бесконечной длине приёмной решётки: $\sigma_0^2 = \sigma_1^2|_{M\to\infty}$. В этом случае можно воспользоваться явным видом матрицы $\boldsymbol{A}^\dagger\boldsymbol{A} \simeq \left\|R_F^{-2}\pi J_0(\pi|n-k|)\right\|$ [9], где $J_0(x)$ — функция Бесселя нулевого порядка, $R_{\mathrm{F}} = (\lambda R/2)^{1/2}$ — радиус первой зоны Френеля. Расчёт собственных значений этой матрицы показал, что $\mu_{\max} \simeq 2{,}67\,R_{\mathrm{F}}^{-2}\sqrt{N}$ и $\mu_{\min} \simeq 2\,R_{\mathrm{F}}^{-2}$; при этом величина $\sigma_0^2\varepsilon_0^{-4}R_{\mathrm{F}}^{-4}S_0$ с ростом N монотонно возрастает и достаточно быстро выходит на константу, равную 0,153 [9–10].

Для приёмной решётки конечных размеров поведение μ_i $(i = 1, \ldots, N)$ исследовалось в [10] в зависимости от M и смещения центра излучателя $d = 2x_0/\lambda$. Показано, что при $d = 0$ с уменьшением M до некоторого значения M_0 величина $\mu_{\min}$ практически не изменяется, а величина $\mu_{\max}$ монотонно уменьшается; при этом σ_1^2 незначительно возрастает. При $M < M_0$ и дальнейшем уменьшении M вначале одно собственное значение $\mu_{\min}$ начинает быстро приближаться к 0, а затем начинает возрастать количество собственных значений $\mu_i \simeq 0$; при этом, естественно, σ_1^2 резко возрастает. В качестве иллюстрации на рис. 1а приведены μ_i для различных M* $(d = 0)$. Аналогично ведут себя μ_i при возрастании $|d|$ и фиксированном M: с увеличением $|d|$ количество $\mu_i \simeq 0$ возрастает (см. рис. 1б, где приведены μ_i для различных d при $M = 192$), так что при больших $|d|$ остаётся только одно собственное значение, отличное от 0 (фактически это соответствует уже расположению излучателя в дальней зоне приёмной решётки).

Приведённый выше анализ показывает, что каждому конкретному набору параметров (R, λ, N, d) отвечает определённая характерная длина приёмной решётки $M_0\lambda/2$ (например, в рассмотренном случае $M_0 \simeq 84$). При меньших длинах приёмной решётки дисперсия σ_1^2 резко возрастает и получение несмещённых оценок уже оказывается проблематичным (соответствующие расчёты проделаны в [9–10]).

*Здесь и в дальнейшем результаты численных расчётов приводятся для $N = 12$, $\lambda = 6$ м, $R = 8\lambda$.

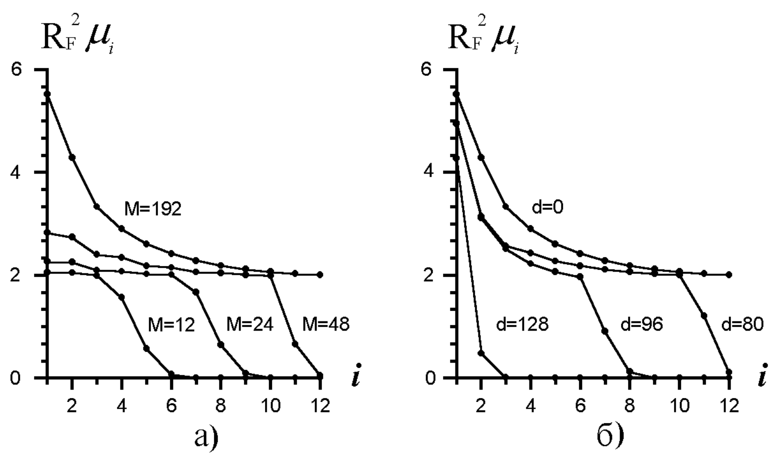


Рис. 1.

Для движущегося излучателя уравнение (10) следует рассматривать как обычную систему из N^2 линейных уравнений относительно неизвестных элементов матрицы $\boldsymbol{G}$. Для представления этой системы в каноническом виде ниже приводится специальная процедура, существенно использующая эрмитовость входящих в систему матриц $\boldsymbol{B_s}$, $\boldsymbol{G}$, $\boldsymbol{Z}$, где $\boldsymbol{Z}$ обозначает правую часть (10) [10].

Пусть $\mathcal{H}$ — линейное пространство эрмитовых матриц. Размерность пространства $\mathcal{H}$ равна N^2, где N — размер составляющих $\mathcal{H}$ матриц, и в $\mathcal{H}$ имеется естественный ортонормированный базис, состоящий из матриц $\boldsymbol{E}^{(n)}$, $\boldsymbol{F}^{(n,j)}$, $\boldsymbol{H}^{(n,j)}$ с элементами

$$\begin{cases} E_{kl}^{(n)} = \delta_{nk}\delta_{nl}, & n = 1, ..., N\,, \\ F_{kl}^{(n,j)} = \dfrac{1}{\sqrt{2}}(\delta_{nk}\delta_{jl} + \delta_{jk}\delta_{nl}), & 1 \le n < j \le N\,, \\ H_{kl}^{(n,j)} = \dfrac{i}{\sqrt{2}}(\delta_{nk}\delta_{jl} - \delta_{jk}\delta_{nl}), & 1 \le n < j \le N\,,\ \ k,l = 1, ..., N\,, \end{cases} \tag{15}$$

где δ_{nk} — символ Кронекера. Перепишем (10) в виде линейного уравнения в пространстве $\mathcal{H}$: $\mathcal{A}(\boldsymbol{G}) = \boldsymbol{Z}$, где $\mathcal{A}$ — симметрический оператор в $\mathcal{H}$ вида $\mathcal{A}(\boldsymbol{G}) = \sum\limits_s \boldsymbol{B_s G B_s}$.

Так как $\boldsymbol{G}$ — эрмитова матрица, её разложение по базису (15) имеет вид

$$\boldsymbol{G} = \sum_k G_{kk}\boldsymbol{E}^{(k)} + \sum_{k>l} \sqrt{2}\,\mathrm{Re}\,G_{kl}\,\boldsymbol{F}^{(k,l)} + \sum_{k>l} \sqrt{2}\,\mathrm{Im}\,G_{kl}\,\boldsymbol{H}^{(k,l)}\,, \tag{16}$$

где $1 \le l < k \le N$. Аналогичное разложение следует и для $\boldsymbol{Z}$.

Установим теперь взаимно однозначное соответствие между эрмитовыми матрицами и векторами из пространства $\mathbf{R}^{N^2}$.

Пусть $r_k = (k-1)(k-2)/2,\ k = 2, \ldots, (N+1)$, и для пары $(k,l),\ 1 \le l < k \le N$,

$$j(k,l) = r_k + l\,. \tag{17}$$

Тогда j есть взаимно однозначное соответствие между множеством пар чисел $\{(k,l),\ 1 \le l < k \le N\}$, определяющим позицию элемента в матрице, и множеством чисел $\{1, \ldots, N(N-1)/2\}$, определяющих позицию этого же элемента в векторе. Следовательно, $N \times N$ матрице $\boldsymbol{G}$ (16) будет отвечать $N^2 \times 1$

вектор $\mathbf{g}$ с элементами

$$\mathrm{g}_m(\boldsymbol{G}) = \begin{cases} \operatorname{Re} G_{k(m)l(m)}, & 1 \le m \le N(N-1)/2, \\ \operatorname{Im} G_{k(m')l(m')}, & N(N-1)/2 < m \le N(N-1), \\ & m' = m - N(N-1)/2, \\ G_{m''m''}, & N(N-1) < m \le N^2, \\ & m'' = m - N(N-1) \end{cases} \tag{18}$$

из пространства $\mathbf{R}^{N^2}$.

Для построения обратного отображения заметим, что для любого $j \in \{1, \ldots, N(N-1)/2\}$ существует единственное $k \in \{2, \ldots, (N+1)\}$, такое, что $r_k < j \le r_{k+1}$. Обозначим это число $k(j)$. Тогда

$$(k,l)(j) = \Big(k(j), j - r_{k(j)}\Big)\,. \tag{19}$$

Следовательно каждому вектору $\mathbf{g}$ будет отвечать эрмитова матрица $\boldsymbol{G}$ с элементами G_{kl}:

$$G_{kl}\,(\mathbf{g}) = \begin{cases} \mathrm{g}_{j(k,l)} + i\, \mathrm{g}_{j(k,l)+N(N-1)/2}\,, & 1 \le l < k \le N\,, \\ \\ \mathrm{g}_{N(N-1)+k}\,, & k = l,\ 1 \le k \le N\,. \end{cases} \tag{20}$$

Заметим, что преобразования (18), (20) — взаимно обратные. Таким образом, установлено взаимно однозначное соответствие $\boldsymbol{G} \rightleftharpoons \mathbf{g}$. Аналогичные преобразования имеют место и для $\boldsymbol{Z} \rightleftharpoons \mathbf{z}$. Окончательно систему (10) можно переписать в виде

$$\sum_{r=1}^{N^2} \left(\sum_s \mathcal{A}^s_{mr} \right) \mathrm{g}_r = \mathrm{z}_m\,, \qquad m = 1, \ldots, N^2, \tag{21}$$

где

$$\mathcal{A}^s_{mr} = \begin{cases} \operatorname{Re}\Big(B^s_{i(m),k(r)} B^s_{l(r),j(m)} + B^s_{i(m),l(r)} B^s_{k(r),j(m)}\Big)\,, \\ \qquad\qquad 1 \le m \le N(N-1)/2,\ 1 \le r \le N(N-1)/2, \\ \operatorname{Im}\Big(B^s_{i(m),l(r)} B^s_{k(r),j(m)} - B^s_{i(m),k(r)} B^s_{l(r),j(m)}\Big)\,, \\ \qquad\qquad 1 \le m \le N(N-1)/2,\ N(N-1)/2 < r \le N(N-1), \\ \operatorname{Re}\Big(B^s_{i(m),k(r)} B^s_{k(r),j(m)}\Big)\sqrt{2}, \\ \qquad\qquad 1 \le m \le N(N-1)/2,\ N(N-1) < r \le N^2, \\ \operatorname{Re}\Big(B^s_{i(m),k(r)} B^s_{l(r),j(m)} - B^s_{i(m),l(r)} B^s_{k(r),j(m)}\Big)\,, \\ N(N-1)/2 < m \le N(N-1),\ N(N-1)/2 < r \le N(N-1), \\ \operatorname{Im}\Big(B^s_{i(m),k(r)} B^s_{k(r),j(m)}\Big)\sqrt{2}, \\ \qquad\qquad N(N-1)/2 < m \le N(N-1),\ N(N-1) < r \le N^2, \\ B^s_{j(m),k(r)} B^s_{k(r),j(m)}, \\ \qquad\qquad N(N-1) < m \le N^2,\ N(N-1) < r \le N^2 \end{cases} \tag{22}$$

($B^s_{j,k}$ обозначает $j,\ k$-й элемент матрицы $\boldsymbol{B_s}$) или в каноническом виде

$$\mathcal{A}\,\mathbf{g} = \mathbf{z}\,, \tag{23}$$

где $\mathcal{A} = \|\mathcal{A}_{mr}\|$, $\mathcal{A}_{mr} = \sum\limits_s \mathcal{A}^s_{mr}$. Легко видеть, что $\mathcal{A}$ есть симметрическая $N^2 \times N^2$ матрица оператора $\boldsymbol{G} \to \boldsymbol{B_s G B_s}$ в $\mathcal{H}$ в базисе (15) и для её обращения при решении системы (23) уже можно использовать стандартные процедуры.

Из (23) получим оценку $\hat{\mathbf{g}}$:

$$\hat{\mathbf{g}} = \mathcal{A}^{-1}\mathbf{z}\,, \tag{24}$$

где $\mathbf{z}$ соответствует матрице $\boldsymbol{Z} = \sum\limits_s \boldsymbol{u}_s\boldsymbol{u}_s^\dagger - \sum\limits_s \boldsymbol{B}_s$. Легко показать, что оценка $\hat{\mathbf{g}}$ — несмещённая: $E\{\hat{\mathbf{g}}\} = \mathbf{g}$, если $\mathcal{A}^{-1}$ существует. Оценка $\hat{\boldsymbol{G}}$ получается из $\hat{\mathbf{g}}$ преобразованием (20).

Будем характеризовать погрешность оценивания $\boldsymbol{G}$, как и в случае неподвижного источника, средней дисперсией σ_2^2, обусловленной только помехой. После несложных, но громоздких выкладок получаем для изотропной помехи

$$\sigma_2^2 = N^{-2}\varepsilon_0^4 \mathrm{tr}\,\mathcal{A}^{-1}\,. \tag{25}$$

Исследуем зависимость дисперсии σ_2^2 от длины траектории движения L при заданном размере приёмной решётки M для источника, движущегося в свободном пространстве, полагая, что сам источник и его траектория движения параллельны приёмной решётке, а центральная точка траектории находится напротив центра решётки. Заметим, что нулевая длина траектории $L = 0$ соответствует неподвижному источнику; в этом случае $\mathrm{tr}\,\mathcal{A}^{-1} = S^{-1}\left[\mathrm{tr}\,(\boldsymbol{A}^\dagger\boldsymbol{A})^{-1}\right]^2$ и при $S = S_0$ уравнение (25) переходит в (14).

Вначале будем полагать число независимых отсчётов S постоянным, не связанным с длиной траектории. На рис. 2а показаны зависимости $\sigma_2^2(l)/\sigma_0^2$ от безразмерной длины $l = 2L/\lambda$ при различных M (σ_0^2 — дисперсия для приёмной решётки бесконечной длины, использованная здесь в качестве нормировки).

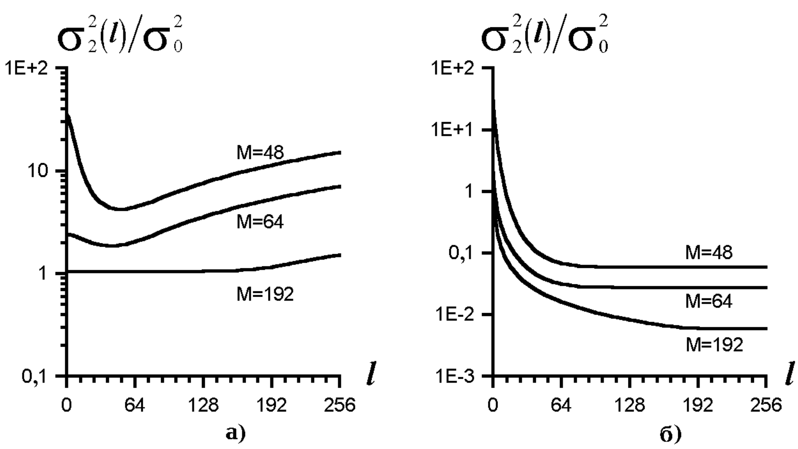


Рис. 2.

Как видно из графиков, при хорошей обусловленности матрицы системы ($M > M_0$) дисперсия монотонно возрастает с увеличением длины траектории (нижний график на рис. 2а). Такое поведение достаточно очевидно, если учесть, что матрица $\boldsymbol{G}$ даже при $l = 0$ достаточно хорошо обусловлена и дисперсия её элементов σ_1^2 только возрастает при смещении центра источника относительно центра приёмной решётки (см. поведение μ_i — рис. 1б). В этом случае, при фиксированном общем времени наблюдения источника выгоднее располагать его неподвижно и симметрично относительно приёмной решётки, чем перемещать относительно неё. Для плохой обусловленности матрицы системы ($M \lesssim M_0$) при $l = 0$ хорошо определяются только некоторые группы элементов матрицы $\boldsymbol{G}$; по мере увеличения l реконструируются с малой дисперсией уже другие группы элементов $\boldsymbol{G}$. Таким образом, вначале наблюдается уменьшение дисперсии σ_2^2 — фактор более полного определения матрицы $\boldsymbol{G}$ является преобладающим (верхние графики на рис. 2а). При дальнейшем увеличении l дисперсия снова начинает возрастать — определяющим является фактор постоянства S, т. е. в этом случае существует некоторая оптимальная длина траектории движения [10].

При априорно заданном движении источника число независимых отсчётов S пропорционально длине траектории и оба фактора: снижение дисперсии из-за увеличения S и изменения дисперсии за счёт различных взаимных расположений приёмной решётки и источника при движении последнего, — присутствуют одновременно. На рис. 2б представлены зависимости $\sigma_2^2(l)/\sigma_0^2$ для этого случая. Как следует из рис. 2б, фактор снижения дисперсии из-за увеличения S является определяющим. Кривые с ростом l выходят на константу при удалении источника за границу дальней зоны приёмной решётки.

3. РЕКОНСТРУКЦИЯ ИЗЛУЧАТЕЛЯ ПРИ НЕИЗВЕСТНОЙ ПОМЕХЕ

При неизвестной ковариационной матрице помехи $\boldsymbol{K}$ к (10) добавляется второе уравнение правдоподобия, которое получается аналогично (6) после дифференцирования (4) по $\boldsymbol{K}^{-1}$:

$$\left(\frac{\partial \ln w}{\partial \boldsymbol{K}^{-1}}\right)^T = \sum_s [-\boldsymbol{K} + \boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger - \boldsymbol{A}_s\boldsymbol{G}\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger + \ldots$$
$$\ldots + \boldsymbol{p}_s\boldsymbol{p}_s^\dagger - \boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{p}_s\boldsymbol{p}_s^\dagger - \boldsymbol{p}_s\boldsymbol{p}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger + \ldots] = 0\,. \qquad (26)$$

После его линеаризации, используя те же предположения и процедуры, что и в предыдущем разделе, получаем с учётом (10) систему из двух уравнений для неизвестных $\boldsymbol{G}$ и $\boldsymbol{K}$:

$$\sum_s (\boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s)\boldsymbol{G}(\boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s) =$$
$$= \sum_s \boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{p}_s\boldsymbol{p}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s - \sum_s \boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s\,, \qquad (27)$$

$$\boldsymbol{K} = \hat{\boldsymbol{T}} - \frac{1}{S}\sum_s \boldsymbol{A}_s\boldsymbol{G}\boldsymbol{A}_s^\dagger\,, \qquad (28)$$

где $\hat{\boldsymbol{T}} = (1/S)\sum_s \boldsymbol{p}_s\boldsymbol{p}_s^\dagger$. После подстановки в (27) оценки $\hat{\boldsymbol{K}}$, следующей из (28), получаем, в отличие от (10), уже нелинейное уравнение относительно $\boldsymbol{G}$. Его можно, однако, трансформировать в линейное, используя следующее свойство (27): это уравнение даёт несмещённую оценку $\boldsymbol{G}$ при замене $\boldsymbol{K}$ на произвольную невырожденную эрмитову матрицу $\boldsymbol{K}_0$ с одновременной заменой последнего слагаемого $\sum_s \boldsymbol{A}_s^\dagger\boldsymbol{K}^{-1}\boldsymbol{A}_s$ в правой части (27) на $\sum_s \boldsymbol{A}_s^\dagger\boldsymbol{K}_0^{-1}\boldsymbol{K}\boldsymbol{K}_0^{-1}\boldsymbol{A}_s$. Подставляя $\hat{\boldsymbol{K}}$ в изменённую правую часть (27), получаем уже линейное уравнение относительно $\boldsymbol{G}$

$$\sum_s \left(\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s - \frac{1}{S}\sum_{s'} \boldsymbol{D}_{ss'}\boldsymbol{G}\boldsymbol{D}_{ss'}^\dagger\right) = \boldsymbol{X}\,, \qquad (29)$$

где

$$\boldsymbol{B}_s = \boldsymbol{A}_s^\dagger\boldsymbol{K}_0^{-1}\boldsymbol{A}_s, \qquad \boldsymbol{D}_{ss'} = \boldsymbol{A}_s^\dagger\boldsymbol{K}_0^{-1}\boldsymbol{A}_{s'},$$
$$\boldsymbol{X} = \sum_s \boldsymbol{A}_s^\dagger\boldsymbol{K}_0^{-1}(\boldsymbol{p}_s\boldsymbol{p}_s^\dagger - \hat{\boldsymbol{T}})\boldsymbol{K}_0^{-1}\boldsymbol{A}_s\,.$$

Решение (29) обладает важным свойством несмещённости при любых $\boldsymbol{K}_0$ и $\boldsymbol{K}$; различие между этими матрицами влияет лишь на дисперсию оценки $\hat{\boldsymbol{G}}$. *

Физический смысл проделанных операций состоит в следующем. При построении уравнения для оценки второго момента распределения случайных источников на излучателе мы должны, во—первых,

*При больших различиях между априорно выбранной $\mathbf{K}_0$ и реальной матрицей ковариации помехи $\mathbf{K}$ дисперсия оценки $\mathbf{G}$ соответственно возрастает.

скомпенсировать составляющую, отвечающую фоновой помехе (второе слагаемое в правой части (27)) и, во—вторых, использовать $\boldsymbol{K}^{-1}$ в матричных окружениях $\boldsymbol{p}_s\boldsymbol{p}_s^\dagger$ и $\boldsymbol{G}$ для "выбеливания" сигнальных векторов. Последняя операция, как известно, позволяет весьма эффективно подавлять сосредоточенные по углу прихода помехи, обеспечивая большой выигрыш при обработке. При более-менее изотропной помехе, однако, адаптация не даёт большого выигрыша и наиболее важную роль играет компенсация фоновой помехи. В этом случае, очевидно, целесообразно использовать детерминированную матрицу $\boldsymbol{K}_0$ подходящей структуры. При дальнейшем анализе будем полагать $\boldsymbol{K}_0 = \mathbf{I}$,** при этом $\boldsymbol{B}_s = \boldsymbol{A}_s^\dagger\boldsymbol{A}_s$, $\boldsymbol{D}_{ss'} = \boldsymbol{A}_s^\dagger\boldsymbol{A}_{s'}$ и $\boldsymbol{X} = \sum_s \boldsymbol{A}_s^\dagger\left(\boldsymbol{p}_s\boldsymbol{p}_s^\dagger - \hat{\boldsymbol{T}}\right)\boldsymbol{A}_s$.

Нетрудно видеть, что с учётом эрмитовости матрицы $\boldsymbol{D}_{ss'}\boldsymbol{D}_{ss'}^\dagger$ процедура решения уравнения (10) легко обобщается и на уравнение (29). Вводя, как и в предыдущем разделе, симметрический оператор $\mathcal{B} \in \mathcal{H}$ вида

$$\mathcal{B}(\boldsymbol{G}) = \sum_s \left(\boldsymbol{B}_s\boldsymbol{G}\boldsymbol{B}_s - \frac{1}{S}\sum_{s'}\boldsymbol{D}_{ss'}\boldsymbol{G}\,\boldsymbol{D}_{ss'}^\dagger\right), \tag{30}$$

представим систему (29) в виде, анологичном (21):

$$\sum_{r=1}^{N^2}\left(\sum_s \mathcal{B}_{mr}^s\right)\mathrm{g}_r = \mathrm{x}_m\,, \qquad m = 1,\ldots,\, N^2, \tag{31}$$

или в виде $\mathcal{B}\mathbf{g} = \mathbf{x}$, где $\mathcal{B} = \|\mathcal{B}_{mr}\|$, $\mathcal{B}_{mr} = \sum_s \mathcal{B}_{mr}^s$, а вектор $\mathbf{x}$ соответствует матрице $\boldsymbol{X}$. Элементы $\mathcal{B}$ получаются из элементов $\mathcal{A}$, если в (22) произвести замену

$$B_{j(m),k(r)}^s B_{k(r),j(m)}^s \longrightarrow B_{j(m),k(r)}^s B_{k(r),j(m)}^s - S^{-1}\sum_{s'} D_{j(m),k(r)}^{s,s'}(D_{k(r),j(m)}^{s,s'})^*\,, \tag{32}$$

где $D_{j,k}^{s,s'}$ обозначают элементы матриц $\boldsymbol{D}_{ss'}$. Аналогично (24) получаем (если существует $\mathcal{B}^{-1}$) несмещённую оценку вектора $\hat{\mathbf{g}}$:

$$\hat{\mathbf{g}} = \mathcal{B}^{-1}\mathbf{x}\,. \tag{33}$$

Оценка $\boldsymbol{G}$ получается из $\hat{\mathbf{g}}$ преобразованием (20).

Погрешность оценивания, как и ранее, можно характеризовать компонентой средней дисперсии σ_3^2, обусловленной только помехой. В случае изотропной помехи $\boldsymbol{K} = \varepsilon_0^2\mathbf{I}$ σ_3^2 находится аналогично (25):

$$\sigma_3^2 = N^{-2}\varepsilon_0^4 \mathrm{tr}\,\mathcal{B}^{-1}\,. \tag{34}$$

Исследуем для движущегося источника поведение (34) в случае, когда S пропорционально длине траектории l. Численный анализ показывает, что существенная вначале разница между σ_3^2 и σ_2^2 с увеличением длины траектории уменьшается, и при больших l $\sigma_3^2 \sim \sigma_2^2$ (отметим, что $\sigma_3^2 > \sigma_2^2$). В качестве иллюстрации на рис. 3 приведены зависимости $\sigma_3^2(l)/\sigma_0^2$ (сплошная кривая) и $\sigma_2^2(l)/\sigma_0^2$ (пунктир) от l при $M = 64$.

Таким образом, даже при сравнительно короткой траектории движения источника средняя дисперсия оценки элементов $\boldsymbol{G}$ достаточно слабо зависит от того, известна ковариационная матрица помехи или нет.

**Отметим, что при этом условии уравнение (29) совпадает с уравнением, полученным в [11] методом наименьших квадратов.

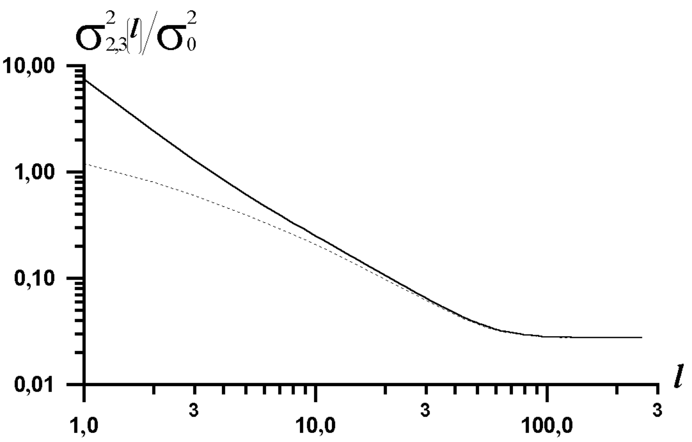


Рис. 3.

ЗАКЛЮЧЕНИЕ

Предложен метод реконструкции матрицы ковариации протяжённого широкополосного движущегося источника по результатам наблюдения его поля (в присутствии фоновой помехи) с помощью антенной решётки, находящейся на малых дистанциях от источника. Получено линейное матричное уравнение для искомой матрицы ковариации, дающее несмещённую оценку, близкую к асимптотически эффективной при малых отношениях сигнал/помеха; предложен способ его сведения к канонической системе линейных уравнений для элементов матрицы ковариации. Исследовано поведение дисперсии оценки в зависимости от длины приёмной решётки, минимальной дистанции до источника и т.п.; тем самым определены условия измерений, при которых возможно решение данной обратной задачи с приемлемой точностью.

В случае произвольной неизвестной фоновой помехи получены линеаризованные уравнения правдоподобия для матриц ковариации и помехи и источника (физически раздельное определение характеристик помехи и источника возможно за счёт движения последнего в предположении стационарности помехи). Поскольку при исключении неизвестной матрицы ковариации помехи уравнение для источника оказывается нелинейным, предложен способ его сведения к линейному, обеспечивающему несмещённую оценку вне зависимости от вида помехи. Способ заключается в дополнительном введении некоторой априорной матрицы, выбеливающей сигнал; причём, если различие между априорной и реальной матрицами помехи незначительно, то полученная оценка будет близка к оптимальной. В общем случае необходимо непосредственное решение полученных в работе линеаризованных уравнений как для сигнала от источника, так и для помехи. Использование для выбеливания полученной оценки ковариации помехи, в случае помехи с сильной угловой анизотропией, по-видимому, может существенно повысить помехозащищённость метода, как это имеет место, например, в адаптивных антенных решётках. Свойства таких решений, однако, могут оказаться нетривиальными; их анализ может явиться предметом дальнейших исследований.

Работа поддержана Российским фондом фундаментальных исследований — проекты № 96–02–19457 и № 96–02–19462.

ЛИТЕРАТУРА

1. Костылев В. И. //Изв. вузов. Радиоэлектроника, 1988. Т. 31. № 7. С. 35.
2. Кремер И. Я., Костылев В. И. //Изв. вузов. Радиоэлектроника, 1985. Т. 28. № 12. С. 63.
3. Кремер И. Я., Костылев В. И. //Изв. вузов. Радиофизика, 1988. Т. 31. № 1. С. 12.
4. Кремер И. Я., Костылев В. И. //Радиотехника, 1984. № 11. С. 71.
5. Fiks I. Sh., Sidorovskaya N. A., Turchin V. I. //J. Phys. IV. Colloque C5, suppl. au J. de Phys. III. V. 4. P. C5-1109.
6. Fiks I. Sh., Sidorovskaya N. A., Turchin V. I. — In: Abstracts of III Intern. Cong. On Air- and Structure-Borne Sound and Vibration. June 13-15, 1994, Monreal, Canada. /Ed. by Malcolm J.Crocker. P. 1701.
7. Постоенко Ю. К., Сидоровская Н. А., Турчин В. И., Угриновский Р. А., Фикс Г. Е., Фикс И. Ш. //Формирование акустических полей в океанических волноводах: Сб. научн. тр. — Н. Новгород: Изд-во ИПФ РАН, 1994. С. 247.
8. Gershman A. B., Turchin V. I., Ugrinovsky R. A. //IEE Elec. Lett., 1992. V. 28. № 18. P. 1677.
9. Фикс И. Ш. //Препринт ИПФ РАН № 402. — Н. Новгород, 1996. – 26 с.
10. Вдовичева Н. К., Фикс И. Ш. //Препринт ИПФ РАН № 434. — Н. Новгород, 1997. – 21 с.
11. Вдовичева Н. К., Фикс И. Ш. //Препринт ИПФ РАН № 446. — Н. Новгород, 1997. – 17 с.
12. Burg J. P., Luenderger D. J., and Wenger D. L. — In: Proc. IEEE, September 1982. V. 70. P. 963.
13. Караваев В. В., Сазонов В. В. Статистическая теория пассивной локации. — М.: Радио и связь, 1987. – 238 с.

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

Поступила в редакцию
10 марта 1998 г.

REMOTE SENSING OF MOVING WIDE-BAND SOURCES

N. K. Vdovicheva, V. I. Turchin, I. Sh. Fiks

An estimation technique focused on the reconstruction of covariance matrix of random elementary sources described an extended noise emitter is considered. The linearized maximum likelihood equations for the covariance matrix estimated are obtained in case of apriori known and unknown noise background. The variance of estimates of matrix elements is investigated.

УДК 621.396.96

ОБНАРУЖЕНИЕ СЛАБЫХ СИГНАЛОВ НА ФОНЕ РАССЕЯННЫХ В НЕОДНОРОДНОЙ СРЕДЕ АКТИВНЫХ ПОМЕХОВЫХ ПОЛЕЙ. I. ГАУССОВО ПРИБЛИЖЕНИЕ

М. А. Островский

В рамках гауссова приближения рассматривается оптимальная процедура обнаружения слабых сигналов на фоне рассеянных активных помеховых полей. Наибольшее внимание уделяется специфическим вопросам обнаружения и, в частности, задаче адаптивного восстановления поля первичной волны на входах многоканального обнаружителя.

ПОСТАНОВКА ЗАДАЧИ

В [1] исследовано влияние среды распространения на статистические характеристики поля активных помех, действующего на линейную (N+1)-элементную антенную решётку (АР). Показано, что, вследствие рассеяния на неоднородностях, пространственный коэффициент корреляции помехи уменьшается, а высшие кумулянтные функции возрастают. В результате денормализации снижается качество подавления помехи в корреляционных автокомпенсаторах и, соответственно, сокращается зона наблюдения РЛС в помеховых условиях. Данное явление наблюдается в ряде частотных диапазонов: ВВЧ — из-за отражений падающей волны от подстилающей поверхности, в УВЧ и S — из-за рассеяния радиоволн в облаках гидрометеоров и оптически ненаблюдаемых объектов (ангел-эхо). Таким образом, поле активных помех, сформированное излучением волн передатчиком помехопостановщика и прошедшее слой неоднородной среды, приобретает в точке приёма, наряду с известными, реверберационные свойства, во многом сходные со свойствами пассивной помехи.

На основе анализа модели помехи при использовании допущения о статистической независимости её временных отсчётов в различных элементах разрешения по времени запаздывания Δt и периодах повторения T в [1] найдено выражение многомерной плотности вероятности (ПВ) пространственно—временных значений помехи $\vec{\eta}$ в импульсном объёме РЛС

$$W_{(N+1)MB}(\vec{\eta}) = \prod_{\nu=0}^{B-1} \prod_{\mu=0}^{M-1} \prod_{i=-N/2}^{N/2} \omega_1 \left\{ \sum_{m=-N/2}^{N/2} \psi(i-m) \int_{-\infty}^{\infty} h(z,m) \times \right.$$
$$\left. \times\, \eta(t-\nu\Delta t-\mu T-z,\, m) dz \right\}, \qquad (1)$$

где B — база сложного зондирующего сигнала РЛС, M — количество импульсов в сигнальном пакете, $\psi(i)$ — импульсная характеристика пространственного обеляющего фильтра, $h(z,m)$ — импульсная характеристика временного фильтра, восстанавливающего первичную волну, $\omega_1\{\cdot\}$ — одномерная ПВ пространственно обелённых и восстановленных отсчётов помехи.

Существо предложенной модели состоит в доказательстве пространственно—временного фильтрующего действия среды распространения на воздействие белого негауссова поля $\xi(t,i)$, являющегося межэлементным градиентом произведения полей диэлектрической проницаемости среды и первичной волны $\eta_0(t,i)$. Безусловно, поле $\xi(t,i)$ является искусственно введённым (модельным), т. к. характеризует лишь градиент реально существующего рассеянного поля $\eta(t,i)$, т. е. ту его малую часть, которая статистически не взаимодействует с аналогичными полями, действующими на иные элементы АР. Однако его введение позволяет, во-первых, исследовать причины возможной денормализации помехи [1] и, во-вторых, относительно просто получить выражение для её многомерной ПВ (1).

Поскольку количество рассеяний, участвующих в образовании указанного градиента, может оказаться недостаточным для выполнения условий центральной предельной теоремы, одномерный закон распределения $\omega_1(\xi)$ в большинстве случаев отличается от гауссова. Последнее не может не сказаться на структуре оптимального обнаружителя сигналов, представляющего собой, в общем случае, достаточно сложное нелинейное инерционное устройство. Вместе с тем, представляет интерес исследование особенностей оптимального обнаружения в условиях рассеяния помеховых радиоволн в более простых, но достаточно легко физически интерпретируемых случаях распределения помех. Одним из них является случай гауссова распределения, исследуемый в данной работе.

ОПТИМАЛЬНОЕ ОБНАРУЖЕНИЕ СЛАБЫХ ДЕТЕРМИНИРОВАННЫХ СИГНАЛОВ

Пусть на вход обнаружителя действует выборка $\vec{\mathbf{y}}$ аддитивной смеси слабого детерминированного сигнала $\vec{\mathbf{S}}$ и активной помехи $\vec{\eta}$, имеющая размерность $(N+1)MB$. Выборка характеризуется логарифмической функцией правдоподобия, получаемой путём смещения логарифма распределения (1) под действием сигнала

$$\ln W_{(N+1)MB}(\vec{\mathbf{y}}-\vec{\mathbf{S}}) = \sum_{\nu=0}^{B-1}\sum_{\mu=0}^{M-1}\sum_{i=-N/2}^{N/2} \ln\omega_1\Bigg\{\sum_{m=-N/2}^{N/2}\psi(i-m)\int\limits_{-\infty}^{\infty} h(z,m)\times$$

$$\times\Big[y(t-\nu\Delta t-\mu T-z,\,m) - S(t-\nu\Delta t-\mu T-z,\,m)\Big]dz\Bigg\}.$$

Если выполняется условие регулярности ПВ, в смысле существования её логарифмических производных, последняя может быть разложена в ряд по степеням сигнала в точке $S=0$

$$\ln W_{(N+1)MB}(\vec{\mathbf{y}}-\vec{\mathbf{S}}) = \ln W_{(N+1)MB}(\vec{\mathbf{y}}) + \sum_{\nu=0}^{B-1}\sum_{\mu=0}^{M-1}\sum_{i=-N/2}^{N/2}\sum_{n=1}^{\infty}\frac{f_n(\vec{\mathbf{y}})}{n!}\times$$

$$\times\Bigg\{\sum_{k=-N/2}^{N/2}\psi(i-k)\int\limits_{-\infty}^{\infty} h(x,k)S(t-\nu\Delta t-\mu T-x,\,k)dx\Bigg\}^n, \tag{2}$$

где $f_n(\vec{\mathbf{y}}) = (-1)^n\dfrac{\partial^n\ln\omega_1[\xi(t-\nu\Delta t-\mu T,\,i)]}{\partial\xi^n(t-\nu\Delta t-\mu T,\,i)}$ — коэффициент разложения, $\xi(t-\nu\Delta t-\mu T,\,i) = \sum\limits_{m=-N/2}^{N/2}\psi(i-m)\int\limits_{-\infty}^{\infty}h(z,m)y(t-\nu\Delta t-\mu T-z,\,m)dz$ — результат пространственно—временной фильтрации принятой выборки.

Хотя, строго говоря, разложение случайной функции в (2) носит формальный характер, т. к. сходимость ряда может обеспечиваться только в среднестатистическом смысле, этот приём довольно часто используется в литературе при асимптотических представлениях функций и отношений правдоподобия [2]. В случае выполнения условия $(N+1)MB \gg 1$ наибольший практический интерес представляет обнаружение слабых (пороговых) сигналов, когда разложение (2) может быть ограничено линейным членом

$$\ln W_{(N+1)MB}(\vec{\mathbf{y}}-\vec{\mathbf{S}}) \approx \ln W_{(N+1)MB}(\vec{\mathbf{y}}) + \\ + \sum_{\nu=0}^{B-1}\sum_{\mu=0}^{M-1}\sum_{i=-N/2}^{N/2} f_1(\vec{\mathbf{y}}) \sum_{k=-N/2}^{N/2} \psi(i-k) \int_{-\infty}^{\infty} h(x,k)S(t-\nu\Delta t-\mu T-x,\,k)dx. \tag{3}$$

Вычитая из (3) логарифм ПВ принятой выборки $\vec{y}$, получим выражение логарифма отношения правдоподобия (ЛОП) и соответствующее решающее правило обнаружения

$$\ln\Lambda \approx \sum_{\nu=0}^{B-1}\sum_{\mu=0}^{M-1}\sum_{i=-N/2}^{N/2}\left[\sum_{k=-N/2}^{N/2}\psi(i-k)\int_{-\infty}^{\infty}h(x,k)S(t-\nu\Delta t-\mu T-x,k)dx\right]\times \\ \times\ f_1\left[\sum_{m=-N/2}^{N/2}\psi(i-m)\int_{-\infty}^{\infty}h(z,m)y(t-\nu\Delta t-\mu T-z,m)dz\right]\geq \ln\Lambda_0\,, \tag{4}$$

где $\ln\Lambda_0$ — порог обнаружения.

Обнаружитель (рис. 1) представляет собой пару многоканальных (по числу элементов АР) каскадов восстановления поля первичной волны (ВПВ) и пространственных обеляющих фильтров (ПОФ), разделённых многоканальным нелинейным преобразователем (НП), диаграммо–образующую схему (ДОС), выполняющую функцию пространственного накопления сигналов, внутрипериодный согласованный фильтр (СФ), когерентный накопитель сигнальной пачки (КН) и пороговое устройство (ПУ).

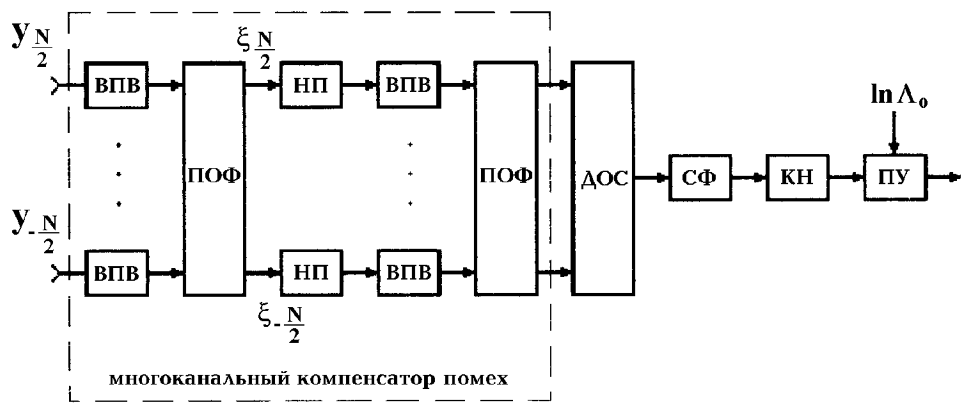


Рис. 1.

Многоканальная часть схемы на входах ДОС реализует функцию компенсации негауссовых активных помех. Существенным отличием её от известной для гауссовых нерассеянных помех [3] является наличие в каждом канале приёма пары комплексно–сопряжённых восстанавливающих первичную волну фильтров и безынерционных нелинейных преобразователей с характеристикой $f_1(\xi) = -\dfrac{\partial \ln\omega_1(\xi)}{\partial\xi}$. Пара комплексно–сопряжённых ВПВ и ПОФ обязана своим происхождением корреляционному характеру записи ЛОП (4), при которой требуется точное согласование характеристик ДОС сигнальному полю на выходах первого ПОФ. На практике же обычно стремятся обеспечить согласование характеристики ДОС с сигнальным полем на входе обнаружителя [3]. При этом в схему обнаружителя на выходах НП приходится добавлять точно такие же каскады ВПВ-ПОФ с зеркально–симметричными импульсными характеристиками.

В случае выполнения условий центральной предельной теоремы характеристика НП трансформируется в линейную $f_1(\xi) = \xi/\langle\xi^2\rangle$, пара обеляющих фильтров образует, так называемый, "переобеляющий"пространственный фильтр [3], а передаточная характеристика ВПВ преобразуется в квадратичную $H(\omega, i)H^*(\omega, i)$. Таким образом, учёт рассеивающих свойств среды приводит к необходимости восстановления поля первичной волны как в случае гауссовой, так и негауссовой помехи. Исследуя первую задачу, остановимся более детально на особенностях восстановления случайного первичного поля, т. е. поля, создаваемого постановщиком активных шумовых помех в идеальной среде. Все остальные элементы обнаружителя (рис. 1) достаточно подробного обсуждались в работе [4].

АДАПТИВНОЕ ВОССТАНОВЛЕНИЕ ПОЛЯ ПЕРВИЧНОЙ ВОЛНЫ

Проблеме восстановления радиополей уделяется большое внимание в радиосвязи, гидро- и метеолокации, в системах дистанционного зондирования протяжённых объектов, при воспроизведении радиолокационных и радиояркостных изображений, в задачах природно—ресурсных исследований Земли и планет [5—7]. В большинстве известных работ эта проблема раскрывается либо в предположении об априорной известности порождающего поля, если требуется восстановить характеристики канала рассеяния [5, 6], либо при использовании различных модельных описаний флуктуаций среды для восстановления первичного поля [7]. В активной радиолокации, в условиях действия активных шумовых помех задача восстановления первичного поля существенно усложняется из-за неизвестности статистик как порождающего поля, так и поля диэлектрической проницаемости среды распространения. Поэтому решаемая в работе задача требует самостоятельного исследования.

Примем за начало координат положение первичного источника, удалённого от центрального элемента АР на величину r_0 и расположенного под углом Θ_0 по отношению к её нормали. Пусть ансамбль неоднородностей образован совокупностью L независимых точечных рассеивателей с диэлектрической проницаемостью ε_l, ориентацией Θ_l и разностью длин путей распространения первичной и рассеянной l-й неоднородностью волн Δr_l. Тогда, в соответствии с теорией однократного рассеяния, при условии расположения антенны РЛС во Фраунгоферовой зоне, поле падающей на i-й элемент АР волны $y(t, i) = \eta(t, i)$ представляет весовую сумму первичных волн [1]

$$\eta(t, i) = \eta_0(t, i) + \sum_{l=1}^{L} \varepsilon_l \eta_0 \left(t - \frac{\Delta r_l}{c},\, i\right) \exp\left[j \cdot 2\pi \frac{id}{\lambda}(\sin\Theta_l - \sin\Theta_0)\right], \tag{5}$$

где d — шаг линейной АР, λ — длина волны РЛС.

Из (5) видно, что рассеянное поле $\eta(t, i)$ связано с первичным линейным оператором, который, в принципе, может быть обращён. Действительно, переходя к частотному изображению (5) и выражая спектр первичного поля через спектр рассеянного, а затем снова переходя в область оригиналов, получим

$$\eta_0(t, i) = \int_{-\infty}^{\infty} h(z, i)\eta(t - z,\, i)dz, \tag{6}$$

где $h(z, i) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} \dfrac{\exp(j\omega z)d\omega}{1 + \sum_{l=1}^{L} \beta_l \exp(-j\omega\tau_l)}$ — импульсная характеристика i-го восстанавливающего фильтра, $\beta_l = \varepsilon_l \exp\left[j \cdot 2\pi \dfrac{id}{\lambda}(\sin\Theta_l - \sin\Theta_0)\right]$ — амплитудно—фазовый l-й параметр рассеяния, $\tau_l = \dfrac{\Delta r_l}{c}$ — временной l-й параметр рассеяния.

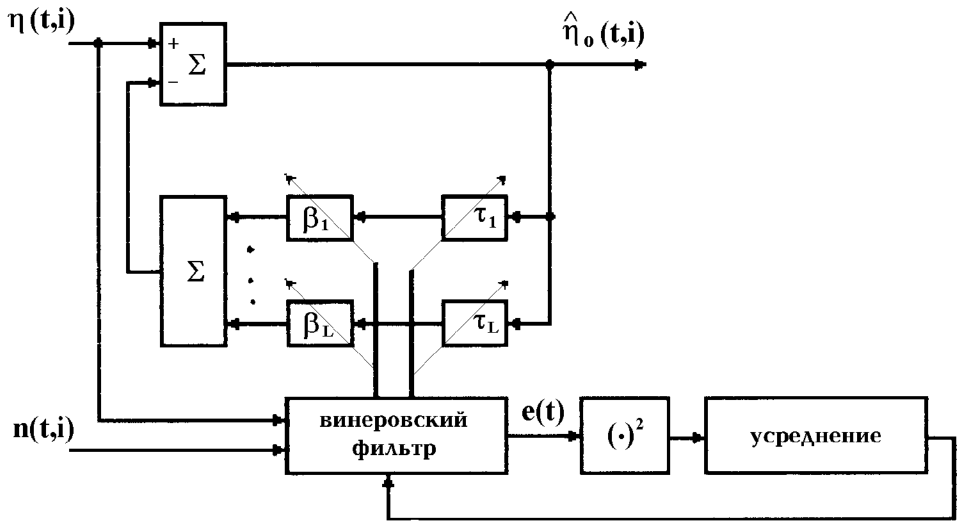


Рис. 2.

Преобразование (6) описывает процесс идеального восстановления поля первичной волны путём поканальной временной фильтрации каждого из пространственных отсчётов помехи. Восстанавливающий фильтр представляет собой рециркулятор (рис. 2), обратная связь которого содержит L суммируемых каналов с элементами задержки на время τ_l и комплексными умножителями на коэффициент β_l (рис. 2). Как несложно убедиться, идеальный восстанавливающий фильтр компенсирует все составляющие помехи, действующие с любых иных направлений кроме Θ_0. Для этого необходима лишь достаточно полная априорная информация об амплитудно—фазовом $\vec{\beta} = \|\beta_1, \ldots, \beta_L\|$ и временном $\vec{\tau} = \|\tau_1, \ldots, \tau_L\|$ параметрах рассеянного поля. В реальных условиях такая информация не доступна наблюдателю, поэтому задачу восстановления приходится решать в классе адаптивных систем с использованием многомерных текущих оценок векторов $\hat{\beta}$ и $\hat{\tau}$. Естественно, что при этом операция восстановления осуществляется не совсем идеально, а вместо первичного поля $\eta_0(t,i)$ на выходе фильтра (рис. 2) формируется его оценка $\hat{\eta}_0(t,i)$.

Адаптивное управление параметрами рециркулятора осуществляется путём минимизации средней мощности сигнала рассогласования $e(t)$. Для его формирования может использоваться, например, хорошо развитый в литературе винеровский подход [2]. В классической постановке, последний сводится к линейной фильтрации суммы рассеянного $\eta(t,i)$ и статистически несвязанного с ним шумового $n(t,i)$ полей, а также к образованию разности между реакцией фильтра и некоторым желаемым полем $\gamma(t,i)$

$$e(t,i) = \int\limits_{-\infty}^{\infty} K(z,i)[\eta(t-z,i) + n(t-z,i)]dz - \gamma(t,i), \tag{7}$$

где $K(z,i)$ — импульсная характеристика винеровского фильтра i-го канала.

При выборе желаемого поля $\gamma(t,i) \equiv n(t,i)$ и совпадении статистических характеристик шумового $n(t,i)$ и первичного $\eta_0(t,i)$ полей, минимизация дисперсии сигнала рассогласования приводит к оптимальной передаточной характеристике винеровского фильтра

$$K_0(\omega,i) = \left[1 + \frac{S_\eta(\omega,i) + S^*_\eta(\omega,i)}{2S_0(\omega,i)}\right]^{-1}, \tag{8}$$

где $S_0(\omega,i)$ — спектральная плотность мощности первичного поля. Находя из (5) спектр мощности рассеянного поля

$$S_\eta(\omega, i) = S_0(\omega, i) \left[1 + \sum_{l=1}^{L} \beta_l \exp(-j\omega\tau_l)\right]^2,$$

убеждаясь в его действительности и подставляя в (8), получим

$$K_0(\omega, i) = \left\{1 + \left[1 + \sum_{l=1}^{L} \beta_l \exp(-j\omega\tau_l)\right]^2\right\}^{-1} =$$
$$= H^2(\omega, i, \vec{\beta}, \vec{\tau})[1 + H^2(\omega, i, \vec{\beta}, \vec{\tau})]^{-1}. \qquad (9)$$

Из (9) следует, что оптимальный винеровский фильтр представляет замкнутую систему, в обратной связи которой размещена пара комплексно сопряжённых ВПВ. При априорной неопределённости параметров рассеянной волны винеровский фильтр так же, как и ВПВ должен быть адаптируемым, т. е. путём минимизации мощности сигнала рассогласования на выходе приближать свою передаточную характеристику $K(\omega, i, \hat{\beta}, \hat{\tau})$ к оптимальной $K_0(\omega, i, \vec{\beta}, \vec{\tau})$. Сформированные при адаптации оценки $\hat{\beta}$ и $\hat{\tau}$ могут использоваться для управления параметрами основного восстанавливающего фильтра (рис. 2).

Указывая на желательную близость временных корреляционных характеристик первичного и шумового полей, отметим, что, вообще говоря, форма первой для нерассеянного поля точно известна в точке приёма. Действительно, первичное поле не зависит от неизвестных параметров среды и, в силу своей относительной широкополосности, определяется частотными характеристиками приёмных трактов РЛС. Это справедливо как для прицельных шумовых помех, ширина спектра которых в 3–5 раз превышает полосу пропускания приёмника, так и для заградительных помех, для которых данное превышение составляет десятки, сотни и более раз [8].

СИНТЕЗ УПРАВЛЕНИЙ ПАРАМЕТРАМИ АДАПТИВНОГО ВОССТАНАВЛИВАЮЩЕГО ФИЛЬТРА

Определим требуемые управления параметрами восстанавливающего фильтра, для чего, применительно к обобщённому векторному параметру $\hat{\mathbf{a}}$, используем одну из простейших модификаций алгоритма стохастической аппроксимации [9]

$$\frac{d\hat{\mathbf{a}}}{dt} = -\Gamma\nabla\{e^2(t, i)\} = -2\Gamma\, e(t, i)\nabla\{e(t, i)\}, \qquad (10)$$

где Γ — постоянный шаг адаптации, $\nabla = \left\| \frac{\partial}{\partial \hat{\mathrm{a}}_l}, \ldots, \frac{\partial}{\partial \hat{\mathrm{a}}_L} \right\|$ — градиент ошибки восстановления по оценке параметра $\hat{\mathbf{a}}$.

Будем считать, что поля $n(t, i)$ и $e(t, i)$ являются узкополосными и обладают свойством дифференцируемости по времени [10]. Для использования в (10) сигнала $e(t, i)$ подставим в (7) формулу передаточной функции винеровского фильтра, получаемую из (9) путём замены параметров рассеяния на их оценки. Пренебрегая в этом промежуточном выражении членами второго порядка малости $0(|\hat{\beta}_l|^2)$, получим приближённую рекуррентную запись

$$e(t, i) \approx \frac{1}{2}[\eta(t, i) - n(t, i)] - \frac{1}{2}\sum_{l=1}^{L} \hat{\beta}_l[n(t - \hat{\tau}_l, i) + e(t - \hat{\tau}_l, i)] -$$
$$- \frac{1}{2}\sum_{l=1}^{L} \hat{\beta}_l^*[n(t + \hat{\tau}_l, i) + e(t + \hat{\tau}_l, i)]. \qquad (11)$$

Уравнения (10) и (11) пригодны для синтеза оптимальных управлений восстанавливающим фильтром, который ниже проведём отдельно для неизвестных амплитудно–фазовых и временных параметров рассеяния.

1. Предположим, что составляющие вектора запаздываний полностью известны: $\vec{a} = \vec{\beta}$. Находя из (11) производную сигнала $e(t,i)$ по p-й компоненте оценки амплитудно–фазового вектора $\hat{\beta}_p$ и подставляя её в (10), получим

$$\frac{d\hat{\beta}_p}{dt} = \Gamma_1 e(t,i)\left[n(t-\tau_p,\, i) + e(t-\tau_p,\, i) + \hat{\beta}_p \frac{de(t-\tau_p,\, i)}{d\hat{\beta}_p}\right], \quad p = 1,\ldots,L\,.$$

Третье слагаемое этого выражения характеризует связь оценки $\hat{\beta}_p$, сформированной в момент времени t, от сигнала в более поздний момент. Поскольку это противоречит принципу причинности, положим его равным нулю. В результате получим необходимое уравнение оптимального управления амплитудно–фазовым параметром восстанавливающего фильтра

$$\frac{d\hat{\beta}_p}{dt} = \Gamma_1 e(t,i)[n(t-\tau_p,\, i) + e(t-\tau_p,\, i)], \quad p = 1,\ldots,L\,. \tag{12}$$

2. Пусть теперь все составляющие амплитудно–фазового вектора рассеяния априорно известны $\vec{a} = \vec{\tau}$. Производя соответствующие дифференцирования (11) по оценке p-й компоненты вектора запаздывания и подставляя его в (10), получим

$$\frac{d\hat{\tau}_p}{dt} = \Gamma_2 e(t,i)\Big\{\beta_p^*\left[\dot{n}(t+\hat{\tau}_p,\, i) + \dot{e}(t+\hat{\tau}_p,\, i)\right] - \beta_p\left[\dot{n}(t-\hat{\tau}_p,\, i) + \dot{e}(t-\hat{\tau}_p,\, i)\right]\Big\}, \tag{13}$$

$$p = 1,\ldots,L\,,$$

где $\dot{e}(t)$ — производная сигнала по текущему времени.

Как видно из (11), текущее значение сигнала $e(t,i)$ зависит не только от задержанных, но и упреждённых по времени своих значений. В результате этого управление (13) оказывается физически не реализуемым. Для сохранения физической реализуемости (13) формально пренебрежём в нём упреждёнными составляющими $\dot{n}(t+\hat{\tau}_p,\, i)$ и $\dot{e}(t+\hat{\tau}_p,\, i)$. В результате получим

$$\frac{d\hat{\tau}_p}{dt} = -\Gamma_2 e(t,i)\beta_p\left[\dot{n}(t-\hat{\tau}_p,\, i) + \dot{e}(t-\hat{\tau}_p,\, i)\right], \quad p = 1,\ldots,L\,. \tag{14}$$

В случае одновременной неопределённости векторов параметров рассеяния $\vec{\tau}$ и $\vec{\beta}$ истинное значение τ_p в (12) заменяется на оценку $\hat{\tau}_p$, а β_p в (14) — на аналогичную оценку $\hat{\beta}_p$. При этом, структурная схема совместного оптимального управления p-ми компонентами параметров восстанавливающего фильтра, как итог проведённого синтеза, принимает вид, показанный на рис. 3.

ТОЧНОСТЬ ОЦЕНОК ПАРАМЕТРОВ РАССЕЯНИЯ

С точки зрения теории оценивания адаптивный восстанавливающий фильтр является оптимальным следящим измерителем параметров рассеяния. Его оптимальность понимается как максимальная (в байесовском смысле) близость установившихся управлений фильтра к истинным значениям этих параметров. Поэтому целесообразно рассматривать определение точностных характеристик ВПВ как классическую задачу теории оценивания.

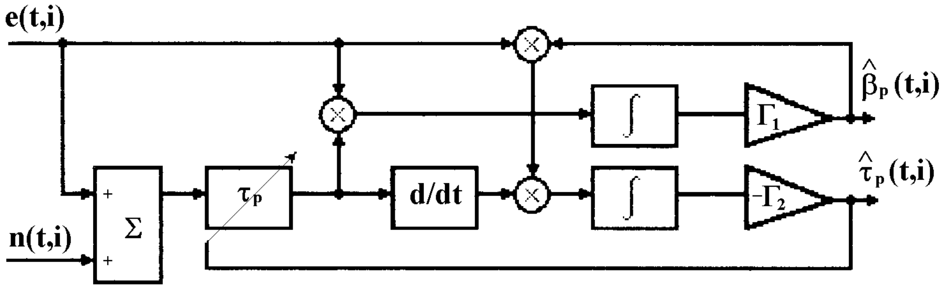


Рис. 3.

Пусть на вход оптимального измерителя в i-ом канале приёма последовательно во времени с интервалом статистической независимости Δt поступает k-мерная выборка $\vec{\eta} = \|\eta_1, \ldots, \eta_k\|$, представляющая собой сумму рассеянной части поля $\vec{\tilde{\eta}}$ и первичной волны $\vec{\eta}_0$, где $\eta_\nu = \eta(\nu\Delta t, i)$. Так как рассеянная часть поля зависит от обобщённого параметра рассеяния $\vec{\mathrm{a}} = \|\mathrm{a}_1, \ldots, \mathrm{a}_L\|$, она в данном случае выступает в роли полезного сигнала. В свою очередь, первичная волна, независящая от этих параметров, выполняет функции помехи. Выборка характеризуется временной функцией правдоподобия (в отличие от пространственно—временной (2))

$$P_k(\vec{\mathrm{h}}/\vec{\mathrm{a}}) = P_k\Big((\eta_1 - \tilde{\eta}_1), \ldots, (\eta_k - \tilde{\eta}_k)\Big) = \prod_{\nu=1}^{k} P_1(\eta_\nu - \tilde{\eta}_\nu),$$

где $P_1(\cdot)$ — одномерная ПВ ν-го временного отсчёта принимаемой выборки, $k \le MB$ — объём поступающей выборки, а нижняя граница оценки её параметров даётся неравенством Рао—Краммера

$$\sigma_a^2 = \frac{1}{\left\langle \left[\nabla \sum\limits_{\nu=1}^{k} \ln P_1(\eta_\nu - \tilde{\eta}_\nu) \right]^2 \right\rangle}\,. \tag{15}$$

Полагая статистическую независимость отдельных составляющих вектора параметров, найдём из (15) максимальную дисперсию ошибок амплитудно—фазового и временного параметров рассеянного поля.

1. Пусть $\vec{\mathrm{a}} = \vec{\beta}$. Поскольку для гауссовой ПВ помехи (5) имеет место соотношение

$$\frac{\partial \ln P_1(\eta_\nu - \tilde{\eta}_\nu)}{\partial \vec{\beta}} = \frac{1}{\langle \eta_0^2 \rangle}\, \eta_0(\nu\Delta t)\, \eta_0^*(\nu\Delta t - \tau)\,,$$

максимальная дисперсия оценки параметра равна

$$\left(\sigma_\tau^2\right)_{\max} = \frac{\langle \eta_0^2 \rangle^2}{\sum\limits_{\nu_1=1}^{k} \sum\limits_{\nu_2=1}^{k} \langle \eta_0(\nu_1\Delta t)\, \eta_0^*(\nu_2\Delta t)\, \eta_0^*(\nu_1\Delta t - \tau)\, \eta_0(\nu_2\Delta t - \tau) \rangle} =$$

$$= \frac{\langle \eta_0^2 \rangle^2}{\sum\limits_{\nu_1=1}^{k} \sum\limits_{\nu_2=1}^{k} \{ R_0[(\nu_1 - \nu_2)\Delta t]\, R_0[(\nu_2 - \nu_1)\Delta t] + R_0(\tau)\, R_0(-\tau) \}}\,.$$

Учитывая чётность временной корреляционной функции первичного поля $R_0(\tau) = \langle \eta_0^2 \rangle \rho_0(\tau)$, а также абсолютную некоррелированность его отсчётов при $\nu_1 \ne \nu_2$, имеем

$$\left(\sigma_\tau^2\right)_{\max} = \frac{1}{k[1 + k\rho_0^2(\tau)]}\,. \tag{16}$$

Анализируя (16), отметим, что задача восстановления первичной волны имеет практический смысл, когда интервал запаздывания рассеянной части поля τ пренебрежимо мал по сравнению с интервалом разрешения по времени запаздывания Δt. В противном случае рассеянная часть поля обладает слабой временной корреляцией с первичным полем и эффективная компенсация первой попросту невозможна. При этом справедливо приближённое соотношение $k\rho_0^2(\tau) \approx k \gg 1$, преобразующее (16) в зависимость

$$\left(\sigma_\tau^2\right)_{\max} \approx \frac{1}{k^2}\,. \tag{17}$$

2. Пусть $\vec{a} = \vec{\tau}$. Тогда максимальное значение дисперсии ошибки оценивания параметра запаздывания равно

$$\left(\sigma_\tau^2\right)_{\max} = \frac{\langle \eta_0^2\rangle^2}{k^2\langle\beta\beta^*\rangle\,\langle \eta_0(\nu\Delta t)\,\eta_0^*(\nu\Delta t)\,\dot\eta_0(\nu\Delta t-\tau)\,\dot\eta_0^*(\nu\Delta t-\tau)\rangle}\,.$$

Учитывая, что $\langle\beta\beta^*\rangle = \langle\varepsilon^2\rangle$, а для гауссовых узкополосных процессов справедливы соотношения [10]

$$\langle \dot\eta_0(\nu\Delta t-\tau)\,\dot\eta_0^*(\nu\Delta t-\tau)\rangle = -R_0''(0) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\omega^2 S_0(\omega)\,d\omega\,,$$

$$\langle \eta_0(\nu\Delta t)\,\dot\eta_0^*(\nu\Delta t-\tau)\rangle = \langle \eta_0^*(\nu\Delta t)\,\dot\eta_0(\nu\Delta t-\tau)\rangle = \frac{dR_0(\tau)}{d\tau}\,,$$

перепишем последнее выражение в виде

$$\left(\sigma_\tau^2\right)_{\max} = \frac{1}{k^2\langle\varepsilon^2\rangle\left\{\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\omega^2 S_0^0(\omega)\,d\omega + [\rho_0'(\tau)]^2\right\}}\,,$$

где $S_0^0(\omega) = \int\limits_{-\infty}^{\infty}\rho_0(\tau)\mathrm{e}^{-j\omega\tau}d\tau$ — нормированная спектральная плотность мощности первичного поля. Оценка составляющих знаменателя последнего выражения, например, для колокольной формы коэффициента временной корреляции $\rho_0(\tau) = \exp\left[-3{,}56\left(\frac{\tau}{\Delta t}\right)^2\right]$, показывает, что в практически интересном случае $\left|\frac{\tau}{\Delta t}\right| \ll 1$ величина второго слагаемого $[\rho_0'(\tau)]^2$ пренебрежимо мала по сравнению с первой. При этом последнее выражение приобретает окончательный вид

$$\left(\sigma_\tau^2\right)_{\max} \approx \frac{1}{k^2\langle\varepsilon^2\rangle\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\omega^2 S_0^0(\omega)\,d\omega}\,. \tag{18}$$

КАЧЕСТВО ВОССТАНОВЛЕНИЯ ПЕРВИЧНОЙ ВОЛНЫ

Для оценки качественных показателей восстановления первичной волны введём ошибку

$$E(t,i) = \hat\eta_0(t,i) - \eta_0(t,i), \tag{19}$$

где $\hat{\eta}_0(t,i) = \int\limits_{-\infty}^{\infty} \hat{h}(z,i,\hat{\beta},\hat{\tau})\,\eta(t-z,\,i)\,dz$ — выходное поле адаптивного восстанавливающего фильтра порядка L_1 с импульсной характеристикой $\hat{h}$, $\eta_0(t,i) = \int\limits_{-\infty}^{\infty} h(z,i,\vec{\beta},\vec{\tau})\eta(t-z,\,i)\,dz$ — выходное поле идеального восстанавливающего фильтра порядка L с характеристикой (6).

Входящая в выражение (19) импульсная характеристика $\hat{h}(z,i,\hat{\beta},\hat{\tau})$ даже в установившемся состоянии является случайной функцией, поскольку зависит не от истинных значений параметров рассеяния $\vec{\beta}$ и $\vec{\tau}$, а от их случайных оценок. Учитывая это, вычислим дисперсию ошибки восстановления

$$\begin{aligned}
\langle E^2\rangle &= \iint\limits_{-\infty}^{\infty} \left\langle \hat{h}(z_1,i,\hat{\beta},\hat{\tau})\hat{h}^*(z_2,i,\hat{\beta},\hat{\tau})\eta(t-z_1,\,i)\eta^*(t-z_2,\,i)\right\rangle dz_1dz_2 - \\
&- \iint\limits_{-\infty}^{\infty} \left\langle \hat{h}(z_1,i,\hat{\beta},\hat{\tau})\eta(t-z_1,\,i)\eta^*(t-z_2,\,i)\right\rangle h^*(z_2,i,\vec{\beta},\vec{\tau})\,dz_1dz_2 - \\
&- \iint\limits_{-\infty}^{\infty} h(z_1,i,\vec{\beta},\vec{\tau}) \left\langle \hat{h}^*(z_2,i,\hat{\beta},\hat{\tau})\eta(t-z_1,\,i)\eta^*(t-z_2,\,i)\right\rangle dz_1dz_2 + \\
&+ \iint\limits_{-\infty}^{\infty} h(z_1,i,\vec{\beta},\vec{\tau})h^*(z_2,i,\vec{\beta},\vec{\tau}) \left\langle \eta(t-z_1,\,i)\eta^*(t-z_2,\,i)\right\rangle dz_1dz_2\,.
\end{aligned}$$

Вследствие значительной инерционности управлений адаптивного ВПВ, его импульсная характеристика формируется не мгновенно, а спустя некоторое время, существенно превышающее интервал взаимной статистической зависимости действующего поля $\vec{\eta}$ ($k \gg 1$). При этом флуктуации случайной импульсной характеристики $\hat{h}$ практически перестают зависеть от флуктуаций поля $\vec{\eta}$, а последнее выражение в результате приобретает вид

$$\begin{aligned}
\langle E^2\rangle &= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_\eta(\omega)\Big[\left\langle \hat{H}(\omega,i,\hat{\beta},\hat{\tau})\hat{H}^*(-\omega,i,\hat{\beta},\hat{\tau})\right\rangle - \\
&- \left\langle \hat{H}(\omega,i,\hat{\beta},\hat{\tau})\right\rangle H^*(-\omega,i,\vec{\beta},\vec{\tau}) - H(\omega,i,\vec{\beta},\vec{\tau}) \left\langle \hat{H}^*(-\omega,i,\hat{\beta},\hat{\tau})\right\rangle + \\
&+ H(\omega,i,\vec{\beta},\vec{\tau})H^*(-\omega,i,\vec{\beta},\vec{\tau})\Big]d\omega\,.
\end{aligned} \tag{20}$$

Как следует из (20), дисперсия ошибки восстановления зависит от первых двух моментов распределения передаточной функции адаптивного ВПВ. Для их нахождения будем считать ошибки измерения параметров рассеяния $\delta\beta_p = \hat{\beta}_p - \beta_p$, $\delta\tau_p = \hat{\tau}_p - \tau_p$ несмещёнными и малыми. Разлагая передаточную функцию $\hat{H}(\omega,i,\hat{\beta},\hat{\tau})$ в степенной ряд по этим ошибкам относительно значения $\hat{H}(\omega,i,\vec{\beta},\vec{\tau})$ и ограничивая разложение линейным членом

$$\hat{H}(\omega,i,\hat{\beta},\hat{\tau}) = \left[1 + \sum_{l=1}^{L_1}\hat{\beta}_l \mathrm{e}^{-j\omega\hat{\tau}_l}\right]^{-1} \approx \hat{H}(\omega,i,\vec{\beta},\vec{\tau})\Big[1-$$

$$- \hat{H}^*(-\omega,i,\vec{\beta},\vec{\tau})\sum_{l=1}^{L_1}(\delta\beta_l)\mathrm{e}^{-j\omega\tau_l} + j\omega\hat{H}^*(-\omega,i,\vec{\beta},\vec{\tau})\sum_{l=1}^{L}(\delta\tau_l)\beta_l\mathrm{e}^{-j\omega\tau_l}\Big],$$

в результате соответствующих усреднений получим

$$\begin{cases} \langle \hat{H}(\omega, i, \hat{\beta}, \hat{\tau}) \rangle = \hat{H}(\omega, i, \vec{\beta}, \vec{\tau}), \\ \langle \hat{H}^*(-\omega, i, \hat{\beta}, \hat{\tau}) \rangle = \hat{H}^*(-\omega, i, \vec{\beta}, \vec{\tau}), \\ \left\langle \hat{H}(\omega, i, \hat{\beta}, \hat{\tau}) \hat{H}^*(-\omega, i, \hat{\beta}, \hat{\tau}) \right\rangle = \\ = \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau})[1 + \sigma_\beta^2 L_1 \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau}) + \omega^2 \langle \varepsilon^2 \rangle L_1 \sigma_\tau^2 \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau})], \end{cases} \tag{21}$$

где $\hat{H}(\omega, i, \vec{\beta}, \vec{\tau}) = \left[1 + \sum\limits_{l=1}^{L_1} \beta_l \exp(-j\omega\tau_l) \right]^{-1}$ — передаточная характеристика ВПВ, отличающаяся от идеальной (6) только своим порядком $L_1 \le L$.

Если теперь подставить выражения для дисперсий оценок (17) и (18) в (21), а затем полученное выражение использовать в (20), окончательно имеем

$$\begin{aligned} \langle E^2 \rangle &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} S_\eta(\omega) \left\{ \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau}) \left[1 + \frac{L_1}{k^2} \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau}) + \right.\right. \\ &+ \left. \frac{L_1}{k^2} \frac{\omega^2 \hat{H}^2(\omega, i, \vec{\beta}, \vec{\tau})}{\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \omega^2 S_0^0(\omega) d\omega} \right] - \hat{H}(\omega, i, \vec{\beta}, \vec{\tau}) H^*(-\omega, i, \vec{\beta}, \vec{\tau}) - \\ &- \left. H(\omega, i, \vec{\beta}, \vec{\tau}) \hat{H}^*(-\omega, i, \vec{\beta}, \vec{\tau}) + H^2(\omega, i, \vec{\beta}, \vec{\tau}) \right\} d\omega \,. \end{aligned} \tag{22}$$

Для анализа (22) рассмотрим несколько характерных ситуаций.

1. Предположим, что среда не содержит неоднородностей ($L = = L_1 = 0$). При этом передаточные характеристики как идеального H, так и реального $\hat{H}$ ВПВ тождественно обращаются в единицу. Тогда из (22) получаем совершенно очевидный результат

$$\langle E^2 \rangle = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} S_\eta(\omega) \times 0 d\omega = 0,$$

подтверждающий, что прошедшая однородную среду и падающая на АР волна совпадает с первичной.

2. Пусть среда неоднородна ($L > 0$), а восстановление первичной волны в РЛС не производится ($L_1 = 0$). Тогда $\hat{H} \equiv 1$, а из (22) получаем

$$\begin{aligned} \langle E^2 \rangle &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} S_\eta(\omega) [1 - H^*(-\omega, i, \vec{\beta}, \vec{\tau}) - H(\omega, i, \vec{\beta}, \vec{\tau}) + H^2(\omega, i, \vec{\beta}, \vec{\tau})] d\omega = \\ &= \langle \eta^2 \rangle + \langle \eta_0^2 \rangle - \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} S_0(\omega) \left[\frac{1}{H(\omega, i, \vec{\beta}, \vec{\tau})} + \frac{1}{H^*(-\omega, i, \vec{\beta}, \vec{\tau})} \right] d\omega = \\ &= \langle \eta^2 \rangle - \langle \eta_0^2 \rangle - \sum_{l=1}^{L} \beta_l R_0(-\tau_l) - \sum_{l=1}^{L} \beta_l^* R_0(\tau_l) = \end{aligned}$$

$$= \langle \eta^2 \rangle - \langle \eta_0^2 \rangle - 2\langle \eta_0^2 \rangle \sum_{l=1}^{L} \rho_0(\tau_l)\varepsilon_l \cos\varphi_l,$$

где $\varphi_l = 2\pi \dfrac{id}{\lambda}(\sin\Theta_l - \sin\Theta_0)$ — начальная фаза l-го рассеяния. Если неоднородности образуют достаточно большой ансамбль ($L \gg 1$), а начальные фазы элементарных рассеяний распределены равномерно в симметричных пределах, то величина суммы в последнем выражении мало отлична от нуля. При этом дисперсия ошибки восстановления приближённо равна разности дисперсий рассеянной и первичной волны:

$$\langle E^2 \rangle \Big|_{\substack{L_1=0,\\ L\neq 0}} \approx \langle \eta^2 \rangle - \langle \eta_0^2 \rangle . \tag{23}$$

Известны количественные оценки данной дисперсии, которые в S-диапазоне по экспериментальным данным [1] не превышают уровня 20 дБ по отношению к мощности первичной волны.

3. Пусть среда неоднородна ($L > 0$), а порядок ВПВ совпадает с числом неоднородностей ($L = L_1$, $\hat{H} = H$). Это вовсе не означает, что порядок ВПВ возрастёт до нескольких десятков, сотен и даже тысяч, поскольку в исследуемой задаче практическое значение имеют лишь немногочисленные и наиболее интенсивные рассеяния. В этом случае из (22) получим

$$\langle E^2 \rangle \Big|_{L=L_1\neq 0} = \frac{L}{k^2} \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) H^2(\omega, i, \vec{\beta}, \vec{\tau}) \left[1 + \frac{\omega^2}{\frac{1}{2\pi} \int_{-\infty}^{\infty} \Omega^2 S_0^0(\Omega) d\Omega} \right] d\omega .$$

Входящий в последнее выражение квадрат передаточной характеристики ВПВ H^2 является суммой единичной и быстро флуктуирующей, по сравнению с $S_0(\omega)$, функцией частоты. Его флуктуации увеличиваются с увеличением числа L. Отсюда ясно, что вклад частотных флуктуаций этой составляющей в величину интеграла несуществен и последнее выражение можно переписать

$$\langle E^2 \rangle \Big|_{L=L_1\neq 0} \approx \frac{L}{k^2} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} S_0(\omega) d\omega + \langle \eta_0^2 \rangle \right] = \frac{2L}{k^2} \langle \eta_0^2 \rangle . \tag{24}$$

Сравнение (23) и (24) позволяет оценить эффективность восстановления первичной волны с помощью логарифмического показателя

$$q = 10 \lg \frac{\langle E^2 \rangle \big|_{L_1=0,\, L\neq 0}}{\langle E^2 \rangle \big|_{L_1=L\neq 0}} \approx (-23 + 20 \lg k - 10 \lg L) \ \ [\text{дБ}]. \tag{25}$$

Восстановление следует считать эффективным, если $q > 0$, что соответствует условию

$$k \geq 14{,}4 \sqrt{L}\,,$$

связывающему нижнюю границу требуемого объёма накопления с числом наиболее интенсивных рассеивателей.

На рис. 4 изображена зависимость логарифмического показателя эффективности ВПВ (25) от объёма накопления k (или безразмерного времени адаптации) и числа доминирующих неоднородностей L. Из рисунка видно, что точность адаптивного восстановления первичной волны возрастает при увеличении времени адаптации и уменьшается при росте числа неоднородностей. В частности, при бесконечно

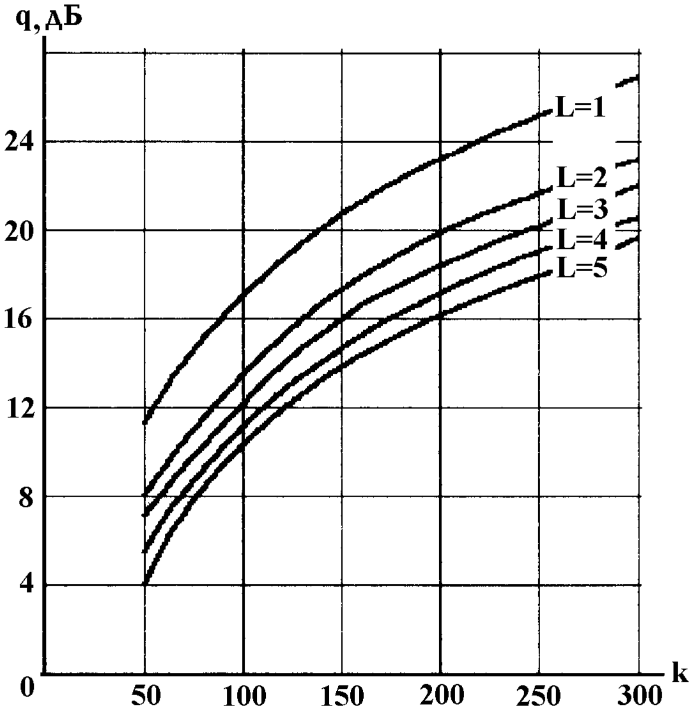


Рис. 4.

большом времени адаптации ошибка восстановления стремится к нулю, а эффективность адаптивного ВПВ сближается с эффективностью идеального. При наличии рассеяния изображённый на рисунке показатель q фактически характеризует дополнительный по отношению к ПОФ выигрыш в подавлении гауссовых активных помех. Таким образом, выше была показана необходимость и практическая возможность увеличения защищённости РЛС от рассеянных в неоднородной среде гауссовых помеховых полей путём восстановления из них первичной волны.

ЛИТЕРАТУРА

1. Островский М. А. // Изв. ВУЗов. Радиофизика, 1997. Т. 40. № 6. С. 769.
2. Миддлтон Д. Введение в статистическую теорию связи. Т. 2. — М.: Сов. радио, 1962.
3. Ширман Я. Д., Манжос В. Н. Теория и техника обработки радиолокационной информации на фоне помех. — М.: Радио и связь, 1981.
4. Островский М. А., Пахомов Ю. И., Дряхлов Ю. А. // Изв. ВУЗов. Радиофизика, 1990. Т. 33. № 5. С. 580.
5. Монзинго Р. А., Миллер Т. У. Адаптивные антенные решетки: Введение в теорию. — М.: Радио и связь, 1986.
6. Уидроу Б., Стирнз С. Адаптивная обработка сигналов. — М.: Радио и связь, 1989.
7. Фалькович С. Е., Пономарев В. И., Шкварко Ю. В. Оптимальный прием пространственно–временных сигналов в радиоканалах с рассеянием. — М.: Радио и связь, 1989.
8. Защита от радиопомех /Под ред. М. В. Максимова. — М.: Сов. радио, 1976.
9. Цыпкин Я. З. Основы теории обучающихся систем. — М.: Наука. 1970.
10. Тихонов В. И. Статистическая радиотехника. — М.: Сов. радио. 1966.

ВЗРКУ, г. Н. Новгород, Россия

Поступила в редакцию
31 октября 1997 г.

DETECTION OF WEAK SIGNALS ON THE BACKGROUND OF ARTIFICIAL NOISES, SCATTERED BY IRREGULAR MEDIUM.
I. GAUSSIAN APPROXIMATION

M. A. Ostrovsky

We have synthesized optimal detectors of weak signals registered on the background of artificial noises. The problem of adaptive restoration of a primary wave in the multichannel detector has been considered in detail.

УДК 621.391.244

ОЦЕНКА ИМПУЛЬСНЫХ ХАРАКТЕРИСТИК СВЕРХШИРОКОПОЛОСНЫХ СИСТЕМ

В. И. Кошелев, В. Т. Сарычев, С. Э. Шипилов

Предлагаются временной и частотный алгоритмы восстановления импульсных характеристик линейных сверхширокополосных систем по конечным наборам цифровых данных, представляющих входной и выходной сигналы. Приводятся результаты апробации алгоритмов численным моделированием при наличии аддитивного гауссовского шума в выходном сигнале. Показано, что наибольшей помехоустойчивостью при оценке импульсных характеристик обладают системы с шумоподобными сигналами на входе.

ВВЕДЕНИЕ

Развитие физики и техники сверхширокополосного излучения за последнее время позволило излучать в пространство электромагнитные импульсы длительностью единицы наносекунд и мощностью до 1 ГВт [1, 2]. Благодаря этому становится перспективным использование таких импульсов для радиолокации как воздушных, так и подземных объектов. Когда принятый сигнал представляется конечным набором цифровых данных, возникает ряд проблем с определением свойств зондируемых объектов.

Известно, что свойства линейных систем, определяющие преобразование сигналов, отражаются их импульсными характеристиками (ИХ). Если прямая задача заключается в нахождении выходного сигнала по известным ИХ и входному сигналу, то обратная задача состоит в оценке ИХ по значениям отсчётов сигнала на входе и выходе рассматриваемой линейной системы. В условиях полной информации (входной и выходной сигналы известны на всей временной оси) задача легко решается на основе Фурье—преобразований. Действительно, связь между входным сигналом $X(t)$, выходным $Y(t)$ и ИХ $h(\tau)$ определяется интегральным соотношением

$$Y(t) = \int\limits_0^\infty h(\tau) X(t-\tau)\,\mathrm{d}\tau\,. \tag{1}$$

В общем случае ИХ $h(\tau)$ является линейным оператором. Совершая над левой и правой частями выражения (1) Фурье—преобразования, легко получить

$$h(\omega) = Y(\omega)/X(\omega), \tag{2}$$

где $X(\omega)$, $Y(\omega)$ и $h(\omega)$ — Фурье—образы соответствующих функций.

Приведённые выражения предполагают знание входного и выходного сигналов на всей вещественной оси, в то время как для практиков интересны ситуации, когда эти сигналы представлены конечными наборами цифровых данных на ограниченных временных интервалах, причём частота дискретизации соизмерима с полосой частот сигнала, что свойственно сверхширокополосным системам. Переход от непрерывных сигналов к дискретизированным сопряжён с рядом трудностей принципиального характера. Основные из них — нарушение принципа причинности, появление ложных боковых лепестков и нулей в спектральных оценках, конечность полосы частот, отсутствие единственности решения как при оценивании спектров $X(\omega)$, $Y(\omega)$, так и при оценивании ИХ. Кроме того, возможная неэквидистантность при дискретизации сигналов порождает дополнительные проблемы.

Цель работы — *разработка помехоустойчивых алгоритмов оценки импульсных характеристик линейных сверхширокополосных систем при произвольных способах дискретизации входного и выходного сигналов.*

1. ОЦЕНКА КОМПЛЕКСНЫХ СПЕКТРОВ ПРИ НЕЭКВИДИСТАНТНОЙ ДИСКРЕТИЗАЦИИ

В общем случае результат дискретизации сигнала представляется двумя последовательностями: $(X_1, \ldots, X_N)$ и $(t_1, \ldots, t_N)$ (уровни сигнала и соответствующие им значения аргумента). Без нарушения общности начало отсчёта аргумента можно выбирать таким образом, чтобы $t_1 = 0$. Значения $(X_1, \ldots, X_N)$ далее рассматриваются как компоненты N-мерного вектора $\mathbf{X}$.

Кроме исходной последовательности $(t_1, \ldots, t_N)$ значений аргументов вводится виртуальная последовательность $(\tau_1, \ldots, \tau_M)$, где $M \le N$. Соответственно этим последовательностям определяются векторы $\mathbf{e}(\omega)$ и $\mathbf{f}(\omega)$ согласно выражениям

$$\begin{gathered} e_n(\omega) = \exp(-i\omega t_n/T), \quad 1 \le n \le N, \\ f_m(\omega) = \exp(-i\omega \tau_m/T), \quad 1 \le m \le M, \\ T = (\tau_M - \tau_1)/(M-1). \end{gathered} \tag{3}$$

Здесь и далее используется безразмерная частота ω, переход к размерной частоте $\omega_\text{р}$ производится соотношением $\omega = \omega_\text{р} \cdot T$.

Комплексный спектр (КС) $X(\omega)$ определяется в виде линейной относительно неизвестного M-мерного вектора $\mathbf{F}$ модели

$$X(\omega) = \mathbf{F} \cdot \mathbf{f}^*(\omega), \tag{4}$$

где * означает комплексное сопряжение.
Этому КС соответствует модельный N-мерный вектор

$$\hat{\mathbf{X}} = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} X(\omega) \mathbf{e}(\omega) \, \mathrm{d}\omega\,. \tag{5}$$

После интегрирования по частоте выражение (5) можно записать в следующем виде:

$$\hat{\mathbf{X}} = \mathbf{F}\mathbf{E}\,. \tag{6}$$

Для вещественных сигналов элементы матрицы $\mathbf{E}$ определяются согласно выражению

$$E_{mn} = \frac{T}{\pi(\tau_m - t_n)} \sin\left(\frac{\pi(\tau_m - t_n)}{T}\right). \tag{7}$$

Если размерности векторов $\hat{\mathbf{X}}$ и $\mathbf{F}$ одинаковы ($M = N$), то компоненты $\mathbf{F}$ могут быть найдены путём непосредственного решения векторного уравнения (6). При $M < N$ оценка компонент $\mathbf{F}$ основывается на поиске минимума функционала среднеквадратичной невязки

$$\Phi = |\hat{\mathbf{X}} - \mathbf{X}|^2. \tag{8}$$

Минимум невязки обеспечивается следующим решением в общем случае переопределённой системы M линейных уравнений (6):

$$\mathbf{F} = \mathbf{X}\mathbf{E}^{-1}, \tag{9}$$

где $\mathbf{E}^{-1}$ — псевдообратная матрица $\mathbf{E}$.

В случае, когда $M = N$ и $\tau_k = t_k = t_1 + (k-1)T$, т. е. при эквидистантной дискретизации матрица $\mathbf{E}$ является диагональной единичной матрицей. Следовательно, согласно (9), $\mathbf{F} = \mathbf{X}$ и комплексный спектр (4) представляет собой классическое дискретное преобразование Фурье.

После вычисления компонент вектора $\mathbf{F}$ можно для любого значения аргумента t найти значение $\hat{X}(t)$ согласно выражению

$$\hat{X}(t) = \frac{T}{\pi} \sum_{m=1}^{M} F_m \sin\left(\frac{\pi(\tau_m - t)}{T}\right) \Big/ (\tau_m - t). \tag{10}$$

При $M = N$ и $\tau_k = t_k$ это выражение совпадает с интерполяционной формулой Котельникова для конечных эквидистантных наборов данных.

2. ГАШЕНИЕ ПРЕДВЕСТНИКОВ

Принципу причинности отвечает условие $h(\tau) = 0$ при $\tau < 0$ [3]. Кроме того, в настоящей работе рассматриваются случаи, когда этому же условию удовлетворяют входной и выходной сигналы. Модельные сигналы $\hat{X}(t)$, $\hat{Y}(t)$ и ИХ, построенные на основе Фурье-преобразования по ограниченной полосе частот, такому условию удовлетворять не будут, т. е. у них появятся предвестники. Можно предложить, по крайней мере, два пути борьбы с предвестниками.

а) При неизменной полосе частот сигнала, определяемой частотой его дискретизации, повышается детальность описания спектра путём увеличения размерности вектора $\mathbf{F}$. Интервал значений индексов компонент этого вектора расширяется в отрицательную область: $-\hat{M} \le m \le M$. В качестве функционала, минимизация которого позволяет оценить значения компонент вектора $\mathbf{F}$, предлагается использовать следующее выражение:

$$\Phi = \frac{1}{T} \int_{-\infty}^{0} |\hat{X}(t)|^2 \, \mathrm{d}t + |\hat{\mathbf{X}} - \mathbf{X}|^2. \tag{11}$$

Вектор $\mathbf{F}$, минимизирующий этот функционал, определяется линейным векторным уравнением

$$\mathbf{F}(\mathbf{G} + \mathbf{E}\mathbf{E}^{\mathrm{H}}) = \mathbf{X}\mathbf{E}^{\mathrm{H}}, \tag{12}$$

где $\mathbf{E}^{\mathrm{H}}$ — транспонированная матрица $\mathbf{E}$, а элементы матрицы $\mathbf{G}$ определяются следующими интегралами:

$$G_{jm} = \frac{T}{\pi^2} \int_{-\infty}^{0} \frac{\sin[\pi(\tau_j - t)] \sin[\pi(\tau_m - t)]}{(\tau_j - t)(\tau_m - t)} \, \mathrm{d}t\,. \tag{13}$$

Матрица $\mathbf{E}$, как и ранее, определяется выражением (7), но значения её индексов m при аргументе τ пробегают указанный выше расширенный до отрицательных значений интервал.

Предварительные исследования показали, что подобный способ подавления предвестников не является эффективным. Увеличение размерности вектора $\mathbf{F}$ приводит к незначительному понижению уровня предвестников ($\sim 10\%$), при этом вычислительные трудности существенно возрастают.

б) Расширение полосы частот до интервала $(0, \infty)$ является радикальным средством подавления предвестников. Вместо выражения (10) для модельной функции используется интегральное Фурье-преобразование

$$\hat{X}(t) = \frac{1}{\pi} \int_{0}^{\infty} [a_x(\omega) \cos(\omega t) + b_x(\omega) \sin(\omega t)] \, \mathrm{d}\omega\,, \tag{14}$$

где коэффициенты Фурье определяются выражениями

$$a_x(\omega) = \frac{1}{\pi}\int\limits_0^\infty \cos(\omega t)\sum_{m=1}^{M} F_m \frac{\sin[\pi(\tau_m - t)/T]}{\tau_m - t}\,\mathrm{d}t\,,$$
$$b_x(\omega) = \frac{1}{\pi}\int\limits_0^\infty \sin(\omega t)\sum_{m=1}^{M} F_m \frac{\sin[\pi(\tau_m - t)/T]}{\tau_m - t}\,\mathrm{d}t\,. \qquad (15)$$

В результате интегрирования получаем

$$a_x(\omega) = \frac{1}{2\pi}\sum_{m=1}^{M} F_m \cdot \begin{cases} \cos(\omega m)\{\pi + \mathrm{Si}[m(\pi+\omega)] + \mathrm{Si}[m(\pi-\omega)]\} + \\ \sin(\omega m)\{\mathrm{Ci}[m(\pi-\omega)] - \mathrm{Ci}[m(\pi+\omega)]\}, \text{ при } \omega < \pi, \\ \cos(\omega m)\{\mathrm{Si}[m(\omega-\pi)] - \mathrm{Si}[m(\pi+\omega)] + \\ \sin(\omega m)\{\mathrm{Ci}[m(\pi+\omega)] - \mathrm{Ci}[m(\omega-\pi)]\}, \text{ при } \omega > \pi, \end{cases}$$

$$b_x(\omega) = \frac{1}{2\pi}\sum_{m=1}^{M} F_m \cdot \begin{cases} \sin(\omega m)\{\pi + \mathrm{Si}[m(\pi+\omega)] + \mathrm{Si}[m(\pi-\omega)]\} + \\ \cos(\omega m)\{\mathrm{Ci}[m(\pi+\omega)] - \mathrm{Ci}[m(\pi-\omega)]\}, \text{ при } \omega < \pi, \\ \sin(\omega m)\{\mathrm{Si}[m(\omega-\pi)] - \mathrm{Si}[m(\pi+\omega)]\} + \\ \cos(\omega m)\{\mathrm{Ci}[m(\pi+\omega)] - \mathrm{Ci}[m(\omega-\pi)]\}, \text{ при } \omega > \pi, \end{cases}$$

где Si и Ci — функции интегрального синуса и косинуса.

Легко убедиться, что функции, Фурье-коэффициенты которых определяются согласно (15), не имеют предвестников.

bf в) Интерполяция Лагранжа. Практическое использование полученных выше (на основе интерполяции Котельникова) выражений коэффициентов Фурье осложняется не только громоздкостью выражений, но и медленным убыванием абсолютных значений этих коэффициентов с ростом частоты. В случае, если частота дискретизации велика, так что сигнал $X(t)$ является монотонной функцией между соседними дискретами аргумента t, кроме того за пределами окна наблюдения сигнал обращается в ноль или становится пренебрежимо малым, то можно использовать простые интерполяционные выражения, например, интерполяцию Лагранжа по нескольким точкам.

При интерполяции Лагранжа по трём точкам выражения коэффициентов Фурье последовательности $X_1, \ldots, X_N$ имеют вид

$$a(\omega) = \frac{1}{2\omega^3}\sum_{n=1}^{N} X_n\, A_n(\omega), \qquad b(\omega) = \frac{1}{2\omega^3}\sum_{n=1}^{N} X_n\, B_n(\omega), \qquad (16)$$

где коэффициенты $A_n(\omega)$ и $B_n(\omega)$ определяются выражениями

$$A_n(\omega) = 6\{\sin[\omega(n{+}1)] - \sin(\omega n)\} + 2\{\sin[\omega(n{-}1)] - \sin[\omega(n{+}2)]\} + \\ + 3\omega\{\cos(\omega n) - \cos[\omega(n{+}1)]\} + \omega\{\cos[\omega(n{+}2)] - \cos[\omega(n{-}1)]\}, \qquad (17)$$

$$B_n(\omega) = 6\{\cos(\omega n) - \cos[\omega(n{+}1)]\} + 2\{\cos[\omega(n{+}2)] - \cos[\omega(n{-}1)]\} + \\ + 3\omega\{\sin(\omega n) - \sin[\omega(n{+}1)]\} + \omega\{\sin[\omega(n{+}2)] - \sin[\omega(n{-}1)]\}. \qquad (18)$$

Аналогичными соотношениями представляются коэффициенты Фурье выходного сигнала $\mathbf{Y}$ и импульсной характеристики $\mathbf{H}$.

3. ОЦЕНИВАНИЕ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ

Интеграл свёртки (1) при интерполяции подынтегральных функций трансформируется в следующее векторное соотношение:

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{D}(\mathbf{X}). \tag{19}$$

Для случая, когда $h(\tau)$ является скалярной функцией и используется интерполяция Лагранжа по трём точкам, элементы матрицы $\mathbf{D}(\mathbf{X})$ определяются выражением

$$D_{jm} = \frac{1}{120}(X_{j-m-3} - 14X_{j-m-2} + 36X_{j-m-1} + 86X_{j-m} + 11X_{j-m+1}). \tag{20}$$

Искомым в выражении (19) является вектор $\mathbf{H}$. Размерности векторов и матрицы в (19) ограничены соотношениями $\dim\mathbf{Y} = \dim\mathbf{H} + \dim\mathbf{X} + 3$ и $\dim\mathbf{D} = \dim\mathbf{Y} \cdot \dim\mathbf{H}$. Таким образом, решение уравнения (19) сопряжено с нахождением псевдообратной матрицы $\mathbf{D}^{-1}$. В данной работе эта матрица находилась с помощью специально разработанной процедуры обращения прямоугольных матриц, основанной на методе ортогонализации Грама-Шмидта.

В общем случае, когда импульсная характеристика имеет вид

$$h(\tau) = h_0(\tau) + h_1(\tau)\frac{\mathrm{d}}{\mathrm{d}\tau} + \ldots, \tag{21}$$

в правой части выражения (19) появятся члены, соответствующие операторам дифференцирования:

$$\mathbf{Y} = \mathbf{H}_0 \cdot \mathbf{D}_0(\mathbf{X}) + \mathbf{H}_1 \cdot \mathbf{D}_1(\mathbf{X}) + \ldots \tag{22}$$

Для нахождения вида элементов матриц $\mathbf{D}_k,\ k > 0$ следует использовать интерполяционные выражения, соответствующие k-й производной.

4. АПРОБАЦИЯ ОЦЕНКИ ИХ

Основной целью данного раздела являются выяснения меры влияния аддитивного шума выходного сигнала $\mathbf{Y}$ на точность восстановления ИХ при известном входном сигнале $\mathbf{X}$.

Исследовалась наиболее простая ситуация, когда в выражении (21) для ИХ доминирует один член. В этом случае выражение (22) сводится к виду (19), где в качестве аргумента матрицы $\mathbf{D}$ используется или вектор $\mathbf{X}$ (если доминирующим для $h(\tau)$ является $h_0(\tau)$), или соответствующая доминирующему члену производная входного сигнала.

Использовались модели сигналов с различной крутизной фронтов. Ширина фронтов модельных сигналов в единицах интервала дискретизации определялась значением параметра Ng. Для очень крутых фронтов с шириной, равной интервалу дискретизации, $Ng = 1$.

На рис. 1 (кривые 1–4) представлены сигналы с очень крутыми фронтами. Сигналы с менее крутыми фронтами ($Ng = 3$) отличались от приведённых на рисунке только в пределах трёх крайних отсчётов, где они по линейному закону обращались в нуль к границам окна. Кроме того, использовался шумоподобный сигнал, каждый отсчёт которого генерировался датчиком случайных чисел, распределённых по нормальному закону. Ширина окна или длительность сигнала $\mathbf{X}$ равнялась 129 отсчётам. Буквой H на рисунке обозначена кривая, представляющая модель ИХ с длительностью в 63 отсчёта.

Процесс моделирования состоял из следующих операций. По известным исходным векторам $\mathbf{X}$ и $\mathbf{H}$ на основе (19) вычислялся незашумлённый вектор $\mathbf{Y}$. После чего к сигналу $\mathbf{Y}$ добавлялся аддитивно белый гауссовский шум. Для полученного сигнала оценка компонент вектора $\mathbf{H}$ производилась тремя различными способами: решением переопределённой системы линейных уравнений (19) (вариант а) и

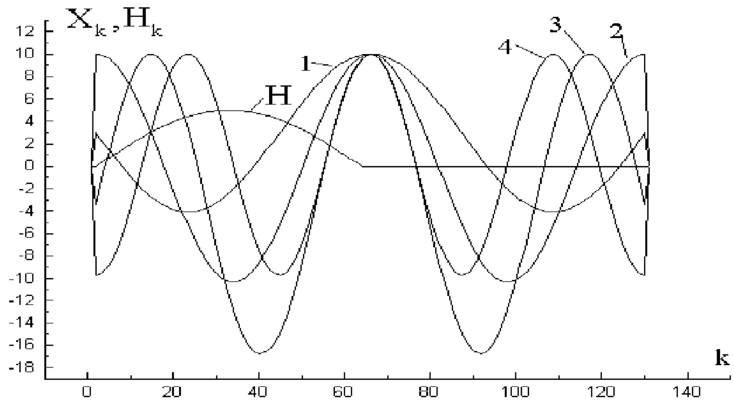


Рис. 1.

на основе Фурье-преобразований, когда спектры сигналов $\mathbf{X}$ и $\mathbf{Y}$ оценивались на основе выражений (16)–(18) (вариант b) и на основе традиционного конечного преобразования Фурье (вариант c). При вычислении $h(\omega)$ на основе выражения (2) использовался стандартный алгоритм регуляризации [4]. Для описания помехоустойчивости алгоритма оценки ИХ использовались коэффициенты μ_h и μ_y, которые определяются выражениями

$$\mu_h = \log_2 \frac{\langle|\delta\mathbf{H}|^2\rangle}{|\mathbf{H}|^2}\,, \qquad \mu_y = \log_2 \frac{\langle|\delta\mathbf{Y}|^2\rangle}{|\mathbf{Y}|^2}\,, \tag{23}$$

где $\langle|\delta\mathbf{H}|^2\rangle$ и $\langle|\delta\mathbf{Y}|^2\rangle$ — вызванные шумом среднеквадратичные отклонения векторов $\mathbf{H}$ и $\mathbf{Y}$.

Коэффициент μ_y при расчёте каждой линии изменялся от -18 до $+2$ с шагом 2. Объём статистики для каждой точки линии был равен 64. Наряду с коэффициетом μ_h оценивался коэффициент

$$\langle\mu_h\rangle = \log_2 \frac{|\langle\mathbf{H}\rangle - \mathbf{H}|^2}{|\mathbf{H}|^2}\,. \tag{24}$$

На рис. 2 и 3 представлены зависимости μ_h и $\langle\mu_h\rangle$ от коэффициента μ_y. Буквенные указатели кривых соответствуют вариантам алгоритмов, которыми оценивались ИХ. Скобки $\langle\rangle$ указывают, что кривая соответствует усреднённому коэффициенту $\langle\mu_h\rangle$. Рис. 2 соответствует сигналу $\mathbf{X}$ (под номером 4 рис. 1) с очень крутыми фронтами (Ng = 1). Рис. 3 соответствует этому же сигналу, но со сглаженными фронтами (Ng = 3).

Оказалось, что результаты расчётов могут быть аппроксимированы семейством параллельных линий. Соответствующее регрессионным линиям выражение имеет вид

$$\mu_h = \mu_y + K_i\,. \tag{25}$$

Причём характер сигнала $\mathbf{X}$ влияет лишь на значение коэффициентов K_i и не изменяет наклон регрессионных линий. Укручение фронтов сигнала $\mathbf{X}$ приводит к опусканию линий регрессий. Действительно, все линии рис. 2 расположены ниже соответствующих линий рис. 3. Этот результат вполне объясним тем, что укручение фронтов ведёт к уширению полосы сигнала, что в свою очередь повышает помехоустойчивость оценок ИХ.

В процессе расчётов обнаружилась зависимость помехоустойчивости оценок ИХ от выбранного метода. Для варианта а) (оценка ИХ на основе решения системы (19)) вне зависимости от характера входного сигнала $\mathbf{X}$ результаты расчётов с хорошей точностью ложились на регрессионные прямые во всём исследуемом диапазоне значений коэффициентов.

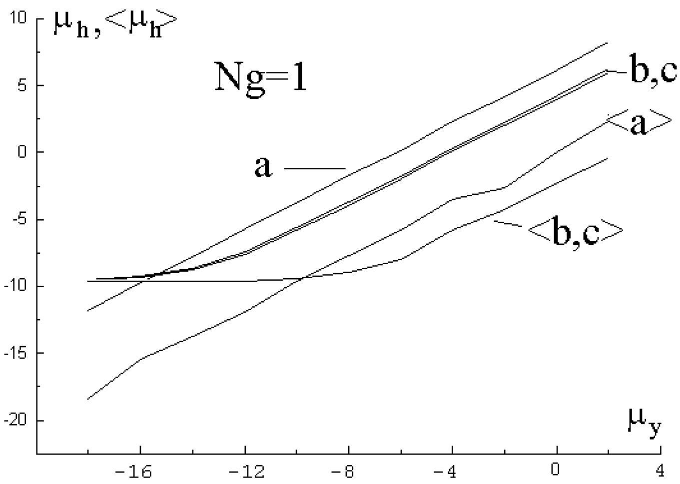


Рис. 2.

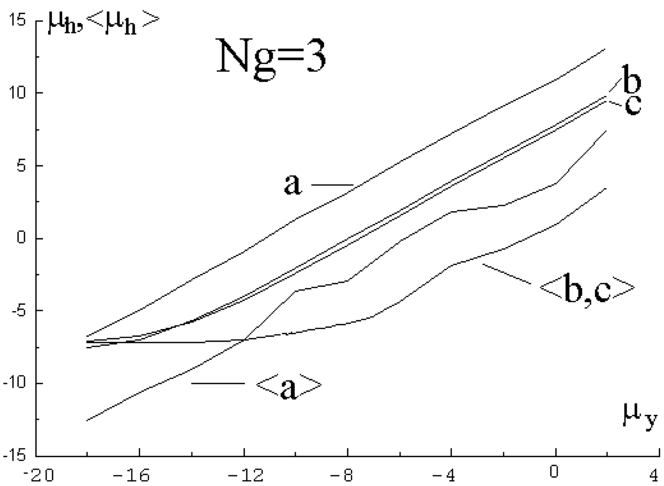


Рис. 3.

Варианты b) и c) мало отличались друг от друга (линии варианта c располагались несколько выше соответствующих линий варианта b). Для обоих этих вариантов наблюдалось появление порога по точности. Как бы не был мал шум сигнала $\mathbf{Y}$, точность оценки коэффициентов μ_h и $\langle\mu_h\rangle$ не превосходила этого порога. Уровень порога определялся характером входного сигнала. Таким образом, значения оценок ИХ вариантами b) и c) подчинялись регрессионому уравнению (25) лишь в области, где значения коэффициентов μ_h и $\langle\mu_h\rangle$ превосходили порог.

Значения коэффициентов K_i и $\langle K_i\rangle$, соответствующих сигналам рис. 1 с индексом i ($i = n$ соответствует шумоподобному $\mathbf{X}$), приводятся ниже в таблице.

Коэффициенты K_i нижней половины таблицы, соответствующие усреднённым оценкам ИХ, меньше соответствующих коэффициентов верхней половины примерно на 6 (исключая шумоподобный сигнал). Объём статистики при усреднении был $N = 64 = 2^6$.

Как видно из таблицы, для шумоподобного сигнала, обладающего самым широким спектром среди других типов сигналов, значения коэффициентов K_i для всех вариантов отрицательны, т. е. шумоподобные сигналы наиболее помехоустойчивы. Даже сглаживание переднего и заднего фронтов шумопо-

Таблица

Ng	ва-ри-ант	K_1	K_2	K_3	K_4	K_n
1	a	13.92 ± 0.05	9.85 ± 0.04	13.15 ± 0.07	6.24 ± 0.06	-1.54 ± 0.05
	b	10.95 ± 0.03	7.68 ± 0.03	10.3 ± 0.03	4.03 ± 0.04	-2.7 ± 0.02
	c	11.08 ± 0.03	7.92 ± 0.03	10.43 ± 0.03	4.26 ± 0.04	-2.69 ± 0.02
3	a	19.65 ± 0.05	15.01 ± 0.06	19.05 ± 0.1	11.14 ± 0.1	-3.73 ± 0.04
	b	14.01 ± 0.05	11.22 ± 0.08	12.92 ± 0.15	7.57 ± 0.07	-4.02
	c	14.74 ± 0.04	11.52 ± 0.04	13.48 ± 0.03	7.92 ± 0.05	-3.95 ± 0.02
1	$\langle a \rangle$	7.78 ± 0.34	3.92 ± 0.38	6.99 ± 0.33	$.14 \pm 0.36$	-7.85 ± 0.4
	$\langle b \rangle$	4.89 ± 0.42	1.78 ± 0.45	4.38 ± 0.4	-1.96 ± 0.51	-8.05 ± 0.82
	$\langle c \rangle$	5.21 ± 0.7	2.085 ± 0.63	4.53 ± 0.38	-1.78 ± 0.5	-8.03 ± 0.83
3	$\langle a \rangle$	13.61 ± 0.5	8.81 ± 0.6	12.69 ± 0.67	5.2 ± 0.68	$-9.81 \pm .35$
	$\langle b \rangle$	8.18 ± 0.94	5.82 ± 1.38	7.72 ± 1.54	2.22 ± 1.28	-9.61 ± 1.0
	$\langle c \rangle$	8.73 ± 0.57	5.78 ± 0.85	7.61 ± 0.53	2.41 ± 1.08	$-9.5 \pm 1.$

добного сигнала не привело к ухудшению помехоустойчивости, как это имело место для других типов входных сигналов.

На рис. 4 и рис. 5 приведены зависимости коэффициентов μ_h и $\langle \mu_h \rangle$ от μ_y для обоих вариантов шумоподобного сигнала (рис. 4 без сглаживания фронтов, рис. 5 — со сглаживанием). Для варианта a) метода оценки ИХ регрессионный характер зависимости коэффициентов μ_h от μ_y сохранился. Тогда как в вариантах b) и c) такая зависимость полностью отсутствовала — методы работали на уровне своих порогов по помехоустойчивости. Преимущество использования варианта a) перед двумя другими в этом случае очевидно.

ВЫВОДЫ

Помехоустойчивость методов оценки ИХ в сильной степени определяется полосой частот входного сигнала — чем шире полоса, тем выше помехоустойчивость.

Выбор метода оценки ИХ следует производить с учётом уровня шума выходного сигнала. Для слабого шума следует отдать предпочтение решению системы линейных уравнений во временном базисе. При высоком уровне шума рекомендуется использовать спектральные методы восстановления ИХ.

Усреднение оценок ИХ, полученных временным алгоритмом, позволяет повысить их точность на величину, пропорциональную объёму выборки во всём исследуемом диапазоне шумов выходного сигнала. Спектральные алгоритмы имеют насыщение по точности оценок ИХ при увеличении объёма выборки.

Наибольшей помехоустойчивостью при оценке импульсных характеристик обладают системы с шумоподобными сигналами на входе.

Работа выполнена при поддержке CRDF. Проект RZ2–501.

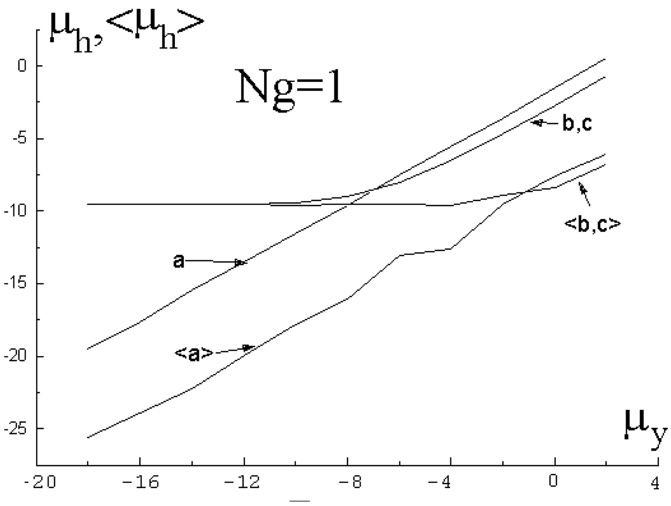


Рис. 4.

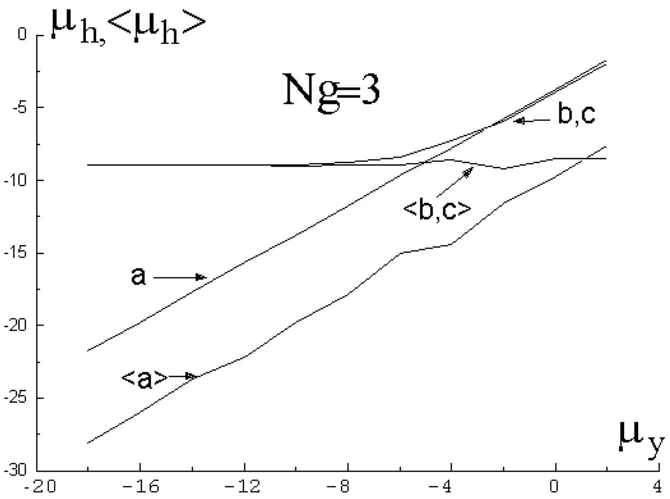


Рис. 5.

ЛИТЕРАТУРА

1. Oicles J. A., Grant J. R., and Herman M. H. //SPIE Proc., 1995. V. 2557. P. 225.
2. Koshelev V. I., Buyanov Y. I., Kovalchuk B. M. et. al. //SPIE Proc., 1997. V. 3158. P. 209.
3. Нуссенцвейг Х. М. Причинность и дисперсионные соотношения. — М.: Мир, 1976. – 462 с.
4. Численные методы решения некорректных задач /Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. — М.: Наука, 1990. – 232 с.

Институт сильноточной
электроники РАН, г. Томск, Россия

Поступила в редакцию
13 февраля 1998 г.

ESTIMATION OF IMPULSE RESPONSE OF ULTRAWIDEBAND SYSTEMS

V. I. Koshelev, V. T. Sarychev, S. E. Shipilov

Time and frequency domain algorithms of impulse response (IR) recovery of linear ultrawideband systems by limited sequence of numerical date presenting an input and output are suggested.

The results of algorithms approbation by numerical simulation with the presence of the additive Gaussian noise in the output signal are given.

Is shown that the highest noise immunity at the IR estimation have the systems with noise-like signals at the input.