
СОДЕРЖАНИЕ

Том 67, номер 2, 2022

ОБЗОР

Магнитные наночастицы в медицине: успехи, проблемы, достижения

Ю. А. Кокшаров, С. П. Губин, И. В. Таранов, Г. Б. Хомутов, Ю. В. Гуляев

99

ЭЛЕКТРОДИНАМИКА И РАСПРОСТРАНЕНИЕ РАДИОВОЛН

Развитие метода восстановления эффективной частоты соударений электронов в ионосферной плазме

А. С. Крюковский, Д. С. Лукин, Е. В. Михалёва, Д. В. Растягаев

117

Переходная зона в высокочастотной задаче дифракции на импедансной границе со скачком кривизны. Метод Кирхгофа и метод пограничного слоя

Е. А. Злобина, А. П. Киселев

130

АНТЕННО-ФИДЕРНЫЕ СИСТЕМЫ

Синтез двухзеркальной бифокальной цилиндрической системы с минимальными аберрациями

В. А. Калошин, Ви Ут Нам

140

СТАТИСТИЧЕСКАЯ РАДИОФИЗИКА

Компенсация отражений от местных предметов в импульсно-доплеровских метеорадиолокаторах

Б. М. Вовшин, В. К. Битюков

149

ТЕОРИЯ И МЕТОДЫ ОБРАБОТКИ СИГНАЛОВ

Статистический синтез алгоритма формирования радиолокационных изображений объектов с учетом декорреляции импульсов

В. Н. Лагуткин

157

Исследование помехоустойчивости глубокой сверточной нейронной сети при обнаружении транспортных средств на аэрофотоснимках Земли

О. В. Ильина, М. В. Терешонок

166

РАДИОФИЗИЧЕСКИЕ ЯВЛЕНИЯ В ТВЕРДОМ ТЕЛЕ И ПЛАЗМЕ

Исследование электрических и структурных параметров шунгита по данным электросилового спектроскопии

И. В. Антонец, Е. А. Голубев, В. Г. Шавров, В. И. Щеглов

174

Аномальные токи, индуцируемые во внешней цепи изменениями параметров образца

С. Г. Дмитриев

181

ПРИМЕНЕНИЕ РАДИОТЕХНИКИ И ЭЛЕКТРОНИКИ В БИОЛОГИИ И МЕДИЦИНЕ

Интеллектуальный анализ характеристик пульсовой волны методами структурной декомпозиции сигналов фотоплетизмографии

Л. В. Лабунец, А. Б. Борзов, Н. Ю. Макарова

185

ЭЛЕКТРОННАЯ И ИОННАЯ ЭМИССИЯ

Полевая эмиссия электронов многоострийных катодных матриц в сильных импульсных электрических полях

Р. К. Яфаров

197

ФИЗИЧЕСКИЕ ЯВЛЕНИЯ В ЭЛЕКТРОННЫХ ПРИБОРАХ

Исследование возможности применения фотозарядового эффекта в электронике

П. С. Мартьянов, Д. В. Чуриков

203

ХРОНИКА

К 90-летию Анатолия Васильевича Вашковского

209

УДК 53.097

МАГНИТНЫЕ НАНОЧАСТИЦЫ В МЕДИЦИНЕ: УСПЕХИ, ПРОБЛЕМЫ, ДОСТИЖЕНИЯ

© 2022 г. Ю. А. Кокшаров^{a, b}, С. П. Губин^c, И. В. Таранов^{b, *}, Г. Б. Хомутов^{a, b}, Ю. В. Гуляев^b

^aФизический факультет МГУ им. М.В. Ломоносова,
Ленинские горы, дом 1, стр. 2, Москва, 119991 Российская Федерация

^bИнститут радиотехники и электроники им. В.А. Котельникова РАН,
ул. Моховая, 11, стр. 7, Москва, 125009 Российская Федерация

^cИнститут общей и неорганической химии им. Н.С. Курнакова РАН,
Ленинский просп., 31, Москва, 119991 Российская Федерация

*E-mail: i.v.t@bk.ru

Поступила в редакцию 20.07.2021 г.

После доработки 04.08.2021 г.

Принята к публикации 11.08.2021 г.

Представлен анализ современного состояния исследований, связанных с разработкой, развитием и практическим применением методов биомедицинской радиоэлектроники и наномедицины, включающих применения магнитных наночастиц. Подчеркивается важная роль рациональных физических научных подходов и экспериментальных методов при создании эффективных и безопасных средств терапии, адресной управляемой доставки лекарственных препаратов и диагностики включая пространственную визуализацию, использующих магнитные наночастицы. Приведены примеры успешного практического применения магнитных наночастиц в медицине на основе этих методов, проведен анализ основных проблем и перспектив этого направления науки.

DOI: 10.31857/S0033849422020073

ВВЕДЕНИЕ

Разработка новых высокоэффективных средств диагностики и терапии имеет в настоящее время стратегическое значение, что, в частности, показала продолжающаяся пандемия коронавируса. Повышение эффективности традиционной медицины связано с созданием и внедрением в практику новых подходов к диагностике и терапии, включающих наряду с биохимическими методами различные физические воздействия и средства диагностики. Введение в организм коллоидных наночастиц и функциональных наноструктурированных систем может обеспечить такие возможности визуализации и диагностики. Кроме того, это может позволить реализовать избирательное дистанционное внешнее физическое воздействие в локальных областях организма и обеспечить непосредственный терапевтический эффект (например, гипертермию) или управляемую адресную доставку и высвобождение лекарственных препаратов в целевой области. Применение функциональных наночастиц и других коллоидных нанообъектов и наноструктур (систем, состоящих из нанообъектов) в медицине — сложнейшая мультидисциплинарная научная и практическая задача, решение которой далеко от своего заверше-

ния. В предлагаемом обзоре основное внимание уделено медицинскому применению магнитных наночастиц (МНЧ). Несмотря на очевидные и зачастую впечатляющие успехи фундаментальных исследований в этой области (см., например, обзоры и монографии, а также оригинальные статьи [1–54]), при реальном практическом биомедицинском применении МНЧ и композиций на их основе возникают те же проблемы, что и при применении других видов лекарственных нанопрепаратов. Главные проблемы, обуславливающие медленное внедрение нанолекарств в реальную клиническую практику, связаны прежде всего со сложностью и многофакторностью процессов взаимодействия неорганических наночастиц с различными биологическими структурами организма человека. Эти взаимодействия и их последствия не всегда можно предсказать, проводя доклинические исследования (*in vitro* и на животных моделях). В последнее время стало очевидно, что для дальнейшего прогресса в изучении поведения магнитных наночастиц непосредственно в организме человека особенно необходимы разработка и внедрение новых радиоэлектронных технологий их пространственной визуализации.

Относительно медленный прогресс наномедицины уже начал порождать сомнения в принципах

альной возможности преодолеть существующие проблемы наномедицины в обозримом будущем [7, 55–66]. Недостатки практически ориентированных медицинских методов, использующих наночастицы и другие нанообъекты и наноструктуры, обусловлены, в частности, большой сложностью таких методов [7], их недостаточной надежностью (невоспроизводимостью результатов) [67], эффективностью, много меньшей ожидаемой [57, 60, 66, 68], и другими причинами [69]. Так, с 1995 по 2018 г. только 15 “пассивных” антираковых лекарственных наноносителей были одобрены для практического применения, и ни один “активный” наноноситель не прошел успешно клинические испытания, необходимые для широкого медицинского использования [61, 66, 70]. В период с 1985 по 2005 г. было опубликовано около 3 млн научных статей по онкологической наномедицине, но только 3% заявленных методик дошло до стадии клинических испытаний [59]. Как следствие, крупный научный фонд США National Cancer Institute (NCI), который в 2005 г. с энтузиазмом начал финансировать 10-летний Cancer Nanotechnology Plan, в 2019 г. прекратил финансирование программы Centers of Cancer Nanotechnology Excellence (CCNEs) [65].

Все сказанное в полной мере относится и к методам, использующим МНЧ [71–76]. МНЧ и магнитные коллоидные системы на их основе отличаются от других видов лекарственных коллоидных наноносителей расширенными возможностями управляемой доставки к больному органу (с помощью внешнего магнитного поля [1, 77–85], а также возможностью их самостоятельного использования в качестве терапевтических средств, например в методе гипертермии [86–93], диагностических средств, например в магниторезонансной томографии [94–99], и мультифункциональных (тераностических) средств [100–102]. Однако химические, биологические, медицинские проблемы, возникающие на пути к практическому применению, сходны для всех видов коллоидных наноносителей [103].

Возникает вопрос, почему же, несмотря на очевидные достижения в понимании фундаментальных основ наномедицины [31, 41, 104, 105] и разнообразных современных высокоточные экспериментальные методы изучения нанообъектов [26, 106–109], ее практическое применение тормозится. В данном обзоре мы постарались ответить на этот вопрос, а также обсудили возможное будущее наномедицины, уделив особое внимание особенностям и перспективам медицинского использования магнитных наночастиц, являющихся функциональными объектами ряда методов биомедицинской радиоэлектроники.

1. ОБЩИЕ СВЕДЕНИЯ О НАНОМЕДИЦИНЕ

В настоящее время наномедицина достигла уровня развития, когда ее начинают рассматривать как отдельную область наук о жизни [110, 111]. Истоки наномедицины принято относить к работам немецкого врача и исследователя П. Эрлиха, высказавшего идею создания лекарства, избирательно и специфически действующего на больной орган и не затрагивающего здоровые ткани. Такое идеальное лекарство Эрлих назвал “магической пулей” [112]. Сейчас этот подход к созданию средств лекарственной терапии онкологических заболеваний принято называть избирательной токсичностью [113], причем избирательное действие предполагается на клеточном или даже внутриклеточном уровне [114]. Эрлих создал [115] первое лекарство для химиотерапии (и сам предложил этот термин) и ввел важное понятие терапевтического индекса (отношение минимальной уже действующей и максимальной еще безопасной концентраций лекарства (рис. 1 [116]), используемого до сих пор для количественной оценки относительной безопасности лекарства [116]. Эрлих считал, что действие лекарства на клетку объясняется последовательностью обычных химических реакций, а наиболее эффективные лекарства должны иметь низкую молекулярную массу [113]. Последнее предположение впоследствии полностью подтвердилось для искусственных химически синтезируемых лекарств, к которым относится большинство (более 80%) современных фармакологических препаратов [114].

Для любого лекарства важным является вопрос дозы. Для того чтобы лекарство подействовало, его концентрация в крови достаточно долго должна быть в пределах, задающих терапевтический индекс (см. рис. 1). Если лекарственное средство не капсулировано в носителе (не имеет защитной оболочки, отделяющей его от жидкой биологической среды организма), его концентрация в крови после введения в организм сначала быстро возрастает, а затем быстро убывает (см. рис. 1а) вследствие метаболизации и вывода из организма [116]. Такая фармакинетическая зависимость приводит, как правило, к необходимости увеличения вводимой дозы лекарства и, как следствие, превышения максимально нетоксичной концентрации (см. рис. 1а). Постепенное высвобождение терапевтических молекул из носителя лекарства, благодаря специальной оболочке, позволяет избежать этой проблемы (см. рис. 1б).

Идея замедленного поступления лекарства в кровь сыграла важную роль в истории наномедицины [117]. В 60-е годы XX в. появились макроскопические носители лекарств, которые благодаря своей оболочке медленно выделяли содер-

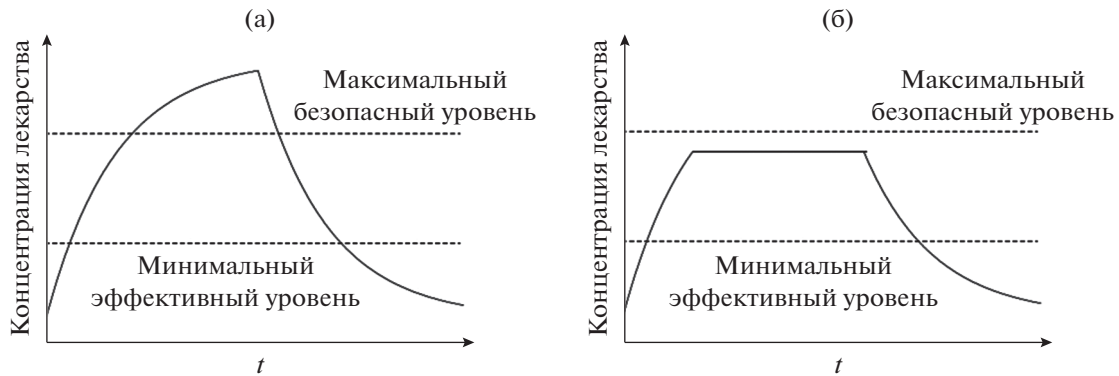


Рис. 1. Фармакинетическая кривая лекарства для носителя без защитной оболочки (а) и с оболочкой (б). Адаптировано из [116, р. 1025]. “Плато” на кривой (б) обеспечивает требуемую эффективность лекарства без риска превышения максимального безопасного уровня концентрации.

жимое наружу. Первыми материалами для таких носителей были парафин (технология Spansules [118]), силиконовая резина [119] и сополимер полиэтиленвинилацетат [120]; использовались они, в частности, для лечения катаракты (препарат Ocuser[®]) и контрацепции (Norplant[®]). В 80-е годы появились биоразлагаемые полимеры, в частности, (полилактидгликолид, PLGA), на основе которых стали делать пористые микрочастицы, содержащие терапевтическое вещество. По этому принципу был создан, к примеру, антираковый препарат Decareptyl[®], одобренный в Европе в 1986 г. Другой биоразлагаемый полимер полиэтиленгликольтетрафталат (PEG-T) стал основой препарата Locteron[®] для доставки α -интерферона, используемого с середины 90-х годов прошлого века до сих пор. Дальнейшее уменьшение размера лекарственных нано-носителей произошло, когда были синтезированы конъюгаты полимеров, в частности полигидроксипропилметакриламид (PHРМА) и стирол-малеиновый ангидрид (SMA), с терапевтическими молекулами, например молекулой доксирубицина [121, 122].

Еще одним важным для наномедицины событием стало открытие в 1984 г. Х. Маэда эффекта “усиленного проникновения и удержания” коллоидных частиц в активно растущих опухолях (“enhanced permeation and retention”, или эффект EPR) [123]. Конъюгат полимер–лекарство (SMA, конъюгированный с противораковым пептидным препаратом, неокарциностатином), который Маэда назвал “SMANCS”, был дополнительно “помечен” молекулой красителя. Оказалось, что краситель накапливается в опухолевой ткани в заметно большей степени, чем в здоровой. Маэда предположил, что быстро формирующаяся сосудистая сеть в опухолях имеет поры большего размера (около 800 нм), чем нормальная, а лимфодренажная система не способна работать эффективно (рис. 2). Это и приводит к повышенной концен-

трации коллоидного нанолейкарства в опухолевой ткани.

До настоящего времени эффект EPR является основой почти всех методик “пассивной” доставки лекарств. Одним из недостатков эффекта EPR с точки зрения применения в наномедицине является трудность удаления наноносителей, уже освободивших лекарство, из опухоли, с тем чтобы освободить место новым носителям лекарства [124]. Поиск новых полимерных конъюгатов привел к открытию и последующему активному использованию наноносителей на основе полимерных мицелл [125–127]. Хорошо известно, что амфифильные блок-сополимеры самопроизвольно собираются в полимерные мицеллы диаметром несколько десятков нанометров в водных средах. Такие полимерные мицеллы имеют уникальную структуру “ядро–оболочка”, в которой внутрен-

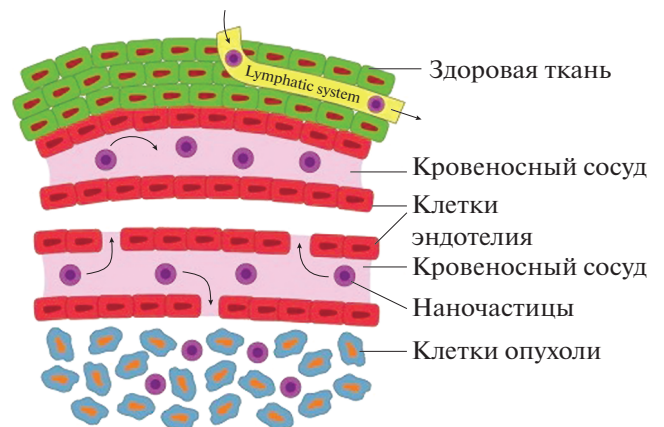


Рис. 2. Схема, поясняющая эффект ЭПР. Адаптировано из [68]. Проницаемость для наноносителей кровеносных сосудов в опухоли гораздо выше, чем в нормальной ткани, благодаря “пористой” структуре стенок сосудов, вызванной, возможно, ускоренным ростом опухоли.

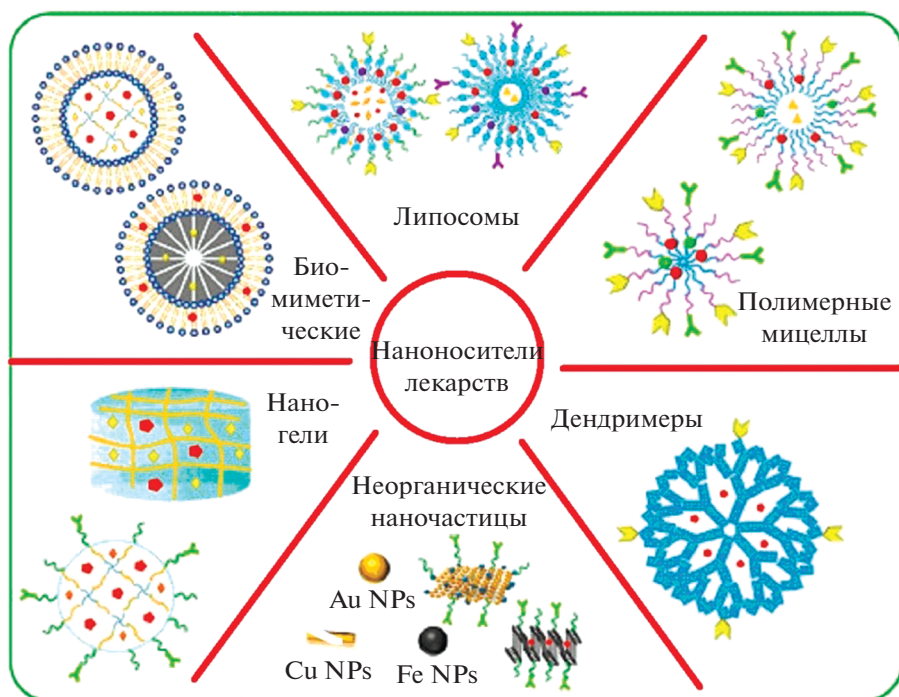


Рис. 3. Классификация нанообъектов, используемых для создания наноносителей лекарств. Адаптировано из [139].

няя часть может служить наноконтейнером для лекарственных средств, в том числе гидрофобных. Внешняя оболочка мицеллы образована гидрофильными полимерами, такими, например, как полиэтиленгликоль (PEG). Полимерные мицеллы имеют ряд преимуществ, таких как простое приготовление, эффективная загрузка лекарственного средства без химической модификации исходного лекарственного средства, контролируемое высвобождение лекарственного средства, увеличение времени циркуляции в кровотоке. Для полимерных мицелл также характерен эффект EPR.

Одновременно с созданием полимерных мицелл возрос интерес к биомиметическим везикулам — липосомам в качестве возможных наноносителей лекарств. Липосомы активно изучались с 60-х годов XX в., так как рассматривались достаточно адекватными искусственными моделями клеточных мембран [128]. Липосомы представляют собой небольшие искусственные везикулы сферической формы с мембраной, состоящей из фосфолипидных бислоев [2, 17, 129–132]. Они могут быть изготовлены из природных нетоксичных фосфолипидов и холестерина в виде одного или нескольких концентрических бислоев, способных инкапсулировать гидрофильные и гидрофобные лекарственные средства. Размер липосом зависит от их состава и способа получения и варьируется от ≈ 10 нм до ≈ 2.5 мкм [133]. Во многих одобренных к практическому использованию наномедицинских методиках используют именно

липосомы [134, 135]. Это обусловлено, в частности, возможностью получать липосомы с различными размерами, их биосовместимостью, биоразлагаемостью, низкими токсичностью и иммуногенностью [128, 136]. Большой интерес и перспективы с точки зрения избирательного управления высвобождением капсулированного лекарства внешними дистанционными физическими радиоэлектронными воздействиями (переменные электромагнитные поля [51, 52], импульсы электрического поля [53]) имеют наноконструктивные липосомы, мембраны которых содержат суперпарамагнитные наночастицы оксидов железа.

В настоящее время в наномедицине кроме липосом и полимерных мицелл для создания наноносителей лекарств используют дендримеры, наногели, неорганические наночастицы (в том числе магнитные), биомиметические частицы, а также различные сложные наноконструкции на их основе [23, 137–139]. Перечислим некоторые наиболее типичные и полезные для медицинских применений особенности нанообъектов [17, 21, 140], показанных на рис. 3.

Неорганические наночастицы (включая магнитные): повышают скорость и предел растворимости лекарственных средств, их биодоступность; они относительно малотоксичны; магнитные наночастицы могут управляться внешним магнитным полем.

Наночастицы на основе полимеров: многофункциональны, повышают стабильность лекарствен-

ного препарата, улучшают проникновение лекарства в ткани и клетки.

Липосомы и мицеллы: существенно увеличивают количество переносимого лекарственного средства и снижают его токсичность, защищают переносимое лекарство от деградации, увеличивают время циркуляции лекарства в кровеносной системе.

Биомиметические: отличаются повышенной биосовместимостью и низкой иммуногенностью.

Дендримеры: повышают однородность нанолекарства по размеру и форме.

Наногели: высокая полезная нагрузка, допускают различные управления процессом высвобождения лекарства.

Отметим, что перечисленные особенности отдельных нанообъектов в той или иной степени присущи и всем другим их типам. В данном обзоре подробно остановимся на медицинском использовании МНЧ, которые относятся к неорганическим наночастицам.

2. ОСОБЕННОСТИ МЕДИЦИНСКОГО ПРИМЕНЕНИЯ МАГНИТНЫХ НАНОЧАСТИЦ

Основные физико-химические свойства МНЧ подробно описаны в литературе (см., например, [80–82, 100, 141]). Главной особенностью, отличающей МНЧ от других видов наноносителей лекарств, является возможность воздействовать на них внешним электромагнитным полем и использовать в различных методах и подходах биомедицинской радиоэлектроники. Это воздействие может обеспечить управление движением наночастиц и их аккумуляцию в целевых локальных областях организма, а при необходимости и существенный нагрев (в методе гипертермии). С точки зрения классификации магнитных свойств МНЧ являются ферро- или ферритмагнетиками, проявляющими благодаря однодоменности суперпарамагнитные свойства при достаточно высоких температурах (выше так называемой температуры блокировки) [142]. Для медицинских применений наиболее пригодны МНЧ на основе оксидов железа – магнетита (Fe_3O_4) и маггемита ($\gamma\text{-Fe}_2\text{O}_3$) [143–145], поскольку нанофазные оксиды железа малотоксичны, они широко распространены в живых системах и могут метаболизироваться клетками [145]. Биогенные магнитные наночастицы оксида железа (магнетит и маггемит) в настоящее время найдены в самых разных живых организмах и играют важную физиологическую роль, позволяя организму ориентироваться в магнитном поле Земли (магнитосомы у бактерий, насекомых, рыб, птиц и пр.) а также являются следствием и проявлением патологий, в частности, у человека

их присутствие коррелирует с нейродегенеративными заболеваниями [17].

Часто МНЧ на основе оксидов железа в литературе называют SPIONs (superparamagnetic iron oxide nanoparticles), имея в виду характерное свойство этими наночастиц – суперпарамагнетизм [141, 146]. В последние десятилетия SPIONs становятся все более популярными благодаря их многочисленным биомедицинским применениям, таким как терапия рака с помощью гипертермии, опосредованной магнитным полем, адресная доставка лекарств, заместительная терапия железом, а также MPT [147–149]. Хотя SPIONs, как правило, считаются биосовместимыми и обладают низким цитотоксическим потенциалом [150], очевидно, что эти частицы в организме человека являются все же инородными телами и могут оказывать различные эффекты на иммунную систему, которые приводят к гиперчувствительности, иммуносупрессии или, наоборот, иммунной стимуляции [151, 152].

Цели поиска новых лекарственных средств с использованием нанотехнологий, основным или вспомогательным элементом которых являются МНЧ, многообразны. К ним можно отнести [153–155]:

- улучшение так называемого терапевтического индекса путем повышения эффективности и/или снижения токсичности;
- направленная доставка лекарств, специфичная к типу ткани, или даже на клеточном или субклеточном уровне;
- улучшение фармацевтических свойств терапевтических молекул – стабильности, растворимости, времени циркуляции в организме, повышение концентрации лекарства в больном органе;
- возможность управляемого высвобождения лекарства [156];
- возможность направленной внутриклеточной доставки терапевтических средств на основе биологических макромолекул – ДНК, малые интерферирующие РНК, мРНК и др. [157];
- одновременную доставку (co-delivery) сложных терапевтических средств для повышения эффективности и преодоления лекарственной резистивности, возможность более точного подбора соотношения компонентов сложного лекарства с учетом конкретного пациента (так называемая “персонализированная медицина”) [158];
- возможность успешного транцитоза лекарств, т.е. преодоления эндотелиального (например, гематоэнцефалического) и эпителиального (гастроэнтерологического) барьеров;
- возможность более точных методов диагностики (визуализации);

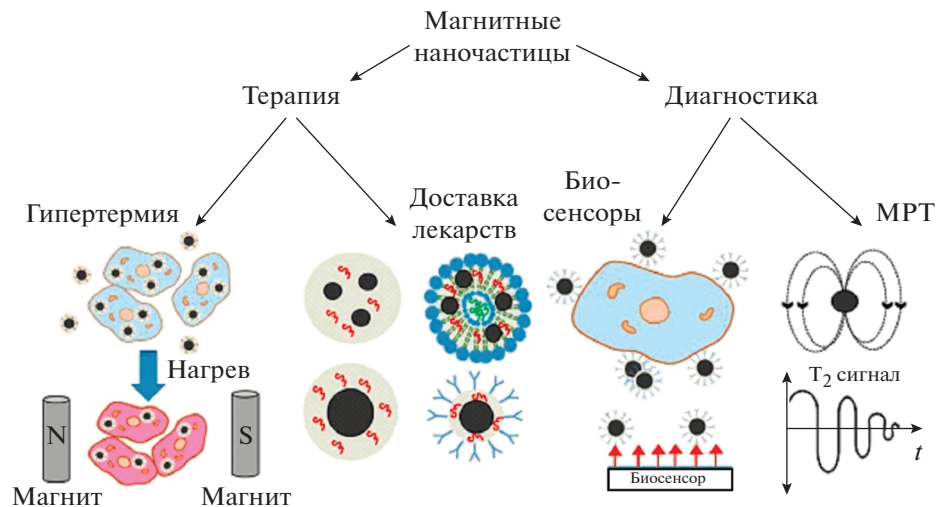


Рис. 4. Наглядное представление основных видов медицинского использования МНЧ. Адаптировано из [159]. Показаны схемы двух основных терапевтических методов (гипертермии и направленной доставки лекарств), а также факторы, ключевые для использования МНЧ в диагностических целях: специфичность связывания для биосенсоров и влияние магнитного поля МНЧ на время магнитной релаксации T_2 в методе МРТ.

– возможность одновременной диагностики и терапии, в том числе в режиме реального времени, так называемая тераностика;

– возможность использования собственных специфических свойств наночастиц для терапии, например, гипертермия с применением МНЧ.

Обычно медицинское применение МНЧ рассматривается либо как терапевтическое либо как диагностическое (рис. 4) [159]. Однако МНЧ могут быть и многофункциональными, т.е. одновременно и диагностическими и терапевтическими (т.е. тераностическими) препаратами [100, 160–162]. Для каждой конкретной практической медицинской задачи МНЧ модифицируются, чтобы улучшить уже имеющиеся или получить особые специфические свойства, т.е. функционализируются. Это необходимо, в частности, потому, что наночастица лекарства до достижения “мишени” должен свободно находиться в биологических жидкостях сложного химического состава (кровь, лимфа, цитоплазма и пр.). При достижении “мишени”, носитель должен специфически провзаимодействовать с ней, что также требует модификации МНЧ. Если МНЧ используются с диагностической целью, радиоактивные или флуоресцентные молекулы должны быть заранее прикреплены к поверхности МНЧ. На рис. 5 схематично показаны возможные виды функционализации МНЧ [163, 164].

Важной задачей является контролируемое высвобождение лекарства из защитной оболочки или везикулы тогда, когда наночастица достиг мишени [117, 165]. В идеале, функционализация МНЧ должна позволить управлять процессом высвобождения лекарств, используя различные виды стимуляции [166] и, кроме того, дать воз-

можность учитывать индивидуальные особенности пациента, т.е. применять персонализированную наномедицину [70, 167].

3. УСПЕХИ МЕДИЦИНСКОГО ПРИМЕНЕНИЯ МАГНИТНЫХ НАНОЧАСТИЦ

Примеры успешного медицинского применения МНЧ можно условно разделить на две группы: 1) внедрение метода, использующего МНЧ, в реальную медицинскую практику; 2) разработку метода или его отдельной части, обещающего преимущества по сравнению с аналогичными методами, используемыми в практической деятельности медиков, но пока не разрешенного к клиническому использованию. Примеров первого рода гораздо меньше, чем второго [42, 134, 168–172]. Причины этого обсуждаются в следующем разделе, здесь же отметим, что одной из основных причин медленного внедрения МНЧ в медицинскую практику являются строгие критерии допуска новых лекарственных средств и методик, ужесточающиеся с каждым годом. Наиболее известными организациями, принимающими или отклоняющими новые лекарства, являются Американское управление по санитарному надзору за качеством пищевых продуктов и медикаментов (US Food and Drug Administration, FDA) и Европейское медицинское агентство (EMA) [134, 106, 107, 162, 173]. Кроме того, многие аспекты медицинских устройств и приборов регулирует Международная организация по стандартизации (ISO) (www.iso.org). Чтобы дойти до финальной стадии реального медицинского применения

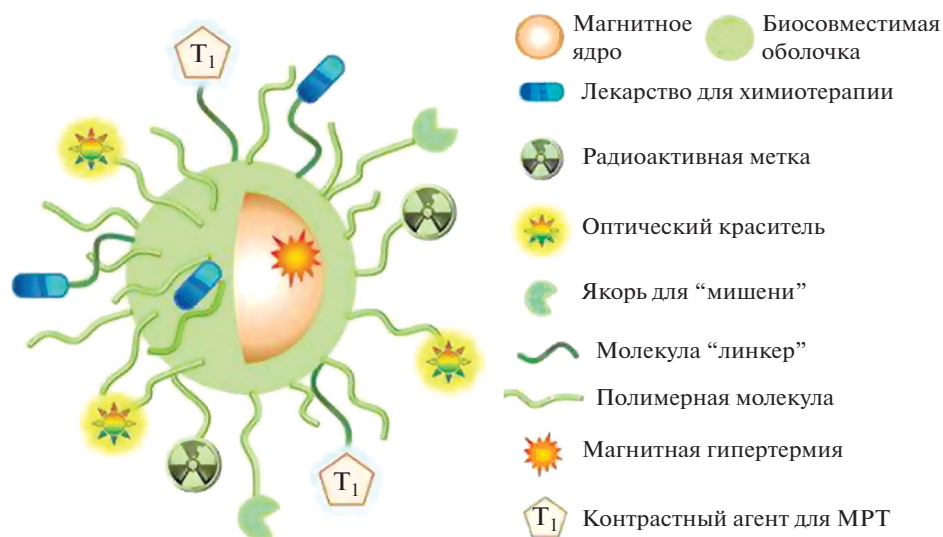


Рис. 5. Возможные виды функционализации МНЧ. Адаптировано из [164]. “Линкер” – искусственно синтезированная двухцепочечная молекула ДНК. “Якорь” – лиганд, специфичный для конкретной мишени (молекулы или молекулярного комплекса).

новое лекарственное средство должно пройти несколько стадий [69, 174, 175]. Первая стадия включает фундаментальные исследования, завершающиеся детальной разработкой научных основ всех элементов нового метода. В процессе реализации второй стадии (доклинических испытаний) необходимо *in vitro* и *in vivo* (на адекватной предлагаемому к лечению органу человека модели животного) определить множество характеристик нового лекарственного метода, например, особенности метаболизма лекарства, его относительную эффективность в сравнении с другими лекарствами, оптимальную дозу, оптимальный метод доставки, побочные эффекты, в частности цитотоксичность, возможное взаимодействие с другими лекарствами. Организация и проведение доклинических испытаний должны соответствовать рекомендациям “good laboratory practices” [176–178]. Стадия клинических испытаний включает четыре фазы. Первая обычно задействует менее 20 здоровых добровольцев, длится несколько месяцев и ставит основной целью найти безопасную дозу. Вторая фаза клинического исследования проходит с участием добровольцев-пациентов (обычно несколько сотен) и ставит основной целью найти оптимальную дозу. Третья фаза во многом дублирует вторую, но число добровольцев может достигать нескольких тысяч человек. Цель третьей фазы – подтвердить результаты двух предыдущих фаз и выявить максимально возможное число побочных эффектов. На основе результатов третьей фазы новое лекарство регистрируется. На четвертой фазе (пострегистрационной) число добровольцев еще увеличивается, проверяется статистическая достоверность данных об эффективности и безопасности,

полученных на предыдущих этапах исследования. Вторая, третья и четвертая фазы клинических испытаний могут длиться по несколько лет. Нередко от рождения концепции нового лекарственного средства до его коммерциализации проходит 10–20 лет [140]. На протяжении всего процесса клинических испытаний и после их завершения результаты всех исследований, а также патентная информация и информация о предполагаемом производителе передаются в уполномоченному выдавать разрешение на практическое использование организацию, например, FDA. После завершения клинических испытаний в случае одобрения после тщательной экспертизы, в частности, сравнения эффективности и безопасности уже существующих лекарств и методов с предлагаемыми, новое лекарство или метод могут производиться как коммерческий продукт. Но и после начала реального применения этого продукта FDA или аналогичная организация инспектирует заводы производители, выявляет случаи не выявленных ранее побочных эффектов, контролирует выпуск дженериков после окончания действия патента. Нередки случаи отзыва разрешения на практическое использования ранее одобренных лекарственных препаратов [179]. Кроме того, даже внедренные в практику препараты могут оказаться не оправдывающими всех первоначальных, возникших на стадии фундаментальных исследований ожиданий. Так, капсулированный в липосомы доксирубицин (торговая марка Doxil®), одобренный FDA в 1995 г. и входящий в тройку наиболее коммерчески успешных нанопрепаратов, не увеличивает существенно время жизни использующих его пациентов, хотя и существенно снижает

токсичность [61], т.е. повышает качество жизни пациентов.

Наиболее известный пример успешного доведения до коммерческого использования нанолечения с МНЧ, в котором используются именно магнитные свойства наночастиц – препарат для лечения глиобластомы головного мозга NanoTherm[®], производимый компанией MagForce AG (www.magforce.com), одобренный ЕМА в 2011 г. [180–184]. Использование NanoTherm[®] предполагает внутривенное введение покрытых аминокислотной оболочкой наночастиц оксида железа со средним диаметром 12 нм непосредственно в опухоль (для этого требуется предварительная трепанация черепа) и нагрев переменным магнитным полем частотой 100 кГц, создаваемым специальной системой NanoActivator[®]. Магнитная гипертермия сочетается с химиотерапией и радиотерапией. Хотя после внедрения в терапевтическую практику NanoTherm[®] глиобластома головного мозга по-прежнему остается неизлечимым заболеванием, однако среднюю продолжительность жизни после первого диагностирования этой болезни удалось увеличить более чем в два раза [181]. Успех NanoTherm[®] при лечении глиобластомы привел к попыткам его более широкого использования. В настоящее время продолжаются клинические испытания NanoTherm[®] для лечения рака простаты [185] и некоторых других видов онкологии [184, 186].

Другой пример реального успешного медицинского использования МНЧ – препарат Ferumoxytol (коммерческие названия Ferahem[®] в США, Rienso[®] в ЕС) [187–189]. Первоначально он был одобрен FDA в 2009 г. для лечения анемии больных с почечной недостаточностью, для которых обычные препараты были либо токсичны, либо малоэффективны [190]. Ferumoxytol представляет собой коллоидный раствор наночастиц магнетита со средним диаметром около 8 нм, покрытых оболочкой из полусинтетического углевода (размер вместе с оболочкой около 30 нм) [191]. Оболочка изолирует биоактивное железо от компонентов плазмы после инъекции до тех пор, пока частицы не будут поглощены макрофагами ретикулоэндотелиальной системы печени, селезенки и костного мозга. Железо высвобождается из углеводной оболочки в везикулах внутри макрофагов и либо накапливается внутриклеточно в виде ферритина, либо транспортируется трансферрином в плазму для включения в гемоглобин во время эритропоэза [192]. Ferumoxytol вводится внутривенно и обеспечивает достаточно быстрое восстановление уровня железа в организме без существенных побочных эффектов [188].

Хотя наночастицы Ferumoxytol суперпарамагнитные, это их свойство никак не используется при лечении анемии. Однако для целей МРТ нужны именно суперпарамагнитные наночасти-

цы. Поэтому Ferumoxytol начали исследовать для проверки возможности его использования как контрастного агента в диагностике методом МРТ широкого круга болезней: рассеянного склероза, панкреатита, кардиологических и онкологических заболеваний и др. [169, 193, 194]. На сайте www.clinicaltrials.gov Международного реестра клинических исследований Национального института здоровья США некоторые такие исследования отмечены как законченные (на ранних фазах). Вопрос, будут ли эти исследования продолжены и будет ли Ferumoxytol использоваться в качестве контрастного агента для МРТ, остается открытым. Схожий с Ferumoxytol препарат Ferumoxtran-10 (другое название Ferrotran, коммерческие названия – Sinerem[®], Combidex[®]), в котором оболочка наночастиц магнетита образована полисахаридом декстраном, изначально предназначался для применения в МРТ-диагностике рака лимфатических узлов и был разрешен к применению в некоторых европейских странах [169]. Однако в 2005 г. в FDA и 2007 г. в ЕМА заявки были отозваны из-за появившихся сообщений о побочных эффектах. В настоящее время почти в половине всех МРТ-анализов с использованием контрастных агентов применяются различные соединения гадолиния [97, 99]. Однако из-за накопления ионов гадолиния в тканях (в том числе в мозге) эти соединения могут быть небезопасными, особенно при неоднократном использовании [195]. Поэтому вероятность использования Ferumoxytol и других видов SPIONs в качестве замены гадолиний-содержащих агентов в МРТ остается достаточно большой [196–199], даже несмотря на сообщения о возможных побочных эффектах при использовании Ferumoxytol в качестве лекарства от анемии [200] и в целом не вполне ясную ситуацию со степенью токсичности МНЧ на основе оксидов железа [201]. Здесь, по-видимому, нужно сделать выбор: либо вообще отказаться от использования SPIONs из-за возможных побочных эффектов [202], либо все же использовать SPIONs в методе МРТ для определенных типов онкологии, где другие виды контрастных агентов неприемлемы [203, 204]. Второе решение, с нашей точки зрения, выглядит предпочтительней.

У МНЧ есть уникальное свойство, отличающее их от других видов нанопрепаратов – ими можно управлять с помощью внешнего магнитного поля, что делает их незаменимыми в методах “направленной доставки лекарств”. Однако успехи клинического применения МНЧ в этой области пока очень скромные [171, 205]. Единственный на сегодняшний день пример направленной доставки лекарства с помощью МНЧ – работы [1, 77], не вышедшие, однако, за рамки второй фазы клинических испытаний. Это обусловлено, в частности, ограниченной применимостью использованного в [1, 77] метода – только к поверх-

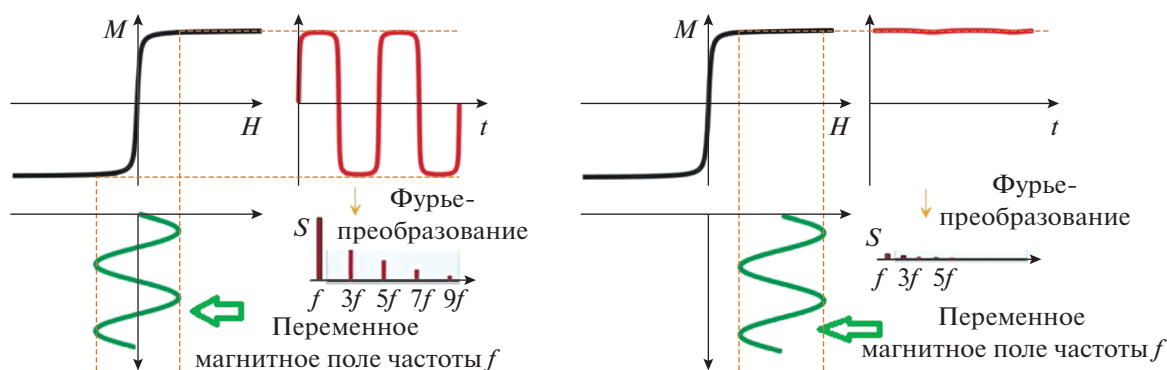


Рис. 6. Принцип метода MRI, позволяющего визуализировать пространственное расположение суперпарамагнитных МНЧ, для которых поле магнитного насыщения отличается как от макроскопических ферромагнетиков, так и от парамагнетиков. Адаптировано из [207]. Повернутая на 90° синусоида показывает зависимость от времени переменной компоненты внешнего магнитного поля. Постоянная компонента магнитного внешнего поля отсутствует на левой панели и положительна на правой панели, где величина суммарного внешнего поля лежит в области насыщения намагниченности МНЧ. Регистрируемый сигнал магнитного отклика МНЧ и амплитуды S его фурье-гармоник показаны столбиками на правой стороне каждой панели.

ностным опухолям. Это заведомо снижает число добровольцев-пациентов и не дает полноценно провести статистически достоверные клинические испытания с достаточно большим числом участников, что требуется для одобрения методики регулятором. В работах [1, 77] SPIONs диаметром около 100 нм, химически связанные с эпирубицином [206], вводили в кровеносные сосуды вблизи опухоли на поверхности лица или шеи пациента. Постоянный магнит, создающий магнитное поле с индукцией 0.8 Тл, помещали вблизи опухоли, чтобы он притягивал к ней МНЧ с лекарством. Процедура применялась к 14 пациентам, для которых предыдущие этапы лечения по традиционным схемам не привели к заметному успеху. Кроме того, все опухоли были неоперабельны. Хотя благодаря применению метода направленной доставки доксирубицина с помощью МНЧ некоторых успехов в лечении поверхностных опухолей удалось достичь, в целом пока полученные результаты не вполне убедительны [1].

В литературе можно найти многочисленные примеры исследований МНЧ, потенциально интересных для практического медицинского применения, но не дошедшие до клинических испытаний, т.е. проводившихся либо на культуре ткани, либо на животных (см., например, недавний обзор [159]).

Крайне важным при использовании МНЧ в терапевтических или диагностических медицинских методах является прямой контроль пространственного расположения наночастиц внутри организма пациента. Особенно это актуально в методах с управляемым высвобождением лекарства в целевой области организма, которое нужно производить только после того, как МНЧ достигнет органа-«мишени». Решающая эту задачу экспери-

ментальная радиоэлектронная методика, получившая название MPI (magnetic particle imagine), была предложена в 2005 г. [207]. В основе метода лежит специфическое свойство МНЧ – суперпарамагнетизм. В отличие от макроскопических ферромагнетиков и парамагнетиков, насыщение намагниченности которых при комнатной температуре происходит в очень малых (~1 Э) и очень больших (~1 МЭ) магнитных полях соответственно, намагниченность суперпарамагнетиков выходит на плато во внешних магнитных полях с напряженностью величиной около 1 кЭ, легко достижимой с помощью обычных лабораторных источников магнитного поля. Зависимость намагниченности M от напряженности H внешнего магнитного поля, характеризующаяся участками насыщения, на которых намагниченность практически не изменяется, показана на рис. 6. Сканер MPI содержит систему из шести катушек, создающих ортогональные друг другу квазистатические магнитные поля, сумма которых обеспечивает смещение (т.е. сканирование) в пространстве области нулевого поля (field-free point, FFP), на которое накладывается гармонически меняющееся переменное магнитное поле с амплитудой, позволяющей проходить нелинейный участок кривой намагничивания (левая панель рис. 6). Приемная катушка сканера регистрирует зависимость намагниченности от времени, которая будет нелинейной, только если в область FFP попали МНЧ. В этом случае фурье-преобразование сигнала приемной катушки даст заметные амплитуды гармоник, на частотах, кратных частоте переменного магнитного поля. Если МНЧ находятся вне области FFP, соответствующий сигнал приемной катушки будет почти постоянен и амплитуда фурье-гармоник будет мала (правая панель рис. 6).

Метод МРІ получил дальнейшее развитие в работах [208, 209]. В работе [209] описана пилотная установка, позволяющая применять МРІ к исследованиям на мозге человека. Состояние теории и техники эксперимента МРІ на 2021 г. подробно описано в обзоре [186].

Завершая этот раздел, в целом можно сделать вывод, что пока применение МНЧ в медицине не соответствует ожиданиям, которые на них возлагали на начальных этапах развития наномедицины [210–213]. Тем не менее за последние 20 лет удалось доказать принципиальную возможность практической реализации магнитной гипертермии [181], направленной магнитной управляемой доставки лекарств и снижения благодаря этому их токсичности [1, 77], способности МНЧ играть роль контрастных агентов в МРТ [191]. Таким образом, особые физические свойства МНЧ, которые рассматриваются как их преимущества при создании новых лекарственных препаратов и методов диагностики и терапии, вполне реальны и обуславливают необходимость и перспективность дальнейших исследований в этой области.

4. ПРОБЛЕМЫ МЕДИЦИНСКОГО ПРИМЕНЕНИЯ МАГНИТНЫХ НАНОЧАСТИЦ

Как отмечалось выше, основная проблема медицинского применения МНЧ, как и других нанолечков, заключается в крайне медленном их практическом внедрении [111, 168, 214–216]. Эта проблема присуща и нанотехнологиям в целом [217]. На то, что проблема шире и очень актуальна, указывает вхождение в научный оборот терминов “кризис трансляции” [218], кризис невоспроизводимости [219, 220] и другие подобные словосочетания, выражающие озабоченность медленным прогрессом в данной области.

Причины, препятствующие широкому практическому использованию (“трансляции”) МНЧ можно условно разбить на три группы [66, 103, 111, 221]: I) научные (фундаментальные и технические); II) клинические (медицинские и социальные); III) маркетинговые (коммерческие).

Группа I включает: 1) недостаточное понимание механизмов взаимодействия МНЧ с биологическими компонентами и структурами организма при доставке лекарства в больной орган [222], а также 2) механизмов аккумуляции МНЧ с лекарством в самом органе; 3) различное поведение МНЧ в экспериментах *in vitro* и *in vivo*; 4) технические сложности реализации метода (или его отдельных частей), использующего МНЧ, препятствующие их использованию где-либо кроме отдельных высокотехнологичных лабораторий [223]; 5) плохая изученность процессов фармакокинетики с участием МНЧ; 6) неоднозначность данных о био-

безопасности [224, 225], 7) возможная неадекватность используемых животных моделей аналогам в организме человека [153, 226].

Группа причин II включает: 1) неприемлемую для многих научных коллективов дороговизну клинических испытаний (сотни миллионов долларов [227]); 2) неверно выбранную “мишень” (например, тип онкологической опухоли) для данного метода; 3) ошибки в выборе дозировки и длительности лечения; 4) плохую воспроизводимость результатов вследствие недостаточно строгого контроля за синтезом и характеристикой используемых МНЧ (т.е. отсутствие достаточно строгих стандартов для МНЧ); 5) возможное изменение свойств МНЧ при переходе от лабораторного синтеза к промышленному (проблема масштабирования); 6) разнообразие и труднопредсказуемость возможных побочных эффектов при проведении 3-й и 4-й фаз клинических испытаний на большой выборке добровольцев.

Проблемы группы III (маркетинга) нанолечков с МНЧ разнообразны, но главная – большая конкуренция, в том числе и из-за появляющихся новых традиционных (не “нано”) лекарственных средств, менее токсичных и более эффективных, чем их предшественники.

Можно схематично изобразить основные препятствия, возникающие на пути внедрения нанопрепаратов в медицинскую практику, в виде “горной гряды” (рис. 7).

Термин “биологические барьеры” (см. рис. 7) обобщает большое количество различных факторов [54]. Как указывалось выше, одним из основополагающих принципов наномедицины является требование, чтобы лекарство не влияло на здоровые ткани организма, что приводит к задаче “капсулирования” (создания защитных оболочек) лекарственных средств. Вместе с тем биологическая среда, воздействующая на лекарство в процессе доставки к больному органу, не должна менять его терапевтические свойства. Однако при поступлении даже капсулированных или поверхностно-модифицированных МНЧ в кровь вследствие опсонизации их поверхность быстро покрывается “биокорной”, состоящей преимущественно из молекул белка и липидов, существенно влияющей на их стабильность, участие в метаболизме, в том числе времени вывода из организма, и на его иммунный ответ [114, 228, 229]. Детали взаимодействия введенных в организм пациента лекарственных препаратов, содержащих МНЧ, и разнообразных биологических структур могут существенно зависеть от индивидуальных особенностей организма и поэтому должны тщательно изучаться при клинических испытаниях на значительном числе добровольцев. Медико-биологическая гетерогенность потенциальных потребителей нанолечков – се-

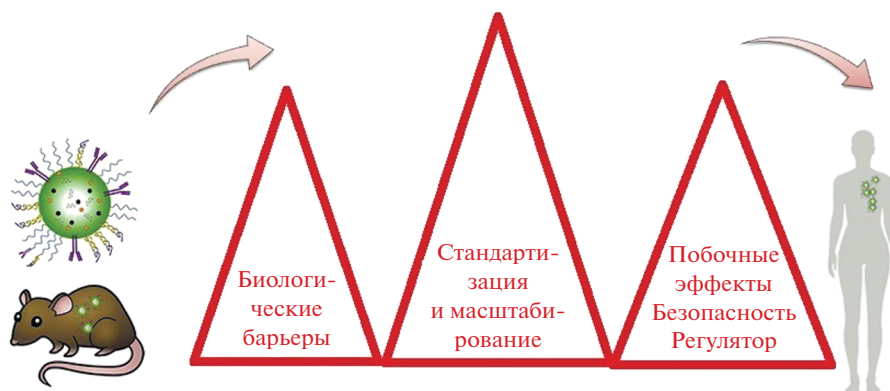


Рис. 7. Схематическое представление основных проблем с переходом от исследований нанолечеств в лаборатории *in vitro* и *in vivo* (на культурах ткани и животных моделях) к полноценному медицинскому применению. Адаптировано из [111]. Высота треугольников-“барьеров” может быть различна для различных видов нанолечеств.

рьезное препятствие на пути его практического внедрения.

Проблема выработки единых стандартов качества для МНЧ, предполагаемых к использованию в медицине, также остается пока нерешенной [230]. Хотя в целом лабораторный синтез наночастиц достиг высокого уровня совершенства [231], проблемой остается “невоспроизводимость” — ситуация, когда опубликованные результаты научного исследования не могут быть воспроизведены в других лабораториях (или, что хуже, даже в той, где проводились исследования) [232]. Особенно важно иметь точные сведения о токсичности МНЧ, если предполагается их медицинское применение. Среди важнейших и наиболее общих вопросов, которые требуют своего разрешения, можно выделить следующие [233].

1. Будут ли наночастицы проявлять сходные свойства и взаимодействия в различных органах и клетках? Зависят ли эти свойства и взаимодействия от возраста человека? Оказывает ли история перемещения частицы в организме какое-либо влияние на эти взаимодействия?

2. Как результаты исследований наночастиц *in vivo* и *in vitro* на клеточных культурах зависят от свойств (и происхождения) этих культур? Результаты исследований на одних и тех же (формально) клеточных культурах могут сильно различаться даже при использовании одних и тех же наноматериалов.

3. Может ли скорость введения частиц *in vivo* влиять на скорость агрегации частиц, приводящей к их локально повышенной концентрации и, как следствие, осаждению, тромбообразованию и тому подобным эффектам. Как скорость введения влияет на поглощение наночастиц и метаболизм в ключевых органах фильтрации, таких как печень, селезенка, капилляры легких и почки, и

не возникнут ли при этом токсичные побочные продукты?

4. Как функционализация наночастиц изменяет их токсичность *in vivo*?

До сих пор мы перечисляли проблемы медицинского применения наночастиц общие для всех видов нанолечеств. Однако магнитные свойства МНЧ обуславливают и некоторые специфические проблемы [41]. Первая из этих проблем имеет фундаментальный характер. Так, направленная доставка лекарственных препаратов с помощью статического магнитного поля возможна только к больным органам, находящимся на поверхности человеческого тела [205]. Как известно [234], частица сферической формы во внешнем магнитном поле испытывает силу, направленную по градиенту квадрата напряженности H^2 . Чтобы локализовать МНЧ в заданной области пространства, как это требуется в методе управляемой доставки лекарств, нужно создать магнитное поле, имеющее в некоторой точке внутри области локализации максимум величины H^2 . Возникает вопрос, возможно ли это. Ответ дает магнитостатическая теорема Томсона [234]: в области, в которой отсутствуют источники рассматриваемого магнитостатического поля квадрат напряженности H^2 , как функция пространственных координат, может иметь в качестве экстремума только минимум, но не максимум. Из теоремы Томсона следует, что суперпарамагнитные наночастицы не могут находиться в устойчивом равновесии под действием только магнитостатических сил.

Это вносит трудности при использовании МНЧ в методах направленной доставки лекарств [235, 236]. Очевидно, что для направленной доставки лекарственных препаратов с помощью МНЧ удобно использовать именно статическое магнитное поле, создаваемое постоянными маг-

нитами или соленоидами [85]. Именно постоянные магниты применялись в клинических испытаниях [77] для лечения поверхностных опухолей.

Проблема может быть в принципе решена, если применять динамическое управление МНЧ с использованием магнитного поля, медленно изменяющегося по определенному алгоритму [235, 237]. Однако это сильно усложняет и без того технически непростой магнитный метод направленной доставки лекарств. Другой путь решения указанной проблемы — использование миниатюрного ферромагнитного стента, внедряемого в опухоль и становящегося центром притяжения МНЧ [238]. Однако при этом теряется важное преимущество магнитного метода доставки лекарств — неинвазивность.

Есть и другие специфические проблемы медицинского использования МНЧ — повышенная склонность к агрегированию из-за магнитных диполь-дипольных взаимодействий [239], необходимость создания значительной степени неоднородности магнитных полей [240]. Первая из указанных проблем принципиально решается соответствующей модификацией поверхности МНЧ, подбором правильной концентрации и магнитных параметров частиц, вторая проблема — техническая, она может быть решена правильной пространственной конфигурацией источников магнитного поля.

5. ПЕРСПЕКТИВЫ МЕДИЦИНСКОГО ПРИМЕНЕНИЯ МАГНИТНЫХ НАНОЧАСТИЦ

Перед наномедициной, частью которой являются методы, использующие МНЧ — управляемые магнитным полем гипертермия, направленная доставка и управляемое высвобождение лекарства, магниторезонансные методы диагностики, в настоящее время стоят сложные задачи [103, 111, 172, 241]. Суть этих задач — преодоление многочисленных проблем, описанных в предыдущих разделах. Чтобы успешно разрешить эти проблемы, необходимо в том числе устранить некоторые перекосы к организации научных исследований. До сих пор большинство исследований проводились коллективами, включающими химиков, физиков, биологов, материаловедов, инженеров и лишь изредка практикующих медиков. Тем самым, уже в начале исследовательского процесса конечная задача — создание нового более эффективного и безопасного по сравнению с существующими лекарственного нанопрепарата, предназначенного для лечения конкретной болезни (а не абстрактных “онкологических заболеваний”), отодвигается на будущее, а основное внимание уделяется получению наноконструкций с очень хорошими физико-химическими свойствами, критерий качества для которых, однако, может быть никак не связан с конечной целью. Эта

особенность современных работ по наномедицине получила название “эффект невидимой гориллы” [63]. На русском языке содержание этого эффекта можно выразить пословицей “не видеть леса за деревьями”.

Другая отрицательная особенность научных работ по наномедицине последних 20–30 лет, требующая преодоления, в англоязычной литературе получила название “publication bias” или “the file drawer problem” [242, 243]. Она состоит в том, что в подавляющем большинстве исследовательских статей сообщают только о положительных результатах, не упоминая о негативных. То же относится к большинству обзорных статей по наномедицине, в которых основное внимание уделяется успехам и в гораздо меньшей степени обсуждаются неудачи или “подводные камни”. В случае поиска новых лекарственных средств и разработки новых методов диагностики и терапии знание об отрицательных или негативных результатах не менее важно, чем о позитивных.

Еще одна важная задача, без решения которой невозможен прогресс наномедицины — стандартизация характеристик нанообъектов для того, чтобы иметь возможность сравнивать как доклинические результаты исследовательских работ, так и, что более важно, результаты клинических испытаний [92, 233]. Необходимо ввести перечень стандартных и обязательных для измерения физико-химических характеристик МНЧ, в частности, степени полидисперсности (разброса по размерам), степени токсичности, патогенности, биоразлагаемости, склонности к агрегации, основных магнитных свойств (для МНЧ) и др. Все наиболее важные характеристики, влияющие на токсичность, биодоступность, иммуногенность должны быть экспериментально измерены (определены) и подробно описаны, чтобы правильно и полно интерпретировать результаты работы и позволить сравнить их с ранее опубликованными данными. Разумеется, такая стандартизация должна учитывать специфику отдельных разделов наномедицины и особенности различных заболеваний.

Научные группы, создающие биомедицинские методы с использованием МНЧ, могут не синтезировать их сами, а воспользоваться уже готовыми. В настоящее время 14 коммерческих фирм предлагают МНЧ научно-исследовательского уровня [92]. Из них один тип частиц (Ferrotran) проходит клинические испытания, и еще несколько типов VivoTrax™, VivoTrax Plus™, FeraSpin™, RCL-01, panomag®, Perimag®, Synomag® готовы к ним.

Во многом успех будущих медицинских применений МНЧ зависит от развития радиоэлектронной технологии МРИ визуализации магнитных наночастиц внутри организма [186, 207]. Невозможно контролировать действие лекарства, адресно до-

ставляемого с помощью МНЧ к больному органу, не имея прямой информации об успешности этой доставки. Кроме того, с использованием технологии МРІ недавно удалось применить тераностический подход на животной модели — визуализация МНЧ сопровождалась процессом гипертермии [244]. Стремление к многофункциональности, сочетанию терапевтического и диагностического эффектов, будет, по-видимому, основной тенденцией будущих исследований, связанных с применением МНЧ в медицине [245–247].

ЗАКЛЮЧЕНИЕ

Проведенный нами анализ успехов и проблем биомедицинских применений МНЧ и использующих такие наночастицы методов биомедицинской радиоэлектроники показывает, что их потенциальные возможности в этой области раскрыты пока далеко не в полной мере. Лишь в единичных случаях и в ограниченном виде удалось на практике реализовать методы магнитной гипертермии и управляемой направленной доставки лекарств. Хотя практическое использование МНЧ для диагностических целей более успешно, но и здесь возникает много проблем, связанных с неизбежными побочными эффектами. В научном сообществе постепенно созревает понимание необходимости изменений как в самом подходе к научным исследованиям в этой области, так и в содержании этих исследований. Мультидисциплинарность исследований должна расширяться за счет экспертов в медицине и физиологии. Кроме продолжения исследований по оптимизации процессов направленной доставки и управляемого высвобождения лекарственных препаратов, гипертермии, особенно важно развивать физические методы визуализации МНЧ в живом организме, а также мультифункциональные методы и подходы, эффективно решающие одновременно несколько диагностических и терапевтических задач.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках государственного задания (номер АААА-А19-119041590070-1) и при финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-12-50280).

СПИСОК ЛИТЕРАТУРЫ

1. *Lübbe A.S., Alexiou C., Bergemann C.* // J. Surgical Research. 2001. V. 95. P. 200.
2. *Muzykantov V., Torchilin V.* Biomedical Aspects of Drug Targeting. N.Y.: Springer, 2002.
3. Nanocarrier Technologies Frontiers of Nanotherapy / Ed. M.R. Mozafari. N.Y.: Springer, 2006.
4. Biomedical Applications of Nanotechnology / Eds. V. Labhasetwar, D.L. Leslie-Pelecky. Hoboken: John Wiley & Sons, 2007.
5. Micro and Nano Technologies in Bioanalysis: Methods and Protocols / Eds. J.W. Lee, R.S. Foote. N.Y.: Humana Press, 2009.
6. Oral Controlled Release Formulation Design and Drug Delivery. Theory to Practice / Eds. H. Wen, K. Park. Hoboken: John Wiley & Sons, 2010.
7. *MacEwan S.R., Chilkoti A.* // Angew. Chem. Int. Ed. 2017. V. 56. P. 6712.
8. *Ventola C.L.* // Pharmacy and Therapeutics. 2012. V. 37. № 9. P. 512.
9. *Ventola C.L.* // Pharmacy and Therapeutics. 2012. V. 37. № 10. P. 582.
10. *Ventola C.L.* // Pharmacy and Therapeutics. 2012. V. 37. № 11. P. 631.
11. Multifunctional Nanoparticles for Drug Delivery Applications / Eds. S. Svenson, R.K. Prud'homme. N.Y.: Springer, 2012.
12. *Chauhan V.P., Jain R.K.* // Nature Materials. 2013. V. 12. P. 958.
13. *Karimi Z., Karimi L., Shokrollahi H.* // Materials Science and Engineering C. 2013. V. 33. P. 2465.
14. Drug Delivery Systems. Strategies, Technologies and Applications / Ed. Y. Yeo. Hoboken; New Jersey: John Wiley & Sons, 2013.
15. *Etheridge M.L., Campbell S.A., Erdman A.G. et al.* // Nanomedicine: Nanotechnology, Biology, Medicine 2013. V. 9. P. 1.
16. Nanoparticulates as Drug Carriers / Ed. V.P. Torchilin. L.: Imperial College Press, 2006.
17. Handbook of Nanobiomedical Research: Fundamentals, Applications, and Recent Developments. Frontiers in Nanobiomedical Research: V. 3. / Ed. V.P. Torchilin. Singapore: World Sci. Publ. Company, 2014.
18. *Park K.* // J. Control. Release. 2014. V. 190. P. 3.
19. Protein and Peptide Nanoparticles for Drug Delivery / Ed. R. Donev. Oxford: Acad. Press, 2015.
20. Nanoparticles in the Lung: Environmental Exposure and Drug Delivery / Eds. A. Tsuda, P. Gehr. Boca Raton: CRC Press, 2015.
21. *Hillery A.M., Park K.* Drug Delivery. Fundamentals and Applications. Boca Raton: CRC Press, 2016.
22. *Bhatia S.* Natural Polymer Drug Delivery Systems. N.Y.: Springer, 2016.
23. Nanomedicines Design, Delivery and Detection / Ed. M. Braddock. Cambridge: Royal Society of Chemistry, 2016.
24. *Agotegaray M.A., Lassalle V.L.* Silica-coated Magnetic Nanoparticles: An Insight into Targeted Drug Delivery and Toxicology. N.Y.: Springer, 2017.
25. *Kumar P., Srivastava R.* Nanomedicine for Cancer Therapy: From Chemotherapeutic to Hyperthermia-Based Therapy. N.Y.: Springer, 2017.
26. Characterization of Nanoparticles Intended for Drug Delivery / Ed. S.E. McNeil. N.Y.: Humana Press, 2018.
27. *Hui H., Yi X., Hou F. et al.* // ACS Nano. 2019. V. 13. P. 7410.

28. Ferrari E., Soloviev M. Nanoparticles in Biology and Medicine: Methods and Protocols. N.Y.: Humana Press, 2020.
29. Nanoparticles in Medicine / Ed. A.K. Shukla. N.Y.: Springer, 2020.
30. Lammers T., Kiessling F., Hennink W.E., Storm G. // J. Control. Release. 2012. V. 161. P. 185.
31. Lammers T., Ferrari M. // Nano Today. 2020. V. 31. P. 100853.
32. Martins J.P., das Neves J., de la Fuente M. et al. // Drug Delivery and Translational Research. 2020. V. 10. P. 726.
33. Mucosal Delivery of Drugs and Biologics in Nanoparticles / Eds. P. Muttill, N.K. Kunda. N.Y.: Springer, 2020.
34. Gerrard J.A., Domigan L.J. Protein Nanotechnology, Protocols, Instrumentation and Applications. N.Y.: Springer, 2020.
35. Antibacterial Drug Discovery to Combat MDR. Natural Compounds, Nanotechnology and Novel Synthetic Sources / Eds. I. Ahmad, S. Ahmad, K.P. Rumbaugh. N.Y.: Springer, 2020.
36. Nanostructured Biomaterials for Regenerative Medicine / Eds. V. Guarino, M. Iafisco, S. Spriano. N.Y.: Elsevier, 2020.
37. Nanoscience in Medicine. V. 1 / Eds. Daima H.K., Navya P.N., Ranjan S. et al. N.Y.: Springer, 2020.
38. Model Organisms to Study Biological Activities and Toxicity of Nanoparticles / Eds. B. Siddhardha, M. Dyavaiah, K. Kasinathan. N.Y.: Springer, 2020.
39. Bueno J. Preclinical Evaluation of Antimicrobial Nanodrugs. N.Y.: Springer, 2020.
40. Richardson J.J., Caruso F. // Nano Lett. 2020. V. 20. P. 1481.
41. Lane L.A. // Appl. Phys. Rev. 2020. V. 7. ID 011316. <https://doi.org/10.1063/1.5052455>
42. Anik M.I., Hossain M.K., Hossain I. et al. // Nano Select. 2021. V. 2. P. 1.
43. Mirza Z., Karim S. // Seminars in Cancer Biology. 2021. V. 69. P. 226.
44. Nanomaterials for Application in Medicine and Biology / Eds. M. Giersig, G.B. Khomutov. Dordrecht: Springer, 2008.
45. Khomutov G.B., Kim V.P., Koksharov Yu.A. et al. // Colloids and Surfaces. A: Physicochemical and Engineering Aspects. 2017. V. 532. P. 26.
46. Antipina M.N., Sukhorukov G.B. // Adv. Drug Delivery Rev. 2011. V. 63. P. 716.
47. Handbook of Clinical Nanomedicine: Nanoparticles, Imaging, Therapy, and Clinical Applications / Eds. R. Bawa, G.F. Audette, I. Rubinstein. Boca Raton: CRC Press, 2016.
48. Гуляев Ю.В., Черепенин В.А., Вдовин В.А. и др. // РЭ. 2015. Т. 60. № 10. С. 1051.
49. Кислов В.В., Колесов В.В., Таранов И.В. // РЭ. 2002. Т. 47. № 11. С. 1385.
50. Gorin D.A., Shchukin D.G., Koksharov Yu.A. et al. // Proc. of SPIE. 2007. V. 6536. № 6. P. 653604.
51. Amstad E., Kohlbrecher J., Muller E. et al. // Nano Lett. 2011. V. 11. P. 1664.
52. Tai L.A., Tsai P.J., Wang Y.C. et al. // Nanotechnol. 2009. V. 20. № 1. P. 135101.
53. Гуляев Ю.В., Черепенин В.А., Таранов И.В., Вдовин В.А. и др. // РЭ. 2021. Т. 66. № 1. С. 82.
54. Nanostructured Biomaterials for Overcoming Biological Barriers / Eds. M.J. Alonso, N.S. Csaba. Cambridge: Royal Soc. Chemistry, 2012.
55. Bae Y.H., Park P. // J. Control. Release. 2011. V. 153. P. 198.
56. Cancer Targeted Drug Delivery. An Elusive Dream / Eds. Y.H. Bae, R.J. Mersny, K. Park. N.Y.: Springer, 2013.
57. Kwon I.K., Lee S.C., Han B., Park K. // J. Control. Release. 2012. V. 164. P. 108.
58. Kratz F., Warnecke A. // J. Control. Release. 2012. V. 164. P. 221.
59. Venditto V.J., Szoka F.C., Jr. // Adv. Drug Delivery Rev. 2013. V. 65. P. 80.
60. Wilhelm S., Tavares A.J., Dai Q. et al. // Nature Rev. Mater. 2016. V. 1. <https://doi.org/10.1038/natrevmats.2016.14>
61. Salvioni L., Rizzuto M.A., Bertolini J.A. et al. // Cancers. 2019. V. 11. <https://doi.org/10.3390/cancers11121855>
62. Allen C., Park K. // J. Controlled Release. 2019. V. 302. P. 203.
63. Park K. // J. Controlled Release. 2016. V. 240. P. 2.
64. Park K. // J. Controlled Release. 2019. V. 298. P. 213.
65. Park K. // J. Controlled Release. 2019. V. 305. P. 221.
66. He H., Liu L., Morin E.E. et al. // Acc. Chem. Res. 2019. V. 52. P. 2445.
67. Leroux J.-C. // Angew. Chem. Int. Ed. 2017. V. 56. P. 15170.
68. Torrice M. // ACS Cent. Sci. 2016. V. 2. P. 434.
69. Foulkes R., Man E., Thind J. et al. // Biomater. Sci. 2020. V. 8. P. 4653.
70. Rosenblum D., Peer D. // Cancer Lett. 2014. V. 352. P. 126.
71. Krishnan K.M. // IEEE Trans. 2010. V. Magn-46. P. 2523.
72. Krishnan K.M. Fundamentals and Applications of Magnetic Materials. Oxford: Oxford Univ. Press, 2016.
73. Clavijo-Jordan V., Kodibagkar V.D., Beeman S.C. et al. // WIREs Nanomedicine and Nanotechnology. 2012. V. 4. P. 345.
74. Laurent S., Saei A.A., Behzadi S. et al. // Expert Opinion on Drug Delivery. 2014. V. 11. P. 1449.
75. Xiong F., Huang S., Gu N. // Drug Development and Industrial Pharmacy. 2018. V. 44. P. 697.
76. Tong S., Zhu H., Bao G. // Materials Today. 2019. V. 31. P. 86.
77. Lübbe A.S., Bergemann C., Riess H. et al. // Cancer Research. 1996. V. 56. P. 4686.
78. Alexiou Ch., Schmidt A., Klein R. et al. // J. Magn. Mater. 2002. V. 252. P. 363.
79. Pankhurst Q.A., Connolly J., Jones S.K., Dobson J. // J. Phys. D: Appl. Phys. 2003. V. 36. P. R167.

80. *Gubin S.P., Koksharov Y.A., Khomutov G.B., Yurkov G.Y.* // Russ. Chem. Rev. 2005. V. 74. P. 489.
81. *Magnetic Nanoparticles* / Ed. S.P. Gubin. Hoboken: John Wiley & Sons, 2009.
82. *Reddy L.H., Arias J.L., Nicolas J., Couvreur P.* // Chem. Rev. 2012. V. 112. P. 5818.
83. *Tietze R., Zaloga J., Unterweger H. et al.* // Biochemical and Biophysical Research Commun. 2015. V. 468. P. 463.
84. *Ye H., Shen Z., Yu L. et al.* // Proc. R. Soc. A. 2018. V. 474. <https://doi.org/10.1098/rspa.2017.0845>
85. *Liu Y.-L., Chen D., Shang P., Yin D.-C.* // J. Control. Release. 2019. V. 302. P. 90.
86. *Hyperthermia in Cancer Treatment: A Primer* / Eds. G.F. Baronzio, E.D. Hager. N.Y.: Springer, 2006.
87. *Dennis C.L., Ivkov R.* // Int. J. Hyperthermia. 2013. V. 29. P. 715.
88. *Kozissnik B., Bohorquez A.C., Dobson J., Rinaldi C.* // Int. J. Hyperthermia. 2013. V. 29. P. 706.
89. *Raikher Yu.L., Stepanov V.I.* // J. Magn. Magn. Mater. 2014. V. 368. P. 421.
90. *Nedyalkova M., Donkova M., Romanova J. et al.* // Advances in Colloid and Interface Sci. 2017. V. 249. P. 192.
91. *Fatima H., Charinpanitkul T., Kim K.-S.* // Nanomaterials. 2021. V. 11. <https://doi.org/10.3390/nano11051203>
92. *Rubia-Rodríguez I., Santana-Otero A., Spassov S. et al.* // Materials. 2021. V. 14. <https://doi.org/10.3390/ma14040706>
93. *Shawa S.K., Kailashiya J., Gangwar A. et al.* // Appl. Surface Sci. 2021. V. 560. <https://doi.org/10.1016/j.apsusc.2021.150025>
94. *Weissleder R., Bogdanov A., Neuwelt E.A., Papisov M.* // Advanced Drug Delivery Rev. 1995. V. 16. P. 321.
95. *Okuhata Y.* // Advanced Drug Delivery Rev. 1999. V. 37. P. 121.
96. *Nanotechnology for Biomedical Imaging and Diagnostics. From Nanoparticle Design to Clinical Applications* / Ed. M.Y. Berezin. Hoboken: John Wiley & Sons, 2015.
97. *Wahsner J., Gale E.M., Rodríguez-Rodríguez A., Caravan P.* // Chem. Rev. 2019. V. 119. P. 957.
98. *Mirza S., Ahmad M.S., Shah M.I.A., Ateeq M.* Metal Nanoparticles for Drug Delivery and Diagnostic Applications / Eds. M.R. Shah, M.I.S. Ullah. N.Y.: Elsevier, 2019. P. 189.
99. *Chakravarty S., Shapiro E.M.* Modern Techniques in Biosensors / Eds. G. Dutta, A. Biswas, A. Chakrabarti. N.Y.: Springer, 2021. P. 135.
100. *Magnetic Nanomaterials. Applications in Catalysis and Life Sciences* / Eds. S.H. Bossmann, H. Wang. Cambridge: Royal Soc. of Chemistry, 2017.
101. *Gould P.* // Materials Today. 2004. V. 7. P. 36.
102. *Varadan V.K., Chen L., Xie J.* Nanomedicine Design and Applications of Magnetic Nanomaterials, Nanosensors and Nanosystems. Hoboken: John Wiley & Sons, 2008.
103. *Hua S., de Matos M.B.C., Metselaar J.M., Storm G.* // Frontiers in Pharmacology. 2018. V. 9. <https://doi.org/10.3389/fphar.2018.00790>
104. *Cardoso V.F., Francesko A., Ribeiro C. et al.* // Adv. Healthcare Mater. 2018. V. 7. <https://doi.org/10.1002/adhm.201700845>
105. *Nia H.T., Munn L.L., Jain R.K.* // Science. 2020. V. 370. <https://doi.org/10.1126/science.aaz0868>
106. *Biomedical Nanotechnology Methods and Protocols* / Ed. S.H. Petrosko. N.Y.: Humana Press, 2011.
107. *Biomedical Nanotechnology Methods and Protocols* / Eds. S.H. Petrosko, E.S. Day. N.Y.: Humana Press, 2017.
108. *Алексеев К.В., Блынская Е.В., Кедик С.А.* Фармацевтическая нанотехнология. М.: Ин-т фармацевт. технологий, 2012.
109. *Chemistry of Nanomaterials Fundamentals and Applications* / Eds. T.I. Awan, A. Bashir, A. Tehseen. Amsterdam: Elsevier, 2020.
110. *Moghim S.M., Hunter A.C., Murray J.C.* // The FASEB J. 2005. V. 19. P. 311.
111. *Wu L.-P., Wang D., Li Z.* // Mater. Sci. & Engineering C. 2020. V. 106. <https://doi.org/10.1016/j.msec.2019.110302>
112. *Witkop B.* // Proc. Amer. Philos. Soc. 1999. V. 143. P. 540.
113. *Альберт А.* Избирательная токсичность. Физико-химические основы терапии. М.: Медицина, 1989.
114. *Immune Aspects of Biopharmaceuticals and Nanomedicines* / Eds. R. Bawa, J. Szebeni, T.J. Webster, G.F. Audette. Singapore: Pan Stanford Publ., 2018.
115. *Winau F., Westphal O., Winau R.* // Microbes and Infection. 2004. V. 6. P. 786.
116. *Ratner B.D., Hoffman A.S., Schoen F.J., Lemons J.E.* Biomaterials Science. An Introduction to Materials in Medicine. N.Y.: Acad. Press, 2012.
117. *Hoffman A.S.* // J. Control. Release. 2008. V. 132. P. 153.
118. *Helfand W.H., Cowen D.L.* // Pharmacy in History. 1983. V. 25. P. 3.
119. *Folkman J., Long D.* // J. Surgical Research. 1964. V. 4. P. 139.
120. *Langer R., Folkman J.* // Nature. 1976. V. 263. P. 797.
121. *Duncan R., Kopecek J.* // Adv. Polymer Sci. 1984. V. 57. P. 51.
122. *Duncan R.* // Nature Rev. 2003. V. 2. P. 347.
123. *Maeda H.* // J. Personalized Medicine. 2021. V. 11. <https://doi.org/10.3390/jpm11030229>
124. *Ngoune R., Peters A., von Elverfeldt D. et al.* // J. Controlled Release. 2016. V. 238. P. 58.
125. *Birrenbach G., Speiser P.P.* // J. Pharm. Sci. 1976. V. 65. P. 1763.
126. *Nishiyama N., Kataoka K.* // Pharmacol. Ther. 2006. V. 112. P. 630.
127. *Kreuter J.* // Intern. J. Pharmaceutics. 2007. V. 331. P. 1.
128. *Lasic D.D., Martin F.J.* Stealth Liposomes. Boca Raton: CRC Press, 2018.

129. Liposomes. Methods and Protocols / Ed. G.G.M. D'Souza. N.Y.: Humana Press, 2017.
130. *Bitounis D., Fanciullino R., Iliadis A., Ciccolini J.* // ISRN Pharmaceutics. 2012. <https://doi.org/10.5402/2012/738432>
131. Liposome Technology. Liposome Preparation and Related Techniques / Ed. G. Gregoriadis. N.Y.: Informa Healthcare, 2006.
132. *Yang F., Jin C., Jiang Y. et al.* // Cancer Treatment Rev. 2011. V. 37. P. 633.
133. *Malam Y., Loizidou M., Seifalian A.M.* // Trends in Pharmacological Sci. 2009. V. 30. P. 592.
134. *Bobo D., Robinson K.J., Islam J. et al.* // Pharm Res. 2016. V. 33. P. 2373.
135. *Rommasi F., Esfandiari N.* // Nanoscale Res.Lett. 2021. V. 16. P. 95.
136. *Bozutto G., Milinari A.* // Intern. J. Nanomedicine. 2015. V. 10. P. 975.
137. *Ganguly K., Chaturvedi K., More U.A. et al.* // J. Controlled Release. 2014. V. 193. P. 162.
138. *Sevastre A.-I., Horescu C., Baloi S.C. et al.* // Coatings. 2019. V. 9. <https://doi.org/10.3390/coatings9100628>
139. *Mu W., Chu Q., Liu Y., Zhang N.* // Nano-Micro Lett. 2020. V. 12. P. 142.
140. *Farjadian F., Ghasemi A., Gohari O. et al.* // Nano-medicine. 2019. V. 14. P. 93.
141. *Gubin S.P., Koksharov Yu.A.* // Inorg. Mater. 2002. V. 38. P. 1085.
142. Magnetic Nanoparticles in Biosensing and Medicine / Eds. N.J. Darton, A. Ionescu, J. Llandro. Cambridge: Cambridge Univ. Press, 2019.
143. *Wu W., Wu Z., Yu T. et al.* // Sci. Technol. Adv. Mater. 2015. V. 16. № 2. P. 023501. <https://doi.org/10.1088/1468-6996/16/2/023501>
144. *McNamara K., Tofail S.A.M.* // Adv. Phys.: X. 2017. V. 2. P. 54.
145. *Geppert M., Himly M.* // Frontiers in Immunology. 2021. V. 12. <https://doi.org/10.3389/fimmu.2021.688927>
146. *Neuberger T., Schopf B., Hofmann H. et al.* // J. Magn. Magn. Mater. 2005. V. 293. P. 483.
147. *Gupta A.K., Gupta M.* // Biomaterials. 2005. V. 26. P. 3995.
148. *Martinkova P., Brtnicky M., Kynicky J., Pohanka M.* // Adv. Healthc. Mater. 2018. V. 7. <https://doi.org/10.1002/adhm.201700932>
149. *Vangijzegem T., Stanicki D., Laurent S.* // Expert. Opin. Drug. Delivery. 2019. V. 16. P. 69.
150. *Patil R.M., Thorat N.D., Shete P.B. et al.* // Biochem. Biophys. Rep. 2018. V. 13. P. 63.
151. *Dobrovolskaia M.A., Shurin M., Shvedova A.A.* // Toxicol. Appl. Pharmacol. 2016. V. 299. P. 78.
152. *Shah A., Dobrovolskaia M.A.* // Nanomedicine. 2018. V. 4. P. 977.
153. *Shi J., Kantoff P.W., Wooster R., Farokhzad O.C.* // Nature. 2017. V. 17. P. 20.
154. *Anderson S.D., Gwenin V.V., Gwenin C.D.* // Nanoscale Research Lett. 2019. V. 14. P. 188.
155. *Stueber D.D., Villanova J., Aponte I. et al.* // Pharmaceutics. 2021. V. 13. <https://doi.org/10.3390/pharmaceutics13070943>
156. *Longo R., Gorrasi G., Guadagno L.* // Nanomaterials. 2021. V. 11. <https://doi.org/10.3390/nano11040848>
157. *Moghimi S.M., Peer D., Langer R.* // ASC Nano. 2011. V. 5. P. 8454.
158. *Guchet X.* // Nanomedicine (Lond.). 2015. V. 10. P. 3167.
159. *Farzin A., Etesami S.A., Quint J. et al.* // Adv. Healthcare Mater. 2020. <https://doi.org/10.1002/adhm.201901058>
160. *Chen F., Ehlerding E.B., Cai W.* // J. Nuclear Medicine. 2014. V. 55. P. 1919.
161. *Gobbo O.L., Sjaastad K., Radomski M.W. et al.* // Theranostics. 2015. V. 5. P. 1249.
162. *Dadfar S.M., Roemhild K., Drude N.I. et al.* // Adv. Drug Delivery Rev. 2019. V. 138. P. 302.
163. *Hola K., Markova Z., Zoppellaro G. et al.* // Biotechnology Adv. 2015. V. 33. P. 1162.
164. *Zoppellaro G.* Magnetic Nanoheterostructures / Eds. S.K. Sharma, Y. Javed. N.Y.: Springer, 2020. P. 57.
165. *Liu J.F., Jang B., Issadore D., Tsourkas A.* // WIREs Nanomed. Nanobiotechnol. 2019. V. 11. <https://doi.org/10.1002/wnan.1571>
166. *Sun Y., Davis E.* // Nanomaterials. 2021. V. 11. <https://doi.org/10.3390/nano11030746>
167. *Zhang Y., Cheng M., Cao J. et al.* // Nanoscale. 2019. V. 11. <https://doi.org/10.1039/c8nr09347d>
168. *Singh D., McMillan J.M., Kabanov A.V. et al.* // Nanomedicine. 2014. V. 9. P. 501.
169. *Anselmo A.C., Mitragotri S.* // The AAPS J. 2015. V. 17. P. 1041.
170. *Anselmo A.C., Mitragotri S.* // Bioengineering & Translational Medicine. 2016. V. 1. P. 10.
171. Clinical Applications of Magnetic Nanoparticles / Ed. N.T.K. Thanh. N.Y.: CRC Press, 2018.
172. *Mitchell M.J., Billingsley M.M., Haley R.M. et al.* // Nature Rev. 2021. V. 20. P. 101.
173. *Soetaert F., Korangath P., Serantes D. et al.* // Adv. Drug Delivery Rev. 2020. V. 163–164. P. 65.
174. *Machin D., Day S., Green S. et al.* Textbook of Clinical Trials. Hoboken: John Wiley & Sons, 2004.
175. *Das R., Masa J.A., Kalappattil V. et al.* // Nanomaterials. 2021. V. 11. P. 1380.
176. *Baldeshwiler A.M.* // Qual. Assur. J. 2003. V. 7. P. 157.
177. *Adamo J.E., Bauer G., Derro M. et al.* // Acad. Med. 2012. V. 87. P. 279.
178. *Zielinska A., Costa B., Ferreira M.V. et al.* // Int. J. Environ. Res. Public Health. 2020. V. 17. <https://doi.org/10.3390/ijerph17134657>
179. *Wysowski D.K., Swartz L.* // Arch. Intern. Med. 2005. V. 165. P. 1363.
180. *Maier-Hauff K., Rothe R., Scholz R. et al.* // J. Neuro-Oncology. 2007. V. 81. P. 51.
181. *Maier-Hauff K., Ulrich F., Nestler D. et al.* // J. Neuro-Oncology. 2011. V. 103. P. 317.

182. *Thiesen B., Jordan A.* // Intern. J. Hyperthermia. 2008. V. 24. P. 467.
183. Nanoparticles for Biomedical Applications Fundamental Concepts, Biological Interactions and Clinical Applications / Eds. E.J. Chung, L. Leon, C. Rinaldi. N.Y.: Elsevier, 2019.
184. *Etemadi H., Plieger P.G.* // Adv. Therap. 2020. V. 3. <https://doi.org/10.1002/adtp.202000061>
185. *Johannsen M., Gneveckow U., Thiesen B. et al.* // Eur. Urol. 2007. V. 52. P. 1653.
186. *Rivera-Rodriguez A., Rinaldi-Ramos C.M.* // Annual Rev. Chem. Biomol. Eng. 2021. V. 12. P. 163.
187. *Danielson B.G.* // J. Amer. Soc. Nephrol. 2004. V. 15. P. S93.
188. *Auerbach M., Ballard H.* // Hematology Am. Soc. Hematol. Educ. Program. 2010. V. 2010. P. 338.
189. *Rund D.* // British J. Hematology. 2021. V. 193. P. 466.
190. *Pai A.B., Garba A.O.* // J. Blood Medicine. 2012. V. 3. P. 77.
191. *Spinowitz B.S., Schwenk M.H., Jacobs P.M. et al.* // Kidney Intern. 2005. V. 68. P. 1801.
192. *Scott L.J., Lyseng-Williamson K.F., McCormack P.L.* // Drugs and Therapy Perspectives. 2013. V. 29. P. 223.
193. *Simon G.H., von Vopelius-Feldt J., Fu Y. et al.* // Investigative Radiology. 2006. V. 41. P. 45.
194. *Toth G.B., Varallyay C.G., Horvath A. et al.* // Kidney Intern. 2017. V. 92. P. 47.
195. *Harvey H.B., Gowda V., Cheng G.* // J. Amer. Coll. Radiol. 2020. V. 17. P. 546.
196. *Crist R.M., Dasa S.S.K., Liu C.H.* // WIREs Nanomed Nanobiotechnol. 2020. V. 13. <https://doi.org/10.1002/wnan.1665>
197. *Elhalawani H., Awan M.J., Ding Y. et al.* // Scientific Data. 2020. V. 7. P. 63. <https://doi.org/10.1038/s41597-020-0392-z>
198. *Daldrup-Link H.E., Theruvath A.J., Rashidi A. et al.* // Pediatric Radiology. 2021. <https://doi.org/10.1007/s00247-021-05098-5>
199. *Scheenen T.W.J., Zamecnik P.* // Investigative Radiology. 2021. V. 56. P. 42.
200. *McCulley L., Gelperin K., Bird S., Harris S. et al.* // Amer. J. Hematol. 2016. V. 91. P. E496.
201. *Chrishtop V.V., Mironov V.A., Prilepskii A.Y. et al.* // Nanotoxicology. 2021. V. 15. P. 167.
202. *Moghimi S.M.* // J. Control. Release. 2014. V. 190. P. 556.
203. *Fortuin A.S., Brüggemann R., van der Linden J. et al.* // WIREs Nanomed Nanobiotechnol. 2018. V. 10. <https://doi.org/10.1002/wnan.1471>
204. *Czarniecki M., Pesapane F., Wood B.J. et al.* // Transl. Androl. Urol. 2018. V. 7 (Suppl 4). P. S453.
205. *Shapiro B., Kulkarni S., Nacev A. et al.* // WIREs Nanomed. Nanobiotechnol. 2015. V. 7. P. 446.
206. *Bonadonna G., Gianni L., Santoro A. et al.* // Ann. Oncol. 1993. V. 4. P. 359.
207. *Gleich B., Weizenecker J.* // Nature. 2005. V. 435. P. 1214.
208. *Nikitin M.P., Orlov A.V., Sokolov I.L. et al.* // Nanoscale. 2018. V. 10. P. 11642.
209. *Graeser M., Thieben F., Szwargulski P. et al.* // Nature Commun. 2019. V. 10. <https://doi.org/10.1038/s41467-019-09704-x>
210. *Widder K.J., Senyei A.E., Scarpelli D.G.* // Proc. Soc. Exp. Biol. Med. 1978. V. 58. P. 141.
211. *Widder K.J., Senyei A.E., Ranney D.F.* // Chemotherapy. 1979. V. 16. P. 213.
212. Controlled Drug Delivery. Fundamentals and Applications / Eds. J.R. Robinson, V.H.L. Lee. Boca Raton: CRC Press, 1987.
213. *Ruuge E.K., Rusetski A.N.* // J. Magn. Magn. Mater. 1993. V. 122. P. 335.
214. *Weiss P.S.* // ASC Nano. 2015. V. 9. P. 3397.
215. *Luxenhofer R., Barz M., Schillmeier M.* // Nanomedicine (Lond.) 2014. V. 9. P. 2083.
216. *Barz M., Luxenhofer R.* // Nanomedicine (Lond.). 2015. V. 10. P. 3089.
217. *Möhwald H., Weis P.S.* // ASC Nano. 2015. V. 9. P. 9427.
218. *Pourgashtasbi G.* // Nanomedicine (Lond.) 2015. V. 10. P. 3121.
219. *Baker M.* // Nature. 2016. V. 533. P. 452.
220. *Bhattacharjee S.* // Int. J. Pharmacokinet. 2017. V. 2. P. 15.
221. *Metselaar J.M., Lammers T.* // Drug Delivery and Translational Res. 2020. V. 10. P. 721.
222. *Behzadi S., Serpooshan V., Tao W. et al.* // Chem. Soc. Rev. 2017. V. 46. P. 4218.
223. *Lammers T.* // Intern. J. Pharmaceutics. 2013. V. 454. P. 527.
224. *Nel A., Xia T., Mädler L., Li N.* // Sci. 2006. V. 311. P. 622.
225. *Najahi-Missaoui W., Arnold R.D., Cummings B.S.* // Int. J. Mol. Sci. 2021. V. 22. <https://doi.org/10.3390/ijms22010385>
226. *Eliasof S., Lazarus D., Peters C.G. et al.* // PNAS. 2013. V. 10. P. 15127.
227. *Moore T.J., Zhang H., Anderson G., Alexander G.C.* // JAMA Intern. Med. 2018. V. 178. P. 1451.
228. *Friberg S., Nyström A.M.* // Future Oncology. 2015. V. 11. P. 1961.
229. *Villanueva-Flores F., Castro-Lugo A., Ramírez O.T., Palomares L.A.* // Nanotechnology. 2020. V. 31. <https://doi.org/10.1088/1361-6528/ab5bc8>
230. *Socoliuc V., Peddis D., Petrenko V.I. et al.* // Magnetochemistry. 2020. V. 6. <https://doi.org/10.3390/magnetochemistry6010002>
231. *Majidi S., Sehrig F.Z., Farkhani S.M. et al.* // Nanomedicine, Biotechnology. 2014. V. 44. P. 722.
232. The Problem with Science: The Reproducibility Crisis and What to Do About It / Ed. R.B. Bausell. N.Y.: Oxford Univ. Press, 2020.
233. *Hofmann-Antenbrink M., Grainger D.W., Hofmann H.* // Nanomedicine: Nanotechnology, Biology, Medicine. 2015. V. 11. P. 1689.
234. *Zangwill A.* Modern Electrodynamics. Cambridge: Cambridge Univ. Press, 2013.
235. *Nacev A., Weinberg I.N., Stepanov P.Y. et al.* // Nano Lett. 2015. V. 15. P. 359.

236. *Shapiro B.* // J. Magn. Magn. Mater. 2009. V. 321. P. 1594.
237. *Komae A., Shapiro B.* // Proc. Am. Contr. Conf. Baltimore, 30 June–02 July, 2010. N.Y. IEEE, 2010. P. 6543.
238. *Avilés M.O., Chen H., Ebner A.D. et al.* // J. Magn. Magn. Mater. 2007. V. 311. P. 306.
239. *Hoshiar A.K., Le T.-A., Amin F.U., Kim M.O. et al.* // J. Magn. Magn. Mater. 2017. V. 427. P. 181.
240. *Li D., Ren Y.* // IEEE Trans. 2018. V. AS-28. № 6. P. 4402107.
241. *Crommelin D.J.A., Florence A.T.* // Intern. J. Pharmaceutics. 2013. V. 454. P. 496.
242. *Kennedy D.* // Sci. 2004. V. 305. P. 451.
243. *Ashkarran A.A., Swann J., Hollis L., Mahmoudi M.* // Trends in Biotechnology. 2021. V. 39. P. 425.
244. *Gaudet J.M., Weber M., Mark A. et al.* // Proc. Am. Ass. Canc. Res. Ann. Meet. 2021, 10–15 Apr, 17–21 May. Philadelphia: AACR; Cancer Res / 2021; 81(13_Suppl): Abstract nr 2801.
245. *Albanese A., Tang P.S., Chan W.C.W.* // Annual Rev. Biomed. Eng. 2021. V. 14. P. 1.
246. *Bondarenko O., Mortimer M., Kahru A. et al.* // Nano Today. 2021. V. 39.
<https://doi.org/10.1016/j.nantod.2021.101184>
247. *Chouhan R.S., Horvat M., Ahmed J. et al.* // Cancers. 2021. V. 13.
<https://doi.org/10.3390/cancers13092213>

ЭЛЕКТРОДИНАМИКА И РАСПРОСТРАНЕНИЕ РАДИОВОЛН

УДК 537.87;621.371;517.958

РАЗВИТИЕ МЕТОДА ВОССТАНОВЛЕНИЯ ЭФФЕКТИВНОЙ ЧАСТОТЫ СОУДАРИЙ ЭЛЕКТРОНОВ В ИОНОСФЕРНОЙ ПЛАЗМЕ

© 2022 г. А. С. Крюковский^{а, *}, Д. С. Лукин^а, Е. В. Михалёва^а, Д. В. Растягаев^а

^а *Российский новый университет,
ул. Радио, 22, Москва, 105005 Российская Федерация*

^{*} *E-mail: kryukovsky56@yandex.ru*

Поступила в редакцию 16.08.2021 г.

После доработки 16.08.2021 г.

Принята к публикации 10.09.2021 г.

На основе данных об ослаблении и запаздывании частотно-модулированных радиосигналов при вертикальном зондировании ионосферы Земли развит метод восстановления эффективной частоты соударений электронов. Методом расширенной бихарактеристической системы выполнено численное моделирование как в случае однослойной, так и двухслойной модели. Построены лучевые траектории частотно-модулированных декаметровых сигналов в координатах время-высота. Исследовано поглощение сигналов от частоты, амплитудные характеристики сигналов, зависимости максимальной высоты отражения и времени прихода от частоты, исследованы зависимости вертикальной компоненты волнового вектора от высоты и времени. Разработанная методика обработки экспериментальных данных позволяет получить зависимость эффективной частоты соударений от высоты с хорошей точностью.

DOI: 10.31857/S0033849422020085

ВВЕДЕНИЕ

В связи с влиянием ионосферы Земли на работу радиосистем различного назначения: систем радиолокации, радиосвязи и радионавигации (позиционирования), актуальной задачей является диагностика и контроль ионосферы, а также постоянный мониторинг экстремальных явлений в атмосфере [1–4]. Математическое моделирование распространения радиоволн в возмущенной ионосферной плазме [5], привязанное непосредственно к оперативным данным наклонного и вертикального зондирования [6–8], является перспективным методом решения указанной задачи.

В данной работе рассмотрен метод восстановления эффективной частоты соударений электронов в ионосферной плазме, основанный на измерении задержки отраженных от ионосферы сигналов и амплитудных характеристик линейно частотно-модулированных (ЛЧМ) сигналов ионозонда вертикального зондирования (ВЗ) [6, 7], так как ионозонды ВЗ являются одним из наиболее эффективных и апробированных средств диагностики ионосферы Земли [9].

1. РАСШИРЕННАЯ БИХАРАКТЕРИСТИЧЕСКАЯ СИСТЕМА

Рассмотрим подход к определению эффективной частоты соударений электронов в ионосферной

плазме, основанный на решении расширенной бихарактеристической системы Лукина [10–12]:

$$\frac{d\vec{r}}{dt} = \frac{2\bar{k}c^2 - \omega^2 \partial \epsilon / \partial \bar{k}}{\partial(\epsilon \omega^2) / \partial \omega}, \quad \frac{d\bar{k}}{dt} = -\frac{\omega^2 \partial \epsilon / \partial \vec{r}}{\partial(\epsilon \omega^2) / \partial \omega}, \quad (1)$$

$$\frac{d\vec{r}_\zeta}{dt} = \frac{\partial}{\partial \zeta} \left(\frac{2\bar{k}c^2 - \omega^2 \partial \epsilon / \partial \bar{k}}{\partial(\epsilon \omega^2) / \partial \omega} \right), \quad (2)$$

$$\frac{d\bar{k}_\zeta}{dt} = -\frac{\partial}{\partial \zeta} \left(\frac{\omega^2 \partial \epsilon / \partial \vec{r}}{\partial(\epsilon \omega^2) / \partial \omega} \right),$$

$$\frac{d\vec{r}_\eta}{dt} = \frac{\partial}{\partial \eta} \left(\frac{2\bar{k}c^2 - \omega^2 \partial \epsilon / \partial \bar{k}}{\partial(\epsilon \omega^2) / \partial \omega} \right), \quad (3)$$

$$\frac{d\bar{k}_\eta}{dt} = -\frac{\partial}{\partial \eta} \left(\frac{\omega^2 \partial \epsilon / \partial \vec{r}}{\partial(\epsilon \omega^2) / \partial \omega} \right).$$

В формулах (1)–(3) введены следующие обозначения: ω – круговая частота, $\omega = 2\pi f$ (f – рабочая частота); $\epsilon(\vec{r}, \bar{k}, \omega)$ – эффективная диэлектрическая проницаемость среды распространения; $\bar{k}(t)$ – волновой вектор; $\vec{r} = (x, y, z)$ – декартовы координаты луча; t – время группового запаздывания сигнала; (η, ζ) – углы выхода луча.

Первые шесть уравнений (1) определяют координаты луча и волновой вектор. Для их находде-

ния необходимо дополнить систему (1) начальными условиями:

$$\begin{aligned} k_x(0) &= \frac{\omega}{c} \sqrt{\epsilon_0} \sin \zeta, & k_y(0) &= \frac{\omega}{c} \sqrt{\epsilon_0} \cos \zeta \cos \eta, \\ k_z(0) &= \frac{\omega}{c} \sqrt{\epsilon_0} \cos \zeta \sin \eta, & \vec{r}(0) &= 0. \end{aligned} \quad (4)$$

В формулах (4) ϵ_0 – значение эффективной диэлектрической проницаемости среды в источнике, который расположен в начале координат. Для того чтобы избежать особенностей при вычислениях вертикального луча, нами выбрана параметризация волнового вектора, для которой вертикальному лучу соответствуют $\zeta = 0$, $\eta = \pi/2$ [13].

Для вертикального луча из (4) находим

$$\begin{aligned} k_x(0) &= 0, & k_y(0) &= 0, & k_z(0) &= \frac{\omega}{c} \sqrt{\epsilon_0}, \\ x(0) &= y(0) = z(0) = 0. \end{aligned} \quad (5)$$

Следующие 12 уравнений (см. (2), (3)) определяют скорости изменений производных координат и компонент волновых векторов по угловым координатам η и ζ . Для этой группы уравнений начальные условия имеют вид

$$\begin{aligned} k'_{x\zeta}(0) &= \frac{\omega}{c} \sqrt{\epsilon_0} \cos \zeta, & k'_{x\eta}(0) &= 0, \\ k'_{y\zeta}(0) &= -\frac{\omega}{c} \sqrt{\epsilon_0} \sin \zeta \cos \eta, \\ k'_{y\eta}(0) &= -\frac{\omega}{c} \sqrt{\epsilon_0} \cos \zeta \sin \eta, \\ k'_{z\zeta}(0) &= -\frac{\omega}{c} \sqrt{\epsilon_0} \sin \zeta \sin \eta, \\ k'_{z\eta}(0) &= \frac{\omega}{c} \sqrt{\epsilon_0} \cos \zeta \cos \eta, \\ \vec{r}'_{\zeta}(0) &= 0, & \vec{r}'_{\eta}(0) &= 0. \end{aligned} \quad (6)$$

В случае нормального падения получаем

$$\begin{aligned} k'_{x\eta}(0) &= k'_{y\zeta}(0) = k'_{z\zeta}(0) = k'_{z\eta}(0) = 0, \\ k'_{x\zeta}(0) &= \frac{\omega}{c} \sqrt{\epsilon_0}, & k'_{y\eta}(0) &= -\frac{\omega}{c} \sqrt{\epsilon_0}. \end{aligned} \quad (7)$$

Решение расширенной бихарактеристической системы позволяет найти расходимость D , которая вычисляется как корень из отношения якобианов \hat{J}_0 и \hat{J} [12, 14]:

$$D = \sqrt{\frac{\hat{J}_0}{\hat{J}_j}}, \quad \hat{J} = \begin{vmatrix} x_t & y_t & z_t \\ x'_\zeta & y'_\zeta & z'_\zeta \\ x'_\eta & y'_\eta & z'_\eta \end{vmatrix}, \quad \hat{J}_0 = \hat{J}|_{r=r_0}. \quad (8)$$

Если среда изотропная, т.е. $\partial\epsilon/\partial\vec{k} = 0$, то в системе (1) первая группа уравнений упрощается:

$$\frac{d\vec{r}}{dt} = \frac{2c^2}{\partial(\epsilon\omega^2)/\partial\omega} \vec{k}. \quad (9)$$

Известно, что в предположении изотропности среды (ионосферной плазмы) эффективную диэлектрическую проницаемость можно представить в виде

$$\epsilon = 1 - X, \quad X = \left(\frac{\omega_p}{\omega} \right)^2 = \frac{4\pi e^2 N(\vec{r})}{m_e \omega^2}, \quad (10)$$

где e и m_e – заряд и масса электрона, c – скорость света, ω_p – плазменная частота, а $N(\vec{r})$ – электронная концентрация. С учетом (10) формула (9) преобразуется и имеет вид

$$\frac{d\vec{r}}{dt} = \frac{c^2}{\omega} \vec{k}. \quad (11)$$

Тогда выражение для якобиана \hat{J} принимает вид

$$\hat{J} = \frac{c^2}{\omega} \begin{vmatrix} k_x & k_y & k_z \\ x'_\zeta & y'_\zeta & z'_\zeta \\ x'_\eta & y'_\eta & z'_\eta \end{vmatrix} = \frac{c^2}{\omega} J. \quad (12)$$

Очевидно, что выражение для расходимости D теперь можно представить как

$$D = \sqrt{\frac{J_0}{J_j}}. \quad (13)$$

Предположим, что эффективная диэлектрическая проницаемость среды зависит только от высоты z , т.е. $\epsilon(z, \omega)$. Тогда из системы (1) следует, что k_x и k_y вдоль траектории постоянны, а так как для вертикального луча начальные значения k_x и k_y равны нулю (см. (5)), то находим, что $k_x = k_y = 0$. Поэтому выражение для J можно переписать в следующем виде:

$$J = k_z \begin{vmatrix} x'_\zeta & y'_\zeta \\ x'_\eta & y'_\eta \end{vmatrix}. \quad (14)$$

Из системы (1) находим, что $x(t) = y(t) = 0$, и остаются два уравнения относительно $z(t)$ и $k_z(t)$, образующие подсистему:

$$\frac{dz}{dt} = \frac{c^2}{\omega} k_z, \quad \frac{dk_z}{dt} = -\frac{1}{2} \omega \frac{\partial\epsilon}{\partial z}, \quad (15)$$

с начальными условиями

$$k_z(0) = \frac{\omega}{c} \sqrt{\epsilon_0}, \quad z = 0. \quad (16)$$

Численное решение системы (14) с учетом (5) позволяет найти $z(t)$ и $k_z(t)$.

Для вычисления якобиана J необходимо знать производные x и y по начальным углам выхода лу-

чей ζ и η (14). Эти производные находятся как решения системы (2), (3):

$$\frac{d\vec{r}_\rho}{dt} = \frac{c^2 \partial \vec{k}}{\omega \partial \rho}, \quad \frac{d\vec{k}_\rho}{dt} = -\frac{1}{2} \omega \frac{\partial}{\partial \rho} \left(\frac{\partial \varepsilon}{\partial \vec{r}} \right), \quad \rho = (\zeta, \eta), \quad (17)$$

с начальными условиями (6) или (7). В предположении, что горизонтальные градиенты отсутствуют, можно установить, что

$$k'_{x\zeta} = \frac{\omega}{c} \sqrt{\varepsilon_0}, \quad k'_{y\zeta} = 0, \quad k'_{x\eta} = 0, \quad k'_{y\eta} = -\frac{\omega}{c} \sqrt{\varepsilon_0}. \quad (18)$$

2. МЕТОД ДИАГНОСТИКИ ЭФФЕКТИВНОЙ ЧАСТОТЫ СОУДАРЕНИЙ

Амплитуда A зондирующего сигнала связана с расходимостью D и поглощением ψ формулой [13, 14]

$$A = E_0 D \exp[-\psi], \quad (19)$$

в которой величина E_0 – значение амплитуды поля вне ионосферы на расстоянии r_0 от источника излучения, расположенного в начале координат. В случае изотропного излучателя получаем

$$E_0 = \sqrt{30W}/r_0. \quad (20)$$

При необходимости в выражение (20) могут быть включены диаграмма направленности, коэффициент усиления и другие параметры. В формуле (20) W – мощность источника излучения, и если мощность выражена в киловаттах, а r_0 – в метрах, то размерность E_0 будет В/м. При вычислениях принято, что $W = 1$ кВт.

Предполагается, что из данных ВЗ можно получить зависимость задержки $2t_m$ отразившегося от ионосферы сигнала от частоты f , а также зависимость амплитуды A принятого сигнала от частоты зондирования f . Параметр $t_m(f)$ – время прихода сигнала из излучателя в точку отражения от ионосферы.

Считая, что зависимость электронной концентрации $N(z)$ от высоты z известна, т.е. уже восстановлена по зависимости задержки сигнала от частоты излучения, можно, опираясь на формулы разд. 1, определить расходимость D (13).

Тогда из (19) находим поглощение

$$\psi = -\ln \frac{A}{DE_0} \equiv L(f). \quad (21)$$

Поглощение ψ можно найти по формуле [15, 16]

$$\psi = -\frac{\omega}{2} \int_0^t \varepsilon_2 dt, \quad (22)$$

где ε_2 – мнимая часть диэлектрической проницаемости среды. Мнимую часть диэлектрической

проницаемости среды можно приближенно представить в виде [9, 15]

$$\varepsilon_2 \cong -X \frac{Z}{1+Z^2}, \quad (23)$$

где

$$Z = v_e/\omega, \quad X = \frac{4\pi e^2 N(z)}{m_e \omega^2}, \quad (24)$$

а v_e – эффективная частота соударений электронов.

Предполагая, что эффективная частота соударений на высотах выше 80 км не превышает 10^6 с⁻¹, в знаменателе выражения (23) можно пренебречь величиной Z^2 и считать, что

$$\varepsilon_2 \approx -XZ. \quad (25)$$

Приравняв (21) и (22) и исключив ψ , можно для определения $v_e(z)$ получить интегральное уравнение типа Вольтерра второго рода:

$$-\frac{\omega}{2} \int_0^t \varepsilon_2 dt = L. \quad (26)$$

С учетом (24) и (25) уравнение (26) нетрудно переписать в виде

$$\frac{1}{\omega^2} \int_0^{t_m} G dt = L, \quad G = v_e V_c N, \quad V_c = \frac{4\pi e^2}{m_e}. \quad (27)$$

Для решения уравнения (27) применим метод итераций [13, 14] и найдем эффективную частоту соударений v_e как функцию высоты z . Следует подчеркнуть, что высота z зависит от группового времени t и вычисляется вдоль лучевой траектории отдельно для каждой частоты. Поэтому функция $z(t)$ также зависит от частоты.

Рассмотрим алгоритм решения уравнения (27). Так как ниже определенной высоты ионосферная плазма отсутствует, то электронная концентрация равна нулю. Следовательно, до некоторой частоты f_0 , для которой время t_m равно t_{m0} , интеграл в левой части (27) равен нулю:

$$\int_0^{t_{m0}} G dt = 0. \quad (28)$$

Рассмотрим частотный интервал (f_0, f_n) , принадлежащий частотному интервалу (f_{\min}, f_{\max}) , на котором осуществляется зондирование, причем $f_{\min} < f_0 < f_n < f_{\max}$. Будем считать, что f_{\max} совпадает с максимальной плазменной частотой ионосферного слоя. Разобьем интервал (f_0, f_n) на n частей: $f_0, f_1, \dots, f_j, \dots, f_n$. Предположим, что на каждом j -м интервале (f_{j-1}, f_j) функция G постоянная и равна G_j . Пусть t_{mj} – время прихода сигнала с частотой f_j из передатчика в точку отражения от ионосферы.

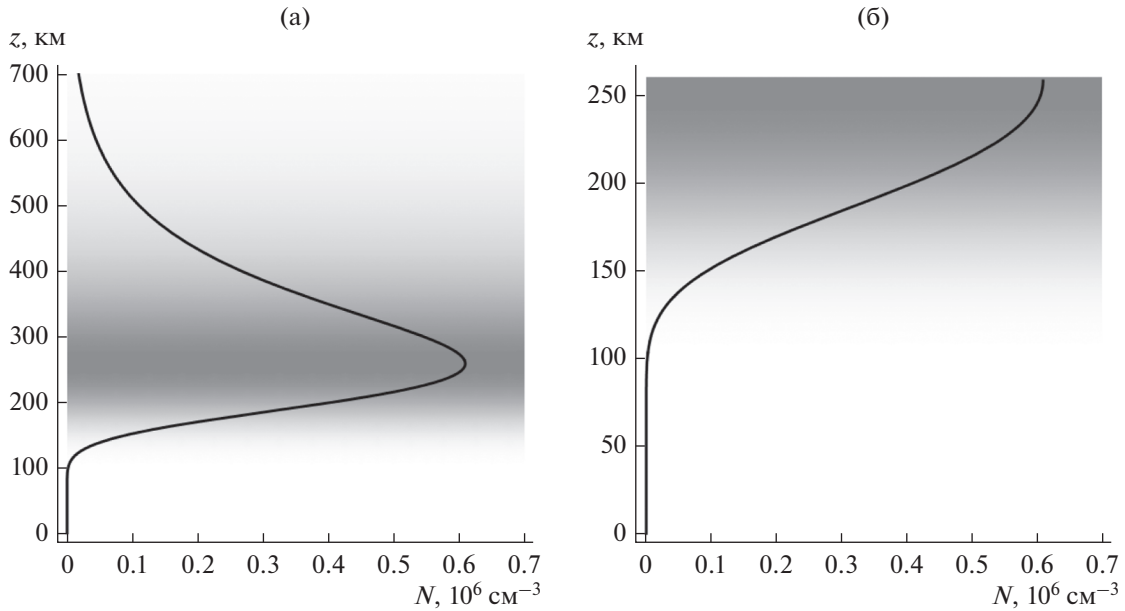


Рис. 1. Зависимость электронной концентрации N от высоты z : а – весь ионосферный слой, б – фрагмент.

Они известны из решения бихарактеристической системы. Тогда из (27) нетрудно получить

$$G_1 = L_1 \frac{\omega_1^2}{t_{m1} - t_{m0}}, \quad G_2 = L_2 \frac{\omega_2^2}{t_{m2} - t_{m1}} - G_1 \frac{t_{m1} - t_{m0}}{t_{m2} - t_{m1}},$$

$$G_j = L_j \frac{\omega_j^2}{t_{m,j} - t_{m,j-1}} - \sum_{i=1}^{j-1} G_i \frac{t_{m,i} - t_{m,i-1}}{t_{m,j} - t_{m,j-1}}, \quad (29)$$

$$j \geq 2.$$

Формулы (29) можно упростить и получить

$$G_j = \frac{L_j \omega_j^2 - L_{j-1} \omega_{j-1}^2}{t_{m,j} - t_{m,j-1}}, \quad j \geq 2. \quad (30)$$

После вычислений G_j , зная, что каждая из этих величин это произведение эффективной частоты соударений электронов $N(z)$, электронной концентрации в точке отражения и константы V_c , можно легко найти зависимость эффективной частоты соударений v_e от высоты z .

3. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ. ОДНОСЛОЕВАЯ МОДЕЛЬ

Приведем примеры численной реализации рассмотренных выше алгоритмов. Далее на основе приведенных выше формул рассчитана амплитуда поля в точке приема, совпадающей с точкой излучения сигнала, для разных частот, восстановлена эффективная частота соударений электронов в плазме v_e и выполнено сравнение модельных и расчетных значений.

Рассмотрим однослойную модель ионосферы. Зависимость электронной концентрации от высоты приведена на рис. 1 и показана как фон, так и линией. Весь ионосферный слой представлен на рис. 1а, а та часть ионосферы, которая доступна сигналам станции ВЗ, – на рис. 1б.

Модель зависимости эффективной частоты соударений электронов v_e от высоты z , взятая из экспериментальных данных, показана на рис. 2а и 2б широкой белой линией [13]. Как и на рис. 1, плотность электронной концентрации показана фоном.

Аппроксимация эффективной частоты соударений v_e , полученная методом наименьших квадратов, на рис. 2а показана тонкой линией. Она может быть описана выражением

$$\lg v_e = a_0 + a_1 z + a_2 z^2 + \frac{b_1}{z}, \quad (31)$$

в котором $a_0 \approx -1.02$, $b_1 \approx 498.9$, $a_1 \approx 0.0080$, $a_2 \approx -8.04 \times 10^{-6}$. По данным ВЗ зависимость $v_e(z)$ может быть восстановлена не выше максимума слоя F2, т.е. только в нижней части ионосферы, показанной на рис. 2б. Зависимость эффективной частоты соударений от высоты в этой области проще. Аппроксимацию можно представить гиперболической функцией:

$$\lg v_e = a + b/z, \quad (32)$$

в которой $a \approx 0.617$, $b \approx 416.18$. В формулах (31) и (32) высота z выражена в километрах.

Рассмотрим численные решения системы бихарактеристических уравнений (15), (16) (рис. 3, 4).

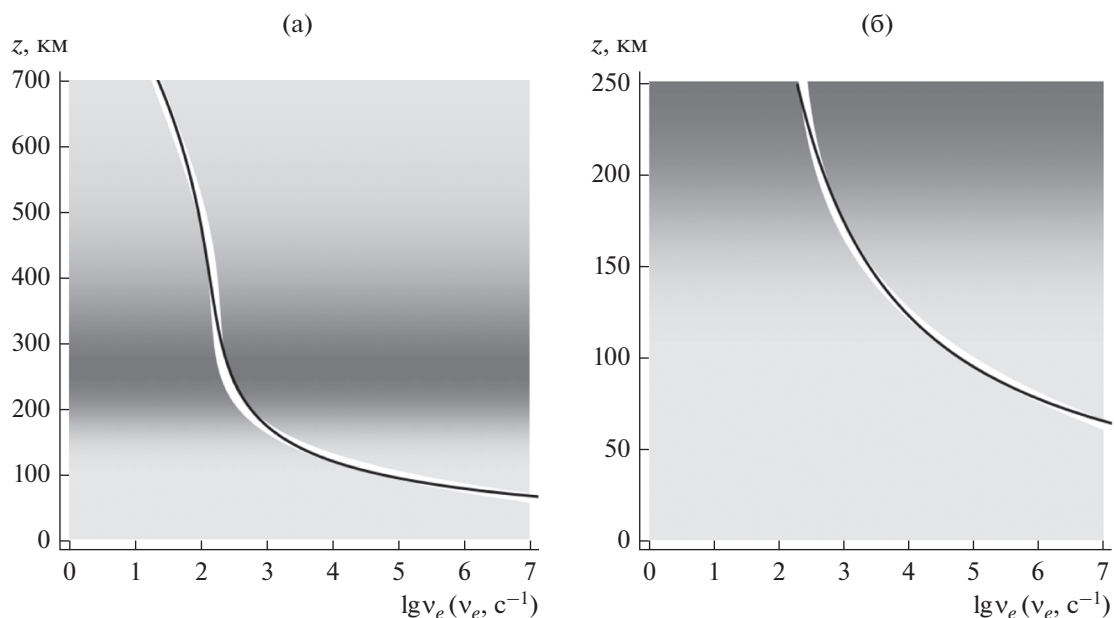


Рис. 2. Моделирование зависимости $\lg v_e$ от высоты: а – полный профиль, б – фрагмент, белая линия – экспериментальные данные, темная линия – результат моделирования.

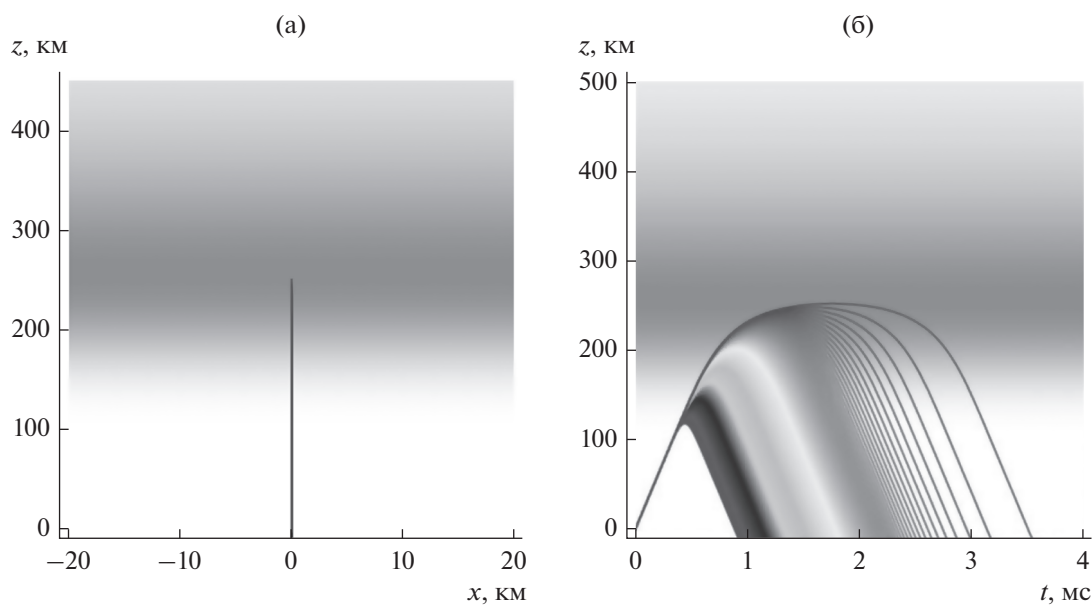


Рис. 3. Лучевые траектории для однослойной модели: а – в плоскости (x, z) , б – в плоскости (t, z) на частотах от 1 до 6.993 МГц (снизу вверх).

На рис. 3 показаны лучевые траектории. В плоскости (x, z) все лучевые траектории образуют вертикальную прямую линию (см. рис. 3а), а в координатах (t, z) каждая частота от 1 до 6.993 МГц имеет свою траекторию и показана своим оттенком серого цвета (рис. 3б). Из рисунка видно, что с возрастанием частоты растет время распространения сигнала.

Зависимости вертикальной компоненты волнового вектора k_z от времени t и высоты z показаны на рис. 4. Частота возрастает от 1 до 6.993 МГц на рис. 4а слева направо, а на рис. 4б – снизу вверх.

При построении рис. 4 волновой вектор k_z нормирован на начальное значение волнового вектора $k_0 = k_z(0)$. Отметим, что величина k_0 яв-

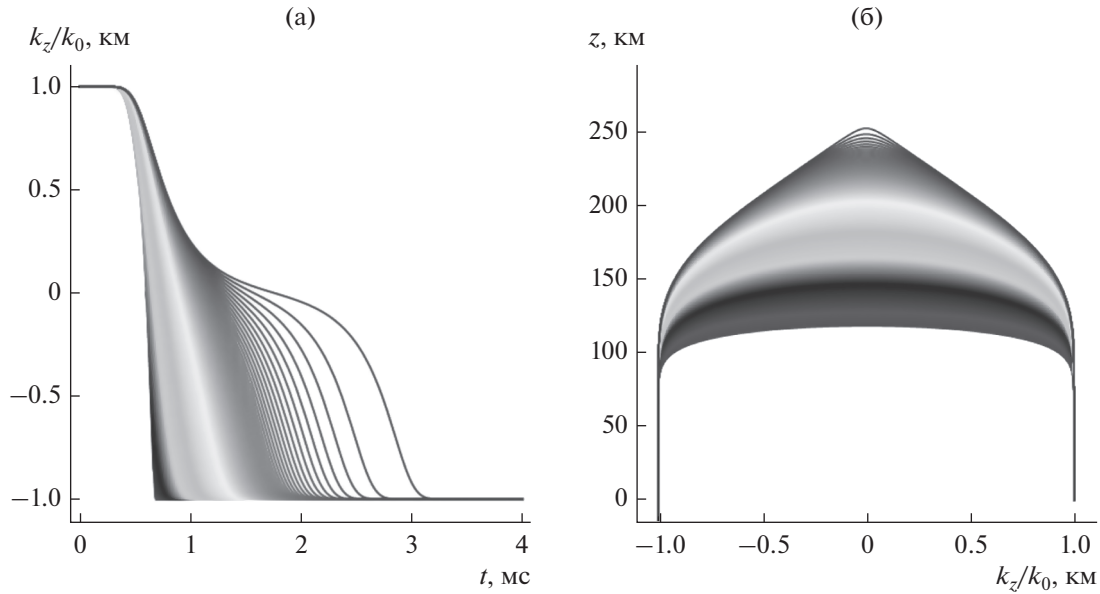


Рис. 4. Зависимость k_z от времени (а) и высоты (б) для однослойной модели на частотах от 1 до 6.993 (а – слева направо, б – снизу вверх).

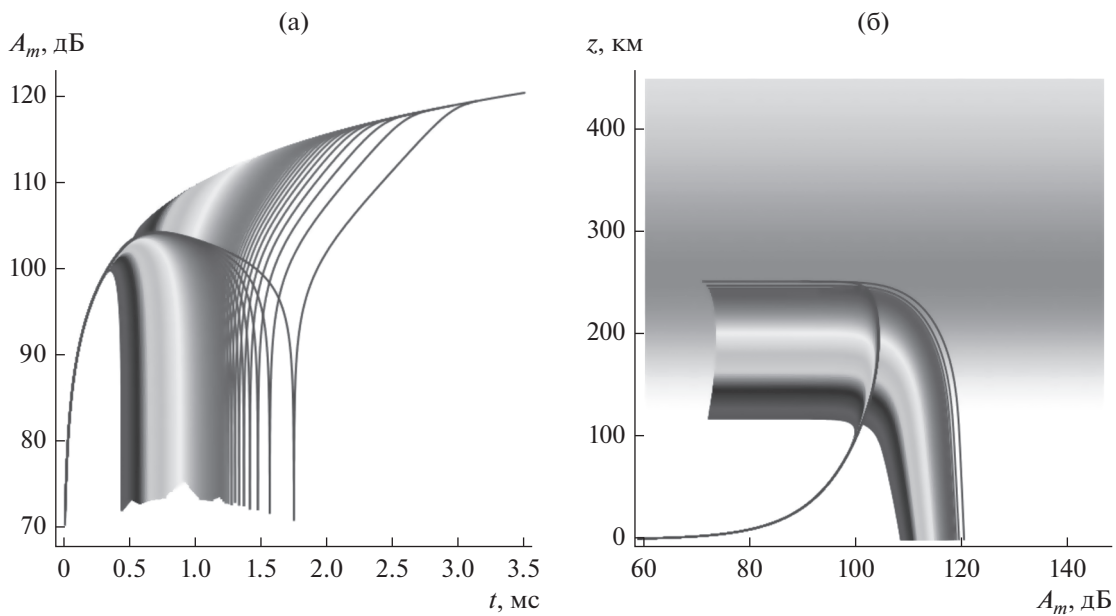


Рис. 5. Зависимость амплитуды сигнала от времени (а) и высоты (б) для однослойной модели.

ляется функцией частоты (16). Из рис. 4 видно, что значение нормированного волнового вектора меняется при распространении в ионосфере, уменьшаясь до 0 в точке отражения, а после смены знака волнового вектора опять по модулю возрастает. Когда отношение k_z/k_0 становится равным 1, луч покидает ионосферу.

Рассмотрим теперь амплитудные характеристики, получаемые из решения расширенной би-

характеристической системы (18). Зависимости ослабления сигнала A_m от времени и высоты

$$A_m = -20 \lg \sqrt{\left| \frac{J_0}{J} \right|} \quad (33)$$

вдоль лучей, соответствующих различным частотам, приведены на рис. 5.

При движении вдоль луча от источника ослабление сначала резко увеличивается (до 100 дБ).

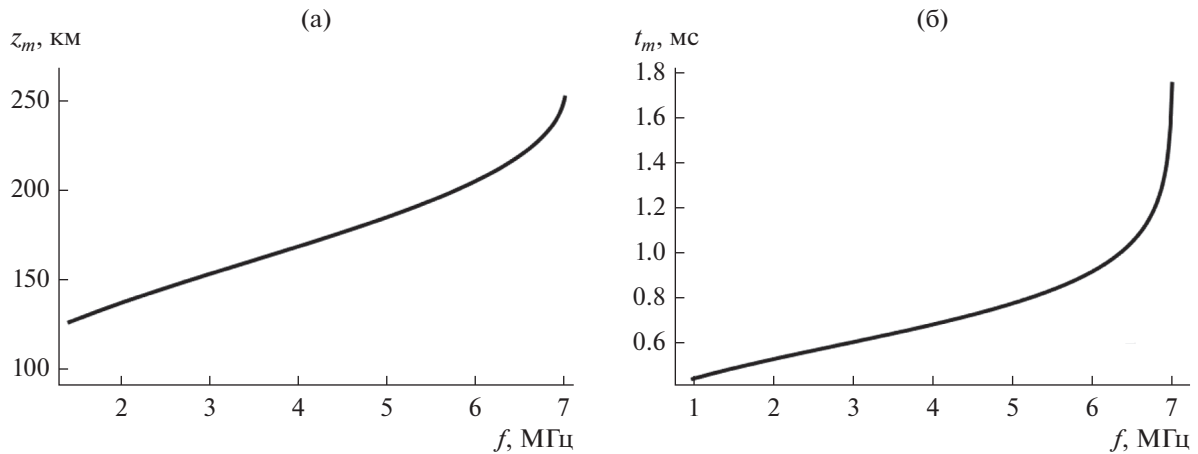


Рис. 6. Зависимость высоты отражения сигнала z_m (а) и времени t_m (б) от частоты f для однослойной модели.

Далее ослабление резко уменьшается при приближении к точке отражения, соответствующей каустике [16–19], а потом опять возрастает, когда луч возвращается в источник излучения. На каустике (в точке отражения) в геометрооптическом приближении амплитуда сигнала обращается в бесконечность, хотя из рис. 5 это не очевидно. Однако более точные расчеты показывают, что при приближении к точке отражения, а также в источнике $A_m \rightarrow -\infty$.

Рассмотрим характеристики радиосигнала, отраженного от ионосферы, в точке приема. Зависимость высоты отражения сигнала z_m от частоты, вычисленная по модельным данным, показана на рис. 6а, а зависимость времени прихода луча в точку отражения t_m от частоты — на рис. 6б. Видно, что и время t_m , и высота отражения z_m сигнала монотонно растут с ростом частоты f . При приближении частоты f к критической f_m (≈ 7 МГц) на-

блюдается особенно быстрый рост как высоты z_m , так и времени t_m .

Используя частоту f как параметр и опираясь на данные рис. 6а и 6б, можно построить зависимость высоты отражения сигнала z_m от задержки t_m (рис. 7). Для однослойной модели эта кривая плавная и взаимно-однозначная.

На рис. 8 приведена зависимость от частоты f поглощения ψ , найденная по формулам (22), (23), при условии, что зависимость логарифма эффективной частоты соударений ν_e от высоты z определяется графиком на рис. 2б. Видно, что поглощение убывает при увеличении частоты и начинает вновь увеличиваться при приближении частоты f к критической.

Зависимости амплитуды сигнала в точке приема A (19) и функции L (21) от частоты f показаны на рис. 9. Амплитуда радиосигнала, отраженного от ионосферы, сначала растет при увеличении ча-

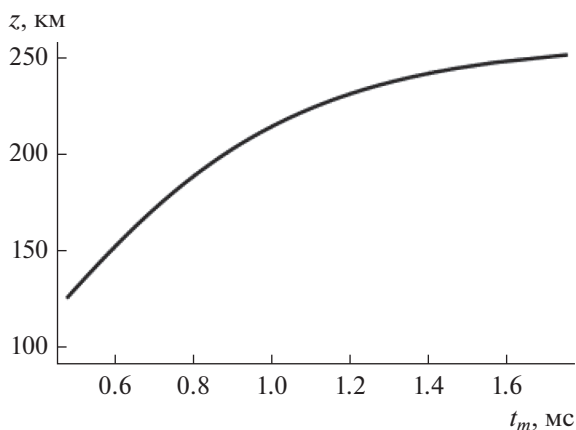


Рис. 7. Функция высоты отражения сигнала z_m от времени t_m .

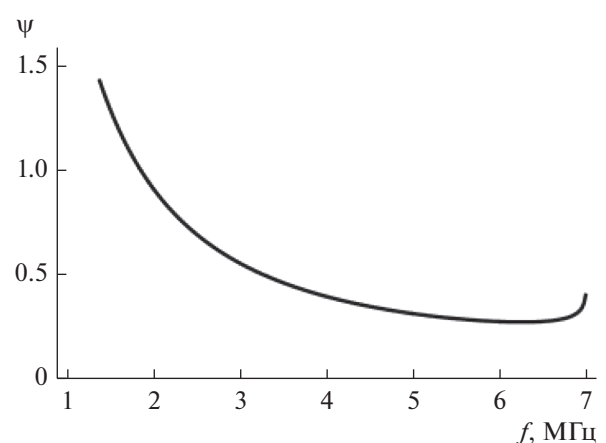


Рис. 8. Зависимость поглощения ψ от частоты f .

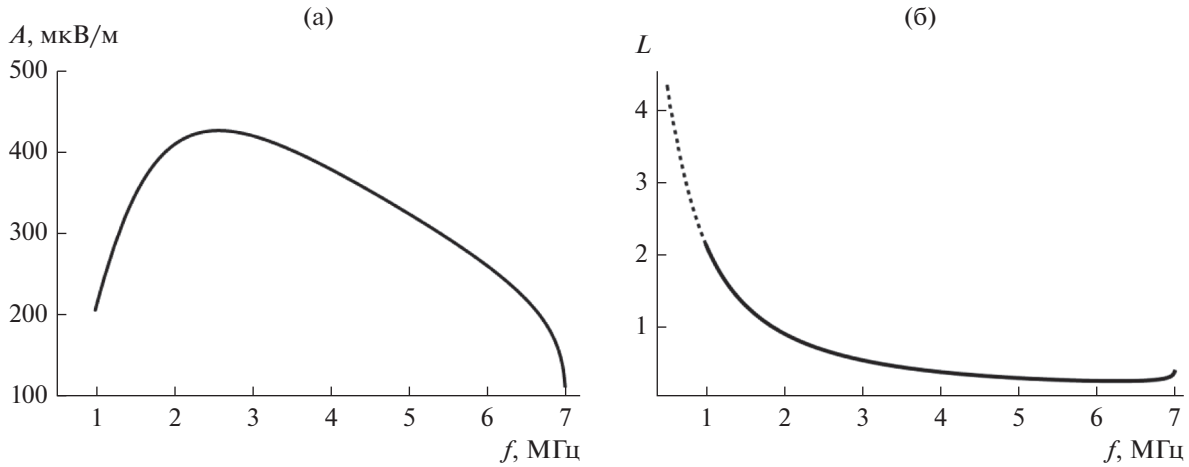


Рис. 9. Зависимость амплитуды A (а) и логарифма амплитудной функции L (б) от частоты f .

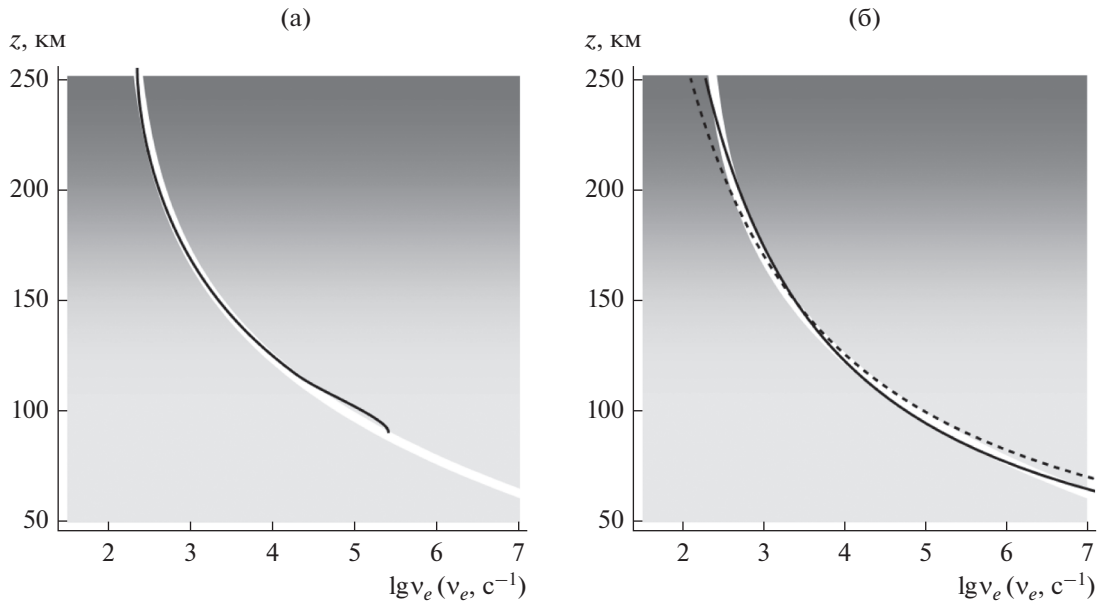


Рис. 10. Частота соударений для однослойной модели: а – сопоставление зависимости от высоты модельной (широкая белая линия) и расчетной (темная линия) частоты соударений, б – аппроксимация эффективной частоты соударений по формулам (32) (сплошная темная линия) и (34) (пунктир).

стоты (до ~ 2.5 МГц), а затем резко убывает (см. рис. 9а). Зависимость функции L (20) от частоты f более плавная (см. рис. 9б), поскольку она пропорциональна логарифму отношения амплитуды сигнала и лучевой расходимости, умноженной на постоянный коэффициент. Участок рис. 9б на частотах меньше 1 МГц получен с помощью экстраполяции (пунктирная линия) и важен для вычислений.

Результаты моделирования зависимости эффективной частоты соударений электронов ν_e от высоты z показаны на рис. 10. Светлой широкой линией показана модельная зависимость эффективной частоты соударений ν_e от высоты, а тон-

кой – расчетные значения. Видно, что модельные и расчетные значения совпадают с хорошей точностью, а сама точность совпадений модельных и расчетных значений увеличивается с высотой. На рис. 10б помимо модельной зависимости ν_e приведены экстраполяционные кривые, построенные по формуле (32) (тонкая линия) и по формуле (34):

$$\lg \nu_e = a + b/z, \quad a \approx 0.164, \quad b \approx 483.5, \quad (34)$$

полученной по расчетным данным (пунктирная линия). Из рис. 10б следует, что совпадение удовлетворительное.

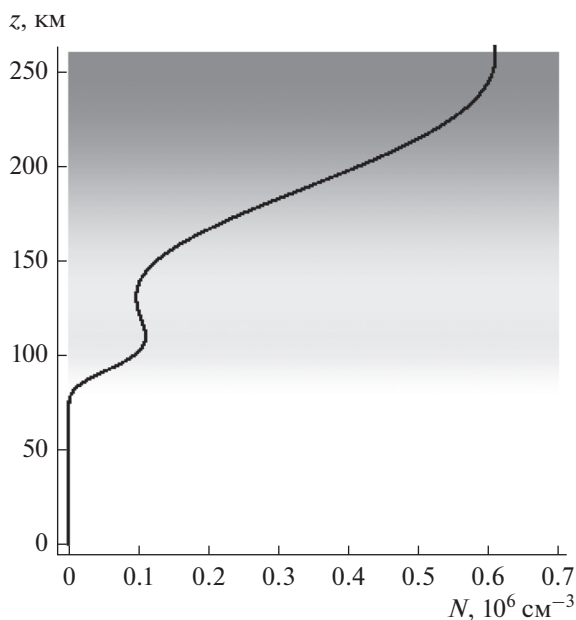


Рис. 11. Зависимость электронной концентрации от высоты для двухслойной модели.

4. РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ. ДВУСЛОЕВАЯ МОДЕЛЬ

Проблемы с восстановлением эффективной частоты соударений возникают тогда, когда имеет место двухслойная модель электронной концентрации или присутствуют локальные неоднородности. На рис. 11 показана используемая в расчетах зависимость электронной концентрации от высоты для двухслойной модели со слоями F и E.

На рис. 12 представлены лучевые траектории для двухслойной модели. Как и на рис. 3б, каждая частота имеет свою траекторию и показана своим оттенком серого цвет: от 1 МГц (нижняя линия) до 6.993 МГц (верхняя). В отличие от рис. 3б, на рис. 12 хорошо видна группа лучей, отразившаяся от нижнего слоя E.

Зависимости вертикальной компоненты волнового вектора k_z/k_0 от времени t и от высоты z показаны на рис. 13. Частота возрастает от 1 до 6.993 МГц на рис. 13а слева направо, а рис. 13б — снизу вверх. Сравнение рис. 4 и 13 позволяет оценить влияние слоя E на волновой вектор. Особенно следует отметить центральную область на рис. 13б, соответствующую волновому межслойному каналу, в которую не проникают лучи.

Рассмотрим особенности характеристик радиосигнала, отраженного от ионосферы, в точке приема в двухслойной модели. Зависимость высоты отражения сигнала от частоты показана на рис. 14. Видно, что, как и в случае однослойной модели, высота отражения z_m сигнала монотонно

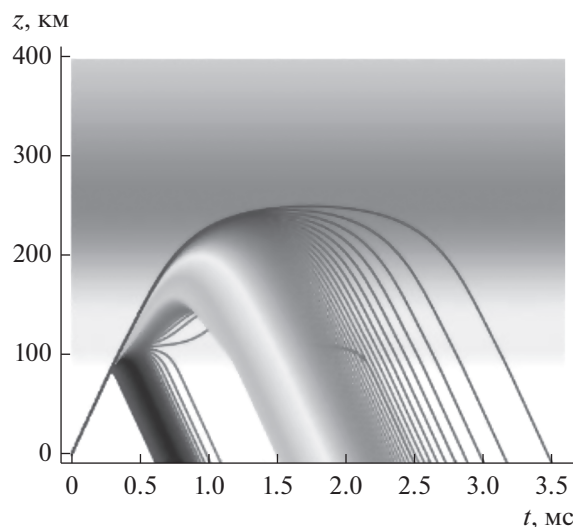


Рис. 12. Лучевые траектории в плоскости (t, z) для двухслойной модели на частотах от 1 до 6.993 МГц (снизу вверх).

растет с ростом частоты f . В районе частоты порядка 3 МГц виден разрыв, обусловленный межслойной долиной.

Однако зависимость времени t_m прихода луча в точку отражения от частоты уже не является монотонной. Особенно резкий рост времени t_m наблюдается при приближении частоты f к критической частоте как слоя F f_m (≈ 7 МГц), так и слоя E (≈ 3 МГц). Резкий рост кривых на рис. 14 в районе межслойной долины и максимума слоя F2 неизбежно приводит к неустойчивости итерационного процесса при восстановлении эффективной частоты соударений.

Как и в случае однослойной модели, исключив частоту f , можно построить график зависимости высоты отражения z_m от задержки t_m (рис. 15). Как и на рис. 14, пунктирная линия на рис. 15 связана с разрывом, обусловленным межслойной долиной, и условно соединяет две части графика. Следует отметить неоднозначность: одному и тому же времени t_m могут соответствовать две различные высоты отражения, что связано с резким увеличением времени t_m при приближении к критическим частотам как слоя E, так и слоя F.

Зависимости ослабления сигнала A_m от времени и от высоты для двухслойной модели, рассчитанные по формуле (33) вдоль лучей, соответствующих различным частотам, приведены на рис. 16.

Сравнивая рис. 5 для однослойной модели и рис. 16 для двухслойной, видим, что для двухслойной модели кривые разделились на две группы. Первая группа кривых соответствует лучам, отразившимся от слоя E. Это левые кривые на рис. 16а и нижние кривые на рис. 16б. Вторая группа кри-

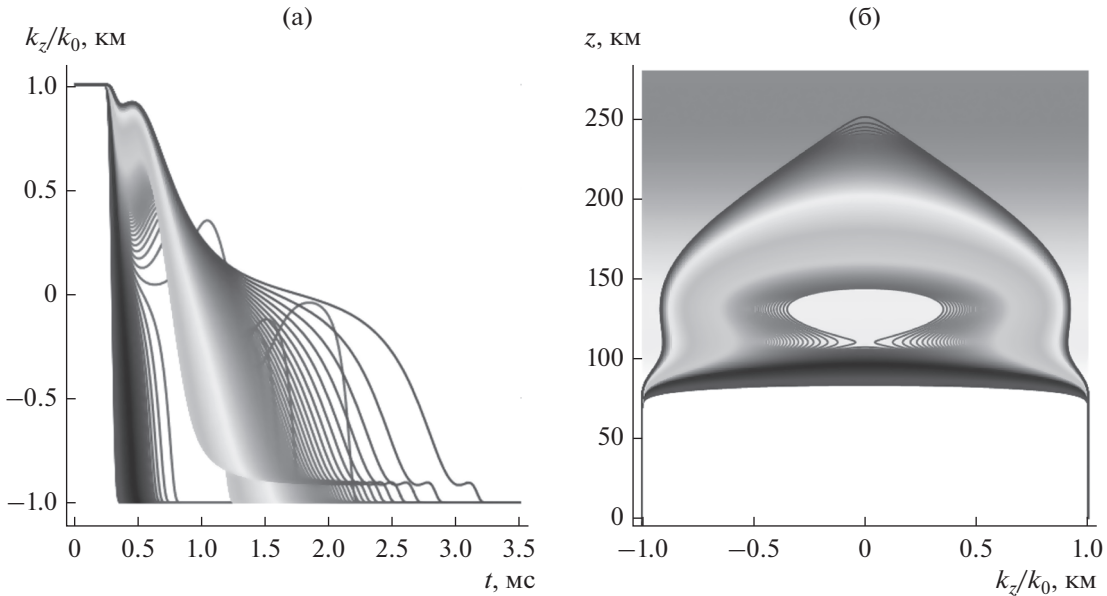


Рис. 13. Зависимость k_z/k_0 от времени (а) и высоты (б) для двуслойной модели на частотах от 1 до 6.993 МГц (а – слева направо, б – снизу вверх).

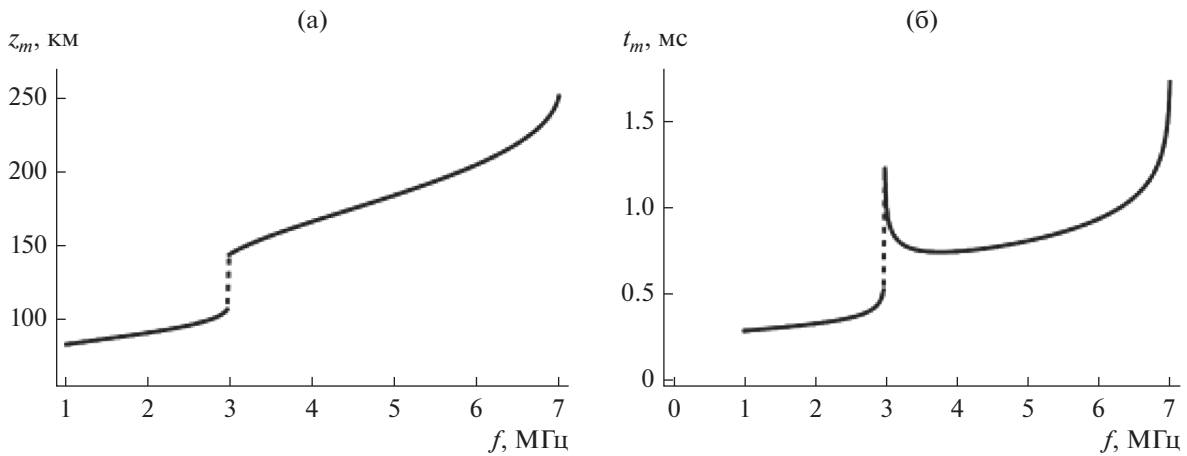


Рис. 14. Зависимость высоты отражения сигнала z_m (а) и времени t_m (б) от частоты f для двуслойной модели.

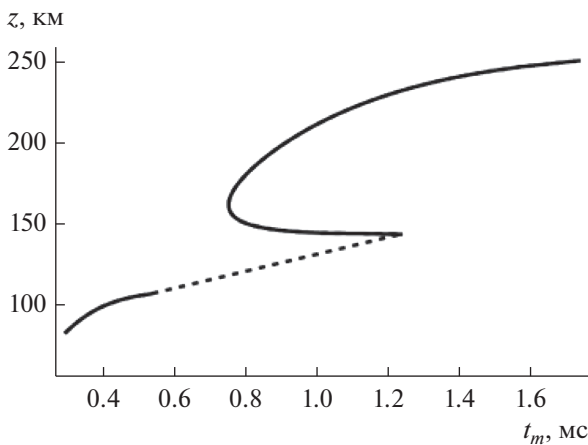


Рис. 15. Зависимость высоты отражения сигнала z_m от времени t_m .

вых соответствует лучам с более высокой частотой, отразившимся от слоя F (правые кривые на рис. 16а и верхние кривые на рис. 16б). Между ними прослеживаются линии, связанные с переходной областью.

Зависимость поглощения ψ от частоты f представлена на рис. 17. В отличие от однослойной модели (см. рис. 8), функция имеет экстремум (максимум) в окрестности критической частоты слоя E и не является монотонной.

Перейдем теперь к восстановлению зависимости эффективной частоты соударений ν_e как функции высоты z . На рис. 18 показана частотная зависимость функции L , используемой для восстановления.

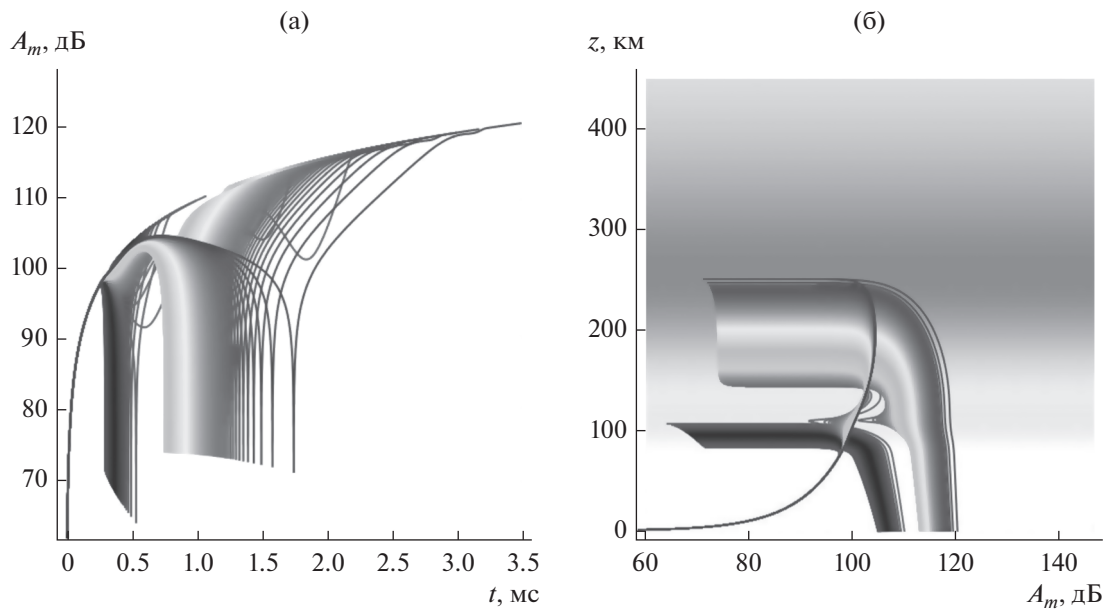


Рис. 16. Зависимость амплитуды сигнала от времени (а) и от высоты (б) для двухслойной модели.

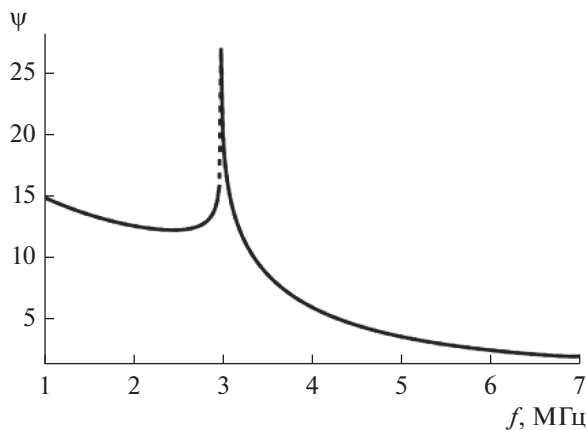


Рис. 17. Зависимость поглощения ψ от частоты f для двухслойной модели.

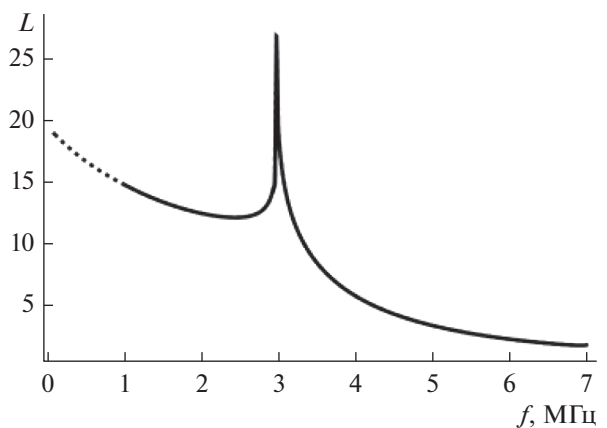


Рис. 18. Зависимость функции L от частоты f для двухслойной модели.

Она, разумеется, соответствует функции ψ , но имеет аналитическое продолжение в область малых частот, на которых измерения недоступны. Наличие слоя E, а главное, межслоевой долины препятствует прямому применению описанного выше алгоритма для восстановления v_e как функции высоты z . Результат восстановления $v_e(z)$ на низких частотах (и малых высотах) представлен на рис. 19.

Для нижней ионосферы расчетные (сплошная черная линия) и модельные значения (белая широкая линия) совпадают с хорошей точностью. Петля на конце черной линии — это следствие приближения (28). Нижняя часть ионосферы яв-

ляется особенно важной областью, поскольку формирует основное поглощение радиоволны.

По данным, представленным на рис. 19, методом наименьших квадратов можно построить экстраполяционную модель:

$$\lg v_e = a + \frac{b}{z}, \quad a \approx 0.443, \quad b \approx 442.8, \quad (35)$$

которая, естественно, отличается и от аналогичной модели для однослойной ионосферы (34) и от аппроксимации исходной зависимости (32). Однако результат экстраполяции оказывается вполне приемлемым.

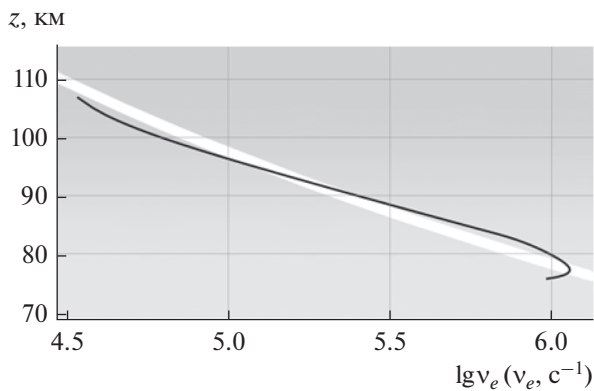


Рис. 19. Сравнение зависимости модельной (широкая белая линия) и расчетной (темная линия) частот соударений электронов от высоты.

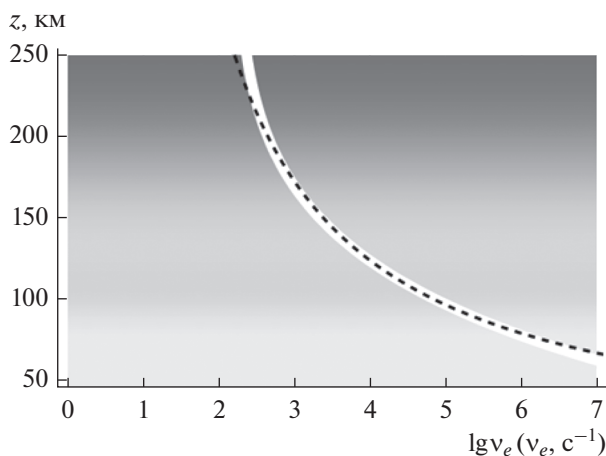


Рис. 20. Сравнение исходной зависимости (широкая белая линия) и экстраполяционной модели (пунктир).

На рис. 20 сопоставлена экстраполяционная модель (пунктирная линия), построенная по расчетным данным, и исходная зависимость.

Очевидно, что совпадение хорошее. На высоких частотах с помощью данных зондирования ЛЧМ возможна дополнительная коррекция.

ЗАКЛЮЧЕНИЕ

Таким образом, в работе развит метод восстановления эффективной частоты соударений электронов в ионосферной плазме на основе данных об ослаблении и запаздывании частотно-модулированных радиосигналов при вертикальном зондировании ионосферы Земли. На основе метода расширенной бихарактеристической системы выполнено численное моделирование как в случае однослойной, так и двухслойной модели.

Рассчитаны и приведены графики лучевых траекторий частотно-модулированных сигналов в координатах время–высота на разных частотах, построены графики зависимостей поглощения сигналов в точке приема от частоты, графики максимальной высоты отражения и времени прихода сигналов в точку отражения от частоты, получены зависимости амплитуды сигнала от времени и высоты вдоль лучей, построены зависимости вертикальной компоненты волнового вектора от высоты и времени. Разработанная методика обработки экспериментальных данных позволяет получить зависимость эффективной частоты соударений от высоты с хорошей точностью.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 20-12-00299).

СПИСОК ЛИТЕРАТУРЫ

1. Куницын В.Е., Терещенко Е.Д., Андреева Е.С. Радиотомография ионосферы. М.: Физматлит, 2007.
2. Иванов Д.В. Методы и математические модели исследования распространения в ионосфере сложных дециметровых сигналов и коррекции их дисперсионных искажений. Йошкар-Ола: Марийский гос. технич. ун-т, 2006.
3. Куркин В.И., Орлов И.И., Попов В.Н. Метод нормальных волн в проблеме коротковолновой радиосвязи. М.: Наука, 1981.
4. Andreeva E.S., Frolov V.L., Kunitsyn V.E. et al. // Radio Sci. 2016. V. 51. № 6. P. 638. <https://doi.org/10.1002/2015RS005939>
5. Крюковский А.С., Куркин В.И., Ларюнин О.А. и др. // РЭ. 2016. Т. 61. № 8. С. 794.
6. Cedrik M., Podlesnyi A., Kurkin V. // 2020 7th All-Russian Microwave Conf. (RMC). Moscow. 25–27 Nov. N.Y.: IEEE, 2020. P. 260. <https://doi.org/10.1109/RMC50626.2020.9312341>
7. Podlesnyi A., Kurkin V., Cedrik M. // 2020 7th All-Russian Microwave Conf. (RMC). Moscow. 25–27 Nov. N.Y.: IEEE, 2020. P. 263. <https://doi.org/10.1109/RMC50626.2020.9312232>
8. Ryabova M.I., Ivanov D.V., Chernov A.A., Ovchinnikov V.V. // 2020 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO-2020). Svetlogorsk. 1–3 Jul. N.Y.: IEEE, 2020. P. 9166094. <https://doi.org/10.1109/SYNCHROINFO49631.2020.9166094>
9. Дэвис К. Радиоволны в ионосфере. М.: Мир, 1973.
10. Казанцев А.Н., Лукин Д.С., Спиридонов Ю.Г. // Космич. исследования. 1967. Т. 5. № 4. С. 593.
11. Крюковский А.С., Лукин Д.С., Растягаев Д.В., Скворцова Ю.И. // РЭ. 2015. Т. 60. № 10. С. 1001.
12. Крюковский А.С., Лукин Д.С., Кирьянова К.С. // РЭ. 2012. Т. 57. № 9. С. 1028.

13. *Kryukovsky A.S., Lukin D.S., Mikhaleva E.V. et al.* // 2020 7th All-Russian Microwave Conf. (RMC). Moscow. 25–27 Nov. N.Y.: IEEE. 2020. P. 211. <https://doi.org/10.1109/RMC50626.2020.9312352>
14. *Крюковский А.С., Лукин Д.С., Михалёва Е.В., Растягаев Д.В.* // Всерос. открытая науч. конф. “Современные проблемы дистанционного зондирования, радиолокации, распространения и дифракции волн”. Муром. 25–27 мая 2021. С. 57.
15. *Бова Ю.И., Крюковский А.С., Лукин Д.С.* // РЭ. 2019. Т. 64. № 1. С. 3.
16. *Крюковский А.С., Лукин Д.С., Бова Ю.И.* // РЭ. 2020. Т. 65. № 12. С. 1160. <https://doi.org/10.31857/S0033849420120128>
17. *Крюковский А.С., Растягаев Д.В., Скворцова Ю.И.* // Вестн. Рос. нового ун-та. Сер. “Управление, вычислительная техника и информатика”. М.: РосНОУ, 2013. № 4. С. 47.
18. *Ипатов Е.Б., Крюковский А.С., Лукин Д.С. и др.* // РЭ. 2014. Т. 59. № 12. С. 1180.
19. *Kryukovskii A.S., Rastyagaev D.V., Lukin D.S.* // Russ. J. Math. Phys. 2009. V. 16. № 2. P. 251.

ЭЛЕКТРОДИНАМИКА И РАСПРОСТРАНЕНИЕ РАДИОВОЛН

УДК 537.874.6

ПЕРЕХОДНАЯ ЗОНА В ВЫСОКОЧАСТОТНОЙ ЗАДАЧЕ ДИФРАКЦИИ НА ИМПЕДАНСНОЙ ГРАНИЦЕ СО СКАЧКОМ КРИВИЗНЫ. МЕТОД КИРХГОФА И МЕТОД ПОГРАНИЧНОГО СЛОЯ

© 2022 г. Е. А. Злобина^{а, *}, А. П. Киселев^{а, б, с, **}^а Санкт-Петербургский государственный университет,
Университетская наб. 7-9, Санкт-Петербург, 199034 Российская Федерация^б Санкт-Петербургское отделение Математического института им. В.А. Стеклова РАН,
наб. р. Фонтанки 27, Санкт-Петербург, 191023 Российская Федерация^с Институт проблем машиноведения РАН,
Большой проспект В.О. 61, Санкт-Петербург 199178, Российская Федерация

*E-mail: ezlobina2@yandex.ru

**E-mail: aleksei.kiselev@gmail.com

Поступила в редакцию 24.08.2021 г.

После доработки 24.08.2021 г.

Принята к публикации 01.09.2021 г.

Исследовано волновое поле в двумерной задаче о некасательном падении высокочастотной волны от точечного источника на импедансную границу со скачком кривизны. Детально описано поле в окрестности предельного луча. Показано, что выражения, найденные в рамках метода Кирхгофа, согласуются с результатами построений методом пограничного слоя. Определены области пригодности полученных выражений, для чего подробно изучена геометрия задачи.

DOI: 10.31857/S0033849422020176

ВВЕДЕНИЕ

Интерес к высокочастотным задачам дифракции на препятствиях, обладающих в каждой точке непрерывной касательной, но имеющих разрывную кривизну, не ослабевает уже шесть десятилетий (см., например, [1–13]) – в сущности, с момента создания геометрической теории дифракции (ГТД) [13–15]. Эта тематика, помимо приложений к радиолокации, привлекает внимание исследователей тем, что для нее нет простой эталонной задачи, из которой можно было бы определить дифракционные коэффициенты¹, являющиеся базовым понятием ГТД.

Определению дифракционных коэффициентов для разных граничных условий как при некасательном, так и при касательном падении посвящены соответственно работы [1–8] и [9–12]. В этих задачах, как и вообще в теории дифракции, дифракционные коэффициенты обращаются в бесконечность при приближении к направлению геометрического отражения от границы в точке ее негладкости – *предельному лучу*. В окрестности предельного луча возникает *переходная зона*, где происходит слияние геометрически отраженной волны

и волны, дифрагированной точкой негладкости, а поле выражается через подходящие специальные функции. Для угловых областей, в простейших случаях, поля в переходных зонах описываются имеющими двухсотлетнюю историю формулами, содержащими интегралы Френеля [15]. Сейчас их обобщения систематически изучаются под именем краевых и угловых катастроф [16, 17]. Исследование переходной зоны в задаче дифракции на разрыве кривизны, которому посвящена данная статья, тоже можно отнести к этой проблематике (соответствующая угловая катастрофа имеет тип B_2 [17]).

В подавляющем большинстве работ [1–12] рассматривалась исключительно двумерная задача и использовался восходящий к работам Френеля метод Кирхгофа [15]. Метод Кирхгофа хорошо зарекомендовал себя при описании полей в переходных зонах, но не дает детального описания дифрагированного поля за их пределами [15]. Полученные в его рамках дифракционные коэффициенты для дифракции на разрыве кривизны не совпадают с выражениями, которые дает строгий метод пограничного слоя (см. [13], также ср. результаты [3] и [6]). Метод пограничного слоя [18] (в этой тематике восходящий к работе [3], имеющей большую эвристическую составляющую)

¹ Так Дж. Б. Келлер назвал диаграммы направленности волн, излучаемых точками негладкости границы [14, 15].

строго описывает поле в окрестности сингулярной точки и дает выражение для дифрагированной волны, пригодное на любых расстояниях [8]. Однако полученное с его помощью выражение для поля в переходной зоне ограничено пока малыми расстояниями [8]. Для некасательного падения плоской волны и условия Дирихле, в рамках последовательного погранслоного подхода, влияние скачка кривизны на волновое поле в переходной зоне описывается выражением [8]

$$\sqrt{\frac{2}{\pi}} \frac{[\alpha] r_1}{\sin \varphi_0} \exp\left(ikr_1 - i\frac{z^2}{2}\right) D_{-3}\left(\sqrt{2}z \exp\left(-i\frac{\pi}{4}\right)\right), \quad (1)$$

где $kr_1 \gg 1$, а $[\alpha]$ – величина скачка кривизны. Здесь r_1 и φ_1 – полярные координаты с центром в точке негладкости границы, $\varphi_0 > 0$ – значение угла падения в точке негладкости, и

$$z = \sqrt{kr_1/2}(\varphi_1 - \varphi_0). \quad (2)$$

Присутствие переменной (2) в аргументе специальной функции типично для описания переходных зон, где сливаются плоская и цилиндрическая волны [15].

Выражение (1) применимо только на малых расстояниях $kr_1 \ll 1$ (здесь κ – характеристический параметр, определяемый геометрией границы, см. (6)). Это обусловлено, во-первых, очевидным фактом линейного роста амплитуды по r_1 . Во-вторых, тем требующим углубления в геометрические рассмотрения обстоятельством, что формула (1) перестает сшиваться с геометрически отраженной волной при $|z| \gg 1$, если не выполняются условия $\kappa r_1 \ll 1$ и $\kappa r_1 z^2 \ll 1$ (подробнее см. [8]).

Вместе с тем с точки зрения метода Кирхгофа представляется, что результат для переходной зоны должен быть другим [6, 13]. Для получения главного члена асимптотики, описывающей влияние разрыва кривизны на уходящее поле, достаточно рассматривать границы, составленные из двух касающихся дуг окружностей разного радиуса. Естественно (и это делалось в работах [6, 13]) разбить кирхгофовский интеграл на две соответствующие части, которые без большого труда выражаются через интегралы Френеля от сложных (не таких, как в дифракции на клине) аргументов, зависящих от кривизны соответствующей дуги. Интеграл Френеля простым образом выражается через функцию параболического цилиндра D_{-1} , в то время как погранслоные рассуждения приводят к специальной функции D_{-3} , см. (1). Разрешение этого (как выяснится, кажущегося) противоречия и составляет один из основных результатов данной работы.

Рассмотрена двумерная задача дифракции высокочастотной волны от точечного источника на

импедансной границе со скачком кривизны. Падающая волна приходит в точку негладкости границы некасательно. Граница предполагается пассивной, т.е. поверхностных волн не возникает. Все рассуждения ведутся в высокочастотном приближении: расстояния от точки негладкости до источника и точки наблюдения много больше длины волны. Основываясь на методе Кирхгофа, мы детально исследуем поле в узкой окрестности предельного луча на умеренных (по сравнению с геометрическим параметром задачи κ) расстояниях. Интересующий нас эффект слияния геометрически отраженной и дифрагированной волн асимптотически описывается суммой двух френелевских выражений. Обращаясь к малым расстояниям, мы аналитически прослеживаем переход френелевской асимптотики в выражение, аналогичное (1).

Поле в переходной зоне вблизи предельного луча уже исследовалось методом Кирхгофа в работах [13] и [6]. Однако в [13] были рассмотрены только граничные условия Дирихле и Неймана, а в [6] – специфическое импедансное условие, при котором возникает поверхностная волна, причем авторов преимущественно интересовало поле в случае слияния поверхностной волны с дифрагированной и отраженной волнами. Кроме того, ранее не обсуждались границы применимости возникающих формул и не давалась геометрическая интерпретация аргументов соответствующих спецфункций.

В данной работе, как и в [8], большое внимание уделяется геометрическим рассуждениям, благодаря которым удастся подробно проследить за сшиванием неравномерных по углу асимптотических формул для окрестности предельного луча с классическими геометрическими формулами, непригодными в его окрестности. Практически для всех полученных выражений мы выписываем соответствующие погрешности, что позволяет точно указать границы их применимости. Помимо этого, геометрический анализ позволил проследить, как аргументы спецфункций зависят от формы границы и кривизны фронта падающей волны.

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим дифракцию на границе C , составленной из дуг двух касающихся окружностей разного радиуса. Точки на контуре будем характеризовать длиной дуги s , отсчитываемой от точки касания O (считаем, что $s > 0$ справа от O , рис. 1). Кривизна границы имеет вид

$$\alpha(s) = \alpha_- + [\alpha]\theta(s), \quad (3)$$

где

$$\theta(s) = \begin{cases} 1, & s > 0, \\ 0, & s \leq 0 \end{cases} \quad (4)$$

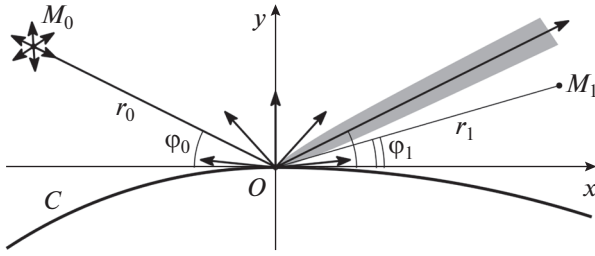


Рис. 1. Дифрагированная волна и переходная зона.

– функция Хевисайда;

$$[\mathfrak{a}] = \mathfrak{a}_+ - \mathfrak{a}_- \tag{5}$$

– амплитуда скачка кривизны (\mathfrak{a}_\pm – константы). Кривизна может быть положительной, отрицательной или знакопеременной. Важным геометрическим параметром задачи является величина

$$\kappa = \max\{|\mathfrak{a}_-|, |\mathfrak{a}_+|\}. \tag{6}$$

Введем декартову систему координат, как показано на рис. 1: начало координат расположено в точке скачка кривизны O , и ось x касается контура.

Предполагается гармоническая зависимость от времени вида $\exp(-i\omega t)$. Здесь ω – круговая частота, связанная с волновым числом k соотношением $\omega/c = k$, и $c = \text{const}$ – скорость распространения волн, которая принимается равной единице: $c = 1$. Полное волновое поле u над контуром C описывается уравнением Гельмгольца

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2 \right) u = -\delta(x - x_0)\delta(y - y_0), \tag{7}$$

где $M_0 = (x_0, y_0)$ – точка источника. Большим параметром в нашей задаче является величина

$$k/\kappa \gg 1. \tag{8}$$

На контуре выполнено импедансное граничное условие:

$$\left(\frac{\partial}{\partial n} + ikg \right) u \Big|_C = 0. \tag{9}$$

Импеданс g предполагается постоянным. В предельных случаях $g = \infty$ и $g = 0$ условие (9) переходит в условие Дирихле и в условие Неймана соответственно. На бесконечности u удовлетворяет условию излучения.

В соответствии с (7) поле u разбивается на падающую и уходящую волны

$$u = u^{\text{пад}} + u^{\text{yx}}, \tag{10}$$

причем падающая волна возбуждается точечным источником колебаний, расположенным в точке M_0 , и имеет вид

$$u^{\text{пад}}(x, y) = G(x, y; x_0, y_0) := \frac{i}{4} H_0^{(1)} \left(k \sqrt{(x - x_0)^2 + (y - y_0)^2} \right). \tag{11}$$

Здесь введено обозначение для удовлетворяющей условию излучения функции Грина для уравнения Гельмгольца в свободном пространстве, $H_0^{(1)}$ – функция Ганкеля. Предполагается, что падающая волна приходит в точку O некасательно.

Согласно ГТД (см., например, [13–15]), уходящее поле в грубом приближении есть сумма волн $u^{\text{отр}}$, геометрически отраженных от гладких частей контура, и волны $u^{\text{диф}}$, дифрагированной точкой негладкости границы:

$$u^{\text{yx}} = u^{\text{отр}} + u^{\text{диф}}. \tag{12}$$

Для некасательного падения лучевой метод (см., например, [19]) дает для $u^{\text{отр}}$ явное выражение, теряющее гладкость на предельном (геометрически отраженном в точке негладкости контура) луче.

Везде далее предполагается, что точка источника $M_0 = (x_0, y_0)$ и точка наблюдения $M_1 = (x_1, y_1)$ лежат на больших по сравнению с длиной волны расстояниях от точки негладкости O :

$$kr_0 \gg 1, \quad kr_1 \gg 1. \tag{13}$$

Тогда дифрагированная волна $u^{\text{диф}}$ является цилиндрической волной:

$$u^{\text{диф}} = A(\varphi_0, \varphi_1; k) \frac{\exp(ik(r_0 + r_1))}{k\sqrt{r_0 r_1}} \times [1 + \mathfrak{E}(r_0, r_1, \varphi_0, \varphi_1; k)]. \tag{14}$$

Здесь $r_{0,1} = |OM_{0,1}|$ – расстояния от точки O до точек источника M_0 и наблюдения M_1 , $\varphi_{0,1}$ – углы между осью Ox и прямыми OM_0 и OM_1 соответственно (см. рис. 1), A – дифракционный коэффициент, а \mathfrak{E} – неравномерная по углам погрешность. Для дифракции на негладких препятствиях (см., например, [3, 8, 15, 20, 21]) характерно обращение A и \mathfrak{E} в бесконечность на предельном луче, где $\varphi_1 = \varphi_0$. В наших рассмотрениях $\varphi_0 > 0$.

В области над границей u^{yx} удовлетворяет однородному уравнению Гельмгольца. Формула Грина (см., например, [15]) позволяет выразить поле u^{yx} в точке M_1 через значения собственно его и его нормальной производной на границе:

$$u^{\text{yx}}(x_1, y_1) = \int_C ds \left(u^{\text{yx}}(X, Y) \frac{\partial G}{\partial n}(X, Y; x_1, y_1) - G(X, Y; x_1, y_1) \frac{\partial u^{\text{yx}}}{\partial n}(X, Y) \right). \tag{15}$$

Здесь $X = X(s)$ и $Y = Y(s)$ – координаты переменной точки $N(s)$ на контуре C , $\partial/\partial n$ – производная по внутренней нормали к границе области. Метод Кирхгофа основан на подстановке в (15) геометрических аппроксимаций u^{yx} и $\partial u^{yx}/\partial n$ (см., например, [15]).

Нас интересует уходящее поле u^{yx} , во-первых, в области, где $kr_1 \sim 1$, и, во-вторых, в области, где $kr_1 \ll 1$.

2. ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ИНТЕГРАЛА КИРХГОФА

2.1. Поле и его нормальная производная на границе

Согласно традиционному кирхгофовскому подходу [15], значения уходящего поля и его нормальной производной на границе принимаются равными соответственно:

$$\begin{aligned} u^{yx}(s) &= R(s)u^{\text{пад}}(s), \\ \frac{\partial u^{yx}}{\partial n}(s) &= -R(s)\frac{\partial u^{\text{пад}}}{\partial n}(s), \end{aligned} \quad (16)$$

где $u^{yx}(s) := u^{yx}(X(s), Y(s))$, а $R(s)$ – коэффициент отражения плоской волны от соответствующей плоской границы. Для импедансного граничного условия (9) он равен [22]

$$R(s) = \frac{\cos \alpha(s) - g}{\cos \alpha(s) + g}, \quad (17)$$

где $\alpha(s)$ – значение угла падения в точке контура $N(s)$. В предельных случаях Дирихле ($g = \infty$) и Неймана ($g = 0$) $R = -1$ и $R = 1$ соответственно. В отличие от работ [6, 21] мы предполагаем, что при $s = 0$ коэффициент отражения $R(s)$ не обращается в нуль или бесконечность.

2.2. Виртуальные лучи

Обозначим расстояние от произвольной точки контура $N(s) = (X(s), Y(s))$ до точки источника $M_0 = (x_0, y_0)$ через $l_0(s)$, а до точки наблюдения $M_1 = (x_1, y_1)$ через $l_1(s)$ (рис. 2):

$$l_{0,1}(s) := |NM_{0,1}| = \sqrt{(X(s) - x_{0,1})^2 + (Y(s) - y_{0,1})^2}. \quad (18)$$

Отметим, что $r_{0,1} = l_{0,1}(0)$. Пользуясь асимптотической функцией Ганкеля [23]

$$H_0^{(1)}(L) = \sqrt{2/\pi L} \exp(iL - i\pi/4) \left(1 + O(L^{-1})\right), \quad (19)$$

$L \gg 1,$

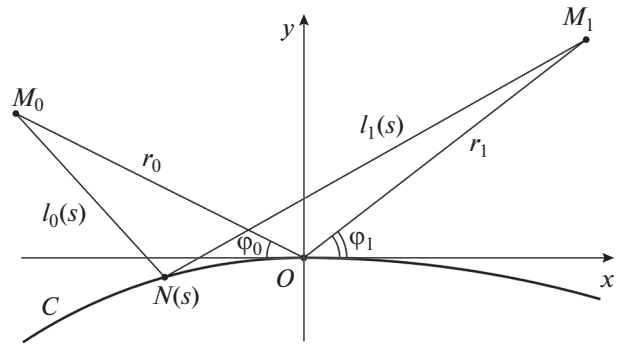


Рис. 2. Виртуальные лучи.

получим при $kl_{0,1}(s) \gg 1$

$$\begin{aligned} G(X, Y; x_{0,1}, y_{0,1}) &= \\ &= \frac{\exp(ikl_{0,1}(s) + i\pi/4)}{\sqrt{8\pi kl_{0,1}(s)}} \left(1 + O\left(\frac{1}{kl_{0,1}(s)}\right)\right), \\ \frac{\partial G}{\partial n}(X, Y; x_{0,1}, y_{0,1}) &= \\ &= ik \frac{\partial l_{0,1}}{\partial n}(s) G(X, Y; x_{0,1}, y_{0,1}) \left(1 + O\left(\frac{1}{kl_{0,1}(s)}\right)\right). \end{aligned} \quad (20)$$

Подынтегральная функция в (15) принимает вид

$$\begin{aligned} &-\frac{R(s)}{8\pi} \left(\frac{\partial l_0}{\partial n}(s) + \frac{\partial l_1}{\partial n}(s)\right) \frac{\exp(ik\tau(s))}{\sqrt{l_0(s)l_1(s)}} \times \\ &\times \left(1 + O\left(\frac{1}{kl_0(s)}\right) + O\left(\frac{1}{kl_1(s)}\right)\right). \end{aligned} \quad (21)$$

Здесь τ – длина ломаной $M_0N(s)M_1$:

$$\tau(s) = l_1(s) + l_0(s). \quad (22)$$

Таким образом, интеграл Кирхгофа ((15) с подстановкой (16) и (20)) имеет вид суммы вкладов виртуальных лучей $M_0N(s)M_1$. Поле в точке наблюдения, лежащей вблизи предельного луча, формируется в результате синфазного сложения виртуальных лучей, у которых точки $N(s)$ лежат вблизи O .

2.3. Разложения при малых s

Выражение для длины звена $N(s)M_{0,1}$ виртуального луча зависит от положения точки $N(s)$ относительно O . Обозначим ее $l_{0,1}^+$ в случае $s > 0$ и $l_{0,1}^-$ в случае $s < 0$. Раскладывая $l_{0,1}(s)$ (см. (18)) при малых s справа и слева от нуля до квадратич-

ных членов, в результате прямого вычисления получим²

$$l_0^\pm(s) = r_0 + s \cos \varphi_0 + \frac{s^2}{2} \left(\alpha_\pm \sin \varphi_0 + \frac{\sin^2 \varphi_0}{r_0} \right) + O \left(\frac{s^3}{r_0^2} (1 + \kappa^2 r_0^2) \right). \quad (23)$$

Выражения для l_1^\pm получаются заменой индекса 0 на 1 и знака плюс перед линейным по s членом на минус. Обозначения $r_{0,1}$ и $\varphi_{0,1}$ введены после (14), см. также рис. 2. Для длины виртуального луча (22) получаем

$$\begin{aligned} \tau_\pm(s) &= l_0^\pm(s) + l_1^\pm(s) = \\ &= r_0 + r_1 + s(\cos \varphi_0 - \cos \varphi_1) + \\ &+ \frac{s^2}{2} \left(\alpha_\pm (\sin \varphi_0 + \sin \varphi_1) + \frac{\sin^2 \varphi_0}{r_0} + \frac{\sin^2 \varphi_1}{r_1} \right) + \\ &+ O \left(s^3 \left(\kappa^2 + \frac{1}{r_0^2} + \frac{1}{r_1^2} \right) \right). \end{aligned} \quad (24)$$

Здесь следует взять индекс $+$ ($-$), если точка $N(s)$ лежит справа (слева) от O . Из (18) нетрудно получить, что

$$\partial l_{0,1}(s) / \partial n|_{s=0} = -\sin \varphi_{0,1}. \quad (25)$$

Подынтегральная функция (21) при малых s переписывается в виде

$$\begin{aligned} &\frac{R(0) \sin \varphi_0 + \sin \varphi_1}{8\pi \sqrt{r_0 r_1}} \exp(ik\tau_\pm(s)) \times \\ &\times \left[1 + O \left(\frac{1}{kr_0} + \frac{1}{kr_1} \right) + O \left(s \left(\kappa + \frac{1}{r_0} + \frac{1}{r_1} \right) \right) \right]. \end{aligned} \quad (26)$$

2.4. Замена переменной в интеграле Кирхгофа

Для точки наблюдения M_1 , расположенной вблизи предельного луча, $\varphi_1 \approx \varphi_0$ (см. рис. 2), и выражение (24) мгновенно переписывается в виде

$$\begin{aligned} \tau_\pm(s) &= r_0 + r_1 + s \sin \varphi_0 \delta\varphi + \\ &+ \frac{s^2}{2\rho} \sin^2 \varphi_0 \left(\frac{2\alpha_\pm \rho}{\sin \varphi_0} + 1 \right) + \epsilon(s), \end{aligned} \quad (27)$$

где $\epsilon(s) = O(s(\delta\varphi)^2) + O(s^2 \delta\varphi(\kappa + r_0^{-1} + r_1^{-1})) + O(s^3(\kappa^2 + r_0^{-2} + r_1^{-2}))$. Здесь введены важные для дальнейшего изложения обозначения

$$\delta\varphi = \varphi_1 - \varphi_0, \quad \rho = r_0 r_1 / (r_0 + r_1). \quad (28)$$

Величина ρ часто встречается при описании полей в областях слияния двух цилиндрических волн (см., например, [24, 25]). Очевидно, что если κr_1

много меньше или порядка единицы, то $\kappa\rho$ тоже много меньше или порядка единицы. Везде далее

$$\kappa\rho \gg 1, \quad (29)$$

что равносильно (13)³.

Перейдем в интеграле Кирхгофа к безразмерной переменной

$$p = s\sqrt{k/\rho} \sin \varphi_0 \exp(-i\pi/4). \quad (30)$$

Поскольку падающая волна приходит в точку O некасательно, $\sin \varphi_0 > 0$. Фаза подынтегральной функции $ik\tau_\pm$ принимает вид

$$\begin{aligned} ik\tau_\pm &= ik(r_0 + r_1) - (\xi p + J_\pm p^2/2) + \\ &+ O((p\xi + p^2)\delta\varphi) + O(p^3/\sqrt{k\rho}). \end{aligned} \quad (31)$$

Здесь введены обозначения

$$\xi = \sqrt{k\rho} \delta\varphi \exp(-i\pi/4), \quad (32)$$

$$J_\pm = 1 + 2\alpha_\pm \rho / \sin \varphi_0, \quad (33)$$

см. (28). В случае выпуклого контура $2\alpha_\pm \rho / \sin \varphi_0 > 0$, и $J_\pm \sim 1$ как при $\kappa\rho \ll 1$, так и при $\kappa\rho \sim 1$. Если же контур имеет вогнутый участок, например $\alpha_- < 0$, то J_- обращается в нуль при $\rho = -\sin \varphi_0 / 2\alpha_-$, что соответствует каустике. В таком случае мы ограничимся рассмотрением области, где $\rho + \sin \varphi_0 / 2\alpha_\pm$ не мало, и тогда $J_\pm \sim 1$.

2.5. Разбиение интеграла и упрощение подынтегрального выражения

Следуя стандартной схеме (см., например, [6, 13]), представим интеграл Кирхгофа в виде суммы двух интегралов по участкам $p < 0$ и $p > 0$, на которых подынтегральная функция гладкая. Возникшие интегралы исследуем с помощью традиционных приемов (см., например, [26]): упрощаем подынтегральное выражение для малых p , удерживая квадратичные члены в фазе и главные члены в амплитуде, переносим малые члены из фазы в амплитуду и распространяем интегрирование на промежутки $(-\infty, 0)$ и $(0, +\infty)$ соответственно.

Таким образом, учитывая (26), (30) и (31), получаем выражение для уходящего поля вблизи предельного луча

$$u^{yx} \approx \frac{R(0) \exp(i\pi/4) \exp(ik(r_0 + r_1))}{4\pi \sqrt{k(r_0 + r_1)}} (I_- + I_+), \quad (34)$$

² Здесь использованы выведенные в Приложении выражения (П.4) для координат точки, лежащей на границе.

³ Это видно из равносильности неравенства $1/\kappa\rho = 1/\kappa r_0 + 1/\kappa r_1 \ll 1$ паре неравенств $1/\kappa r_0 \ll 1$ и $1/\kappa r_1 \ll 1$.

где

$$I_{\pm} = \int_{-\infty}^0 dp \exp\left(\pm \xi p - J_{\pm} \frac{p^2}{2}\right) (1 + \epsilon(p)), \quad (35)$$

причем

$$\epsilon(p) = O(\delta\varphi) + O(1/k\rho) + O(p\sqrt{k\rho}(\delta\varphi)^2) + O(p^2\delta\varphi) + O(p^3/\sqrt{k\rho}).$$

3. ДЕТАЛЬНОЕ ИССЛЕДОВАНИЕ ИНТЕГРАЛА КИРХГОФА

3.1. Асимптотика уходящего поля вблизи предельного луча при $k\rho \sim 1$

Рассмотрим сначала поле в области, где $k\rho \sim 1$. Перейдем в интегралах I_{\pm} к переменным $p\sqrt{J_{\pm}/2} \mp \xi/\sqrt{2J_{\pm}}$, после чего интегралы легко переписываются в виде

$$I_{\pm} = \int_{-\infty}^0 dp \exp\left(\pm \xi p - J_{\pm} \frac{p^2}{2}\right) (1 + \epsilon(p)) = \sqrt{\frac{2\pi}{J_{\pm}}} \exp\left(\frac{\xi^2}{2J_{\pm}}\right) \Phi\left(\mp \frac{\xi \exp(i\pi/4)}{\sqrt{2J_{\pm}}}\right) (1 + \epsilon), \quad (36)$$

где Φ – интеграл Френеля [15],

$$\Phi(t) = \frac{\exp(-i\pi/4)}{\sqrt{\pi}} \int_{-\infty}^t dq \exp(iq^2), \quad (37)$$

а $\epsilon = O(\sqrt{k\rho}(\delta\varphi)^2) + O(1/\sqrt{k\rho})$. Оценка погрешности получена тем же способом, что и в [8].

Введем обозначение для выражения, входящего в аргументы интегралов Френеля в (36):

$$\zeta = \xi \exp(i\pi/4)/\sqrt{2} = \sqrt{k\rho/2} \delta\varphi, \quad (38)$$

см. (32). Величина ζ характерна для аргументов спецфункций, описывающих поля в областях слияния двух цилиндрических волн. Когда источник удаляется на бесконечность вдоль фиксированного направления ($k\rho_0 \gg 1$, $\varphi_0 = \text{const}$), ζ переходит в величину (2), типичную для областей слияния цилиндрической и плоской волн.

Формула (34) для уходящего поля принимает вид

$$u^{yx} = \frac{R(0)}{2\sqrt{2\pi}} \frac{\exp\left(ik(r_1 + r_0) + \frac{i\pi}{4}\right)}{\sqrt{k(r_1 + r_0)}} \times \left(A_- \Phi\left(\frac{\zeta}{\sqrt{J_-}}\right) + A_+ \Phi\left(-\frac{\zeta}{\sqrt{J_+}}\right) \right) (1 + \epsilon), \quad (39)$$

где ϵ определено после (37), а

$$A_{\pm} = \exp(-i\zeta^2/J_{\pm})/\sqrt{J_{\pm}}. \quad (40)$$

Похожая формула была получена в работе [6], в которой авторы не следили за остаточными членами. В предельных случаях Дирихле и Неймана формула согласуется с выражением, полученным в [13] для плоской падающей волны.

В задачах дифракции цилиндрической волны на клине [24] и плоской волны на ленте [25] переменная ζ выступала как аргумент специальных функций, описывавших поля в переходных зонах. В [6], где рассматривалась дифракция цилиндрической волны на криволинейной границе с разрывом кривизны, аргументами таких функций были, как и в (39), отношения $\zeta/\sqrt{J_{\pm}}$. В случае клина и ленты кривизна гладких участков границы равняется нулю, и $J_{\pm} = 1$. Отметим, что, когда $\alpha_{\pm} = 0$, формула (39) описывает поле, отраженное от соответствующей плоской границы.

3.2. Асимптотика уходящего поля при $|\zeta| \gg 1$

Асимптотика интеграла Френеля имеет вид [15]

$$\Phi(t) = \theta(t) - \frac{\exp(it^2 - i\pi/4)}{2\sqrt{\pi}} \left(\frac{i}{t} + \frac{1}{2t^3} + O\left(\frac{1}{t^5}\right) \right), \quad (41)$$

$|t| \gg 1.$

Отметим, что $\theta(t)$ равно нулю при отрицательных t и равно единице для положительных t . При $|\zeta| \gg 1$

$$A_- \Phi\left(\frac{\zeta}{\sqrt{J_-}}\right) = A_- \theta(\zeta) - \frac{\exp(i\pi/4)}{2\sqrt{\pi}\zeta} - \frac{J_- \exp(-i\pi/4)}{4\sqrt{\pi}\zeta^3} + O\left(\frac{1}{\zeta^5}\right),$$

$$A_+ \Phi\left(-\frac{\zeta}{\sqrt{J_+}}\right) = A_+ \theta(-\zeta) + \frac{\exp(i\pi/4)}{2\sqrt{\pi}\zeta} + \frac{J_+ \exp(-i\pi/4)}{4\sqrt{\pi}\zeta^3} + O\left(\frac{1}{\zeta^5}\right). \quad (42)$$

Из формулы (39), в результате сокращения вторых слагаемых в правых частях (42) получим

$$u^{yx} = \frac{R(0)}{2\sqrt{2\pi}} \frac{\exp\left(ik(r_1 + r_0) + \frac{i\pi}{4}\right)}{\sqrt{k(r_1 + r_0)}} \times \left(A_- \theta(\zeta) + A_+ \theta(-\zeta) \frac{J_+ - J_-}{4\sqrt{\pi}\zeta^3} \exp\left(-\frac{i\pi}{4}\right) + O\left(\frac{1}{\zeta^5}\right) \right) \times (1 + \epsilon), \quad |\zeta| \gg 1, \quad (43)$$

где A_{\pm} определены в (40), а погрешность ϵ – после (37). Интересно, что величины J_{\pm} (см. (33)) связаны с геометрическими расхождениями \mathcal{F}_{\pm} цилиндрических волн, отраженных от дуг окружностей с радиусами α_{\pm}^{-1} [19]:

$$\mathcal{F}_{\pm} = r_0 + r_1 + 2\alpha_{\pm} r_0 r_1 / \sin \varphi_0 = (r_0 + r_1) J_{\pm}. \quad (44)$$

С использованием выражений (38) и (44) формула (43) переписывается в геометрических терминах:

$$u^{yx} = \frac{R(0)}{2\sqrt{2\pi}} \exp(ik(r_1 + r_0) + i\pi/4) \times \\ \times [\mathcal{A}_-\theta(\delta\varphi) + \mathcal{A}_+\theta(-\delta\varphi)](1 + \varepsilon) + \\ + \frac{[\mathfrak{a}]R(0)}{2\pi \sin \varphi_0} \frac{\exp(ik(r_1 + r_0))}{(\delta\varphi)^3} \frac{1}{k^2 \sqrt{r_0 r_1}} \times \\ \times (1 + \hat{\varepsilon}), \quad k\rho(\delta\varphi)^2 \gg 1. \quad (45)$$

Здесь

$$\mathcal{A}_\pm = \exp(-ikr_0 r_1 (\delta\varphi)^2 / 2\mathcal{F}_\pm) / \sqrt{k\mathcal{F}_\pm} = \\ = A_\pm / \sqrt{k(r_0 + r_1)}, \quad (46)$$

см. (40), а погрешность имеет вид

$$\hat{\varepsilon} = O(\sqrt{k\rho}(\delta\varphi)^2) + O(1/(k\rho(\delta\varphi)^2)).$$

Первое слагаемое в правой части отлично от нуля для положительных $\delta\varphi$ и соответствует волне, отраженной от контура слева от точки O . Второе слагаемое отлично от нуля при отрицательных $\delta\varphi$ и отвечает волне, отраженной от контура справа от точки O . Третье слагаемое описывает дифрагированную волну, ее диаграмма направленности согласуется с формулами, полученными методом пограничного слоя в [3, 8], где рассматривалась плоская падающая волна.

Остаточные члены в (45) малы при выполнении неравенств (29), $k\rho(\delta\varphi)^4 \ll 1$ и

$$k\rho(\delta\varphi)^2 \gg 1. \quad (47)$$

Последнее условие означает, что точка наблюдения не слишком близка к предельному лучу.

В Приложении показано, что выражения для отраженных волн в (45) согласуются с лучевыми формулами [19] в области, где выполнено неравенство

$$k\rho|\delta\varphi|^3 \ll 1, \quad (48)$$

характеризующее ширину переходной зоны (отсюда вытекает малость $k\rho(\delta\varphi)^4$).

3.3. Волновое поле в области $k\rho \ll 1$

Рассмотрим теперь область, где $k\rho \ll 1$ (но $k\rho \gg 1$, см. (29)). Покажем, что выражение (39), полученное выше методом Кирхгофа, согласуется с выглядящей качественно иначе формулой (1), полученной в [8] методом пограничного слоя. Учитывая выражения (33), перепишем фазы экспонент в (35) в виде

$$-\xi p - J_\pm \frac{p^2}{2} = -\xi p - \frac{p^2}{2} - \frac{\mathfrak{a}_\pm \rho}{\sin \varphi_0} p^2. \quad (49)$$

Поскольку $k\rho \ll 1$, последнее слагаемое мало. Перенесем его в амплитуду и получим

$$I_\pm = \int_0^{+\infty} dp \exp\left(\mp \xi p - \frac{p^2}{2}\right) \times \\ \times \left(1 - \frac{\mathfrak{a}_\pm \rho}{\sin \varphi_0} p^2 + O(p^4 \kappa^2 \rho^2)\right) (1 + \varepsilon(p)) = \\ = \int_0^{+\infty} dp \exp\left(\mp \xi p - \frac{p^2}{2}\right) (1 + \varepsilon(p)) - \\ - \frac{\mathfrak{a}_\pm \rho}{\sin \varphi_0} \int_0^{+\infty} dp p^2 \exp\left(\mp \xi p - \frac{p^2}{2}\right) \times \\ \times (1 + \varepsilon(p) + O(p^2 k\rho)). \quad (50)$$

Сумма интегралов I_\pm принимает вид (поправочные члены не приводим)

$$I_+ + I_- \approx \int_{-\infty}^{+\infty} dp \exp\left(-\xi p - \frac{p^2}{2}\right) - \frac{\mathfrak{a}_-\rho}{\sin \varphi_0} \times \\ \times \int_{-\infty}^{+\infty} dp p^2 \exp\left(-\xi p - \frac{p^2}{2}\right) - \\ - \frac{[\mathfrak{a}]\rho}{\sin \varphi_0} \int_0^{+\infty} dp p^2 \exp\left(-\xi p - \frac{p^2}{2}\right) = \\ = \sqrt{2\pi} \exp\left(\frac{\xi^2}{2}\right) \left(1 - \frac{\mathfrak{a}_-\rho}{\sin \varphi_0} (\xi^2 + 1)\right) - \\ - \frac{2[\mathfrak{a}]\rho}{\sin \varphi_0} \exp\left(\frac{\xi^2}{4}\right) D_{-3}(\xi). \quad (51)$$

Здесь мы явно выделили линейный по амплитуде скачка кривизны $[\mathfrak{a}]$ член и воспользовались интегральным представлением для функции параболического цилиндра [23]:

$$D_\nu(t) = \frac{\exp(-t^2/4)}{\Gamma(-\nu)} \int_0^{+\infty} dp p^{-\nu-1} \exp\left(-tp - \frac{p^2}{2}\right), \quad (52) \\ \nu < 0,$$

а первые два интеграла вычислили, пользуясь тем, что

$$\int_{-\infty}^{\infty} dq \exp(-\beta q^2) = \sqrt{\pi/\beta}$$

(см. [23]). Подставляя (51) в выражение для полного поля (34), получим

$$u^{yx} = \frac{R(0)}{2\sqrt{2\pi}} \frac{\exp(ik(r_1 + r_0) + i\pi/4)}{\sqrt{k(r_1 + r_0)}} \times \\ \times (U_0 + U_1 + V_1). \quad (53)$$

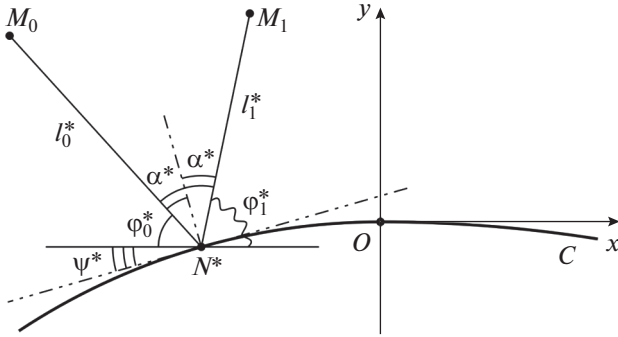


Рис. 3. Геометрическое отражение.

Здесь

$$U_0 = \exp(-i\xi^2)(1 + \varepsilon_1), \quad (54)$$

$$U_1 = \frac{\varkappa - \rho}{\sin \varphi_0} (2i\xi^2 - 1) \exp(-i\xi^2)(1 + \varepsilon_1 + \varepsilon_2), \quad (55)$$

$$V_1 = \sqrt{\frac{2}{\pi}} \frac{[\varkappa] \rho}{\sin \varphi_0} \exp\left(-i\frac{\xi^2}{2}\right) \times D_{-3}\left(\sqrt{2}\xi \exp\left(-i\frac{\pi}{4}\right)\right) (1 + \varepsilon_1 + \varepsilon_2), \quad (56)$$

$\xi = \sqrt{k\rho/2} \delta\varphi$, погрешности имеют вид

$$\varepsilon_1 = O(1/k\rho) + O(k\rho(\delta\varphi)^3),$$

$$\varepsilon_2 = O(k\rho) + O(k\rho^2(\delta\varphi)^2).$$

Поправочные члены малы при выполнении условий (29), (48) и

$$k\rho \ll 1, \quad k\rho^2(\delta\varphi)^2 \ll 1. \quad (57)$$

В Приложении показано, что в области, где справедливы неравенства (29), (48) и (57), первые два слагаемых в (53) соответствуют геометрически отраженной волне. Главный (относительно малого параметра $k\rho$) член, отвечающий U_0 , описывает отраженную волну в главном приближении и представляет собой волну от мнимого (отраженного относительно оси Ox) источника. Члены,

отвечающие U_1 и V_1 , учитывают влияние кривизны границы.

Выражение для линейного по $[\varkappa]$ члена для граничных условий Дирихле ($R(0) = -1$) при удалении источника на бесконечность ($k\rho_0 \gg 1$ с выполнением условий (57)) согласуется с формулой (1), полученной в рамках погранслоевой техники для случая падения плоской волны [8].

ЗАКЛЮЧЕНИЕ

В рамках метода Кирхгофа найдены выражения для уходящего поля в переходной зоне на умеренных и малых расстояниях от границы ($k\rho$ порядка единицы или мало). Детально изучены области их применимости. Установлено, что ширина переходной зоны характеризуется неравенством $k\rho(\delta\varphi)^3 \ll 1$, см. (48). Показано, что в случае, когда точка наблюдения находится на малых расстояниях от границы ($k\rho \ll 1$), найденные асимптотики согласуются с полученными в [8] методом пограничного слоя. Приемы, использованные в работе, могут быть применены в задачах дифракции на препятствиях с кривизной, имеющих более слабые особенности [20, 21].

ПРИЛОЖЕНИЕ

АНАЛИЗ ЛУЧЕВОЙ ФОРМУЛЫ ДЛЯ ОТРАЖЕННОЙ ВОЛНЫ

Проанализируем классическую (см., например, [19]) лучевую формулу для отраженной волны $u^{\text{отр}}$ в окрестности предельного луча на умеренных ($k\rho \sim 1$) и на малых ($k\rho \ll 1$) расстояниях от точки негладкости O . Это позволит определить области пригодности полученных выше формул.

П.1. Формула лучевого метода

В случае гладкой границы лучевой метод [19] дает для геометрически отраженной волны следующую формулу:

$$u^{\text{отр}}(M_1) = \frac{R^*}{2\sqrt{2\pi}} \frac{\exp(ik\tau^* + i\pi/4)}{\sqrt{k\mathcal{F}^*}} \left(1 + O\left(\frac{1}{k\mathcal{F}^*}\right) \right), \quad \mathcal{F}^* = l_0^* + l_1^* + \frac{2\varkappa^* l_0^* l_1^*}{\cos \alpha^*}. \quad (\text{П.1})$$

Здесь $l_{0,1}^*$ — расстояния соответственно от точек источника M_0 и наблюдения M_1 до точки геометрического отражения N^* , $\tau^* = l_0^* + l_1^*$ — значение эйконала в точке M_1 (геометрическое время пробега), α^* — угол падения в точке N^* , \varkappa^* — значение кривизны в точке N^* , R^* — коэффициент отражения в точке N^* (рис. 3). Для границы, кривизна которой имеет скачок в точке O , выражение (П.1)

правильно описывает волны, геометрически отраженные от гладких частей границы справа и слева от O . Однако выражение (П.1) имеет разрыв на предельном луче и поэтому непригодно в узкой его окрестности⁴.

⁴ В работе [27] показано, что в случае разрыва кривизны при подстановке выражения (П.1) в уравнение Гельмгольца на предельном луче возникают дельта-функция и ее производные.

Упростим лучевую формулу (П.1) для точек наблюдения M_1 , лежащих вблизи предельного луча, $|\delta\varphi| \ll 1$, подобно тому как это сделано в [8, 20] для случая падения плоской волны. Умеренные и малые расстояния рассмотрим по отдельности.

П.2. Умеренные расстояния

Пусть сначала $k\rho \sim 1$. Получим выражение для длины дуги s^* , соответствующей точке отражения N^* (см. рис. 3). Ясно, что для точек наблюдения из узкой окрестности предельного луча ($|\delta\varphi| \ll 1$) точка отражения лежит вблизи O , т.е. s^* мало.

Согласно закону геометрического отражения

$$\varphi_0^* + \psi^* = \varphi_1^* - \psi^*, \tag{П.2}$$

где $\varphi_{0,1}^* = \varphi_{0,1}(s^*)$ – углы между осью Ox и прямыми $N_*M_{0,1}$, $\psi^* = \psi(s^*)$ – угол наклона касательной к контуру в точке N^* (см. рис. 3). Из (П.2) следует равенство $\text{tg}\varphi_0^* = \text{tg}(\varphi_1^* - 2\psi^*)$. Это записывается через координаты в виде

$$\begin{aligned} -T(x_0, y_0) &= \frac{T(x_1, y_1) - \text{tg}2\psi^*}{1 + T(x_1, y_1)\text{tg}2\psi^*}, \\ T(x_{0,1}, y_{0,1}) &= \frac{x_{0,1} - X^*}{y_{0,1} - Y^*}. \end{aligned} \tag{П.3}$$

Здесь $X^* = X(s^*)$ и $Y^* = Y(s^*)$ – декартовы координаты точки N^* .

Поскольку граница состоит из дуг окружностей (см. (3)), легко получим:

$$\begin{aligned} \psi^* &= -\alpha^* s^*, \quad X^* = s^* + O(\kappa^2 (s^*)^3), \\ Y^* &= -\alpha^* (s^*)^2 / 2 + (\kappa^2 (s^*)^3). \end{aligned} \tag{П.4}$$

После подстановки (П.4) в (П.3) прямое вычисление, использующее малость s^* , дает

$$\begin{aligned} \sin\delta\varphi + s^* \left(2\alpha^* + \frac{\sin\varphi_0}{r_0} + \frac{\sin\varphi_1}{r_1} \right) + \\ + O(\kappa(r_0 + r_1)(s^*)^2) = 0. \end{aligned} \tag{П.5}$$

$$u^{\text{отр}} = \frac{R(0)}{2\sqrt{2\pi}} \exp\left(ik(r_1 + r_0) + i\frac{\pi}{4} \right) (\mathcal{A}_- + \theta(-\delta\varphi)(\mathcal{A}_+ - \mathcal{A}_-)) (1 + O(k\rho(\delta\varphi)^3)), \tag{П.9}$$

см. (46). Формула (45) согласуется с (П.9) в области, где выполнены неравенства (47) и (48).

П.3. Малые расстояния

Обратимся теперь к случаю, когда точка наблюдения M_1 близка к O , $k\rho \ll 1$, но $k\rho \gg 1$

Учитывая, что $|\delta\varphi| \ll 1$ и пользуясь (44), в результате некоторых выкладок получаем из (П.5):

$$\begin{aligned} s^* &= -\frac{r_0 r_1 \delta\varphi}{\sin\varphi_0} \times \\ &\times \left(\frac{1}{\mathcal{F}_-} + \theta(-\delta\varphi) \left(\frac{1}{\mathcal{F}_+} - \frac{1}{\mathcal{F}_-} \right) \right) (1 + O(\delta\varphi)). \end{aligned} \tag{П.6}$$

Здесь использовано выражение для кривизны контура (3).

Теперь упростим фазу экспоненты в лучевой формуле (П.1). При N^* , близких к O , значение эйконала $\tau^* = l_0^* + l_1^*$ в точке M_1 записывается вполне аналогично (27):

$$\begin{aligned} \tau^* &= r_0 + r_1 + s^* \sin\varphi_0 \delta\varphi + \\ &+ \frac{(s^*)^2}{2r_0 r_1} \sin^2\varphi_0 (\mathcal{F}_- + \theta(s^*)(\mathcal{F}_+ - \mathcal{F}_-)) + \\ &+ O(s^*(\delta\varphi)^2) + O((s^*)^2 \delta\varphi/\rho) + O((s^*)^3/\rho^2). \end{aligned} \tag{П.7}$$

Подставляя сюда (П.6) и снова учитывая малость $\delta\varphi$, получим выражение для фазы экспоненты в (П.1):

$$\begin{aligned} ik\tau^* &= ik(r_1 + r_0) - ik\frac{r_0 r_1}{2} (\delta\varphi)^2 \times \\ &\times \left(\frac{1}{\mathcal{F}_-} + \theta(-\delta\varphi) \left(\frac{1}{\mathcal{F}_+} - \frac{1}{\mathcal{F}_-} \right) \right) + O(k\rho(\delta\varphi)^3). \end{aligned} \tag{П.8}$$

Поправочный член мал при условии (48).

Аналогичным образом упростим выражение в знаменателе формулы (П.1):

$$\mathcal{F}^* = (\mathcal{F}_- + \theta(-\delta\varphi)(\mathcal{F}_+ - \mathcal{F}_-)) (1 + O(\delta\varphi)).$$

Перенесем последнее слагаемое в (П.8) в амплитуду и получим упрощенный вид лучевой формулы при $k\rho \sim 1$ и $|\delta\varphi| \ll 1$:

(см. (29)). Преобразуем формулу (П.9). Очевидно,

$$\frac{r_0 r_1}{\mathcal{F}_\pm} = \frac{\rho}{1 + 2\alpha_\pm \rho / \sin\varphi_0} = \rho - \frac{2\alpha_\pm \rho^2}{\sin\varphi_0} + O(\kappa^2 \rho^3), \tag{П.10}$$

и выражение (П.9) переписывается в виде

$$u^{\text{отр}} \approx \frac{R(0) \exp\left(ik(r_1 + r_0 - \rho(\delta\varphi)^2/2) + i\pi/4\right)}{2\sqrt{2\pi} \sqrt{k(r_0 + r_1)}} \left(1 + (\alpha_- + [\alpha]\theta(-\delta\varphi)) \frac{\rho\left(ik\rho(\delta\varphi)^2 - 1\right)}{\sin\varphi_0} \right). \quad (\text{П.11})$$

Формула (П.11) получена в предположении, что выполнены неравенства (48) и (57)⁵. В главном (относительно параметра $k\rho \ll 1$) приближении отраженная волна представляет собой волну от мнимого источника, соответствующего прямолинейной границе. Следующие члены учитывают влияние кривизны.

Из (П.11) явствует, что слагаемые в (53), отвечающие U_0 и U_1 , в области своей пригодности соответствуют геометрически отраженной волне.

БЛАГОДАРНОСТИ

Авторы благодарны А.В. Попову за полезные обсуждения.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-01-00627).

СПИСОК ЛИТЕРАТУРЫ

1. *Weston V.H.* // IRE Trans. 1962. V. AP-10. № 6. P. 775.
2. *Weston V.H.* // IEEE Trans. 1965. V. AP-13. № 4. P. 611.
3. *Kaminetzky L., Keller J.B.* // SIAM J. Appl. Math. 1972. V. 22. № 1. P. 109.
4. *Senior T.B.A.* // IEEE Trans. 1972. V. AP-20. № 3. P. 323.
5. *Bouche D.* // Ann. Télécommunic. 1992. V. 47. № 9–10. P. 391.
6. *Rogoff Z.M., Kiselev A.P.* // Wave Motion. 2001. V. 33. № 2. P. 183.
7. *Bouche D.P., Aguilera P., Loillier S.* // Proc. Int. Conf. on Electromagnetics in Adv. App. (ICEAA), Verona, 11–15 Sep. 2017. N.Y.: IEEE, 2017. P. 531.
8. *Zlobina E.A., Kiselev A.P.* // Wave Motion. 2020. V. 96. P. 102671.
9. *Попов А.В.* // Тр. V Всес. симпози. по дифракции и распространению волн. Ленинград, 13–17 июня 1970. Л.: Наука, 1971. С. 171.
10. *Michaeli A.* // IEEE Trans. 1990. V. AP 38. № 6. P. 929.
11. *Кирпичникова Н.Я., Филиппов В.Б., Кирпичникова А.С.* // Зап. научн. сем. ПОМИ. 1999. Т. 267. С. 75.
12. *Kirpichnikova A.S., Philippov V.B.* // IEEE Trans. 2001. V. AP-49. № 12. P. 1618.
13. *James G.L.* Geometrical Theory of Diffraction for Electromagnetic Waves. L: IEEE Electromagnetic Ser. 1., 1986.
14. *Keller J.B.* // J. Opt. Soc. Amer. 1962. V. 52. № 5. P. 116.
15. *Боровиков В.А., Кинбер Б.Е.* Геометрическая теория дифракции. М.: Связь, 1978.
16. *Крюковский А.С., Лукин Д.С., Палкин Е.А., Растягаев Д.В.* // РЭ. 2006. Т. 51. № 10. С. 1155.
17. *Крюковский А.С.* Равномерная асимптотическая теория краевых и угловых волновых катастроф. М.: РосНОУ, 2013.
18. *Бабич В.М., Кирпичникова Н.Я.* Метод пограничного слоя в задачах дифракции. Л.: Изд-во ЛГУ, 1974.
19. *Бабич В.М., Булдырев В.С.* Асимптотические методы в задачах дифракции коротких волн. М.: Наука, 1972.
20. *Злобина Е.А., Киселев А.П.* // Алгебра и анализ. 2021. Т. 33. № 2. С. 35.
21. *Злобина Е.А.* // Зап. научн. сем. ПОМИ. 2020. Т. 493. С. 169.
22. *Бреховских Л.М., Гордин О.А.* Акустика слоистых сред. М.: Наука, 1989.
23. *Абрамовиц М., Стиган И.* Справочник по специальным функциям. М.: Наука, 1979.
24. *Уфимцев П.Я.* Теория дифракционных краевых волн в электродинамике. М.: Бином, 2012.
25. *Попов А.В.* // Акуст. журн. 1973. Т. 19. № 4. С. 594.
26. *Эрдейи А.* Асимптотические разложения. М.: Физматлит, 1962.
27. *Злобина Е.А., Киселев А.П.* // Зап. научн. сем. ПОМИ. 2018. Т. 471. С. 113.

⁵ Аналогичные (П.11) выражения были получены в случае отражения плоской волны от границы с негладкой кривизной (как разрывной [8], так и непрерывной [20, 21]). Область их применимости описывалась неравенствами (57), а выполнения (48) не требовалось.

АНТЕННО-ФИДЕРНЫЕ
СИСТЕМЫ

УДК 621.396.67

СИНТЕЗ ДВУХЗЕРКАЛЬНОЙ БИФОКАЛЬНОЙ ЦИЛИНДРИЧЕСКОЙ
СИСТЕМЫ С МИНИМАЛЬНЫМИ АБЕРРАЦИЯМИ© 2022 г. В. А. Калошин^{а, *}, Ви Ут Нам^б^а Институт радиотехники и электроники им. В.А. Котельникова РАН,
ул. Моховая, 11, стр. 7, Москва, 125007 Российская Федерация^б Московский физико-технический институт (национальный исследовательский университет),
Институтский пер., 9, Долгопрудный Московской обл., 141700 Российская Федерация

*E-mail: vak@cplire.ru

Поступила в редакцию 12.01.2021 г.

После доработки 10.03.2021 г.

Принята к публикации 17.03.2021 г.

Развита методика синтеза цилиндрических двухзеркальных бифокальных систем с использованием последовательного нахождения участков зеркал и заданием начального участка вспомогательного зеркала в виде полинома второй и четвертой степени. Один из коэффициентов полиномов определен в результате решения найденного в работе уравнения, которое в общем случае обеспечивает непрерывность вторых производных функций, описывающих поверхности зеркал. Начальный участок главного зеркала найден в результате решения задачи синтеза плоского фронта для центрального положения источника. Определены параметры оптимизации с целью минимизации величины средне-квадратической aberrации при фиксированном расстоянии между зеркалами и угле зрения бифокальной системы. На плоскости этих параметров для углов зрения 50, 70 и 105 град найдены границы области существования решения задачи синтеза и приведены зависимости величины средне-квадратической aberrации от параметров, при этом показано, что ее минимум достигается на границе области существования решения и найден набор параметров, обеспечивающих этот минимум.

DOI: 10.31857/S003384942202005X

ВВЕДЕНИЕ

Бифокальные двухзеркальные системы позволяют расширить угол зрения по сравнению с однозеркальными и апланатическими двухзеркальными системами. В связи с этим синтезу бифокальных систем посвящено большое количество работ [1–14].

В работе [14] на основе известного подхода [1–6] развита методика точного решения задачи синтеза цилиндрических бифокальных двухзеркальных систем в приближении геометрической оптики. В рамках развитой методики осуществлен выбор начальных участков зеркал и обеспечена непрерывность функций, описывающих форму зеркал, и их производных. Непрерывность вторых производных этих функций, которая обеспечивает непрерывность отраженных полей в первом приближении геометрической оптики, была реализована приближенно. Форма начальных участков зеркал при этом задавалась в виде отрезков парабол, один из коэффициентов которых был найден путем численной минимизации скачка второй производной на границах соседних участков зеркал. При этом вопрос о том, обеспечивают ли найденные решения задачи синтеза двухзеркальной системы для заданно-

го угла зрения минимальную среднеквадратическую aberrацию (СКА) остался открытым.

Цель данной работы – дальнейшее развитие методики синтеза и оптимизации цилиндрических бифокальных двухзеркальных систем с целью реализации минимальной СКА в заданном угле зрения. При этом в процессе синтеза задается начальный участок только одного из зеркал.

1. МЕТОДИКА СИНТЕЗА
БИФОКАЛЬНОЙ ДВУХЗЕРКАЛЬНОЙ
СИСТЕМЫ

Рассмотрим задачу синтеза цилиндрической бифокальной двухзеркальной системы, с одной стороны которой расположены два симметричных относительно оси y (рис. 1) фокуса (точки идеальной фокусировки F_1 и F_2) с декартовыми координатами (x_{F_1}, y_{F_1}) и (x_{F_2}, y_{F_2}) . При положении источника цилиндрической волны в каждом из этих фокусов с другой стороны бифокальной системы формируются два симметричных относительно оси y плоских фронта.

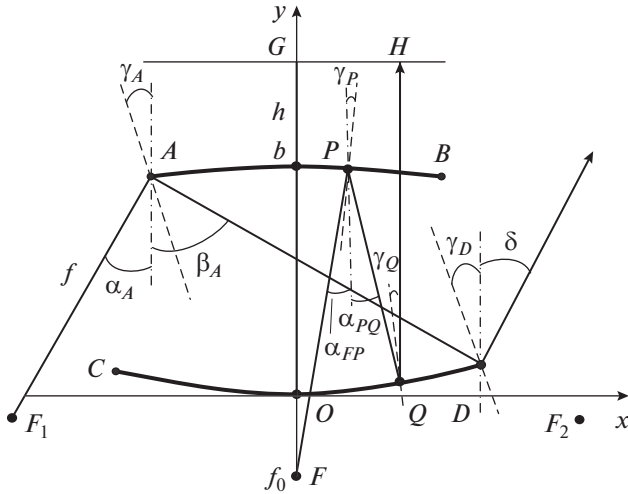


Рис. 1. Начальные участки зеркал.

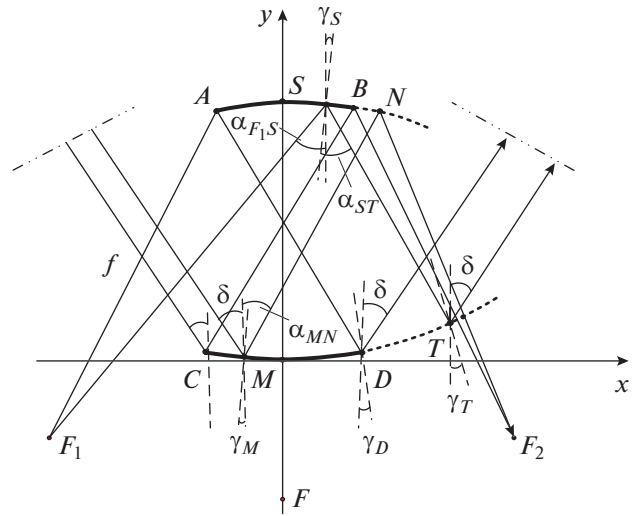


Рис. 2. К определению новых участков зеркал.

Пусть форма вспомогательного (первого) и основного (второго) зеркала описываются неизвестными четными функциями $y_1(x)$ и $y_2(x)$ соответственно, а функция $y_{10}(x)$, описывающая начальный участок вспомогательного зеркала, задана, т.е. на интервале $[-x_0, x_0]$ функция $y_1(x) = y_{10}(x)$ известна. При этом поверхности первого и второго зеркал пересекают ось y в точках $(0, b)$ и $(0, 0)$ соответственно (рис. 1). Потребуем, чтобы лучи из источника, расположенного в точке F_0 с координатами $(0, -f_0)$ после двух отражений от зеркал формировали плоский фронт $y = h$, где h – произвольная постоянная. При этом луч, идущий вдоль оси y , падает на первое зеркало в точке $(0, b)$, отражается от него, падает на второе зеркало в точке $(0, 0)$, снова отражается и снова идет вдоль оси y . Эйконал от источника до фронта этого (осевого) луча имеет вид

$$L_0 = f_0 + 2b + h. \tag{1}$$

Пусть другой луч, выходящий из точки F_0 , падает на первое зеркало в точке P с координатами (x_p, y_p) , отражается от него, падает на второе зеркало в точке Q с координатами (x_q, y_q) и отражается от него параллельно оси y . При этом его эйконал определяется формулой

$$L = \sqrt{x_p^2 + (y_p + f_0)^2} + PQ + h - y_p - PQ \cos(\alpha_{PQ}), \tag{2}$$

где $\alpha_{PQ} = \alpha_{FP} - 2\gamma_p$ – угол между осью y и отрезком PQ ; $\alpha_{FP} = \arctg(x_p / (y_p - y_F))$ – угол между осью y и падающим лучом в точке P ; $\gamma_p = \arctg(y'_0(x_p))$ –

угол между осью y и нормалью первому зеркалу в точке P ; PQ – расстояние от точки P до точки Q .

Потребуем, чтобы все лучи, выходящие из точки F после двух отражений от зеркал, были параллельно оси y и формировали плоский фронт на выходе системы. Для этого необходимо равенство эйконалов всех лучей (от источника до фронта). Потребуем, чтобы эйконалы всех лучей были равны эйконалу центрального луча:

$$\sqrt{x_p^2 + (y_p + f_0)^2} + PQ - y_p - PQ \cos(\alpha_{PQ}) = f_0 + 2b. \tag{3}$$

Решение этого уравнения имеет вид

$$PQ = \frac{f_0 + 2b + y_p - \sqrt{x_p^2 + (y_p + f_0)^2}}{1 - \cos(\alpha_{PQ})}. \tag{4}$$

Зная расстояние от точки P до точки Q и угол α_{PQ} , нетрудно найти координаты точки Q :

$$\begin{aligned} x_Q &= x_p + PQ \sin(\alpha_{PQ}), \\ y_Q &= y_p + PQ \cos(\alpha_{PQ}). \end{aligned} \tag{5}$$

Множество точек Q образует начальный участок второго зеркала.

Для реализации на стыках начальных участков с соседними непрерывности функций, описывающих форму поверхности зеркал и их производных, необходимо, чтобы луч плоской волны, падающей на зеркало под углом к оси y после отражения в точке D попадал в точку A , а после отражения в точке A – в фокус F_1 . Из геометрии, представленной на рис. 2, нетрудно найти координаты этого фокуса, а также фокуса F_2 , учиты-

вая, что он симметричен фокусу F_1 относительно оси y :

$$\begin{aligned} x_{F_1} &= x_A - f \sin(\alpha_{AF_1}), \\ y_{F_1} &= y_A - f \cos(\alpha_{AF_1}), \\ x_{F_2} &= -x_{F_1}; \quad y_{F_2} = y_{F_1}, \end{aligned} \quad (6)$$

где $x_A = -x_0$; $y_A = y(-x_0)$; $\alpha_{AF_1} = \arctg((x_D - x_A)/(y_D - y_A)) + 2 \arctg(y'(x_A))$ – угол между осью y и линией, соединяющей фокус F_1 с краем (точка A) начального участка вспомогательного зеркала; f – расстояние от края начального участка до фокуса.

Рассмотрим луч, который из фокуса F_1 падает на первое зеркало в точке A , отражается и падает на второе зеркало в точке D . Из геометрии рис. 1 следует, что угол выхода луча из системы определяется формулой

$$\delta = \alpha_{AD} - 2 \arctg(y_2'(x_D)), \quad (7)$$

где $y_2'(x_D)$ – первая производная функции $y_2(x)$ в точке D .

Для определения нового участка второго зеркала предложим, что луч из фокуса F_1 падает на на-

чальный участок первого зеркала, отражается от него в точке S с координатами (x_S, y_S) , падает на второе зеркало в точке T с координатами (x_T, y_T) и отражается под углом δ (см. рис. 2). Отсюда получаем угол между осью y и падающим от точки F_1 в точку S лучом $\alpha_{F_1S} = \arctg((x_S - x_{F_1})/(y_S - y_{F_1}))$ и угол $\alpha_{ST} = \alpha_{F_1S} - 2\gamma_S$ между осью y и лучом, отраженным от зеркала в точке S .

Для того чтобы двухзеркальная система формировала на выходе плоский фронт, необходимо равенство эйконолов всех лучей, которые выйдут из фокуса F_1 и после отражения от зеркал идут параллельно (под углом δ к оси Y). Отсюда получаем уравнение

$$\begin{aligned} ST + (x_D - x_S - ST \sin(\alpha_{ST})) \sin(\delta) + \\ + (y_D - y_S + ST \cos(\alpha_{ST})) \cos(\delta) = \\ = f_0 + l_0 - \sqrt{(x_{F_1} - x_S)^2 + (y_{F_1} - y_S)^2}, \end{aligned} \quad (8)$$

где $l_0 = \sqrt{(x_C - x_B)^2 + (y_C - y_B)^2}$; ST – расстояние от точки S до точки T .

Решение этого уравнения имеет вид

$$ST = \frac{f_0 + l_0 - \sqrt{(x_S - x_{F_1})^2 + (y_S - y_{F_1})^2} + (x_S - x_D) \sin(\delta) + (y_S - y_D) \cos(\delta)}{(1 - \sin(\alpha_{ST}) \sin(\delta) + \cos(\alpha_{ST}) \cos(\delta))}. \quad (9)$$

Зная длину ST и угол α_{ST} , можно определять координаты точки T по формулам

$$x_T = x_S + ST \sin(\alpha_{ST}), \quad y_T = y_S - ST \cos(\alpha_{ST}). \quad (10)$$

Множество точек T образует новый участок второго зеркала. При этом функция $y_2(x)$ и ее первая производная также непрерывны на стыке начального участка второго зеркала с его новым (соседним) участком.

Для определения нового участка первого зеркала рассмотрим падение плоской волны на второе зер-

кало. Пусть луч, который падает на начальный участок второго зеркала в точке M с координатами (x_M, y_M) под углом δ к оси y , отражается от второго зеркала и падает на первое зеркало в точке N с координатами (x_N, y_N) , снова отражается и проходит через фокус F_2 (см. рис. 2). Угол между осью y и отраженным лучом в точке M равен $\alpha_{MN} = \delta - 2\gamma_M$.

Приравнивая эйконы лучей, отраженных от разных точек M зеркала, получим уравнение

$$MN + \sqrt{(x_M + MN \sin(\alpha_{MN}) - x_{F_2})^2 + (y_N + MN \cos(\alpha_{MN}) - y_{F_2})^2} - d_C - l_0 - f + d_M = 0, \quad (11)$$

где d_M – расстояние от точки M до фронта волны; d_C – расстояние от точки C до фронта волны.

Решение этого уравнения имеет вид

$$\begin{aligned} MN = \\ = \frac{A^2 - (y_M - y_{F_2})^2 - (x_M - x_{F_2})^2}{2A + 2(x_M - x_{F_2}) \sin(\alpha_{MN}) + 2(y_M - y_{F_2}) \cos(\alpha_{MN})}, \end{aligned} \quad (12)$$

где $A = l_0 + f + 2(x_C - x_M) \sin(\delta) + 2(y_C - y_M) \cos(\delta)$.

Зная расстояние MN и угол α_{MN} , координаты точки N можно найти по формулам

$$\begin{aligned} x_N &= x_M + MN \sin(\alpha_{MN}), \\ y_N &= y_M + MN \cos(\alpha_{MN}). \end{aligned} \quad (13)$$

Множество точек N образует новый участок первого зеркала. При этом функция $y_1(x)$ и ее первая производная непрерывны на стыке начального отрезка с новым.

Для обеспечения непрерывности амплитудного распределения отраженных волн в первом приближении геометрической оптики необходимо, чтобы вторые производные функций, описывающих поверхности зеркал, были непрерывными. Первая производная функции $y_1(x)$ в точке N имеет вид

$$y'_N = \text{tg}(-\gamma_N) = \text{tg}\left(\frac{\alpha_{MN} - \alpha_{F2}}{2}\right). \quad (14)$$

Так как координаты точки N определяются через координаты точки M , вторую производную второго отрезка первого зеркала в точке N определяем дифференцированием его первой производной в точке N по координате x_M :

$$y''_N(x_M) = \frac{d(y'_N)}{d(x_N)} = \left(\frac{d(y'_N)}{d(x_M)}\right) / \left(\frac{d(x_N)}{d(x_M)}\right) = \frac{(\alpha'_{MN}(x_M) - \alpha'_{F2}(x_M))(1 + y'^2_N)}{2x'_N(x_M)}. \quad (15)$$

Для того чтобы вторая производная первого зеркала на стыках была непрерывной, значение второй производной первого участка в точке B должно равняться значению второй производной в точке N , когда точки M и C совпадают. Заменим x_M на $-x_0$ в выражении для второй производной (15) и приравняем его значению второй производной начального отрезка первого зеркала в точке B . В результате получим уравнение

$$\frac{(y_0'^2 + 2T_0y_0' - 1)x'_N - ((1 - y_0'^2)T_0 - 2y_0')y'_N}{f(y_0'^2 + 1)\sqrt{T_0^2 + 1}} + \alpha'_0 + \frac{2y_0''x'_N}{1 + y_0'^2} = 0, \quad (16)$$

где

$$T_0 = \frac{S_b(f_0 + 2b + y_0 - \sqrt{x_0^2 + (y_0 + f_0)^2}) + 2x_0(1 + C_b)}{(1 + C_b)C_b l_x}; \quad S_b = \frac{(y_0'(y_0 + f_0) + x_0)}{(1 + y_0)\sqrt{(y_0 + f_0)^2 + x_0^2}};$$

$$C_b = \frac{((y_0 + f_0) + y_0'x_0)}{(1 + y_0)\sqrt{(y_0 + f_0)^2 + x_0^2}}; \quad \beta'_0 = \frac{y_0 + f_0 - x_0y_0'}{x_0^2 + (y_0 + f_0)^2} - \frac{2y_0''}{1 + y_0'^2};$$

$$l'_x = \frac{\left(-y_0' + \frac{x_0 + (y_0 + f_0)y_0'}{\sqrt{x_0^2 + (y_0 + f_0)^2}}\right)(1 + C_b) + S_b\beta'_0(\sqrt{x_0^2 + (y_0 + f_0)^2} - f_0 - 2b_1 - y_0)}{(1 + C_b^2)},$$

$$\beta_0 = \text{arctg}\left(\frac{x_0}{y_0 + f}\right) - \text{arctg}(y_0'); \quad l_x = \frac{f_0 + 2b_1 + 2b_2 + y_0 - \sqrt{x_0^2 + y_0^2}}{1 + C_b};$$

$$\alpha'_0 = \frac{2\beta'_0}{(1 + l_x C_b \beta'_0 + l'_x S_b)}; \quad S_T = \frac{T_0(1 - y_0'^2) + 2y_0'}{(1 + y_0'^2)\sqrt{1 + T_0^2}};$$

$$C_T = \frac{1 - y_0'^2 - 2T_0y_0'^2}{(1 + y_0'^2)\sqrt{1 + T_0^2}}; \quad A = \sqrt{(x_0 + l_x S_b)^2 + l_x^2 C_b^2} + f;$$

$$A_1 = \frac{S_b - T_0(1 + C_b)}{(1 + C_b)\sqrt{1 + T_0^2}}; \quad B = A^2 + (l_x C_b - f C_T)^2 - (l_x S_b - f S_T)^2;$$

$$B_1 = 2AA_1 + \frac{2(l_x C_b - f C_T)^2 S_b}{1 + C_b} - (l_x S_b - f S_T);$$

$$C_1 = 2A_1 + \frac{2\alpha'(l_x C_b - f C_T - (l_x S_b - f S_T)T_0)}{\sqrt{1 + T_0^2}} - \frac{2S_b}{(1 + C_b)\sqrt{1 + T_0^2}};$$

$$C = 2A + \frac{2(l_x C_b - f C_T)T_0 + 2(l_x S_b - f S_T)}{\sqrt{1 + T_0^2}}; \quad l_0 = \frac{B}{C};$$

$$l'_0 = \frac{B_1 C - B C_1}{C^2}; \quad x'_N = 1 + \frac{l'_0 T_0}{\sqrt{1 + T_0^2}} + \frac{l_0 \alpha'_0}{\sqrt{1 + T_0^2}}; \quad y'_N = \frac{S_b}{1 + C_b} + \frac{l'_0}{\sqrt{1 + T_0^2}} + \frac{l_0 \alpha'_0 T_0}{\sqrt{1 + T_0^2}},$$

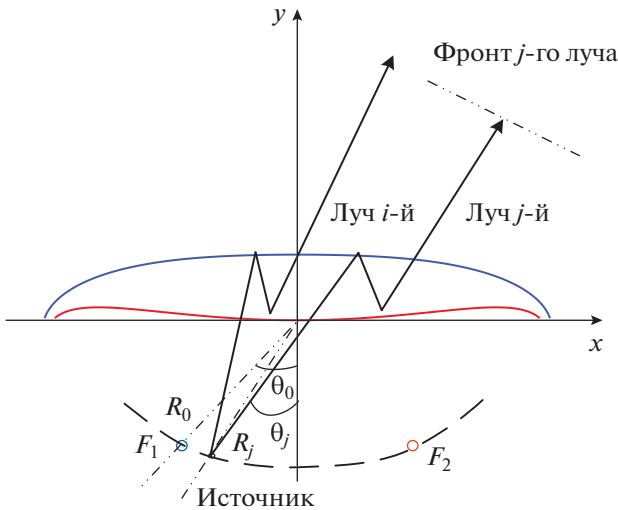


Рис. 3. Геометрия лучей в двухзеркальной системе.

$y_0 = y_1(x_0)$, $y_0' = y_1'(x_0)$; $y_0'' = y_1''(x_0)$ — значения функции $y_1(x)$, ее первой и второй производной соответственно в точке B с координатами (x_0, y_0) .

2. СИНТЕЗ И ОПТИМИЗАЦИЯ БИФОКАЛЬНЫХ ДВУХЗЕРКАЛЬНЫХ СИСТЕМ

Задача синтеза и оптимизации состоит в нахождении формы зеркал и фокальной кривой, обеспечивающие минимальную величину СКА эйконала на выходе двухзеркальной системы, которую будем определять по формуле

$$\sigma = \frac{1}{D} \sqrt{\frac{1}{n} \sum_{i=1}^n (L_i - L_0)^2}, \quad (17)$$

где L_i — эйконал луча с номером i ; n — количество учтенных лучей, D — размер апертуры системы; L_0 — эйконал луча, относительно которого СКА имеет минимальное значение (этот луч будем называть опорным).

Зададим исходные параметры системы: расстояние между зеркалами b , полуразмер начального отрезка x_0 , расстояние f от конца начального отрезка до фокуса F_1 , расстояние p от первого зеркала до фокуса F_0 начальной системы, и используем изложенный выше алгоритм синтеза бифокальной системы с начальным отрезком первого зеркала в виде полинома второго порядка $y(x) = ax^2 + b$ и четвертого $y(x) = a_4x^4 + a_2x^2 + b$. Один из коэффициентов полинома находим из условия непрерывности второй производной функций, описывающих форму зеркал (16). В первом случае, подставляя $y_0 = ax_0^2 + b$; $y_0' = 2ax_0$; $y_0'' = 2a$ в уравнение (16),

находим a . Во втором случае аналогично находим a_4 при заданном значении a_2 , которое в данном случае является дополнительной степенью свободы для оптимизации. Реализуя описанный выше алгоритм синтеза отрезков зеркал m раз, находим форму зеркал, которые состоят из $2m + 1$ отрезков.

Проведем анализ СКА двухзеркальной системы, синтезированной для фиксированного значения угла зрения в зависимости от ее параметров. Для вычисления СКА системы по формуле (17) необходимо найти эйконалы лучей, для определения которых, в свою очередь, необходимо знать направление фронта и величину эйконала L_0 опорного луча, относительно которого будет рассчитываться СКА. Если источник находится в фокусе, углы выхода всех лучей из системы одинаковы и определяются формулой (7). При смещенном положении источника углы выхода лучей будут разные. Найдем k таких лучей и выберем из них несколько опорных, проходящих вблизи центра двухзеркальной системы. По формуле (17) найдем СКА для каждого опорного луча с соответствующим (ортогональным) фронтом (рис. 3) и выберем из полученных величин СКА минимальное значение. Полученное приближенное значение СКА уточняем, меняя угол выхода опорного луча и находя минимум СКА.

Для определения фокальной кривой найдем геометрическое место положений источника (см. рис. 3), которые обеспечивают наименьшую величину СКА. Декартовы и полярные координаты источника связаны формулами $x_{Fj} = -R_j \sin(\theta_j)$, $y_{Fj} = -R_j \cos(\theta_j)$. Задача состоит в том, чтобы найти оптимальную функцию $R_j(\theta_j)$. Для 20 значений угла θ_j находим R_j с использованием стандартной численной процедуры нахождения минимума. Применяя сплайн-интерполяцию, находим функцию $R_j(\theta_j)$ и таким образом получаем фокальную кривую.

Проведем исследование зависимости величины СКА от параметров бифокальной системы. На рис. 4 представлены зависимости СКА от угла зрения системы с разным размером начального участка $2x_0$ в виде параболы, разными расстояниями между зеркалами b_1 и разным числом синтезированных отрезков m . Видно, что СКА медленно уменьшается при увеличении x_0 . Это объясняется тем, что только система из начальных участков имеет фокус в точке $F(0, -f_0)$ на оси x . Новые синтезированные участки уже не обеспечивают точную фокусировку при положении источника в этой точке, их число при заданной апертуре системы увеличивается при уменьшении x_0 и это, соответственно, приводит к увеличению СКА.

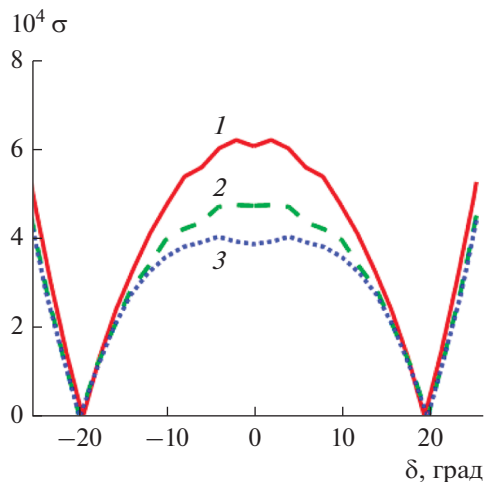


Рис. 4. Зависимость СКА бифокальной двухзеркальной системы с начальным участком первого зеркала в виде параболы в зависимости от угла зрения при $f = 0.6, p = 0.7$ и разных значениях x_0 и b_1 : кривая 1 – $x_0 = 0.026, b_1 = 0.146$; кривая 2 – $x_0 = 0.0455, b_1 = 0.2495$, кривая 3 – $x_0 = 0.065, b_1 = 0.355$.

При увеличении числа синтезированных отрезков зеркал их края приближаются друг к другу. В результате зеркала либо пересекаются (рис. 5а), либо у них появляются точки возврата (рис. 5б). Параметры системы, при которых у зеркала при построении первого нового участка возникают точки возврата или решение задачи синтеза перестает существовать, будем называть критическими.

Далее исследуем СКА бифокальной системы в зависимости от параметров f и $p = f_0 + b$ для трех наборов фиксированных параметров: расстояния b между зеркалами и x_0 (отношение b/x_0 определяет угол зрения). Линии уровня СКА для трех наборов

фиксированных параметров, соответствующих значениям угла зрения 50, 70 и 105 град в зависимости от параметров f и p , показаны на рис. 6а–6в. На рисунках видны границы областей существования решения, а также уменьшение величины СКА по мере приближения параметров системы к этим границам. При этом линии уровней СКА идут почти параллельно границе.

На рис. 7 представлены зависимости СКА от параметра p системы при движении вдоль границы (при критической величине f) для тех же трех наборов фиксированных параметров. Из рисунка видно, что зависимость СКА от p имеет колебательный характер, при этом для больших углов зрения средняя величина СКА при уменьшении p уменьшается. Минимальные величины СКА для исследованного интервала ($p < 2.5$) и углов зрения 50, 70 и 105 град равны $8.0 \times 10^{-6}, 2.2 \times 10^{-5}$ и 4.1×10^{-5} соответственно.

На рис. 8 приведены зависимости СКА бифокальной двухзеркальной системы от угла зрения для трех оптимальных наборов параметров, которые соответствуют трем углам зрения 50, 70 и 105 град. Как видно из рисунка, при увеличении угла зрения примерно в два раза СКА увеличивается в пять раз. При этом полученные в результате минимальные величины СКА в 60 раз меньше СКА двухзеркальных бифокальных систем, синтезированных в работе [13], и близки к СКА трехфокальных систем [14, 15].

Величина апертуры D для разных наборов параметров получается разной. Для анализа полученных результатов удобно считать все величины относительно апертуры системы. Для этого достаточно умножить все геометрические размеры на множитель, равный обратной величине D . В результате для систем с $D = 1$ и углами зрения 50, 70

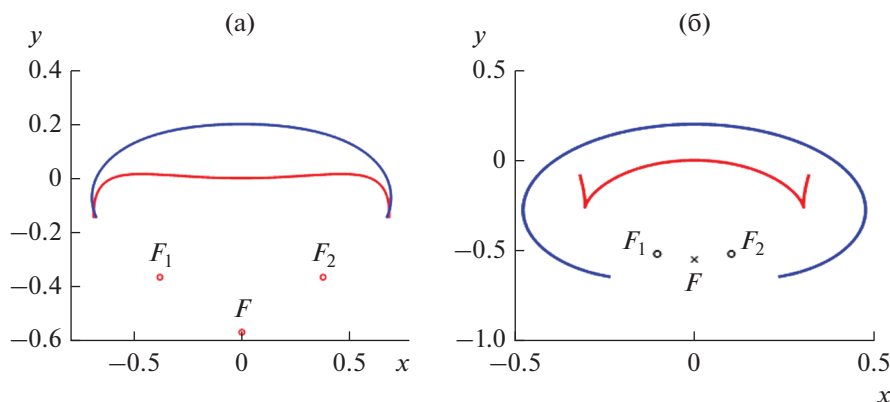


Рис. 5. Геометрия бифокальной двухзеркальной системы с параметрами $n = 1.6, b_1 = 0.2$, а) $x_0 = 0.006, f = 0.65, p = 0.77$; б) $x_0 = 0.035, f = 0.72, p = 0.7479$.

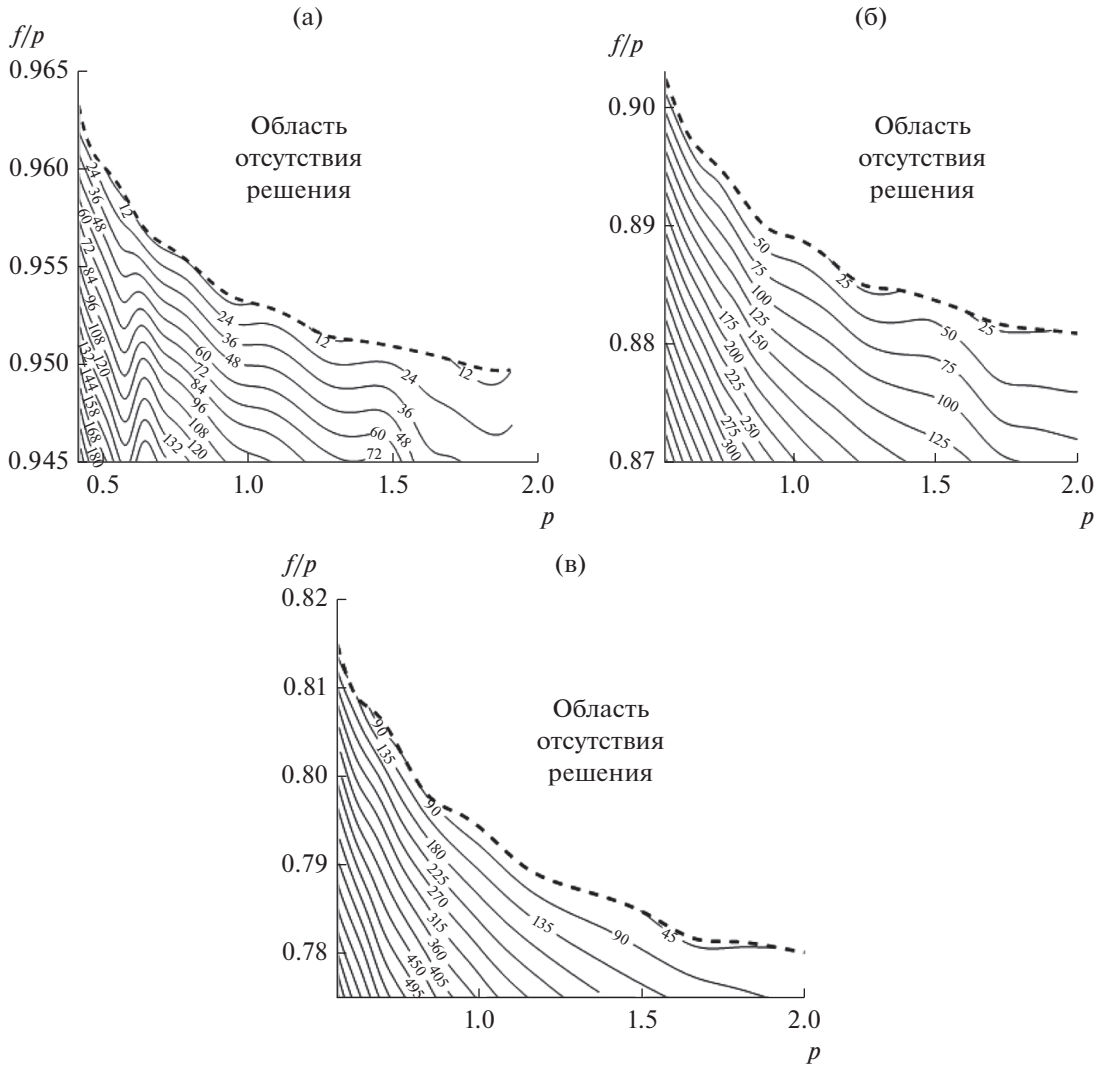


Рис. 6. Линии уровня $10^6\sigma$ бифокальной зеркальной системы в зависимости от параметров p и f/p при $b = 0.1$ и разных углах зрения: а – угол зрения 50 град ($x_0 = 0.0175$, б – угол зрения 70 град ($x_0 = 0.028$), в – угол зрения 105° ($x_0 = 0.042$).

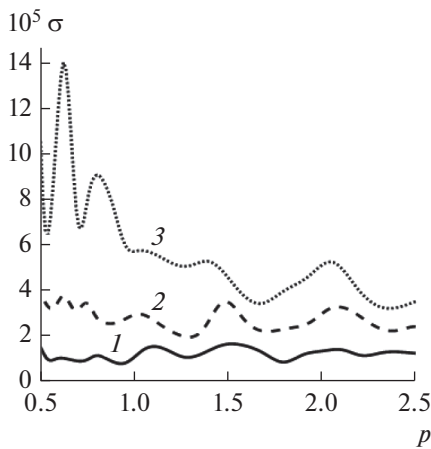


Рис. 7. Зависимости СКА бифокальной зеркальной системы для критических значений f от параметра p при $b = 0.1$: кривая 1 – угол зрения 50 град ($x_0 = 0.0175$), кривая 2 – угол зрения 70 град ($x_0 = 0.028$), кривая 3 – угол зрения 105 град ($x_0 = 0.042$).

и 105 град величина p равна 1.30, 1.14 и 1.41, а расстояние между зеркалами b равно соответственно 0.073, 0.089 и 0.061.

На рис. 9 показаны фокальные кривые синтезированных систем с оптимальными параметрами для углов зрения 50, 70 и 105 град. Видно, что с увеличением угла зрения в два раза угловой размер фокальной кривой увеличивается также в два раза, при этом фокальный размер меняется немонотонно, а различие между минимальным и максимальным размерами составляет 23%.

Дальнейшее исследование показало, что использование дополнительной степени свободы при задании формы первого участка вспомогательного зеркала в виде полинома четвертого порядка не приводит к дополнительному уменьшению СКА.

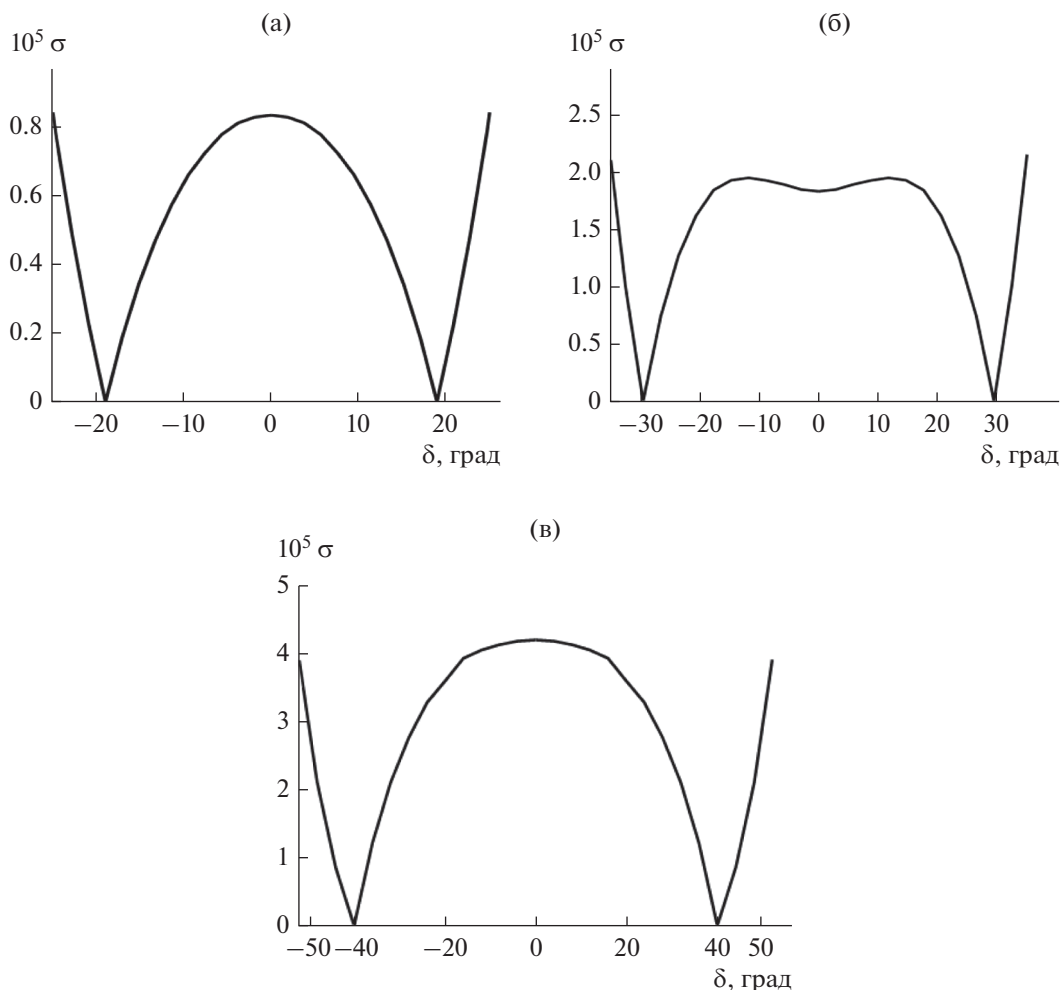


Рис. 8. Зависимость СКА оптимальной бифокальной двухзеркальной системы от угла зрения: а – $x_0 = 0.0175, f = 1.695, p = 1.784$; б – $x_0 = 0.028, f = 1.139, p = 1.287$; в – $x_0 = 0.042, f = 1.793, p = 2.306$.

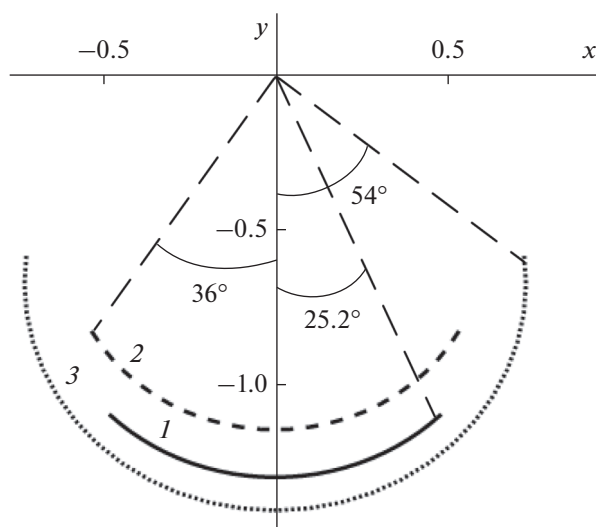


Рис. 9. Фокальные кривые синтезированных двухзеркальных систем: кривая 1 – угол зрения 50 град ($x_0 = 0.0175$), кривая 2 – угол зрения 70 град ($x_0 = 0.028$), кривая 3 – угол зрения 105 град ($x_0 = 0.042$).

ЗАКЛЮЧЕНИЕ

На основе полученных результатов можно сделать следующие выводы.

1. Разработанная методика позволяет синтезировать и оптимизировать двухзеркальные бифокальные цилиндрические системы по минимуму СКА.
2. Синтезированные и оптимизированные в результате разработанной методики бифокальные двухзеркальные системы имеют СКА в десятки раз меньше, чем известные.
3. Увеличение угла зрения в два раза приводит к увеличению СКА в пять раз, при этом различие между минимальным и максимальным продольным размером составляет 23%.
4. Задание формы начального участка вспомогательного зеркала в виде полинома четвертого порядка вместо полинома второго порядка и использование дополнительной степени свободы

не приводит к дополнительному уменьшению СКА.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена за счет бюджетного финансирования в рамках государственного задания по теме 0030-2019-006.

СПИСОК ЛИТЕРАТУРЫ

1. Кинбер Б.Е., Классен В.И., Стеблин В.И. // Волны и дифракция. М., 1981. С. 101.
2. Кинбер Б.Е., Классен В.И., Стеблин В.И. // РЭ. 1983. Т. 23. № 8. С. 1509.
3. Rappaport C.M. // IEEE Trans. 1984. V. AP-32. № 11. С. 1196.
4. Вааз И.Л., Кинбер Б.Е. // РЭ. 1986. Т. 31. № 8. С. 1507.
5. Классен В.И., Кинбер Б.Е., Шишилов А.В., Тоболев А.К. // Антенны. 1987. № 34. С. 324.
6. Бодулинский В.К. // Компьютерная оптика. 1987. № 1. С. 79.
7. Craig W.P., Rappaport C.M., Jeffrey S.M. // IEEE Trans. 1993. V. AP-41. № 11. P. 1481.
8. Shishlov A.V., Shitikov A.M. // Proc. 27 Sci. Conf. on Antenna Theory and Technology. Moscow. 1994. P. 227.
9. Pino A.G., Rappaport C.M., Rubinos J.O., Lorenzo M.E. // IEEE Trans. 1995. V. AP-43. № 10. P. 1022.
10. Lorenzo M.E., Rappaport C.M., Pino A.G. // Proc. IEEE APS Intern. Symp. 2001. V. 2. P. 284.
11. Pino A.G., Lombart N., Gonzalez V.B., Rubino L.O. // IEEE Trans. 2012. V. AP-60. № 9. P. 4119.
12. Plastikov A.N. // IEEE Trans. 2016. V. AP-64. № 7. P. 3251.
13. Калошин В.А., Ле Д.Т. // Журн. радиоэлектроники. 2018. № 9. <http://jre.cplire.ru/jre/sep18/13/text.pdf>.
14. Калошин В.А., Нгием Х.Д., Фролова Е.В. // Журн. радиоэлектроники. 2018. № 1. <http://jre.cplire.ru/jre/jan18/3/text.pdf>.
15. Калошин В.А., Ле Д.Т. // Журн. радиоэлектроники. 2020. № 4. <http://jre.cplire.ru/jre/apr20/4/text.pdf>.

КОМПЕНСАЦИЯ ОТРАЖЕНИЙ ОТ МЕСТНЫХ ПРЕДМЕТОВ
В ИМПУЛЬСНО-ДОПЛЕРОВСКИХ МЕТЕОРАДИОЛОКАТОРАХ© 2022 г. Б. М. Вовшин^{а, *}, В. К. Битюков^а^а МИРЭА – Российский технологический университет,
просп. Вернадского, 78, Москва, 119454 Российская Федерация

*E-mail: boris@eleron.net

Поступила в редакцию 11.09.2021 г.

После доработки 11.09.2021 г.

Принята к публикации 18.09.2021 г.

Рассмотрена актуальная для доплеровских метеорадиолокаторов задача компенсации отражений от местных предметов, искажающих оценки спектральных характеристик метеообразований. Предложена методика статистического синтеза режекторных фильтров с требуемой зоной режекции помех на первом этапе межпериодной обработки сигналов. На примерах фильтров помех с близкой к нулю и расширенной зонами режекции доказана корректность и эффективность предложенной методики синтеза. Для различных ситуаций оценены скоростные характеристики синтезированных фильтров и определены способы управления “провалами” частотных характеристик в зонах воздействия помех.

DOI: 10.31857/S0033849422020140

ВВЕДЕНИЕ

Оценка спектральных параметров метеообразований (МО) является важнейшей задачей, возлагаемой на доплеровские метеорадиолокаторы (ДМРЛ) [1, 2]. Она решается на этапе межпериодной обработки (МПО) эхо-сигналов, отраженных от МО. При наличии мешающих отражений от местных предметов (МП) в виде отдельно стоящих объектов (здания, вышки, трубы и т.д.) или неровностей местности (горы, холмы, земля, леса) результаты оценки спектров МО могут оказаться существенно искаженными, в особенности на малых углах места.

Таким образом, на систему МПО ДМРЛ накладывается дополнительная задача – скомпенсировать или ослабить влияние отражений от МП на результаты доплеровской (спектральной) обработки и тем самым повысить качество информации, выдаваемой ДМРЛ. Одним из возможных вариантов ее решения является применение режекторных фильтров, использующих отличия в частотных свойствах МО и МП, обусловленных различием их радиальных скоростей V_r и характером межпериодных флюктуаций, определяющим ширину доплеровского спектра ΔF_d .

Цель данной работы – статистический синтез и исследования режекторных фильтров с различной глубиной и шириной зон режекции для эффективной компенсации отражений от МП при условии минимальных искажений спектров МО.

1. ОЦЕНИВАНИЕ ПАРАМЕТРОВ МО
В ПРИСУТСТВИИ ПОМЕХ

Известно [3, 4], что оптимальной оценкой вектора параметров $\hat{\Theta} = \{\hat{\Theta}_l\}_{l=1}^L$ является оценка, максимизирующая отношение правдоподобия l или функционал $f(l)$ от него:

$$l(\mathbf{u}, \Theta) = \frac{p_{\text{сп}}(\mathbf{u}, \Theta)}{p_{\text{п}}(\mathbf{u})}; \quad \xi(\mathbf{u}, \Theta) = f[l(\mathbf{u}, \Theta)], \quad (1)$$

где $p_{\text{сп}}(\mathbf{u}, \Theta)$ и $p_{\text{п}}(\mathbf{u})$ – плотности распределений M -мерного вектора входных воздействий $\mathbf{u} = \{u_i\}_{i=1}^M$ по гипотезам наличия H_1 в смеси эхо-сигналов от МО или его отсутствия H_0 , т.е. наличие только шума и, возможно, отражений от МП соответственно.

Надо отметить, что в качестве функционала обычно выбирается функция, монотонно зависящая от l , чаще всего логарифм (\ln). Для анализа уравнения (1) целесообразно воспользоваться наиболее распространенной в практике метеорадиолокации гауссовской аппроксимацией входных воздействий: $\mathbf{u} \sim CN(0, \Phi_1(\Theta))$ и $\mathbf{u} \sim CN(0, \Phi_0)$ для гипотез H_1 и H_0 соответственно. Эта запись означает, что векторы \mathbf{u} являются комплексными C , нормальными N с нулевым средним и корреля-

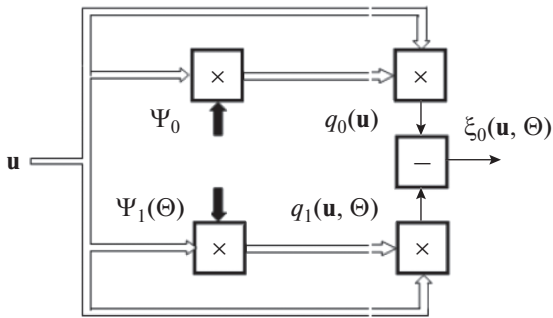


Рис. 1. Схема формирования максимизируемой функции.

ционной матрицей (КМ) Φ . В этом случае отношение правдоподобия можно представить в виде

$$l(\mathbf{u}, \Theta) = c \exp\{\mathbf{u}^* (\Phi_0^{-1} - \Phi_1^{-1}) \mathbf{u}\}, \quad (2)$$

а в роли максимизированной может выступать скалярный функционал векторного аргумента Θ :

$$\begin{aligned} \xi(\mathbf{u}, \Theta) &= c \ln(l(\mathbf{u}, \Theta)) = \\ &= c \xi_0(\mathbf{u}, \Theta) = q_0(\mathbf{u}) - q_1(\mathbf{u}, \Theta), \quad (3) \\ q_0(\mathbf{u}) &= \mathbf{u}^* \Psi_0, q_1(\mathbf{u}, \Theta) = \mathbf{u}^* \Psi_1, \end{aligned}$$

где c – постоянный коэффициент, $\Psi_{0,1} = \Phi_{0,1}^{-1}$ – матрицы, обратные КМ для гипотез H_0 и H_1 соответственно.

На рис. 1 представлена структурная схема получения функции $\xi(\mathbf{u}, \Theta)$, подлежащей максимизации в процессе поиска. Она представляет собой разность двух квадратичных форм вектора входных воздействий с матрицами $\Psi_0(\mathbf{u})$ и $\Psi_1(\mathbf{u}, \Theta)$, обратными КМ Φ_0 и $\Phi_1(\Theta)$. Априори эти матрицы неизвестны и поэтому подлежат оцениванию на этапе адаптации. Вместе с тем известно, что КМ интенсивных отражений от МП имеют достаточно узкий спектр $\Delta F_d < 1$ м/с (даже в X-диапазоне) и в смеси со спектром шума сосредоточены в окрестности нулевой доплеровской частоты $F_d = 0$ (скорость $V_r = 0$). В этой ситуации их можно представить действительными КМ авторегрессионных случайных процессов высокого порядка p с унимодальными спектрами [5, 6].

Учтем, что оптимальное решение поставленной задачи для произвольных КМ Φ_0 и ей обратной Ψ_0 сложно, особенно, если доплеровские спектры отражений от МП достаточно широкие. Поэтому предположения, указанные выше и практически доказанные в [6], позволяют свести ее решение к использованию режекторного фильтра на первом этапе МПО в подавляющем большинстве практических случаев. Структуры и параметры режекторных фильтров могут быть различными, а их

выбор – неоднозначен и, как правило, продиктован представлением о свойствах отражений от МП или базироваться на ранее полученных экспериментальных данных.

2. СТАТИСТИЧЕСКИЙ СИНТЕЗ РЕЖЕКТОРНЫХ ФИЛЬТРОВ

В качестве базового решения для синтеза режекторных фильтров можно использовать известные из литературы [7] фильтры с конечной импульсной характеристикой (КИХ) произвольного m -го порядка. Они обладают важным свойством, которое заключается в том, что на его выходе обеспечивается минимум мощности помехи, если его дискретная импульсная характеристика пропорциональна первой или последней строке $z \times z$ -мерной матрице $\Psi^{(z)} = \Phi^{(z)-1}$, где $\Phi^{(z)}$ – КМ z -мерного вектора помехи на его входе. При этом порядок КИХ-фильтра будет удовлетворять условию $m = z - 1$. Фильтры с такими свойствами в литературе называются фильтрами линейного предсказания с минимальной среднеквадратической ошибкой [8, 9]. Рассмотрим методику синтеза режекторных фильтров на этой основе для произвольного закона зондирования (вобуляции).

Пусть нормированный спектр помехи $s_{\text{норм}}(f)$, подлежащий подавлению, удовлетворяет условию

$$\int_{-1/2}^{1/2} s_{\text{норм}}(f) df = \rho_{pp} = 1, \quad p \in 1, M, \quad (4)$$

где $f = F_d / F_{\text{ср}}$ – доплеровская частота, нормированная к средней частоте зондирования $F_{\text{ср}}$, ρ_{pp} – диагональный элемент нормированной КМ-помехи.

Тогда нормированная $M \times M$ корреляционная матрица помехи $\rho = \{\rho_{pq}\}_{p,q=1}^M$ может быть получена по формуле Винера–Хинчина:

$$\rho = \int_{-1/2}^{1/2} s_{\text{норм}}(f) \mathbf{x}(f) \mathbf{x}^*(f) df, \quad (5)$$

где $\mathbf{x}(f) = \{\exp(j2\pi f \gamma_l)\}_{l=1}^M$, $\gamma_l = \sum_{i=1}^{l-1} T_i / T_{\text{ср}}$ – нормированный к среднему интервалу зондирования $T_{\text{ср}} = 1 / F_{\text{ср}}$ временной интервал между первым и l -м импульсами M -элементной пачки.

Если помеха аппроксимируется процессом авторегрессии высокого порядка p с известной КМ, то вычислений по формуле (5) не требуется. Для спектров гауссовской формы при $p \rightarrow \infty$ можно полагать

$$\rho_{pq} = \rho_1^{|\gamma_p - \gamma_q|}, \quad p, q \in 1, M, \quad (6)$$

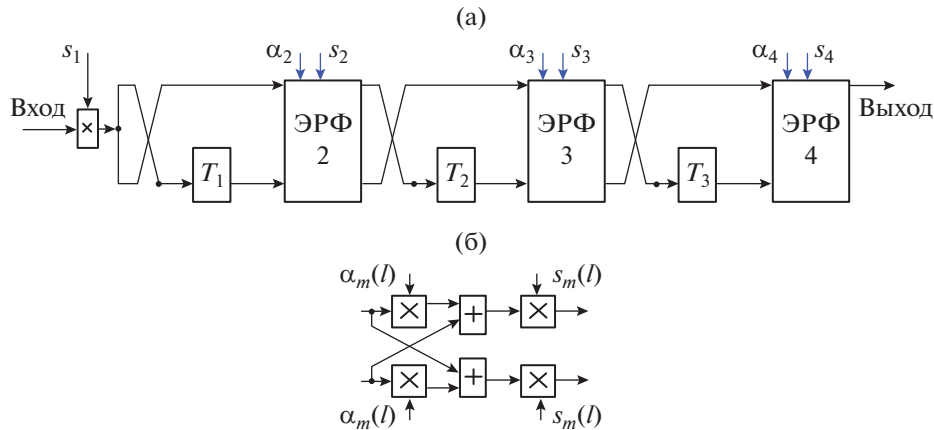


Рис. 2. Схема четырехступенчатого режекторного фильтра (а) и m -й ступени в виде режекторного фильтра (б).

где ρ_1 – коэффициент корреляции (КК) отсчетов помехи, следующих с интервалом T_{cp} .

Тогда КМ аддитивной смеси шума приемника и помехи с относительной интенсивностью κ будет иметь вид

$$\Phi_0 = \{\rho_{pq}\}_{p,q}^M = \mathbf{I}_M + \kappa \rho, \quad (7)$$

где \mathbf{I}_M – единичная $M \times M$ диагональная матрица.

Вспользуемся разложением матрицы, обратной КМ (7), на треугольные сомножители Холецкого $\Psi_0 = \Phi_0^{-1} = \mathbf{H}^* \mathbf{H}$, где $\mathbf{H} = \{h_j\}_{j=1}^z$ – правый сомножитель нижне-верхнего разложения [10]. Тогда z -мерные строки матрицы \mathbf{H} будут представлять собой импульсные характеристики (ИХ) трансверсального КИХ-фильтра m -го порядка. В частности, такой режекторный фильтр может быть также представлен как каскадное соединение элементарных решетчатых фильтров (ЭРФ) [11–13], как показано на рис. 2 для режекторных фильтров 3-го порядка при $z = 4$. Алгоритм определения параметров его настройки сводится к выбору коэффициентов α_m и S_m . Физический смысл параметра α_m заключается в нормировании мощности помехи m -й ступени к единице, а параметр S_m должен декоррелировать выходной процесс m -й ступени с входным.

3. РЕЖЕКТОРНЫЙ ФИЛЬТР ПОМЕХ С БЛИЗКОЙ К НУЛЕВОЙ ШИРИНОЙ СПЕКТРА

В ходе экспериментов, проведенных на первых отечественных ДМРЛ [6, 14], было показано, что коэффициент корреляции отражений от близко стоящих местных предметов (типа труб и мерзлой земли) стремится к единице $\rho_1 \rightarrow 1$, если приемопередающие тракты обладают достаточно

высокой стабильностью. Поэтому ширина спектра межпериодных флуктуаций таких МП достаточно мала [6]. В пределе можно полагать, что она имеет нулевую ширину: $s_{\text{норм}}(f) = \delta(f)$ – дельта-функция. Тогда согласно (5) для нормированной корреляционной матрицы можно записать

$$\rho = \int_{-1/2}^{1/2} \delta(f) \mathbf{x}(f) \mathbf{x}^*(f) df = \mathbf{E} \mathbf{E}^*, \quad (8)$$

где $\mathbf{E} = \{e_{ij}\}_{i=1}^M$, \mathbf{E} – матрица единичного ранга со всеми элементами, равными единице при любом законе зондирования.

В этом случае КМ (7) преобразуется к виду $\Phi_0 = \mathbf{I}_M + \kappa \mathbf{E} \mathbf{E}^*$, а обратная ей матрица равна

$$\Psi_0 = \Phi_0^{-1} = \mathbf{I}_M + \mu \mathbf{E} \mathbf{E}^*, \quad \mu = \kappa / (1 + M \kappa). \quad (9)$$

Тогда треугольную матрицу \mathbf{H} можно представить суммой

$$\mathbf{H} = \mathbf{D1} + \mathbf{D2G}, \quad \mathbf{G} = \{g_{ij}\}_{i,j=1}^M = \begin{cases} 1, & j < i \\ 0, & j \geq i \end{cases}. \quad (10)$$

Для элементов диагональных матриц $\mathbf{D1}$ и $\mathbf{D2}$ справедливы соотношения

$$d1_l = \sqrt{(1 + (l - 1) \kappa) / (1 + l \kappa)} < 1, \\ d2_l = (-\kappa / (1 + (l - 1) \kappa)) d1_l,$$

а \mathbf{G} – нижняя треугольная матрица с нулевой диагональю и единичными поддиагональными элементами.

Согласно этим соотношениям можно показать, что диагональные элементы весового вектора оптимального КИХ-фильтра при $\kappa \gg 1$ и $z \geq 2$ имеют следующий вид:

$$h_{l1} = d1_l = \sqrt{1 / (1 + \kappa)}, \quad h_{zz} = \sqrt{(z - 1) / z}, \quad (11)$$

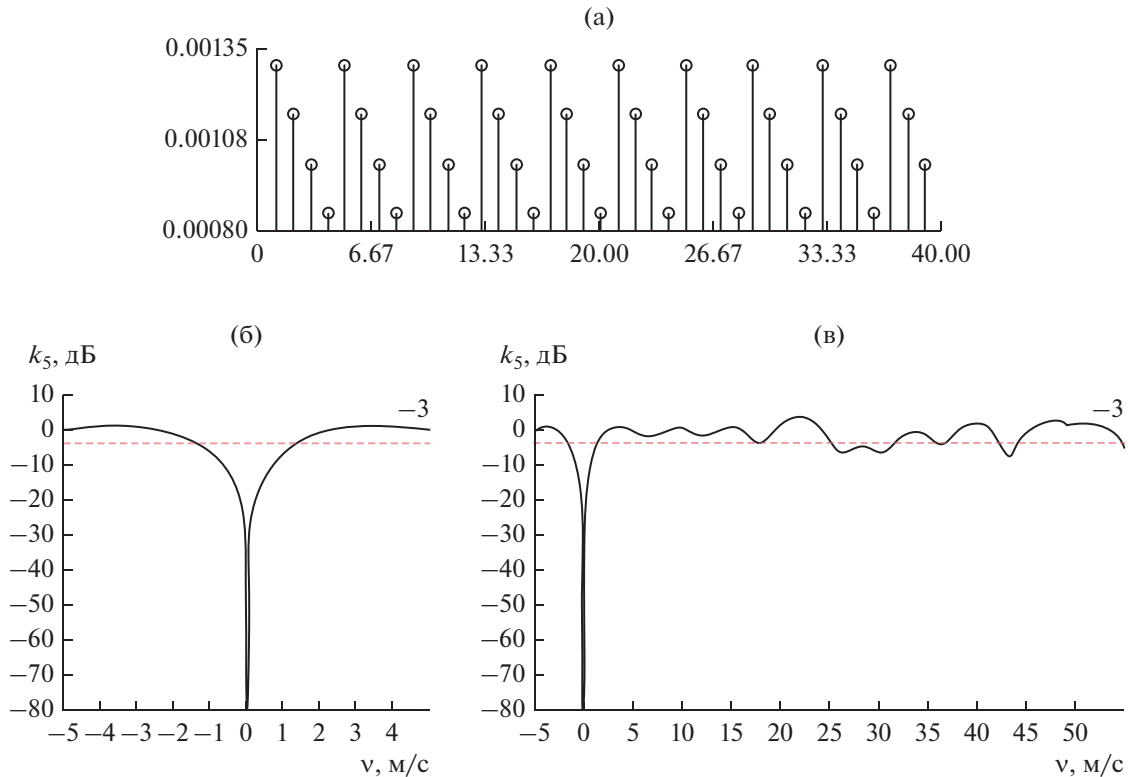


Рис. 3. Закон модуляции 40-элементной пачки импульсов (а) и скоростные характеристики фильтра 4-го порядка (б, в).

а поддиагональные элементы $h_{z,j}$ любой строки матрицы \mathbf{H} равны z -му диагональному элементу матрицы $\mathbf{D2}$.

Свойства матрицы \mathbf{H} заключаются в следующем [10, 14]. Сумма ее диагональных и поддиагональных элементов $\sum_{i=1}^z h_{z,i} \approx 0$ при $\kappa \gg 1$ и $z \geq 2$, а квадрат нормы при тех же условиях — $\mathbf{h}_z^* \mathbf{h}_z = \sum_{i=1}^z h_{z,i}^2 \approx 1$. Эти соотношения позволяют считать, что квадрат модуля частотной характеристики имеет вид

$$K_z(f) = \left| \mathbf{h}_z^* \mathbf{X}_z(f) \right|^2, \quad \mathbf{X}_z(f) = \{x_i(f)\}_{i=1}^z, \quad (12)$$

$$x_i(f) = \exp\{j2\pi f \gamma_i\}$$

и в точке $f = 0$ близок к нулю. Это условие выполняется для фильтров всех порядков $z \gg 2$ с ИХ $\mathbf{h}_{zz}^* = \{h_{z,j}\}_{j=1}^z$ при $z \in 2, M$ и тем точнее, чем выше порядок фильтра.

При этом свойства нормы $\mathbf{h}_z^* \mathbf{h}_z$ означают, что мощность шума на выходе каждого фильтра практически совпадает с мощностью входного некоррелированного шума в смежных z -каналах приема.

На рис. 3а показана 40-элементная модулированная пачка импульсов, а на рис. 3б — соответствующая ей скоростная характеристика (СХ)

$K_z(V_r)$ (для фильтра четвертого порядка) при $\kappa = 10^7$, полученная из (12) заменой аргумента f на $2V_r T_{cp}/\lambda$, где λ — длина волны, T_{cp} — средний период повторения в модулированной пачке импульсов. На рис. 3в показана та же самая СХ, что и на рис. 3б, но в более широком диапазоне. Из рисунка видно, что в точке $V_r = 0$ СХ имеет очень “глубокий” провал, который подавляет неподвижную компоненту любого МП, практически не трогая даже края спектров отражений от МО.

Однако в ряде случаев, например для МП в виде лиственных лесных массивов или травяного покрова земли в летний период, спектр межпериодных флуктуаций может быть достаточно широким, в частности, за счет ветра. Поэтому в этом случае приемлемы режекторные фильтры с большей шириной провала при $V_r = 0$, чем показано на рис. 3.

4. ФИЛЬТРЫ С РАСШИРЕННОЙ ЗОНОЙ РЕЖЕКЦИИ ПОМЕХ

Синтез фильтра с расширенной зоной режекции также можно осуществить по приведенной выше методике. Простейший способ расширения — это

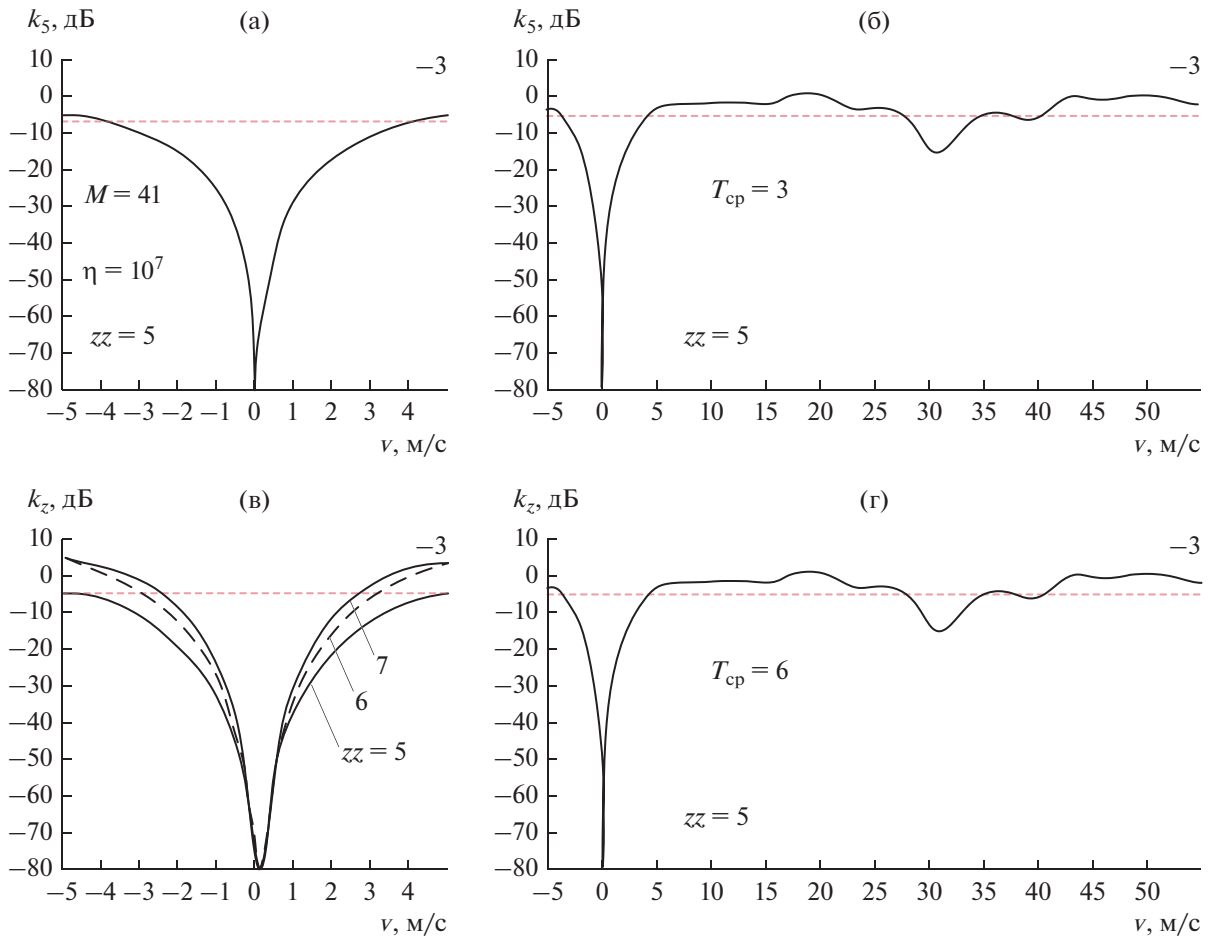


Рис. 4. Скоростные характеристики вариантов фильтров подавления помех с расширенными спектрами: а) $N = 3$, $V_r = f_1\lambda/2$, $V_2 = 0$, $V_3 = -f_1\lambda/2$; б) то же, но в более широком диапазоне; в) $N = 3$, $V_1 = -2f_1\lambda/2$, $V_2 = 0$, $V_3 = \frac{2f_1\lambda}{2}$; г) то же, но в более широком диапазоне.

использовать спектр, составленный из N спектров нулевой ширины

$$s_{\text{норм}}(f) = \sum_{n=1}^N \delta(f - f_n), \quad (13)$$

где f_n – частота n -й компоненты спектра помехи.

Для такой модели спектра нормированная КМ имеет вид

$$\rho = \int_{-1/2}^{1/2} \sum_{n=1}^N \delta(f - f_n) \mathbf{x}(f) \mathbf{x}^*(f) df = \mathbf{X}\mathbf{X}^*, \quad (14)$$

$$\mathbf{X} = \{x(f_i)\}_{i=1}^N.$$

Соответственно, корреляционная матрица Φ_0 смеси помехи и шума и матрица, обратная ей Ψ_0 , вычисляются по формулам

$$\begin{aligned} \Phi_0 &= \mathbf{I}_M + \kappa \mathbf{X}\mathbf{X}^*, \\ \Psi_0 &= \mathbf{I}_M - \kappa \mathbf{X}(\mathbf{I}_N + \kappa \mathbf{X}\mathbf{X}^*)^{-1} \mathbf{X}^* = \mathbf{H}^* \mathbf{H}. \end{aligned} \quad (15)$$

Здесь z -мерные подвекторы строки $\mathbf{h}_z = \{h_{z,j}\}_{j=1}^z$ нижней треугольной $M \times M$ матрицы \mathbf{H} представляют собой импульсные характеристики трансверсальных КИХ-фильтров порядка $m = z - 1$. Эти фильтры должны обеспечить подавление отражений от указанных МП.

На рис. 4 для пачки импульсов с законом вобуляции (см. рис. 3а) представлены скоростные характеристики режкторных фильтров подавления помех со спектром вида (13). При моделировании были приняты следующие параметры: $N = 3$, $V_r = f_1\lambda/2$, $V_2 = 0$, $V_3 = -f_1\lambda/2$ – для рис. 4а; $N = 3$, $V_1 = -2f_1\lambda/2$, $V_2 = 0$, $V_3 = 2f_1\lambda/2$ – для рис. 4в. При этом было выбрано: $f_1\lambda/2 = \Delta V_r = 15 \text{ м/с}$ и $\kappa = 10^7$. На рис. 4б и 4г показаны те же самые СХ, что и на рис. 4а и 4в соответственно, но в увеличенном масштабе.

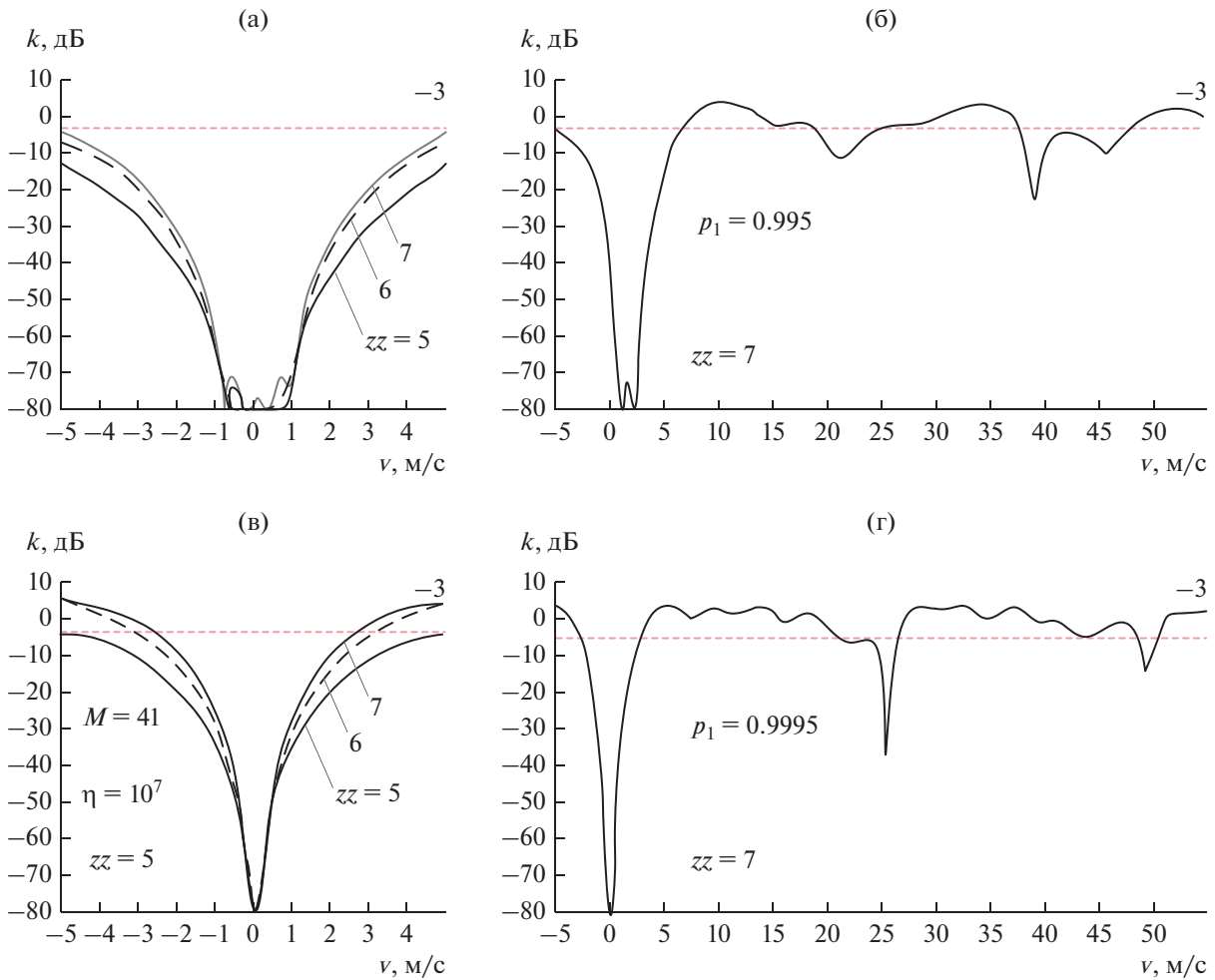


Рис. 5. Скоростные характеристики вариантов фильтров подавления помех с непрерывным спектром и КМ вида (6) при $\rho_1 = 0.995$ (а, б) и 0.9995 (в, г). Рисунки 5б и 5г изображают одну из кривых на рис. 5а и 5в в более широком диапазоне.

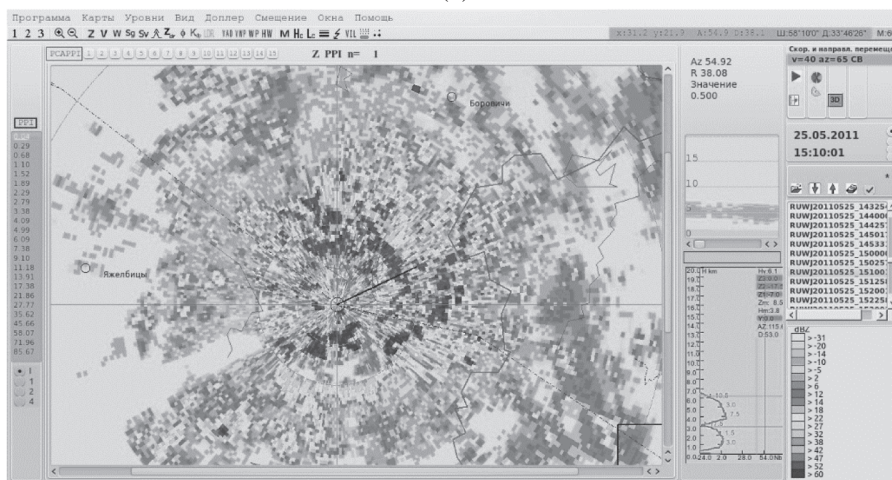
Из рис. 4 видно, что, как и в предыдущем случае, в скоростной характеристике этих фильтров имеются достаточно глубокие “провалы” в окрестности нулевой скорости, но при этом наблюдается существенно более широкая зона режекции по сравнению с рис. 3б. Ширина этой зоны эффективно регулируется значениями ΔV_r , порядком фильтра $m = z - 1$, а также задаваемой интенсивностью помехи κ .

Очевидно, что в качестве исходных здесь могут использоваться не только дискретные спектры вида (13), но и непрерывные спектры. На рис. 5 представлены скоростные характеристики фильтров подавления помех с нормированной КМ вида (6) при $\rho_1 = 0.995$ (рис. 5а, 5б) и $\rho_1 = 0.9995$, (рис. 5в, 5г), $\kappa = 10^7$.

Видно, что в этих случаях также формируется глубокий провал в окрестности $V_r = 0$, ширина и глубина которого могут эффективно регулироваться выбором значения коэффициента корреляции. Причем чем больше значение ρ_1 , тем уже и глубже провал в скоростной характеристике.

На рис. 6а показано одно из экспериментальных конических сечений пространства, иллюстрирующее отражаемость обстановки на малых углах места без использования режекторного фильтра. Видно, что в ближней зоне доплеровского метеорадиолокатора (окрашено темным цветом) из-за влияния местных предметов оценка характеристик МО практически невозможна. После введения КИХ-фильтра 8-го порядка (рис. 6б) отражения от МП в основном подавлены, что создает условия получения корректных оценок параметров МО.

(a)



(б)

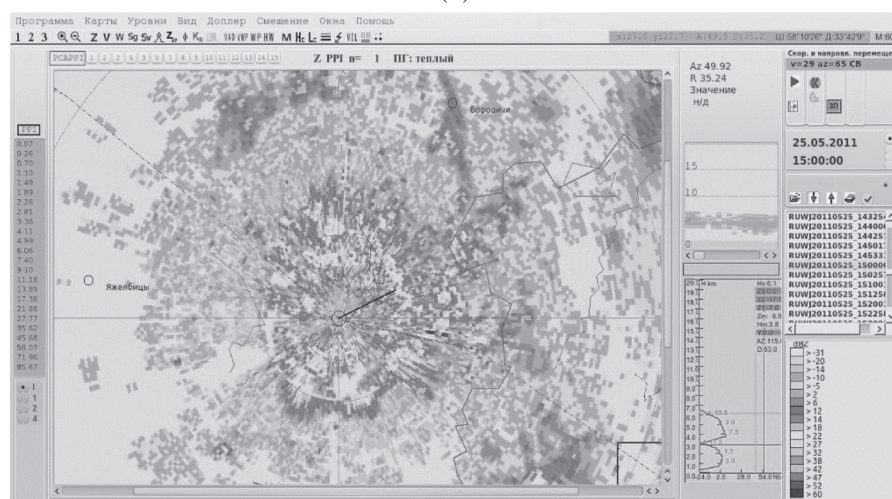


Рис. 6. Экспериментальное свинчатическое сечение пространства доплеровского метеородиолокатора С-диапазона (ДМРЛ-С) на малом угле места при включенном (а) и выключенном (б) режекторном фильтре.

ЗАКЛЮЧЕНИЕ

1. Показано, что задача оценивания параметров метеобразований при наличии отражений от местных предметов сводится к максимизации функционала от отношения правдоподобия. При этом можно сделать обоснованное предположение о гауссовом характере входных воздействий, в которых смесь шума и отражений от местных предметов является помехой, а отражения от метеобразований играют роль полезного сигнала.

2. Методика синтеза межпериодной обработки, оптимальной в статистическом смысле, сводится к отысканию матрицы, обратной корреляционной матрице входного процесса, а оптимальный весовой вектор обработки определяется как строка (столбец) ее треугольного сомножителя разложения Холецкого. В предположении о высокой интенсивности

и узком доплеровском спектре флуктуации местных предметов (при высоком значении коэффициента корреляции) задача упрощается и сводится к выбору параметров неадаптивного режекторного фильтра (КИХ-фильтра).

3. Проведен синтез фильтра с близкой к нулю широкой зоной режекции и определены его импульсная и частотная характеристики. Для 40-элементной вобулированной пачки импульсов представлен конкретный пример частотной характеристики фильтра с глубоким провалом в области нулевой скорости.

4. Синтез фильтра с расширенной зоной режекции проведен на основе представления спектра помехи в виде суммы спектров δ -образной формы. Для тех же условий, что и в п. 3, представлены примеры частотных характеристик с раз-

личной шириной зоны режекции и определены параметры, которые могут управлять шириной и глубиной “провала” в зоне воздействия помех.

Полученные результаты были использованы в первых отечественных доплеровских метеорадиолокаторах С- и Х-диапазонов, где за счет применения режекторных фильтров удалось повысить качество выдаваемой информации.

СПИСОК ЛИТЕРАТУРЫ

1. *Довиак Р., Зрнич Д.* Доплеровские метеорадиолокаторы и метеорологические наблюдения. Л.: Гидрометеиздат, 1988.
2. *Efremov S., Vilegzhanin I., Vovshin B.* // Proc. IRS-2011, Leipzig, Germany, 2011. P. 239.
3. Радиоэлектронные системы. Основы построения и теория. Справочник / Под ред. Я.Д. Ширмана. М.: Радиотехника, 2007.
4. *Сосулин Ю.Г.* Теоретические основы радиолокации и радионавигации. М.: Радио и связь, 1992.
5. *Вовшин Б.М., Жуга Г.А., Лаврукевич В.В., Леховицкий Д.И.* // Прикладная радиоэлектроника. 2007. Т. 6. № 3. С. 21.
6. *Вовшин Б.М., Вылегжанин И.С., Корнеев А.Н., Пушков А.А.* // Успехи совр. радиоэлектроники. 2018. № 2. С. 51.
7. Адаптивные фильтры / Под ред. К. Коузена и П. Гранта. М.: Мир, 1988.
8. *Уидроу Б., Стирнз С.* Адаптивная обработка сигналов. М.: Радиосвязь, 1989.
9. *Репин В.Г., Тартаковский Г.П.* Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио, 1977.
10. *Стренг Г.* Линейная алгебра и ее приложения. М.: Мир, 1980.
11. *Леховицкий Д.И.* // Изв. вузов. Радиофизика, 1992. Т. 35. № 9–10. С. 790.
12. *Вовшин Б.М., Вылегжанин И.С., Корнеев А.Н. и др.* // Успехи совр. радиоэлектроники. 2018. № 6. С. 12.
13. *Третьяков В.А., Куликов Г.В., Лукьянец Ю.Ф.* // Рос. технол. журн. 2020. Т. 8. № 1. С. 34. <https://doi.org/10.32362/2500-316X-2020-8-1-34-42>
14. *Беллман Р.* Введение в теорию матриц. М.: Наука, 1976.
15. *Вовшин Б.М., Вылегжанин И.С., Ефремов В.С., Пушков А.А.* // Сб. статей II Всерос. науч.-техн. конф. “Проблемы военной геофизики и состояния природных сред”. СПб., 2012. Т. 1. С. 370.

ТЕОРИЯ И МЕТОДЫ
ОБРАБОТКИ СИГНАЛОВ

УДК 621.396

СТАТИСТИЧЕСКИЙ СИНТЕЗ АЛГОРИТМА ФОРМИРОВАНИЯ
РАДИОЛОКАЦИОННЫХ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ
С УЧЕТОМ ДЕКОРРЕЛЯЦИИ ИМПУЛЬСОВ

© 2022 г. В. Н. Лагуткин*

Межгосударственная акционерная корпорация “Вымпел”,
ул. Героев Панфиловцев, 10, корп. 1, Москва, 125480 Российская Федерация

*E-mail: vlagutkin@mail.ru

Поступила в редакцию 14.05.2021 г.

После доработки 14.05.2021 г.

Принята к публикации 17.06.2021 г.

С позиций теории статистического синтеза информационных систем рассмотрена задача определения параметров объектов по данным радиолокационных наблюдений с высоким темпом зондирования и изменением ракурса. При разработке параметрической модели сигналов от наблюдаемых объектов использовано приближение локальных отражающих элементов, вращающихся вокруг некоторой оси, и учтена декорреляция принимаемых импульсов, обусловленная неточностями компенсации амплитудных и фазовых изменений, вызванных поступательным движением объекта. Получено общее выражение для отношения правдоподобия как функции относительных амплитуд и координат отражающих элементов и угловой скорости вращения. Показано, что применение критерия максимума функции правдоподобия для определения параметров доминирующих отражающих элементов в частных случаях квазикогерентных или некогерентных импульсов приводит к формированию достаточных статистик в виде квазикогерентных или некогерентных радиолокационных изображений соответственно. Представлены результаты компьютерного моделирования рассмотренных вариантов синтеза радиолокационных изображений при обработке серии импульсов при различных условиях.

DOI: 10.31857/S0033849422020103

ВВЕДЕНИЕ

Исследования методов и условий получения радиолокационных изображений (РЛИ) объектов отражены в большом числе зарубежных и отечественных публикаций, например, [1–6]. Отметим, что в указанных публикациях при построении алгоритмов формирования РЛИ используются известные принципы обработки сигналов от точечных отражателей, что вполне обоснованно с практической точки зрения.

Вместе с тем представляет интерес задача обоснования методов получения РЛИ с позиций теории статистического синтеза информационных систем [7, 8]. В данной работе предпринята попытка такого теоретического обоснования для методов и алгоритмов формирования РЛИ по сериям импульсов, широкополосным или узкополосным, с учетом их декорреляции, т.е. в условиях частичной когерентности или полной некогерентности.

С точки зрения практической целесообразности естественно предположить, что серии радио-

локационных импульсов получены при следующих условиях:

1) интервал изменения ракурса во время наблюдений $\Delta\alpha_p$ существенно больше характерного масштаба изменения диаграммы отражений $\delta\theta_{до} = \lambda/L_{max}$, где λ — длина волны импульсного радиолокатора, L_{max} — максимальный габаритный размер объекта,

2) изменение ракурса за период следования импульсов $\delta\alpha_p$ существенно меньше $\delta\theta_{до}$.

Используемые в работе модели принятых сигналов основаны на хорошо известных физических закономерностях и положениях теории радиолокации [7–11].

1. МОДЕЛЬ ЛОКАЛЬНЫХ ОТРАЖАЮЩИХ
ЭЛЕМЕНТОВ ДЛЯ ПОСЛЕДОВАТЕЛЬНОСТИ
РАДИОЛОКАЦИОННЫХ ИМПУЛЬСОВ

Как известно [9, 10], влияние формы объекта и его ориентации на принятый радиолокационный сигнал с учетом поляризации радиоволн (РВ)

определяется комплексным спектральным коэффициентом отражения

$$\Psi_{\text{пол}}(\omega, \vec{\alpha}) = \mathbf{p}_{\text{Нпр}}^* \Psi_{\text{отр}}(\omega, \vec{k}_1) \mathbf{p}_{\text{Нпад}},$$

где $\mathbf{p}_{\text{Нпад}}$ – вектор-столбец поляризации магнитного поля падающих РВ в связанной с объектом системе координат, $\mathbf{p}_{\text{Нпр}}$ – вектор-столбец поляризации приемной антенны в связанной системе координат, звездочка – знак комплексного сопряжения и транспонирования, $\Psi_{\text{отр}}(\omega, \vec{k}_1)$ – комплексная поляризационная матрица обратного отражения РВ от объекта, \vec{k}_1 – единичный вектор направления наблюдения в связанной с объектом системе координат, вектор $\vec{\alpha}$ отражает влияние на принятый сигнал ориентации объекта и поляризации РВ.

Комплексный спектр нормированной (для компенсации затухания) огибающей принятого сигнала на выходе фильтра, согласованного с зондирующим сигналом [8], можно записать в виде

$$\hat{B}_{\text{псф}}(\Omega, \vec{\alpha}) = \int B_{\text{псф}}(\tau, \vec{\alpha}) \exp(i\Omega\tau) d\tau = |B_0(\Omega)|^2 \Psi_{\text{пол}}(\Omega + \omega_0, \vec{\alpha}), \quad (1)$$

где $B_{\text{псф}}(\tau, \vec{\alpha})$ – нормированная огибающая принятого сигнала на выходе согласованного фильтра, $\hat{B}_0(\Omega)$ – спектр огибающей зондирующего сигнала, ω_0 – опорная частота.

В высокочастотном случае, когда размеры объектов существенно больше длины волны и для расчета характеристик отражения РВ можно использовать приближенные (асимптотические) методы физической теории дифракции [12, 13], комплексная матрица обратного отражения волны от объекта $\Psi_{\text{отр}}(\omega, \vec{k}_1)$ может быть представлена в виде суперпозиции матриц изолированных отражающих элементов

$$\Psi_{\text{отр}}(\omega, \vec{k}_1) \approx \sum_{j=0}^{N-1} \Psi_j(\omega, \vec{k}_1, \beta_j) \exp(i2k_\omega \vec{k}_1 \vec{r}_j),$$

где N – количество отражающих элементов, $\Psi_j(\omega, \vec{k}_1, \beta_j)$ – парциальная поляризационная матрица отражения j -го отражающего элемента, β_j – параметр ракурса наблюдения этого элемента, $k_\omega = \omega/c$ (c – скорость света), \vec{r}_j – координаты условного центра j -го отражающего элемента в связанной системе координат.

Зависимость парциальных матриц отражения от частоты и ракурса наблюдения существенна лишь для элементов с относительно малой кривизной, причем в области углов, близких к углу зеркального отражения. Такие ситуации при обработке сигналов необходимо выделять и анали-

зировать отдельно. На интервалах изменения ракурса вне областей зеркального отражения зависимость парциальных поляризационных матриц от частоты и ракурса наблюдения значительно слабее, при этом количество отражающих элементов N можно считать постоянным. Тогда коэффициент обратного отражения $\Psi_{\text{пол}}(\omega, \vec{\alpha})$ для фиксированной поляризации РВ может быть представлен в виде

$$\Psi_{\text{пол}}(\omega, \vec{\alpha}) \approx \sum_{j=0}^{N-1} \Psi_j \exp(i2k_\omega \vec{k}_1 \vec{r}_j) = \sum_{j=0}^{N-1} \Psi_j \exp(i2k_\omega r_j \cos \theta_j), \quad (2)$$

Ψ_j – комплексный коэффициент отражения j -го отражающего элемента, θ_j – угол между векторами \vec{k}_1 и \vec{r}_j , $r_j = |\vec{r}_j|$.

Выражение (2) представляет собой параметрическую модель принимаемых полезных сигналов. Задачей обработки серии радиолокационных импульсов, принимаемых при наблюдении поворачивающегося относительно радиолокатора объекта, является оценка параметров этой модели: количества доминирующих отражающих элементов, их измеряемых координат в связанной системе координат и комплексных коэффициентов отражения. Измеряемыми являются координаты в продольно-поперечной плоскости, образованной вектором направления наблюдения и векторного произведения вектора поворота и вектора направления наблюдения (см. ниже). Задача восстановления формы объекта по результатам измерений в данной статье не рассматривается.

2. ФУНКЦИЯ ПРАВДОПОДОБИЯ ДЛЯ ПОСЛЕДОВАТЕЛЬНОСТИ ДЕКОРРЕЛИРОВАННЫХ РАДИОЛОКАЦИОННЫХ ИМПУЛЬСОВ

Предположим, что в результате согласованной внутриимпульсной обработки, компенсации радиолокационного затухания и изменений принятых импульсов, вызванных поступательным движением объекта, формируются дискретные нормированные комплексные развертки $\xi_m(\tau_k)$ ($m = 0, \dots, M-1$ – порядковый номер импульса, k – номер отсчета по задержке), которые представляют собой аддитивную смесь нормированных разверток радиолокационных импульсов $B_{\text{псф}}(\tau_k, \vec{\alpha}_m)$ с комплексными коэффициентами c_m и реализаций гауссовского белого (в полосе импульсов) нормированного шума $\eta_m(\tau_k)$

$$\xi_m(\tau_k) = c_m B_{\text{псф}}(\tau_k, \vec{\alpha}_m) + \eta_m(\tau_k),$$

где $\tau_k = k\Delta\tau$, $\Delta\tau = 1/\Delta f_\partial$ – интервал оцифровки развертки, Δf_∂ – ширина полосы импульсов, $k = 0, \dots, K - 1$. Максимальный размер цифровых разверток $K_{\max} \approx (2L_{\max}/c)\Delta f_\partial$. Для узкополосных импульсов $K = 1$, для широкополосных – $K \gg 1$.

Комплексные коэффициенты c_m отражают влияние декорреляции принятых импульсов, обусловленной неточностями оценки и компенсации амплитудных и фазовых изменений, вызванных поступательным движением объекта. При идеально точной компенсации $c_m = 1$. Более реально считать коэффициенты c_m , $m = 0, \dots, M - 1$ неизвестными комплексными величинами, изменяющимися от импульса к импульсу. С теоретической точки зрения учет декорреляции импульсов удобен тем, что позволяет единообразно рассматривать случаи квазикогерентных, частично когерентных и некогерентных импульсов.

Комплексные развертки принимаемых импульсов и их изменение от импульса к импульсу помимо помеховых составляющих – шума и коэффициентов декорреляции – зависят от следующих параметров объекта наблюдения:

- количества доминирующих отражающих элементов объекта,
- параметров отражающих элементов объекта: коэффициентов (амплитуд) отражения и координат в продольно-поперечной плоскости,
- вектора угловой скорости поворота объекта относительно радиолокатора, который в общем случае может быть неизвестным.

Задача синтезируемого алгоритма обработки серии радиолокационных импульсов заключается в статистической оценке указанных параметров. В соответствии с основными положениями теории статистического синтеза алгоритмов обработки ее решение основано на построении отношения правдоподобия и его максимизации по указанным параметрам [7, 8].

Так как $\eta_m(\tau_k)$ – независимые нормальные случайные величины, то из выражения для совместной условной плотности вероятности дискретных разверток

$$f(\{\xi_m(\tau_k)\}/\{B_{\text{псф}}(\tau_k, \bar{\alpha}_m)\}, \{c_m\}) = (2\pi\sigma_\eta^2)^{-MK/2} \times \exp\left\{-\frac{1}{2\sigma_\eta^2} \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} |\xi_m(\tau_k) - c_m B_{\text{псф}}(\tau_k, \bar{\alpha}_m)|^2\right\}$$

получим выражение для оптимизируемой функции правдоподобия

$$L(\{\xi_m(\tau_k)\}/\{B_{\text{псф}}(\tau_k, \bar{\alpha}_m)\}, \{c_m\}) = \frac{f(\{\xi_m(\tau_k)\}/\{B_{\text{псф}}(\tau_k, \bar{\alpha}_m)\}, \{c_m\})}{f(\{\xi_m(\tau_k)\}/\{c_m = 0\})} = \exp\left\{\frac{1}{\sigma_\eta^2} \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} \left[\text{Re}\left(c_m \xi_m^*(\tau_k) B_{\text{псф}}(\tau_k, \bar{\alpha}_m)\right) - |c_m B_{\text{псф}}(\tau_k, \bar{\alpha}_m)|^2/2\right]\right\},$$

где σ_η^2 – дисперсия нормированного шума.

Для широкополосных импульсов функцию правдоподобия, используя (1) и (2), можно записать для комплексных дискретных спектров разверток:

$$L(\{\hat{\xi}_m(\Omega_{k_1})\}/\{\hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m)\}, \{c_m\}) = \exp\left\{\frac{1}{2K\sigma_\eta^2} \sum_{m=0}^{M-1} \sum_{k_1=0}^{K-1} \left[\text{Re}\left(\hat{\xi}_m^*(\Omega_{k_1}) c_m \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m)\right) - |c_m \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m)|^2/2\right]\right\}, \quad (3)$$

где $\Omega_{k_1} = k_1\Delta\Omega$, $\Delta\Omega = 2\pi/(K\Delta\tau)$,

$$\hat{\xi}_m(\Omega_{k_1}) = \sum_{k=0}^{K-1} \xi_m(\tau_k) \exp(i\Omega_{k_1}\tau_k) = \sum_{k=0}^{K-1} \xi_m(\tau_k) \exp\left(i\frac{2\pi}{K}k_1k\right)$$

– дискретный спектр полученных разверток,

$$\hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) = |\hat{B}_0(\Omega_{k_1})|^2 \times \sum_{j=0}^{N-1} \Psi_j \exp\left(i2\frac{\Omega_{k_1} + \omega_0}{c} \bar{k}_j \bar{r}_j(t_m)\right) = |\hat{B}_0(\Omega_{k_1})|^2 \sum_{j=0}^{N-1} \Psi_j \exp\left(i2\frac{\Omega_{k_1} + \omega_0}{c} z_j(t_m)\right), \quad (4)$$

$z_j(t_m) = r_j \cos \vartheta_j(t_m)$, $t_m = m\Delta t$, Δt – период следования импульсов.

Отметим, что для прямоугольных линейно-частотно-модулированных (ЛЧМ) импульсов при больших величинах базы спектр, по существу, ограничен и его амплитуда практически постоянна в пределах полосы шириной Δf_∂ [3].

Для учета изменения углового положения объекта (как совокупности отражающих элементов) относительно радиолокатора в процессе наблюдения можно использовать упрощенную модель, полагая, что объект поворачивается в текущий промежуток времени с некоторой угловой скоростью ω вокруг оси, составляющей с направлением

наблюдения угол χ . Такая модель пригодна для случаев наблюдения как вращающегося объекта, так и стабилизированного объекта, когда изменение ракурса наблюдения обусловлено поступательным движением объекта относительно радиолокатора. При отсутствии дрейфа отражающих элементов для такой модели углового движения имеем

$$\begin{aligned}\cos \vartheta_j(t_m) &= \cos \chi \cos \gamma_j + \sin \chi \sin \gamma_j \cos \varphi_j(t_m), \\ \varphi_j(t_m) &= \varphi_{j0} + w_m \Delta t,\end{aligned}$$

где γ_j – угол между вектором \vec{r}_j и осью вращения, и

$$\begin{aligned}z_j(t_m) &= r_j \cos \vartheta_j(t_m) = \\ &= z'_{jc} + z'_{j0} \cos(w_1 m) - y'_{j0} \sin(w_1 m),\end{aligned}$$

где

$$z'_{jc} = r_j \cos \gamma_j \cos \chi, \quad z'_{j0} = r_j \sin \gamma_j \sin \chi \cos \varphi_{j0},$$

$$y'_{j0} = r_j \sin \gamma_j \sin \chi \sin \varphi_{j0}, \quad w_1 = w \Delta t$$

– в общем случае неизвестные опорные параметры.

Условимся, что ближний к радиолокатору в момент начала наблюдения отражающий элемент имеет номер $j = 0$ и для него $z'_{0c} = 0$.

В случае широкополосных сигналов при $\cos \chi \neq 0$ и небольших изменениях ракурса $\vartheta_j(t_m)$ на интервале времени наблюдения для функции $z_j(t_m)$ с целью уменьшения числа оцениваемых параметров можно использовать аппроксимацию вида

$$z_j(t_m) \approx z_{j0} \cos(w_{1j} m) - y_{j0} \sin(w_{1j} m), \quad (5)$$

где z_{j0} , y_{j0} , w_{1j} – аппроксимирующие параметры, связанные с опорными параметрами соотношениями

$$z'_{j0} = \left(\frac{w_{1j}}{w_1}\right)^2 z_{j0}, \quad y'_{j0} = \frac{w_{1j}}{w_1} y_{j0}, \quad z'_{jc} = z_{j0} - z'_{j0}, \quad (6)$$

причем для отражающего элемента $j = 0$ аппроксимирующие параметры совпадают с опорными,

так как $w_{10} = w_1$, $z'_{0c} = 0$.

Тогда (4) можно представить в виде

$$\begin{aligned}\hat{B}_{\text{псф}}(\Omega_{k_1}, \vec{\alpha}_m) &= \left|\hat{B}_0(\Omega_{k_1})\right|^2 \sum_{j=0}^{N-1} \Psi_j \exp\left[i2 \frac{\Omega_{k_1} + \omega_0}{c} \times \right. \\ &\quad \left. \times (z_{j0} \cos(w_{1j} m) - y_{j0} \sin(w_{1j} m))\right].\end{aligned} \quad (7)$$

В случае узкополосных сигналов или $|\cos \chi| \ll 1$ постоянная составляющая z'_{jc} функции $z_j(t_m)$ не существенна и аппроксимация (5) становится точной, при этом $z_{j0} = z'_{j0}$, $y_{j0} = y'_{j0}$, $w_{1j} = w$.

Отметим, что модель (5) позволяет учесть дрейф отражающих элементов.

Для учета декорреляции импульсов используем для последовательности c_m ортогональное кусочно-постоянное разложение

$$\begin{aligned}c_m &= \sum_{m_1} \tilde{c}_{m_1} \chi_{m_1}(m), \\ \chi_{m_1}(m) &= \begin{cases} 1, & 0 \leq m - m_1 M_1 < M_1 \\ 0, & m < m_1 M_1, m \geq (m_1 + 1) M_1 \end{cases},\end{aligned}$$

где M_1 – характерный интервал изменения последовательности c_m .

В этом случае отношение правдоподобия (3) как функция параметров \tilde{c}_{m_1} достигает максимума при условии

$$\begin{aligned}\tilde{c}_{m_1} &= (1/E_{m_1}) \sum_{m=0}^{M_1-1} \sum_{k_1=0}^{K-1} \hat{\xi}_{m_1 M_1 + m}(\Omega_{k_1}) \hat{B}_{\text{псф}}^*(\Omega_{k_1}, \vec{\alpha}_{m_1 M_1 + m}), \\ E_{m_1} &= \sum_{m=M_1 m_1}^{m=M_1(m_1+1)-1} \sum_{k_1=0}^{K-1} \left|\hat{B}_{\text{псф}}(\Omega_{k_1}, \vec{\alpha}_m)\right|^2.\end{aligned} \quad (8)$$

Интегрирование (усреднение) отношения правдоподобия (8) по \tilde{c}_{m_1} методом перевала с использованием (4), (7), (8) приводит к выражению

$$\begin{aligned}\bar{L}_{\{\tilde{c}_{m_1}\}} &= \left\{ \frac{\{\hat{\xi}_m(\Omega_{k_1})\}}{\{\hat{B}_{\text{псф}}(\Omega_{k_1}, \vec{\alpha}_m)\}} \right\} = \\ &= \exp \left\{ \sum_{m_1} \left[\sum_{m=0}^{M_1-1} \sum_{k_1=0}^{K-1} \hat{\xi}_{m_1 M_1 + m}(\Omega_{k_1}) \times \right. \right. \\ &\quad \left. \left. \times \sum_{j=0}^{N-1} \tilde{\Psi}_{jm_1}^* \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1}}{c} z_j(t_{m_1 M_1 + m'}) \right] \right]^2 \right\}, \\ \tilde{\Psi}_{jm_1} &= \Psi_j / \sqrt{2K\sigma_{\eta}^2 \tilde{E}_{m_1}},\end{aligned} \quad (9)$$

$$\begin{aligned}\tilde{E}_{m_1} &= \sum_{m=0}^{M_1-1} \sum_{k_1=0}^{K-1} \left| \sum_{j=0}^{N-1} \Psi_{jm_1}^* \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1}}{c} z_j(t_{m_1 M_1 + m'}) \right] \right|^2, \\ &\quad \sum_{j=0}^{N-1} |\tilde{\Psi}_{jm_1}|^2 \leq C^2 < \infty, \quad z_j(t_{m_1 M_1 + m'}) = \\ &= z_{j0} \cos(w_{1j}(m_1 M_1 + m')) - y_{j0} \sin(w_{1j}(m_1 M_1 + m')), \end{aligned}$$

которое можно использовать для конструирования схем алгоритмов обработки в различных случаях.

3. ОЦЕНКА ПАРАМЕТРОВ ДОМИНИРУЮЩИХ ОТРАЖАЮЩИХ ЭЛЕМЕНТОВ ОБЪЕКТОВ ПО СЕРИИ ЧАСТИЧНО КОГЕРЕНТНЫХ ИМПУЛЬСОВ

При $M_1 K > 1$ показатель экспоненты в (9) целесообразно преобразовать к виду

$$\ln \bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \hat{\xi}_m(\Omega_{k_1}) \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} = \left| \Phi_{m_1}(z_0, y_0, w_1) \right|^2 = \sum_{m'=0}^{M_1-1} \sum_{k_1=0}^{K-1} \hat{\xi}_{m_1 M_1 + m'}(\Omega_{k_1}) \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1}}{c} \times \right. \quad (12)$$

$$\Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}) = \sum_{m'=0}^{M_1-1} \sum_{k_1=0}^{K-1} \hat{\xi}_{m_1 M_1 + m'}(\Omega_{k_1}) \times \left. \times (z_0 \cos(w_1(m_1 M_1 + m')) - y_0 \sin(w_1(m_1 M_1 + m'))) \right]^2$$

для множества возможных значений w_1 ;

2) некогерентное накопление этих ККРЛИ – получение частично когерентных РЛИ (ЧКРЛИ)

$$G(z_0, y_0, w_1) = \sum_{m_1} \left| \Phi_{m_1}(z_0, y_0, w_1) \right|^2, \quad (13)$$

3) анализ накопленных ЧКРЛИ и выделение доминирующих отражающих элементов объекта на фоне шума с оценкой их аппроксимирующих координат z_{j_0} , y_{j_0} и угловых скоростей w_{1j} по критерию

$$(z_{j_0}, y_{j_0}, w_{1j}) = \arg \max_{z_0, y_0, w_1} G(z_0, y_0, w_1),$$

а также амплитуд $\tilde{\Psi}_{j m_1}$, $j = 0, \dots, N - 1$ по формуле (10);

4) приведение координат доминирующих отражающих элементов к единой угловой скорости по формулам (6).

4. ОЦЕНКА РАДИОЛОКАЦИОННЫХ ПАРАМЕТРОВ ОБЪЕКТОВ ПО СЕРИИ НЕКОГЕРЕНТНЫХ УЗКОПОЛОСНЫХ ИМПУЛЬСОВ С ПЕРЕМЕННОЙ ИЛИ С ПОСТОЯННОЙ НЕСУЩЕЙ ЧАСТОТОЙ

В случае некогерентных импульсов с переменной несущей частотой оптимизируемая функция принимает вид

$$\ln \bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \hat{\xi}_m(\Omega_{k_1}) \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} = \sum_{m_1} \left| \sum_{j=0}^{N-1} \tilde{\Psi}_{j m_1}^* \Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}) \right|^2,$$

$$\Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}) = \sum_{m'=0}^{M_1-1} \sum_{k_1=0}^{K-1} \hat{\xi}_{m_1 M_1 + m'}(\Omega_{k_1}) \times$$

$$\times \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1}}{c} (z_{j_0} \cos(w_{1j}(m_1 M_1 + m')) - y_{j_0} \sin(w_{1j}(m_1 M_1 + m'))) \right].$$

Так как максимум логарифма отношения правдоподобия по $\tilde{\Psi}_{j m_1}$, $j = 0, \dots, N - 1$ при заданном N и $\sum_{j=0}^{N-1} |\tilde{\Psi}_{j m_1}|^2 \leq C^2 < \infty$ достигается при

$$\tilde{\Psi}_{j m_1} = \arg \max_{\{\tilde{\Psi}_{j j}\}} \ln \bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \hat{\xi}_m(\Omega_{k_1}) \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} = \varepsilon \Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}), \quad (10)$$

$$\varepsilon = \left[\frac{1}{\sum_{j=0}^{N-1} \left| \Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}) \right|^2} \right]^{1/2},$$

то для определения параметров доминирующих отражающих элементов $\{z_{j_0}, y_{j_0}, w_{1j}, j = 0, \dots, N - 1\}$ получаем критерий максимума функции

$$f(\{z_{j_0}, y_{j_0}, w_{1j}, j = 0, \dots, N - 1\}) = \max_{\{\tilde{\Psi}_{j j}\}} \ln \bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \hat{\xi}_m(\Omega_{k_1}) \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} = \quad (11)$$

$$= \sum_{j=0}^{N-1} \sum_{m_1} \left| \Phi_{m_1}(z_{j_0}, y_{j_0}, w_{1j}) \right|^2.$$

В соответствии с (11) можно предложить следующий порядок обработки серии импульсов с учетом их декорреляции при $M_1 K > 1$:

1) получение в каждом цикле с номером m_1 достаточных статистик в виде квазикогерентных радиолокационных изображений (ККРЛИ)

$$\bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \xi_m \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} \approx \exp \left\{ \sum_{m_1} \left| \sum_{j=0}^{N-1} \xi_m \tilde{\Psi}_{j m_1}^* \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1(m)}}{c} z_j(t_m) \right] \right|^2 \right\} = \quad (14)$$

$$= \exp \left\{ \sum_{m_1} \left| \xi_m \right|^2 \left| \sum_{j=0}^{N-1} \tilde{\Psi}_{j m_1}^* \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1(m)}}{c} z_j(t_m) \right] \right|^2 \right\}, \quad z_j(t_m) = z'_{j_0} + z'_{j_0} \cos(w_1 m) - y'_{j_0} \sin(w_1 m).$$

Показатель экспоненты (14) можно преобразовать к виду

$$\ln \bar{L}_{\{\tilde{\epsilon}_m\}} \left\{ \left\{ \xi_m \right\} / \left\{ \hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m) \right\} \right\} \approx \sum_{j, j'=0}^{N-1} \tilde{\Psi}_{j m_1}^* \tilde{\Psi}_{j' m_1} \sum_m \exp \left[-i2 \frac{\omega_0 + \Omega_{k_1(m)}}{c} (z_{j m} - z_{j' m}) \right] \left| \xi_m \right|^2 =$$

$$= 2 \operatorname{Re} \sum_{j_1=0}^J \tilde{\Psi}_{j_1} \sum_m \left| \xi_m \right|^2 \exp \left(-i2 \frac{\omega_0 + \Omega_{k_1(m)}}{c} \Delta z_{j_1 m} \right), \quad j_1 = j(j-1)/2 + j', \quad j = \overline{0, N-1}, \quad j' = \overline{1, j}, \quad (15)$$

$$J = N(N-1)/2, \quad \tilde{\Psi}_0 = \sum_{j=0}^{N-1} |\Psi_j|^2 / 2, \quad \Delta z_{0m} = 0, \quad \tilde{\Psi}_{j_1} = \Psi_j^* \Psi_{j'}, \quad \Delta z_{j_1 m} = z_{j m} - z_{j' m}.$$

Максимум (15) по $\hat{\Psi}_{j_1}$, $j_1 = \overline{0, J}$ при заданном N и $\sum_{j_1=0}^J |\hat{\Psi}_{j_1}|^2 \leq C_1^2 < \infty$ достигается при

$$\hat{\Psi}_{j_1} = \sum_m |\hat{\xi}_m|^2 \exp\left(i2 \frac{\omega_0 + \Omega_{k(m)}}{c} \Delta z_{j_1 m}\right),$$

т.е.

$$\begin{aligned} \max_{\{\hat{\Psi}_{j_1}\}} \ln \bar{L}_{\{\hat{\xi}_m\}}(\{\hat{\xi}_m\} / \{\hat{B}_{\text{псф}}(\Omega_{k_1}, \bar{\alpha}_m)\}) = \\ = \sum_{j_1=0}^J \left| \sum_m |\hat{\xi}_m|^2 \exp\left(i2 \frac{\omega_0 + \Omega_{k_1(m)}}{c} \Delta z_{j_1 m}\right) \right|^2, \end{aligned} \quad (16)$$

где изменения относительных координат отражающих элементов $\Delta z_{j_1 m}$ от импульса к импульсу

(по m) определяются аппроксимирующим соотношением, аналогичным (5)

$$\Delta z_{j_1 m} = \Delta z_{j_1 0} \cos(w_{1j_1} m) - \Delta y_{j_1 0} \sin(w_{1j_1} m),$$

где $\Delta z_{j_1 0}$, $\Delta y_{j_1 0}$, w_{1j_1} — аппроксимирующие параметры, связанные с опорными параметрами соотношениями

$$\Delta z'_{j_1 0} = \left(\frac{w_{1j_1}}{w_1}\right)^2 \Delta z_{j_1 0}, \quad \Delta y'_{j_1 0} = \frac{w_{1j_1}}{w_1} \Delta y_{j_1 0}, \quad (17)$$

$$\Delta z'_{j_1 c} = \Delta z_{j_1 0} - \Delta z'_{j_1 0}.$$

В случае некогерентных импульсов с переменной несущей частотой, как видно из (16), достаточной статистикой является некогерентное радиолокационное изображение (НРЛИ)

$$H(\Delta z_0, \Delta y_0, w_1) = \left| \sum_m |\hat{\xi}_m|^2 \exp\left[i2 \frac{\omega_0 + \Omega_{k(m)}}{c} (\Delta z_0 \cos(w_1 m) - \Delta y_0 \sin(w_1 m))\right] \right|^2. \quad (18)$$

Если несущая частота импульсов не меняется, то выражение для НРЛИ принимает вид

$$H_{\text{уп}}(\Delta z_0, \Delta y_0, w_1) = \left| \sum_m |\hat{\xi}_m|^2 \exp\left[i2 \frac{\omega_0}{c} (\Delta z_0 \cos(w_1 m) - \Delta y_0 \sin(w_1 m))\right] \right|^2. \quad (19)$$

В случае некогерентных импульсов синтезируемые цифровые некогерентные радиолокационные изображения (18), (19) для фиксированных угловых скоростей w_1 являются аналогом голограмм интенсивности [14].

В соответствии с (18), (19) можно предложить следующий порядок обработки серии некогерентных импульсов:

1) получение НРЛИ (18) или (19) для множества возможных значений w_1 ;

2) анализ полученных НРЛИ и выделение доминирующих элементов на фоне шума с оценкой их аппроксимирующих параметров $\Delta z_{j_1 0}$, $\Delta y_{j_1 0}$, w_{1j_1} по критерию

$$(\Delta z_{j_1 0}, \Delta y_{j_1 0}, w_{1j_1}) = \arg \max_{z_0, y_0, w_1} H(\Delta z_0, \Delta y_0, w_1);$$

3) приведение координат доминирующих элементов к единой угловой скорости w_1 с оценкой $\Delta z'_{j_1 0}$, $\Delta y'_{j_1 0}$, $\Delta z'_{j_1 c}$ по формулам (17);

4) восстановление параметров отражающих элементов с оценкой их количества J и координат z'_{j_0} , y'_{j_0} , z'_{j_c} , $j = \overline{1, J-1}$ с учетом того, что $z'_{00} = 0$, $y'_{00} = 0$, $z'_{0c} = 0$.

5. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ АЛГОРИТМОВ

Для численного исследования рассмотренных методов обработки серий радиолокационных импульсов разработана компьютерная модель, включающая программу-имитатор радиолокационных измерений в процессе изменения ракурса наблюдения во времени и комплекс алгоритмов обработки серий измерений. Имитатор радиолокационных измерений с учетом аппаратного шума, выполняющий расчет компонентов поляризационной матрицы отражения $S_{EE}(\omega, \theta)$, $S_{HH}(\omega, \theta)$ для осесимметричных объектов, разработан на основе приближений физической теории дифракции [13].

При численной реализации алгоритма оценки параметров принимаемых полезных сигналов в качестве достаточных статистик формируются РЛИ (см. формулы (12), (18), (19)) на прямоугольной сетке в плоскости дальность—поперечное расстояние. Параметры сетки (количество узлов по обеим координатам и размеры ячеек) задает пользователь. При малых суммарных углах поворота объекта, когда $\cos(w_1 M) \approx 1$, $\sin(w_1 M) \approx w_1 M$, формирование РЛИ может быть выполнено с помощью алгоритма быстрого преобразования Фурье.

Для примера на рис. 1–3 представлены результаты моделирования обработки серий из 1000 импульсов, отраженных от идеально проводящего

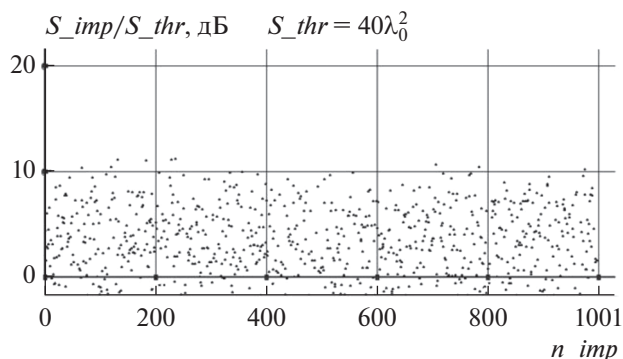


Рис. 1. Пример имитированной последовательности ЭОП.

конечного кругового цилиндра диаметром $30\lambda_0$ и длиной $40\lambda_0$ (λ_0 – длина волны).

При имитации входных данных в примере предполагалось, что за время измерений угол между направлением облучения и осью цилиндра (ракурс наблюдения) равномерно увеличивается с 25° до 75° (со скоростью $0.05^\circ/\text{имп}$) при этом вектор поляризации электрического поля остается в плоскости падения.

Для примера на рис. 1 точками показаны значения отношений (в дБ) эффективной отражающей площади (ЭОП) объекта S_{imp} к пороговому значению ЭОП, эквивалентному аппаратурному шуму, $S_{thr} = 40\lambda_0^2$ в зависимости от номера импульса последовательности.

На рис. 2 показаны результаты цифрового синтеза цифровых частично когерент-

ных РЛИ при $M_1 = 100$, $S_{thr} = 40\lambda_0^2$ для двух типов импульсов:

- а) при использовании широкополосных ЛЧМ-импульсов с относительной девиацией частоты $\delta f = \Delta f_\delta / f_c = 0.1$, $K = 8$;
- б) при использовании узкополосных импульсов.

На рисунке горизонтальная ось – это ось z , вертикальная ось – ось y . Цена деления по обеим осям $10\lambda_0$. Размер изображений по оси z – $75\lambda_0$, по оси y – $100\lambda_0$. Число элементов дискретизации по каждой оси 128.

На цифровом ЧКРЛИ рис. 2а четко выделяются три пятна с приблизительными координатами максимумов в точках $A(36, -17)$, $B(13, 26)$, $C(0, 0)$, соответствующих трем отражающим элементам конечного кругового цилиндра, наблюдаемого под углом 25° в момент начала наблюдения.

На ЧКРЛИ рис. 2б эти отражающие элементы проявляются гораздо менее четко, с существенно худшим разрешением (особенно по продольной оси z).

Моделирование показало, что при усилении шума на порядок, т.е. при $S_{thr} = 400\lambda_0^2$, на широкополосных ЧКРЛИ отмеченные выше пятна также могут быть выделены, однако узкополосные ЧКРЛИ становятся неинформативными.

На рис. 3 показаны результаты синтезированной цифровых некогерентных РЛИ ($S_{thr} = 4\lambda_0^2$) при использовании последовательностей узкополосных импульсов:

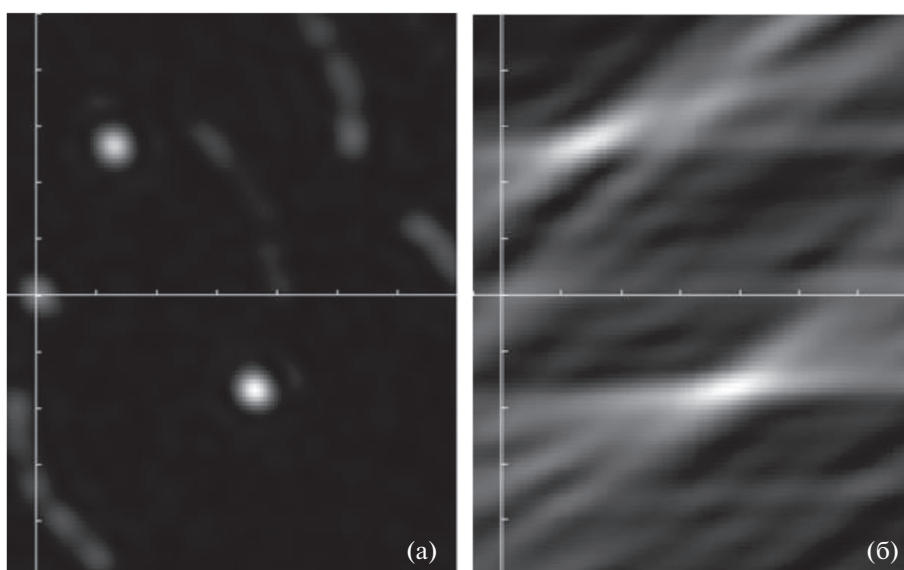


Рис. 2. Синтезированные частично когерентные РЛИ для широкополосных (а) и узкополосных (б) импульсов.

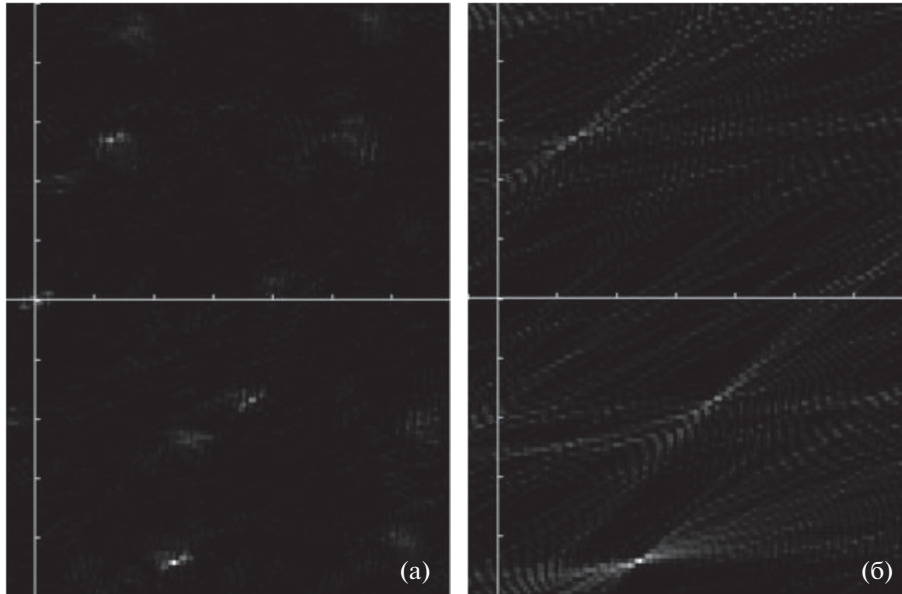


Рис. 3. Синтезированные некогерентные РЛИ для последовательностей узкополосных импульсов с циклически возрастающей частотой (а) и с постоянной частотой (б).

а) с циклически линейно возрастающей частотой в диапазоне с относительной шириной $\delta f = 0.1$, период цикла импульсов $K = 8$;

б) с постоянной частотой.

На НРЛИ рис. 3а выделяются три группы пикселей, сконцентрированных вблизи координат, в соответствии с (25) они равны разностям координат точек A, B, C , выделенных на рис. 2а. На НРЛИ рис. 3б подобные группы пикселей также выделяются, но они менее компактны и более распределены (особенно вдоль продольной оси) по сравнению с рис. 3а.

Моделирование показало, что при усилении шума на порядок, т.е. при $S_{thr} = 40\lambda_0^2$, НРЛИ становятся неинформативными.

ЗАКЛЮЧЕНИЕ

Проведенное теоретическое рассмотрение статистической задачи определения параметров объектов по данным радиолокационных наблюдений с высоким темпом зондирования и изменением ракурса позволяет сделать следующие выводы:

1) для учета изменения углового положения объекта (как совокупности отражающих элементов) относительно радиолокатора в процессе наблюдения можно использовать упрощенную модель (5), полагая, что объект поворачивается в текущий промежуток времени с некоторой угловой скоростью вокруг оси, не совпадающей с линией наблюдения;

2) в случае частично когерентных импульсов применение критерия максимума функции правдоподобия (11) для определения параметров доминирующих отражающих элементов приводит к формированию достаточных статистик в виде квазикогерентных РЛИ (12) в каждом цикле, соответствующем интервалу когерентности, для множества возможных значений угловой скорости вращения и некогерентному накоплению этих ККРЛИ – получению частично когерентных РЛИ (13);

3) в случае некогерентных узкополосных импульсов с переменной или с постоянной несущей частотой достаточными статистиками для определения параметров доминирующих отражающих элементов являются соответствующие некогерентные РЛИ (18), (19);

4) численное исследование рассмотренных методов обработки серий радиолокационных импульсов с получением радиолокационных изображений при различных условиях подтвердило их работоспособность.

СПИСОК ЛИТЕРАТУРЫ

1. *Wehner D.R.* High-Resolution Radar. Boston: Artech House, 1995.
2. *Cheney M., Borden B.* Fundamentals of Radar Imaging. Philadelphia: Soc. for Industrial and Appl. Math. (SIAM), 2009.
3. *Chen V.C., Martorella M.* Inverse Synthetic Aperture Radar Imaging. Principles, Algorithms and Applications. Edison: Sci.Tech. Publ., 2014.

4. Радиолокационные характеристики объектов / Под ред. С.М. Нестерова. М.: Радиотехника, 2015.
5. Курикса А.А. // Труды III Всерос. науч.-техн. конф. Радиотехнич. ин-та им. А.Л. Минца. М.: МГТУ, 2015. С. 238.
6. Казанцев А.А., Перов Д.А., Самородов А.А., Самородов Б.А. // Ural Radio Engineering J. 2018. V. 2. № 2. P. 67.
7. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио, 1977.
8. Ширман Я.Д., Манжос В.Н. Теория и техника обработки радиолокационной информации на фоне помех. М.: Радио и связь, 1981.
9. Справочник по радиолокации. Т. 1. Основы радиолокации. / Под ред. М. Сколника. М.: Сов. радио, 1976.
10. *Radar Cross Section Handbook* / Ed. G.T. Ruck. N.Y.: Plenum Press, 1970.
11. Кобак В.О. Радиолокационные отражатели. М.: Сов. радио, 1975.
12. Keller J.B. // J. Opt. Soc. Amer. 1962. V. 52. № 2. P. 116.
13. Уфимцев П.Я. Теория дифракционных краевых волн в электродинамике. Введение в физическую теорию дифракции. М.: Бином-Лаборатория знаний, 2012.
14. Курикса А.А. // РЭ. 1968. Т. 13. № 5. С. 771.

ТЕОРИЯ И МЕТОДЫ
ОБРАБОТКИ СИГНАЛОВ

УДК 004.85

ИССЛЕДОВАНИЕ ПОМЕХОУСТОЙЧИВОСТИ
ГЛУБОКОЙ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ПРИ ОБНАРУЖЕНИИ
ТРАНСПОРТНЫХ СРЕДСТВ НА АЭРОФОТОСНИМКАХ ЗЕМЛИ

© 2022 г. О. В. Ильина^а, М. В. Терешонок^{а, б, *}

^аМосковский технический университет связи и информатики,
ул. Авиамоторная, 8а, Москва, 111024 Российская Федерация

^бМосковский физико-технический институт (национальный исследовательский университет),
Институтский пер., 9, Долгопрудный Московской обл., 141701 Российская Федерация

*E-mail: m.v.tereshonok@mtuci.ru

Поступила в редакцию 29.06.2021 г.

После доработки 14.09.2021 г.

Принята к публикации 15.09.2021 г.

Исследована помехоустойчивость глубокой сверточной нейронной сети, решающей задачу обнаружения транспортных средств на аэрофотоснимках Земли, исследование позволяет получить оценку качества и корректности работы обученной нейронной сети. Также предложен способ повышения помехоустойчивости системы обнаружения объектов на аэрофотоснимках Земли.

DOI: 10.31857/S0033849422020048

ВВЕДЕНИЕ

В настоящее время в области компьютерного зрения наблюдается тенденция перехода от статистических методов к нейросетевым. Методы глубокого обучения позволяют достичь самых современных результатов по некоторым специфичным проблемам. К задачам компьютерного зрения относятся:

- классификация изображений (присвоение метки класса изображению);
- локализация объектов (определение ограничивающей рамки вокруг одного или нескольких объектов на изображении);
- обнаружение объектов (объединяет задачи классификации и локализации, т.е. определяет ограничивающую рамку вокруг каждого интересующего объекта на изображении и присваивает им метку класса).

Распознавание объектов и их локализация на изображениях – одна из самых фундаментальных и сложных проблем компьютерного зрения. Генерация предложений регионов [1] на изображении, где находится искомый объект, сейчас является активной областью исследований. Приложения обнаружения объектов возникают во многих различных областях, включая обнаружение лиц [2, 3], обнаружение пешеходов для беспилотных автомобилей [4, 5], обнаружение аномалий в медицине [6, 7], обнаружение объектов на аэрофото- и спутниковых снимках [8–10] и т.п. При ис-

пользовании алгоритмов обнаружения объектов важно иметь высокое покрытие интересующих объектов на тестовом изображении, поскольку пропущенные объекты не могут быть восстановлены для их последующей классификации.

Обнаружение объектов на аэрофото- и спутниковых снимках представляет большой практический интерес из-за разнообразия приложений во многих областях, включая мониторинг трафика [11, 12], операции поиска и спасения [13], мониторинг морского судоходства [14] и т.д. На данный момент существуют несколько наборов данных аэрофото- и спутниковых снимков Земли в свободном доступе, которые активно используются для обучения систем обнаружения транспортных средств [15–17]. Такие наборы данных охватывают площадь поверхности до десятков тысяч квадратных километров и могут содержать различное количество транспортных средств на изображениях, которые имеют различные размеры, формы, цвета, яркость и контраст, а также произвольную ориентацию на местности. Качество этих наборов данных может отличаться от качества реальных аэрофото- и спутниковых снимков Земли, например, из-за атмосферных явлений (осадки, туман и т.п.), а также из-за загрязненности, наличия царапин и сколов на объективе съемочной аппаратуры.

Анализ не нейросетевых алгоритмов обнаружения объектов показал, что введение небольших изменений яркости пикселей во входном изобра-

жении вызывает заметное ухудшение качества локализации объектов. Даже изменение одного пикселя изображения уже демонстрирует сильное ухудшение [1]. Однако анализ последних публикаций показывает, что исследований влияния шумовых воздействий на качество работы нейросетевых систем обнаружения и распознавания объектов не проводится. В основном известные исследования систем обнаружения объектов сводятся к повышению метрики точности таких систем на тестовой выборке с фиксированным качеством и уровнем шума. Причем при оценке этих метрик используются тестовые данные с такими же статистическими данными, как и при обучении. В таких работах не исследуется влияние искажений входных изображений на точность обнаружения объектов.

Достижимая точность распознавания фиксированной тестовой выборки не позволяет судить о поведении системы при распознавании изображений с другим уровнем шума. Это обстоятельство ограничивает состоятельность оценки обобщающей способности системы распознавания.

Таким образом, исследование помехоустойчивости глубокой сверточной нейронной сети, решающей задачу обнаружения объектов на изображениях с разным уровнем шума, даст дополнительную информацию о корректности и качестве модели.

Оценка робастности распознавания изображений нейронной сетью в приложении к верификации подписей была проведена в работе [18], в результате была получена зависимость вероятности правильного обнаружения от вероятности ложной тревоги при фиксированном уровне искажений изображения. Выбор архитектуры нейронной сети для распознавания зашумленных изображений был также исследован в [19]. Тем не менее в обеих работах влияние уровня шума в обучающей и контрольной выборках на точность распознавания не было исследовано.

Цель данной работы – исследовать влияние белого гауссова шума, добавленного к значениям яркости пикселей входного изображения, на качество работы глубокой сверточной нейронной сети автоматического обнаружения транспортных средств на аэрофотоснимках Земли.

1. АНАЛИЗ СУЩЕСТВУЮЩИХ МЕТОДОВ ОБНАРУЖЕНИЯ И РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Среди различных архитектур, основанных на глубоких нейронных сетях, сверточная нейронная сеть (СНС) является самым известным и современным методом для задач, связанных с классификацией объектов на изображении. В отличие от задач классификации изображений, обнаруже-

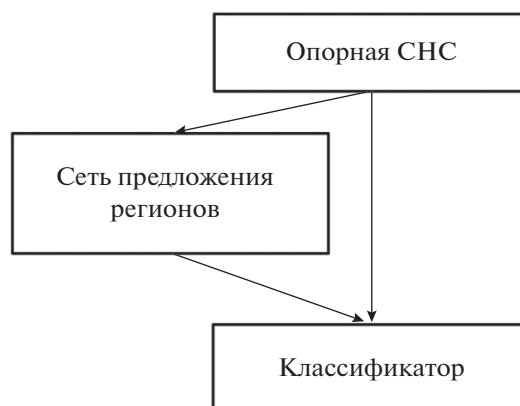


Рис. 1. Упрощенная схема архитектуры двухэтапной системы обнаружения и распознавания объектов на изображении.

ние требует локализации объектов на изображении [20].

Современные алгоритмы обнаружения объектов на изображении можно разделить на две категории: одноэтапные и двухэтапные. К одноэтапным относятся SSD, YOLO, RetinaNet [21], в которых предложение регионов и классификация выполняются за один проход изображения по сети. К двухэтапным сетям относят Faster R-CNN [22], где изображение проходит этап предложения регионов (Region Proposal Network, RPN) и этап классификации объектов, полученных из RPN. Такие системы обладают высокой точностью локализации и распознавания объектов. Таким образом, двухэтапные сети целесообразно использовать при обработке насыщенных объектами статичных изображений с максимальной точностью, а одноэтапные больше подходят для обработки видеопотока с небольшим числом и большим размером объектов в кадре.

Упрощенная схема двухэтапного алгоритма обнаружения и распознавания объектов на изображении на базе глубоких сверточных нейронных сетей Faster R-CNN [22] представлена на рис. 1. Такая архитектура состоит из трех основных блоков: опорная СНС, сеть предложения регионов и классификатор. В качестве опорной СНС используются хорошо зарекомендовавшие себя в задачах распознавания объектов на изображении традиционные сверточные нейронные сети без детекторной части, так как опорная сеть решает задачу извлечения признаков из входного изображения. Полученные с выхода опорной СНС карты признаков поступают на вход сети предложения регионов, которая решает задачу генерации прямоугольников, ограничивающих потенциальные объекты на входном изображении, и задачу бинарной классификации, т.е. задачу предсказания нахождения или отсутствия в предсказанном

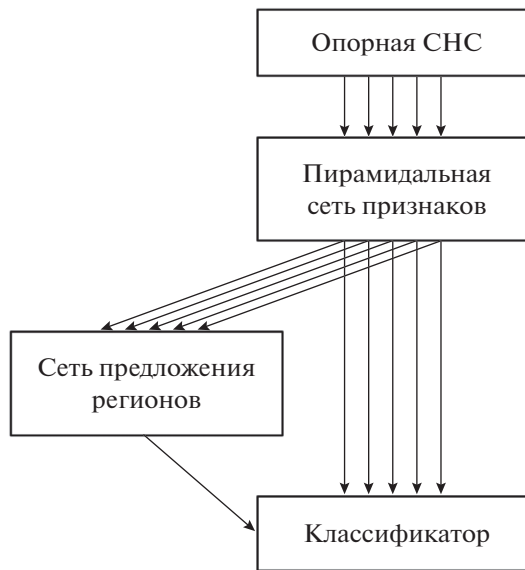


Рис. 2. Упрощенная схема архитектуры системы обнаружения и распознавания объектов на изображении Faster R-CNN с FPN.



Рис. 3. Упрощенная схема архитектуры исследуемой системы обнаружения объектов на входном изображении.

прямоугольнике объекта, причем данная сеть не распознает класс объектов на входном изображении. Полученные с выхода RPN регионы, имеющие высокую вероятность нахождения в них объекта, и полученные с выхода опорной СНС карты признаков поступают в классификатор, где происходит распознавание классов объектов, обнаруженных сетью RPN, а также уточняет координаты входных регионов для более точного обнаружения. Классификатор представляет собой детекторную часть, которая была отсечена от традиционной СНС для получения архитектуры опорной сети.

На выходе двухэтапной системы обнаружения и распознавания объектов (см. рис. 1) получим набор предсказанных системой объектов с информацией об их классе и координатами ограничивающих рамок, в которые эти объекты заключены.

Для решения проблемы обнаружения объектов разных масштабов в статье [23] предложена пирамидальная сеть признаков (Feature Pyramid Network – FPN). Данная сеть является некой надстройкой к существующим архитектурам обнаружения объектов и не зависит от опорных СНС. Выходом такой пирамидальной архитектуры являются карты признаков в нескольких пространственных разрешениях. Упрощенная схема архитектуры системы обнаружения и распознавания объектов на изображении Faster R-CNN с надстройкой FPN представлена на рис. 2.

Система Faster R-CNN с FPN в данный момент занимает лидирующее место в задачах распознавания транспортных средств на аэрофотоснимках Земли [12, 14, 22]. Из архитектуры двухэтапных систем обнаружения и распознавания

объектов на изображении видно, что качество обнаружения объектов в сети RPN сильно влияет на качество обнаружения и распознавания системы в целом, поскольку классификатор работает только с регионами, предложенными ему сетью RPN. Так, в случае потери большого количества истинных объектов на этапе сети предложения регионов, точность обнаружения и распознавания объектов всей системой будет значительно ниже. Поэтому для исследования влияния степени зашумленности аэрофотоснимков Земли на качество обнаружения объектов выбрана часть двухэтапной системы Faster R-CNN с надстройкой FPN, которая решает только задачу обнаружения объектов на изображении. Упрощенная схема архитектуры исследуемой системы изображена на рис. 3.

В качестве входной информации данной системы выступают цветные аэрофотоснимки Земли, в качестве выходной – список координат ограничивающих рамок, где с высокой вероятностью находятся искомые транспортные средства.

2. МЕТРИКИ ОЦЕНКИ КАЧЕСТВА СИСТЕМ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ

Системы обнаружения и распознавания объектов на изображении решают задачи классификации (определяют, присутствует ли объект на изображении и распознают его класс) и локализации (предсказывают координаты ограничивающей рамки вокруг объекта, когда объект присутствует на изображении). Величина полноты обнаружения позволяет оценить способность модели

обнаруживать все релевантные объекты и показывает долю правильных предсказаний от общего числа истинных объектов. Величина точности обнаружения показывает способность модели идентифицировать только релевантные объекты и равна отношению количества правильных обнаружений к общему числу предсказаний [25]. Для вычисления полноты и точности обнаружения каждый предсказанный системой обнаружения ограничивающий прямоугольник сначала должен быть классифицирован как:

– истинно-положительный (true positive, TP): правильное обнаружение ограничивающего прямоугольника;

– ложноположительный (false positive, FP): не-правильное обнаружение несуществующего объекта или несоответствующее обнаружение существующего объекта;

– ложноотрицательный (false negative, FN): необнаруженный истинный ограничивающий прямоугольник.

Приведем уравнения для вычисления метрик полноты R и точности P обнаружения [25]:

$$R = \frac{\sum_{n=1}^S TP_n}{\sum_{n=1}^S TP_n + \sum_{n=1}^{G-S} FN_n}, \quad (1)$$

$$P = \frac{\sum_{n=1}^S TP_n}{\sum_{n=1}^S TP_n + \sum_{n=1}^{N-S} FP_n}, \quad (2)$$

где G – количество всех истинных объектов, N – количество предсказанных системой обнаружения объектов, из которых S объектов предсказаны верно ($S \leq G$), n – номер предсказанного ограничивающего потенциальный объект прямоугольника.

В случае обнаружения объектов на изображении необходимо вычислить метрику отношения площади пересечения к площади объединения истинного и предсказанного прямоугольника (Intersection Over Union, IoU), которая показывает, насколько близки обнаруженные сетью ограничивающие рамки к ограничивающим рамкам истинных объектов. Это измерение выполняется независимо для каждого класса объектов путем оценки степени перекрытия предсказанной и достоверной областей [25]. Идеальное совпадение двух ограничивающих рамок происходит, когда метрика IoU между ними становится равной 1. В случае, когда оба ограничивающих прямоугольника не пересекают друг друга, метрика IoU равна 0. Чем ближе к единице становится IoU, тем лучше считается обнаружение. Поскольку детекторы объ-

ектов также выполняют классификацию каждого ограничивающего прямоугольника, только истинные и обнаруженные ограничивающие рамки одного и того же класса сравниваются метрикой IoU.

После вычисления метрики IoU для всех истинных и всех предсказанных ограничивающих рамок необходимо установить пороговое значение метрики IoU для определения правильности предсказанных ограничивающих рамок. В случае, когда IoU больше или равно заданному пороговому значению, предсказание является истинно-положительным (TP), но в случае, когда IoU меньше заданного порогового значения, то предсказание является ложно положительным (FP).

Оценка качества обнаружения для систем, не зависящих от класса (class-agnostic), сильно отличается от традиционного обнаружения, зависящего от класса, поскольку большинство показателей (точность, матрица ошибок и т.п.) не применимы. В задачах обнаружения объектов общепринято оценивать качество предложений на основе величины полноты обнаружения [1].

3. ОБЗОР НАБОРА ДАННЫХ ДЛЯ ОБУЧЕНИЯ И ОЦЕНКИ ИССЛЕДУЕМОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ ТРАНСПОРТНЫХ СРЕДСТВ

Сравнение известных и находящихся в открытом доступе наборов данных для задач обнаружения транспортных средств на аэрофото- и спутниковых снимках Земли приведено в табл. 1. Из сравнительного анализа видно, что набор, описанный в статье [15], имеет больше исходных изображений и больше категорий объектов, что позволяет обучать и оценивать нейронные сети на более разнообразных данных.

Используемый для обучения и оценки системы обнаружения объектов набор данных содержит пять классов объектов, принадлежащих к категории транспортных средств (легковой автомобиль, грузовой автомобиль, самолет, вертолет, корабль). Интересующие нас объекты в данном наборе данных имеют стороны ограничивающих рамок в диапазоне от 2 до 800 пикселей, а отношения сторон ограничивающих рамок этих объектов варьирует от значения 0.1 до 11. Также на одном изображении может находиться до 1650 транспортных средств, что происходит из-за плотного размещения объектов [15].

4. ИССЛЕДОВАНИЕ ПОМЕХОУСТОЙЧИВОСТИ СИСТЕМЫ ОБНАРУЖЕНИЯ ТРАНСПОРТНЫХ СРЕДСТВ НА АЭРОФОТОСНИМКАХ ЗЕМЛИ

В качестве исследуемой системы обнаружения объектов на аэрофотоснимках Земли выступает глубокая сверточная нейронная сеть, упрощен-

Таблица 1. Сравнение известных наборов данных для задач обнаружения транспортных средств на аэрофото- и спутниковых снимках Земли

Набор данных	Литература	Количество			Размеры изображений, пикс
		классов объектов	объектов	изображений	
DOTA	[15]	14	188 282	2806	от 800 до 4000
NWPU VHR-10	[16]	10	3651	800	~1000
VEDAI	[17]	3	2950	1268	512, 1024

ная архитектура которой изображена на рис. 3. Для обучения и оценки данной системы используется набор данных, представленный в статье [15], который разбивается на обучающую и тестовую выборки. Нами исследуются несколько подходов к обучению системы обнаружения. Первый подход основан на обучении модели обнаружения объектов на исходных (не зашумленных) данных. Также в работе рассмотрены два случая обучения на зашумленных выборках с разной степенью искажения. Зашумленные аэрофотоснимки создаются из исходной обучающей выборки по следующим правилам:

- 1) исходная обучающая выборка случайным образом перемешивается и делится на три части с равным количеством изображений;
- 2) одна подвыборка остается без изменений;
- 3) к изображениям из второй подвыборки добавляется белый гауссов шум по правилу (3);
- 4) к изображениям из третьей подвыборки добавляется белый гауссов шум по правилу (4);

$$Y_2 = X_2 + W_2, \quad (3)$$

$$Y_3 = X_3 + W_3, \quad (4)$$

где X_2, X_3 – исходные изображения, используемые для второй и третьей подвыборок соответственно, представленные в виде трехмерных матриц пикселей, имеющих высоту h пикселей, ширину w пикселей и глубину $d = 3$ пикселя, что соответствует количеству каналов RGB-изображения (диапазон значений яркости в каналах RGB 0...255); W_2, W_3 – матрицы случайных величин с нормальным распределением и нулевой средней величиной, имеющие такие же размеры, что и матрицы X_2 и X_3 соответственно, причем случайные величины матрицы W_2 имеют стандартное отклонение, равное 30 для второго подхода обучения и 100 для третьего подхода обучения (в этих случаях отношения сигнал/шум составляют 4.55 и –5.9 дБ соответственно), а случайные величины матрицы W_3 имеют стандартное отклонение, равное 40 для второго подхода обучения и 200 для третьего подхода обучения (в этих случаях отношения сигнал/шум составляют 2.05 и –11.93 дБ соответственно); Y_2, Y_3 –

изображения второй и третьей подвыборок соответственно.

Шумовое искажение аэрофотоизображений является следствием множества различных процессов. Центральная предельная теорема (ЦПТ) утверждает, что сумма достаточно большого количества некоррелированных случайных процессов с примерно одинаковыми масштабами имеет распределение, близкое к нормальному [26]. Кроме того, множество различных случайных процессов, влияющих на итоговое искажение, имеет различную физическую природу и потому независимо, а следовательно, не коррелировано. Таким образом, автокорреляционная функция итогового искажения имеет вид дельта-функции. Исходя из изложенного выше в качестве шумового воздействия в данной работе используется белый гауссов шум.

Описанные выше варианты обучения (обучение на исходном и частично зашумленном наборах данных соответственно) проводятся на одной и той же архитектуре глубокой СНС (см. рис. 3) и с одинаковыми параметрами обучения.

Для получения кривых помехоустойчивости на вход сети подаются изображения из тестовой выборки с разным уровнем шума и оценивается величина полноты обнаружения для каждого случая. Величина полноты обнаружения вычисляется для всех предсказанных ограничивающих рамок, имеющих наивысшую вероятность нахождения в данной рамке объекта, по формуле (1) при метрике пересечения предсказанных рамок с истинными (IoU) больше 0.5. Для вычисления степени зашумленности в данном исследовании используется величина отношения сигнал/шум (ОСШ), формула которого имеет вид

$$\text{ОСШ} = 20 \times \lg \left(\frac{\sigma_{\text{и}}}{\sigma_{\text{ш}}} \right), \quad (5)$$

где $\sigma_{\text{и}}$ – стандартное отклонение тестового набора данных, характеризующее средний контраст изображения, $\sigma_{\text{ш}}$ – стандартное отклонение аддитивного белого гауссова шума.

На рис. 4 представлены графики зависимости величины полноты обнаружения объектов от

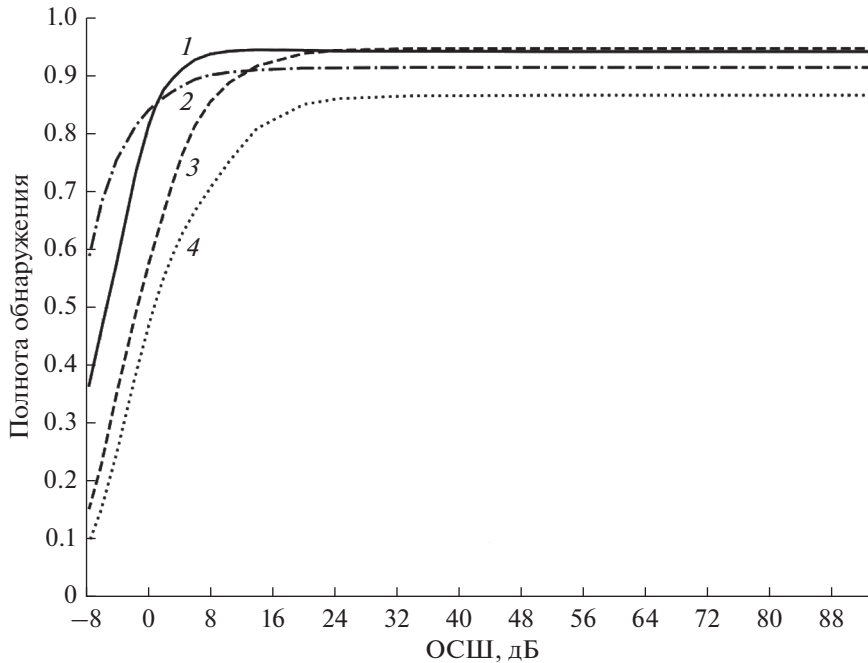


Рис. 4. Графики зависимости величины полноты обнаружения от ОСШ: 1 — для модели, обученной на исходной обучающей выборке; 2 — для модели, обученной на оптимально зашумленной обучающей выборке (треть набора данных осталась без искажений, к изображениям из другой трети набора данных добавлен белый гауссов шум со стандартным отклонением 30, к оставшейся части набора данных добавлен белый гауссов шум со стандартным отклонением 40); 3 — для модели, обученной на избыточно зашумленной обучающей выборке (треть набора данных осталась без искажений, к изображениям из другой трети набора данных добавлен белый гауссов шум со стандартным отклонением 100, к оставшейся части набора данных добавлен белый гауссов шум со стандартным отклонением 200); 4 — для модели, переобученной на исходных наборах данных.

ОСШ для трех различных систем: обученной на исходной обучающей выборке; на оптимально зашумленной обучающей выборке, полученной добавлением зашумленных изображений по правилу (3), (4) с ОСШ = 4.55 и 2.05 дБ; и на избыточно зашумленной обучающей выборке, полученной добавлением зашумленных изображений по правилу (3), (4) с ОСШ = -5.9 дБ и ОСШ = -11.93 дБ. Для сравнения там же представлен график помехоустойчивости системы обнаружения, полученной путем переобучения модели на исходных данных. В этом случае параметры обучения системы обнаружения отличаются от параметров обучения систем, описанных выше.

Для визуализации на рис. 5 представлены образцы изображений с различными значениями ОСШ, используемые для данной оценки.

Графики помехоустойчивости системы обнаружения, представленные на рис. 4, позволяют выделить две области отношений сигнал/шум:

1) “переходная область”, в которой наблюдается заметное (на десятки процентов) повышение полноты обнаружения с ростом ОСШ (-8...8 дБ);

2) “область высоких ОСШ”, в которой полнота обнаружения высока и мало (менее чем на 1%)

меняется, объекты на изображении визуально четко различимы (ОСШ > 8 дБ).

Из графиков, представленных на рис. 4, видно, что при увеличении уровня шума модель, обученная на оптимально зашумленном наборе данных (кривая 2), показывает способность обнаруживать больше релевантных объектов, чем та же модель, но обученная на исходном наборе данных (кривая 1). Необходимо отметить, что при ОСШ, превышающем 20 дБ в проведенном в данной работе эксперименте, модель, обученная на исходном наборе данных, показывает величину полноты обнаружения на 0.004 больше, что является незначительно более высоким значением. Из этого можно сделать вывод, что такая модель лучше работает на неискаженных данных, однако хуже способна противостоять шумовому воздействию. Добавление в обучающую выборку данных с чрезмерно низким ОСШ (кривая 3), т.е. с таким уровнем шума, при котором визуально объекты на изображениях аэрофотоснимках сложно различить, приводит к существенному снижению качества распознавания изображений в области ОСШ выше 0 дБ (в данном эксперименте). При ОСШ ниже 0 дБ наблюдается более высокая помехоустойчивость данной системы по сравнению с системой, обученной на оптимально зашумленных

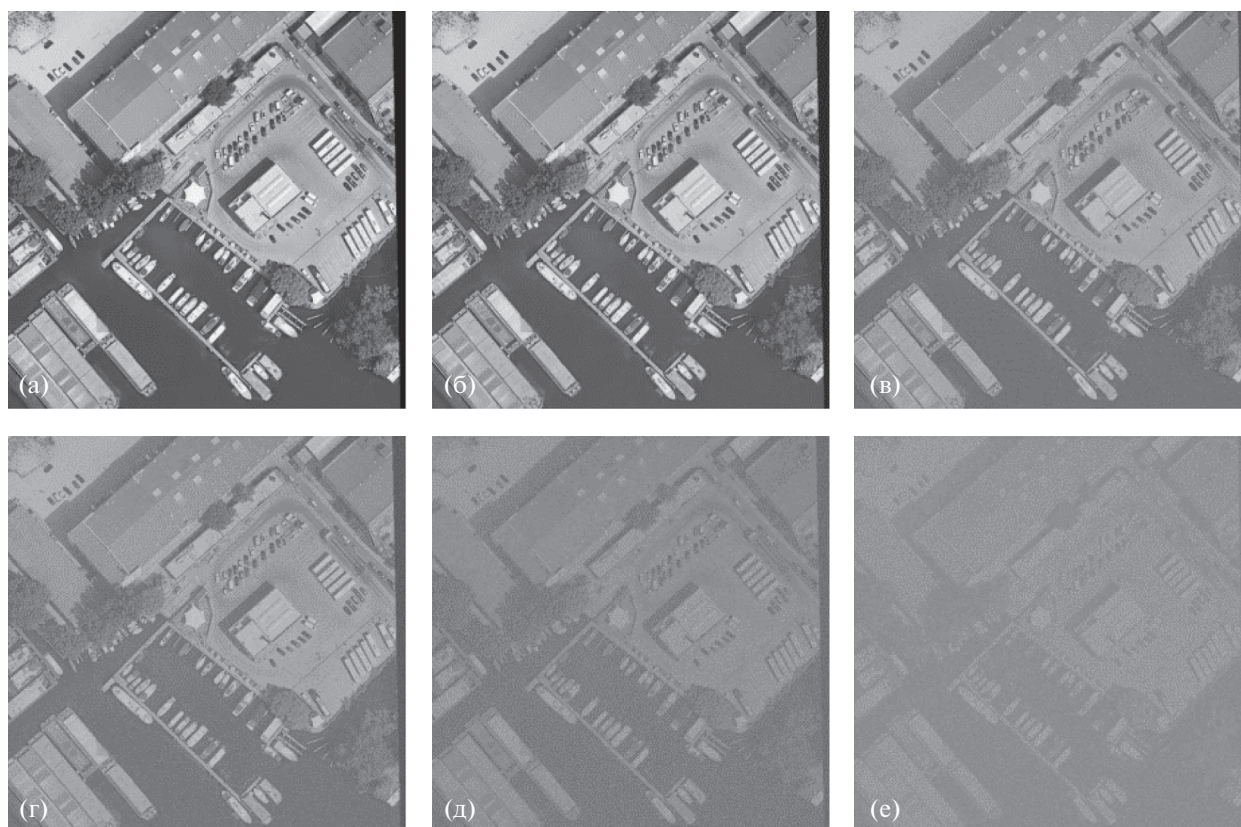


Рис. 5. Образцы изображений из тестовой выборки с различными значениями ОСШ: 94.1 (а), 20.1 (б), 4.55 (в), 2 (г), –5.9 (д), –11.9 дБ (е).

данных (кривая 2). В то же время добавление в обучающую выборку данных с чрезмерно низким ОСШ снижает интегральную (суммарную по всем значениям ОСШ) полноту обнаружения из-за внесения слишком большой неопределенности в весовые коэффициенты нейронной сети. То есть обучение системы обнаружения на наборе данных с чрезмерно низким ОСШ приводит к сдвигу влево и вниз кривой помехоустойчивости. Данное наблюдение свидетельствует о необходимости поиска оптимального уровня неопределенности, которая вводится в обучающую выборку, для получения наивысшей помехоустойчивости системы.

ЗАКЛЮЧЕНИЕ

В данной работе проведено исследование влияния шумового воздействия на величину полноты обнаружения современной архитектуры глубокой сверточной нейронной сети, решающей задачу обнаружения объектов на аэрофотоснимках Земли.

Интегральная полнота обнаружения объектов при различных ОСШ может служить индикатором

истинной обобщающей способности нейронной сети и корректности выбранной модели.

Кроме того, предложен способ повышения помехоустойчивости системы путем ее обучения на частично зашумленном наборе данных. Показано, что чрезмерное искажение обучающих данных приводит к снижению интегральной полноты обнаружения объектов.

Также из данного исследования следует, что повышение качества работы модели путем обучения и оценки исключительно на неискаженных наборах данных не гарантирует, что такая модель даст соответствующий результат при работе с зашумленными входными изображениями. Таким образом, исследование зависимости качества работы сверточных нейронных сетей от уровня шумового воздействия позволяет раскрыть больше информации о корректности и качестве системы обнаружения объектов на аэрофотоснимках при работе в реальных условиях.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 18-72-10118).

СПИСОК ЛИТЕРАТУРЫ

1. *Hosang J., Benenson R., Dollár P., Schiele B.* // IEEE Trans. 2016. V. PAMI-38. № 4. P. 814.
2. *Сидорчук Д., Жижина Е.* // Информационные процессы. 2013. Т. 13. № 2. С. 171.
3. *Li H., Lin Z., Shen X. et al.* // 2015 IEEE Conf. Computer Vision and Pattern Recognition Boston (CVPR). 7–12 Jun. N.Y.: IEEE, 2015. P. 5325.
4. *Dalal N., Triggs B.* // 2005 IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR'05). San Diego. 20–25 Jun. N.Y.: IEEE, 2005. V. 1. P. 886.
5. *Ouyang W., Wang X.* // Proc. IEEE Int. Conf. on Computer Vision. Sydney. 1–8 Dec. 2013. N.Y.: IEEE, 2014. P. 2056.
6. *Strickland R.N., Hahn Hee Il,* // IEEE Trans. 1996. V. MI-15. № 2. P. 218.
7. *Lo S.-C.B., Lou S.-L.A., Lin Jyh-Shyan et al.* // IEEE Trans. 1995. V. MI-14. № 4. P. 711.
8. *Cheng G., Han J.* // ISPRS J. Photogrammetry and Remote Sensing. 2016. V. 117. P. 11.
9. *Long Y., Gong Y., Xiao Z., Liu Q.* // IEEE Trans. 2017. V. GRS-55. № 5. P. 2486.
10. *Chen X., Xiang S., Liu C.-L., Pan C.-H.* // IEEE Geoscience and Remote Sensing Lett. 2014. V. 11. № 10. P. 1797.
11. *Зиядинов В.В., Терешонок М.В.* // T-Comm. 2021. Т. 15. № 4. С. 49.
12. *Sakai K., Seo T., Fuse T.* // IEEE Intelligent Transportation Systems Conf. Auckland. 27–30 Oct. 2019. N.Y.: IEEE, 2019. P. 1776.
13. *Bejiga M.B., Zeggada A., Melgani F.* IEEE Int. Geoscience and Remote Sensing Symp. Beijing. 10–15 July 2016. N.Y.: IEEE, 2016. P. 693.
14. *Yang X., Sun H., Fu K. et al.* // Remote Sensing. 2018. V. 10. № 1. P. 132.
15. *Xia G., Bai X., Ding J. et al.* // IEEE/CVF Conf. Computer Vision and Pattern Recognition. Salt Lake City. 18–23 June 2018. N.Y.: IEEE, 2018. P. 3974.
16. *Cheng G., Zhou P., Han J.* // IEEE Trans. 2016. V. GRS-54. № 12. P. 7405.
17. *Razakarivony S., Jurie F.* // J. Visual Communication and Image Representation. 2016. V. 34. P. 187.
18. *Ferrer M.A., Vargas J.F., Morales A., Ordóñez A.* // IEEE Trans. 2012. V. IFS-7. № 3. P. 966.
19. *Jalalvand A., Demuynck K., De Neve W., Martens J.* // Neurocomputing. 2018. V. 277. P. 237.
20. *Girshick R., Donahue J., Darrell T., Malik J.* // IEEE Conf. Computer Vision and Pattern Recognition. Columbus. 23–28 June 2014. N.Y.: IEEE, 2014. P. 580.
21. *Lin T., Goyal P., Girshick R. et al.* // IEEE Intern. Conf. Computer Vision. Venice. 22–29 Oct. 2017. N.Y.: IEEE, 2017. P. 2380.
22. *Ren S., He K., Girshick R., Sun J.* // IEEE Trans. 2017. V. PAMI-39. № 6. P. 1137.
23. *Qian X., Lin S., Cheng G. et al.* // Remote Sens. 2020. V. 12. № 1. P. 143.
24. *Lin T., Dollár P., Girshick R. et al.* // IEEE Conf. on Computer Vision and Pattern Recognition. Honolulu. 21–26 July 2017. N.Y.: IEEE, 2017. P. 936.
25. *Padilla R., Passos W.L., Dias T.L.B. et al.* // Electronics. 2021. V. 10. P. 279.
26. *Durrett R.* Probability: Theory and Examples. Cambridge: Cambridge Univ. Press, 2019.

ИССЛЕДОВАНИЕ ЭЛЕКТРИЧЕСКИХ И СТРУКТУРНЫХ ПАРАМЕТРОВ ШУНГИТА ПО ДАННЫМ ЭЛЕКТРОСИЛОВОЙ СПЕКТРОСКОПИИ

© 2022 г. И. В. Антонец^а, Е. А. Голубев^б, В. Г. Шавров^с*, В. И. Щеглов^с**

^аСыктывкарский государственный университет им. П. Сорокина,
Октябрьский просп., 55, Сыктывкар, 167001 Российская Федерация

^бИнститут геологии Коми НЦ УРО РАН,

ул. Первомайская, 54, Сыктывкар, 167982 Российская Федерация

^сИнститут радиотехники и электроники РАН им. В.А. Котельникова,
ул. Моховая, 11, корп. 7, Москва, 125009 Российская Федерация

*E-mail: shavrov@cplire.ru

**E-mail: vshcheg@cplire.ru

Поступила в редакцию 11.12.2018 г.

После доработки 08.06.2021 г.

Принята к публикации 27.06.2021 г.

Методом электросилового спектроскопии исследованы пространственные и электрические параметры природного углеродсодержащего шунгита с различным содержанием углерода. Измерены вольт-амперные характеристики, по которым определена зависимость проводимости шунгита от содержания углерода. Показано, что общий ход полученной зависимости успешно коррелирует с подобными данными контактного и СВЧ-методов. Предложена модель структуры шунгита в виде углеродных гранул в кварцевой матрице. Показано, что наиболее близкому к эксперименту значению концентрации углерода около 40% соответствует модель смыкания кубических гранул по диагонали грани куба.

DOI: 10.31857/S0033849422020012

ВВЕДЕНИЕ

Важнейшей задачей техники является создание планарных структур, экранирующих электромагнитное излучение [1–3]. Перспективным материалом для таких структур является природный минерал шунгит, обладающий высокой экранирующей способностью [4]. Другим важнейшим применением шунгита является его использование в качестве составной части (анода) источников тока на основе лития, обладающих весьма высокими энергетическими возможностями в сочетании с крайне малыми габаритами [5, 6]. Главным преимуществом шунгита для перечисленных применений в сравнении с другими, в первую очередь синтетическими углеродсодержащими материалами, является его сравнительно низкая стоимость в сочетании с огромными природными запасами [7].

Шунгит представляет собой горную породу, занимающую промежуточное положение между графитом и антрацитом [8, 9]. За проводящую способность шунгита отвечают графеновые пачки, собранные в более крупные глобулы, далее образующие крупные проводящие слои между непроводящими включениями, состоящими из кварца [10, 11]. При этом интегральную проводимость

шунгита определяет количественное содержание проводящего углерода и его пространственное распределение между непроводящими областями.

В связи с перечисленными задачами важнейшим вопросом становится выяснение структурных и электрических параметров шунгита.

Связь между содержанием углерода и проводимостью шунгита исследовалась в ряде работ, частично обобщенных в [9]. Сильное влияние на проводимость графеносодержащей среды, в том числе в СВЧ-диапазоне, выявлено со стороны структурных особенностей и ориентации графеновых пачек [12, 13].

Можно отметить два уровня структуризации шунгита. Первый – “микроуровень” в масштабе нанометров, отражающий структуру графеновых пачек, собранных в глобулы и ленты, конгломерация которых образует более или менее однородные области, состоящие преимущественно из углерода [7, 14].

Второй – “макроуровень” в масштабе микрометров, где области с преобладающим содержанием углерода перемежаются с областями из кварца, в которых углерод практически отсутствует. Размеры таких областей превышают раз-

меры образований первого уровня на три-четыре порядка и более.

Структура и обусловленные ею электропроводящие свойства шунгита на первом уровне рассматриваются в значительном числе работ, из которых можно отметить, например, [7, 14]. Структура на втором уровне исследована недостаточно, в первую очередь из-за высокой степени хаотичности в пространственном распределении проводящих и непроводящих областей.

Типичным методом исследования структуры шунгита на втором уровне является растровая электронная микроскопия [14–16], позволяющая получить карты пространственного распределения углерода [8, 14]. Однако при этом остаются неизвестными интегральная удельная проводимость шунгита и ее связь с внутренним строением материала. Удельная проводимость достаточно хорошо измеряется контактным методом [17–19], однако связь ее с внутренней структурой шунгита остается неясной.

Перспективным методом одновременного решения обеих задач, т.е. исследования поверхности и измерения проводимости, является метод электросиловой спектроскопии. По-видимому, исторически первое применение этого метода для изучения шунгита отражено в работе [20], а дальнейшее его применение к графиту отмечено в работе [21]. Однако построение на основе этого метода достаточно подробной модели пространственного распределения углерода в шунгите отсутствует.

Цель данной работы – исследовать шунгит методом электросиловой спектроскопии, а также создать модель пространственного геометрического расположения его электропроводящих свойств.

1. МЕТОДИКА ИЗМЕРЕНИЙ И ОБРАЗЦЫ

Статическая проводимость в данной работе была исследована двумя методами. Первый – метод электросиловой спектроскопии с помощью атомно-силового микроскопа [20, 22]. Измерения проводили на сканирующем зондовом микроскопе Интегра Прима (NT-MDT, Зеленоград). Пространственное разрешение, ограничиваемое шумом регистрирующей системы, составляло 0.03 нм. Измерения проводили сканированием по осям x и y на площади 20×20 мкм, а вдоль оси z на 15 мкм.

Второй – контактный метод, позволяющий получить интегральное значение проводимости по довольно большой площади образца (до $1...2$ см²). Методика измерений полностью совпадает с описанной в работах [17–19].

Для полноты получаемых данных, кроме статической измерялась динамическая проводимость на СВЧ в диапазоне 28...40 GHz. Измерения выполняли по отражению и прохождению сигнала СВЧ с

Таблица 1. Содержание углерода в исследованных образцах

Номер образца	Содержание углерода, %	
	весовое	объемное
1	18	24
2	30	41
3	35	47
4	62	73
5	93	96
6	95	97
7	96	97
8	96	97
9	97	98
10	98	98

помощью спектрометра с генератором качающейся частоты ГКЧ-65 и индикатором ЯР-67. Методика измерений на СВЧ описана в работах [18, 19].

Образцы были выполнены в виде полированных пластин из шунгита толщиной 20...50 мкм, площадью до $2...3$ см², наклеенных на стеклянные пластинки толщиной 1 мм. Содержание углерода в образцах определялось методом кулонометрического титрования на анализаторе АН-7529 [9, 14, 20]. Весовое и объемное содержание углерода представлено в табл. 1.

Следует отметить, что для измерения статической проводимости достаточно объективным является только контактный метод, так как при методе силовой спектроскопии значительный вклад в сопротивление вносит контакт зонда микроскопа с поверхностью образца. Кроме того, контактный метод дает значения удельной проводимости, так как геометрические размеры образца составляют единицы миллиметров и легко измеримы. Метод электросиловой спектроскопии позволяет измерить абсолютную величину сопротивления или проводимости образца в точке размещения зонда, а получение удельной величины требует пересчета, затрудняемого некоторой неопределенностью размера контакта из-за сложности геометрического рельефа образца. В результате получаемые этим методом значения сопротивления, как правило, на три порядка превышают значения, определяемые контактным методом.

2. СОПРОТИВЛЕНИЕ И ПРОВОДИМОСТЬ ПО ВОЛЬТ-АМПЕРНЫМ ХАРАКТЕРИСТИКАМ

Для определения статической проводимости были использованы измеренные с помощью сило-

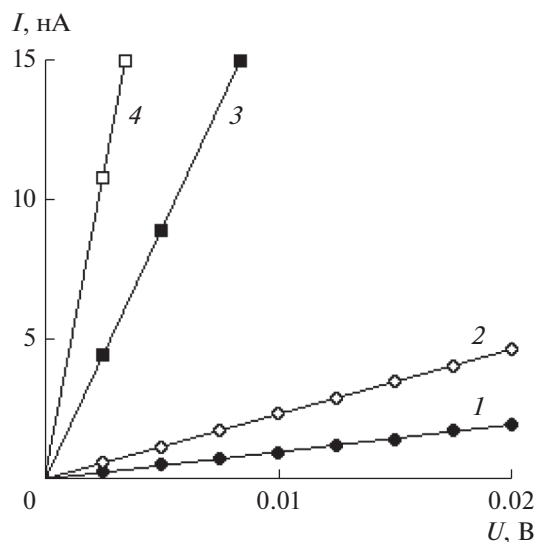


Рис. 1. Вольт-амперные характеристики для образцов шунгита с различным содержанием углерода: 1 – 18% (образец 1); 2 – 30% (образец 2); 3 – 62% (образец 4); 4 – 96% (образец 7).

вого микроскопа вольт-амперные характеристики (ВАХ), т.е. зависимости тока, протекающего через контакт, от приложенного к этому контакту напряжения.

На рис. 1 представлены ВАХ для четырех образцов шунгита (образцы 1, 2, 4 и 7) с различным содержанием углерода. Как видно, ВАХ имеют в основном линейный характер, причем их наклон при повышении концентрации углерода возрастает, что соответствует увеличению объемного содержания проводящего вещества относительно непроводящего.

Результаты измерения сопротивления и проводимости по ВАХ для образцов с различным содержанием углерода приведены в табл. 2. Из таблицы

Таблица 2. Сопротивление и проводимость образцов

Номер образца	V , В	J , нА	R , МОм	R^{-1} , МОм $^{-1}$
1	0.0188	2.22	9.400	0.106
2	0.0196	4.89	4.454	0.224
3	0.0032	10.22	0.348	2.874
4	0.0072	12.89	0.621	1.610
5	0.0124	10.22	1.348	0.742
6	0.0020	13.33	0.167	5.988
7	0.0032	13.33	0.267	3.745
8	0.0036	12.89	0.310	3.226
9	0.0016	14.22	0.125	8.000
10	0.0080	10.22	0.870	1.149

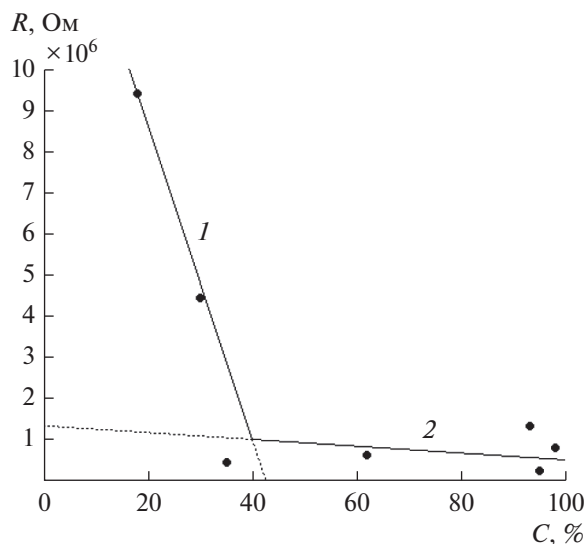


Рис. 2. Зависимость сопротивления образцов 1–6 и 10 от концентрации углерода: точки (каждая точка соответствует отдельному образцу) – эксперимент, кривые 1 и 2 – аппроксимация по формулам (1) и (2) соответственно. Результирующая кривая выделена сплошной линией.

видно, что, несмотря на монотонное увеличение содержания углерода в образце, сопротивление меняется не монотонно, а претерпевает некоторые скачки в ту или иную сторону. Однако все же тенденция уменьшения сопротивления по мере увеличения концентрации углерода прослеживается вполне очевидным образом.

Для большей наглядности зависимости сопротивления от концентрации углерода представлены на рис. 2, где показаны экспериментальные точки, построенные по табл. 2, а также аппроксимирующие их линии.

Из рисунка видно, что сопротивление по мере увеличения концентрации углерода в общем падает. При этом четко прослеживаются две ярко выраженные тенденции: при концентрациях менее 40% линия более крутая, при концентрациях более 40% – более пологая. В соответствии с этим аппроксимирующие линии 1 и 2, соответствующие концентрациям менее 40% и более 40%, построены по следующим формулам:

$$y_1 = K(-0.3570x + 15.36); \quad (1)$$

$$y_2 = K(-0.008x + 1.3), \quad (2)$$

где $K = 10^6$, x соответствует концентрации углерода C (%), y_1 , y_2 соответствуют R (Ом). Результирующая кривая, выделенная сплошной линией, имеет резкий излом вблизи концентрации 40%.

На рис. 3 темными точками представлены полученные по данным табл. 1 зависимости обратного сопротивления, т.е. проводимости образцов

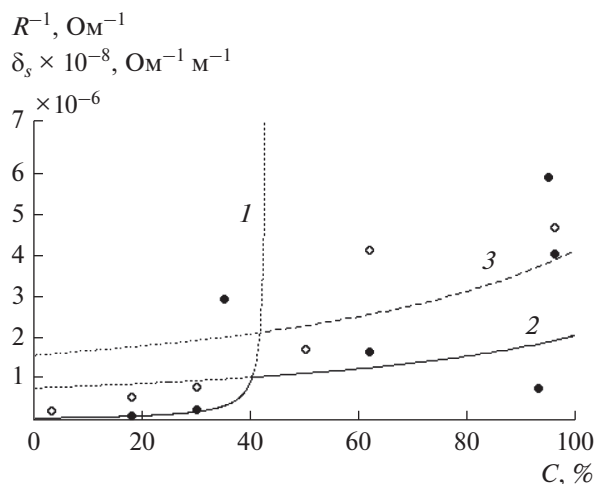


Рис. 3. Зависимость обратного сопротивления (проводимости) – среднее по всем данным для образцов 1–6, 10 (каждая точка соответствует отдельному образцу) от концентрации углерода (темные точки); светлые точки – значения статической проводимости, измеренные контактным методом; кривые 1 и 2 – аппроксимация по формулам (1) и (2) соответственно, кривая 3 построена по данным контактного метода по формуле (2). Результирующая кривая выделена сплошной линией.

от концентрации углерода. На том же рисунке светлыми точками показаны значения статической проводимости, измеренные контактным методом. Для приведения к единству масштаба по вертикальной шкале эти значения умножены на калибровочный коэффициент, равный 10^8 .

Из рисунка видим, что светлые точки довольно близко подходят к темным точкам (различие составляет не более двух-трех раз). То есть измерения по ВАХ вполне пригодны для оценки интегральной удельной проводимости образца в целом. Поэтому далее рассмотрим особенности именно этих данных.

Из расположения темных точек можно видеть, что зависимость претерпевает резкий подъем вблизи концентрации 40%, что соответствует выбросу вверх точки при $C = 35\%$. Второй подъем имеет место при приближении концентрации к 100%. Для аппроксимации экспериментальных точек были использованы обратные кривые, соответствующие формулам (1) и (2). Результирующая зависимость представлена сплошной линией.

С учетом сильного разброса точек вблизи концентрации 100% была построена дополнительная кривая 3, так чтобы она, оставаясь подобной кривой 2, проходила через середину расположения точек при концентрации 93...98%. Для построения такой кривой использовали удвоенную формулу кривой 2 (пунктир).

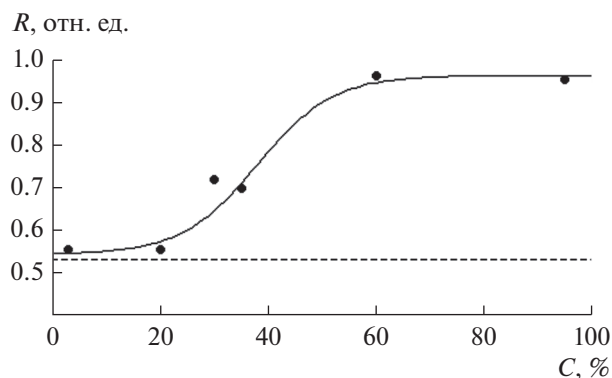


Рис. 4. Зависимость коэффициента отражения СВЧ-сигнала ($f = 27...38$ ГГц) от концентрации углерода: точки – эксперимент (каждая точка соответствует отдельному образцу 1–6), сплошная кривая – аппроксимация, пунктир – коэффициент отражения от стеклянной подложки ($R = 0.53$).

Из рисунка видно, что проводимость резко увеличивается вблизи концентрации 40%, а также слегка возрастает при стремлении концентрации к 100%. То есть можно полагать, что имеются два различных механизма роста проводимости с разной интенсивностью: сильный реализуется при $C = 40\%$ и слабый при $C = 100\%$.

Для дополнительной оценки влияния концентрации углерода на проводимость было исследовано отражение СВЧ-сигнала от исследованных образцов. Соответствующая зависимость коэффициента отражения R (по мощности) от концентрации углерода показана на рис. 4. Аппроксимирующая кривая построена по формуле

$$y = 0.21 \operatorname{th}(0.07x - 2.71) + 0.75, \quad (3)$$

где y соответствует R (отн. ед.).

Из рисунка видно, что за вычетом отражения от стекла, кривая имеет резкий перепад в окрестности концентрации $C = 38\%$. Это означает, что как раз вблизи такой концентрации проводимость резко возрастает.

Приведенные здесь данные относятся к динамической проводимости, которая может превышать статическую на порядок и более [23–25] за счет механизма гранулярных токов [26], проявляющего себя задолго до порога перколяции. То есть результирующая кривая несколько смещается в сторону меньшего содержания углерода. Согласно классическим представлениям [27], порог перколяции имеет место при концентрации где-то около 50%. Таким образом, результаты измерения коэффициента отражения можно считать дополнительным подтверждением факта резкого роста статической проводимости при концентрации углерода около 40%.

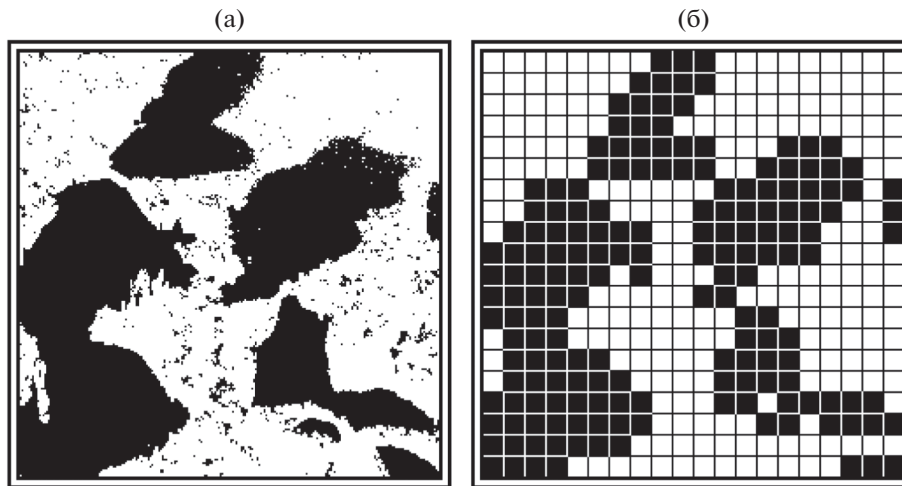


Рис. 5. Карта проводимости шунгита (а) и ее бинарное представление (б) для образца 3: белые области – углерод, черные – кварц (размер области – 20×20 мкм).

При сравнении рис. 2–4 видно, что резкий рост проводимости (или резкое падение сопротивления) приходится на $C = 40\%$, а перепад коэффициента отражения – на $C = 38\%$, т.е. оба резких перепада происходят при одной и той же концентрации углерода. Исходя из совпадения концентраций, можно полагать, что в обоих случаях работает один и тот же механизм. Рассмотрим возможную модель такого механизма на основе геометрического фактора.

В связи с высокой сложностью конфигурации проводящих и непроводящих областей шунгита представляет интерес найти сравнительно простой метод моделирования его структуры на макроуровне (рис. 5).

На рис. 5а представлена карта распределения углерода по поверхности шунгита, составленная по результатам измерения методом электросиловой спектроскопии для образца 3 (см. табл. 1). Весовое содержание углерода в образце составляет 35%, что соответствует объемному содержанию 47%, так что суммарное соотношение на рисунке площадей белых (проводящий углерод) и черных областей (непроводящий кварц) получается примерно равным.

На рис. 5б показано представление карты проводимости на бинарном уровне. Это представление получено путем наложения на исходную карту квадратной сетки с шагом в 1 мкм, после чего единичные квадраты сетки, в которых черный цвет преобладает над белым, окрашивались в черный цвет, а все остальные – в белый.

Из рисунка видно, что бинарное представление отражает пространственную картину распределения проводящего углерода по площади образца с точностью до размера ячейки сетки.

Теперь предположим, что структуру шунгита можно представить в виде совокупности ячеек, соответствующих ячейкам сетки (далее – гранулы). В работах [28, 29] представлены две возможные модели распределения гранул в шунгите, условно названные “кубики с перколяцией” и “песок с жидкостью”. Обе модели предполагают распределение гранул в узлах правильной кубической решетки. Естественно предположить, что проводимость материала в целом будет определяться возможностью смыкания гранул друг с другом.

3. ГЕОМЕТРИЧЕСКАЯ МОДЕЛЬ СМЫКАНИЯ ГРАНУЛ

Итак, предполагаем, что структура состоит из хорошо проводящих замкнутых областей (гранул), каким-то образом распределенных в плохо проводящей среде (матрице). Отношение проводимостей материала гранул и материала матрицы – несколько порядков.

При концентрации меньше 35...45% гранулы разобщены, так что общая проводимость малая – приближается к проводимости материала матрицы. Начиная с концентрации 35...45% гранулы смыкаются, образуются проводящие мостики, общая проводимость резко увеличивается. При концентрации значительно выше 45...55% все гранулы уже сомкнуты, проводимость структуры в целом приближается к проводимости материала гранулы и дальше не увеличивается. То есть эффект полностью аналогичен эффекту перколяции в гранулированных пленках. В качестве простейшей модели можно предположить работу геометрического фактора, т.е. формы гранул. Возможны два простейших варианта: гранулы сферические и гранулы кубические.

3.1. Гранулы сферические

В наиболее простом случае можно предположить, что гранулы имеют сферическую форму и распределены в матрице по узлам кубической решетки. При этом объем одной гранулы радиуса R равен

$$V_s = 4/3 \pi R^3, \quad (4)$$

а объем описанного вокруг нее куба –

$$V_{cs} = 8R^3. \quad (5)$$

Концентрация, при которой гранулы смыкаются, образуя мостики, равна:

$$C_s = V_s/V_{cs} = 0.523, \quad (6)$$

т.е. $C_s = 52\%$.

3.2. Гранулы кубические

Примем теперь, что гранулы имеют форму куба с ребром a , т.е. занимают объем

$$V_a = a^3. \quad (7)$$

Будем предполагать, что проводимость образца в целом определяется проводимостью прямолинейных цепочек гранул, проходящих от одной поверхности образца до другой. При этом будем полагать, что ориентация гранул в следующих друг за другом звеньях цепочки периодически повторяется через одну гранулу, так что достаточно рассмотреть формирование проводимости только одного звена цепочки, состоящего из пары следующих друг за другом двух гранул. Будем считать также, что у первой пары гранул ориентация такова, что плоскость ее куба перпендикулярна оси цепочки, а у второй пары ориентация может меняться. При этих условиях кубические гранулы могут смыкаться в трех следующих вариантах.

1. На плоскость грани одной гранулы приходится плоскость другой гранулы. При этом гранулы смыкаются вплотную, что может произойти только при концентрации $C_c = 100\%$.

2. На плоскость грани одной гранулы приходится ребро куба другой гранулы, т.е. гранулы сомкнуты по диагонали грани куба. При этом куб, описанный вокруг второй грани, имеет объем, определяемый длиной диагонали грани куба исходной гранулы, т.е.

$$V_{cg} = (\sqrt{2})^3 = 2.828a^3. \quad (8)$$

Концентрация при смыкании определяется отношением объема исходной гранулы к объему описанного куба:

$$C_{cg} = \frac{V_a}{V_{cg}} = \frac{a^3}{2.828a^3} = 0.354, \quad (9)$$

т.е. $C_{cg} = 35\%$.

3. На плоскость грани одной гранулы приходится вершина куба другой гранулы, т.е. гранулы сомкнуты по пространственной диагонали куба. При этом куб, описанный вокруг второй гранулы, имеет объем, определяемый длиной пространственной диагонали куба исходной гранулы, т.е.

$$V_{cp} = (\sqrt{3})^3 = 5.196a^3. \quad (10)$$

Концентрация при смыкании определяется отношением объема исходной гранулы к объему описанного куба:

$$C_{cp} = \frac{V_a}{V_{cp}} = \frac{a^3}{5.196a^3} = 0.192, \quad (11)$$

т.е. $C_{cp} = 19\%$.

3.3. Вывод по модели

Таким образом, модель сферических гранул обеспечивает скачок проводимости при концентрации 52%, модель кубических гранул при смыкании по грани куба – 100%, модель кубических гранул при смыкании по диагонали грани куба – 35%, модель кубических гранул при смыкании по пространственной диагонали куба – 19%. Из этих четырех вариантов наилучшим образом с экспериментальным значением 40% согласуется модель кубических гранул, смыкающихся по диагонали грани куба, которая обеспечивает 35%.

В общем же случае, учитывая близкий к хаотическому характер геометрии структуры шунгита, можно полагать, что в формировании его проводимости в той или иной степени принимают участие все четыре модели, однако модель смыкания по диагонали грани куба является преобладающей.

ЗАКЛЮЧЕНИЕ

Методом электросилового спектроскопии исследованы пространственные и электрические параметры природного углеродсодержащего материала шунгита. Измерены вольт-амперные характеристики для ряда образцов шунгита с различным содержанием углерода. Показано, что полученные характеристики имеют линейный характер, причем их наклон с повышением концентрации углерода возрастает. По измеренным характеристикам определено сопротивление контакта между зондом микроскопа и поверхностью образца. Путем сопоставления с данными измерения проводимости контактным методом показано, что значения величины обратного сопротивления контакта прямо пропорционально удельной проводимости образца в целом с калибровочным коэффициентом около трех порядков. Показано, что общий ход зависимости проводимости от концентрации успешно коррелирует с подобными зависи-

мостями, полученными по отражению сигнала СВЧ от образца.

На зависимости проводимости от содержания углерода обнаружен резкий подъем вблизи концентрации 40%, который интерпретирован на основе перколяции углеродных образований в структуре шунгита.

По аналогии с композиционной структурой, представляющей собой хорошо проводящие гранулы, заключенные в плохо проводящую матрицу, предложена модель структуры шунгита в виде углеродных гранул в кварцевой матрице. При рассмотрении различных геометрических моделей смыкания сферических и кубических гранул найдено, что наиболее близкому к эксперименту значению концентрации углерода около 40% соответствует модель смыкания кубических гранул по диагонали грани куба.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках государственного задания при частичной финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-55-53019 ГФЕН_а) а также Российского научного фонда (проект № 21-72-20048, № 21-47-00019) и Правительства Республики Коми (грант 20-42-110004, p_а).

СПИСОК ЛИТЕРАТУРЫ

1. Луцев Л.В., Николайчук Г.А., Петров В.В., Яковлев С.В. // Нанотехника. 2008. № 10. С. 37.
2. Казанцева Н.Е., Рывкина Н.Г., Чмутин И.А. // РЭ. 2003. Т. 48. № 2. С. 196.
3. Антонов А.С., Панина Л.В., Сарычев А.К. // ЖТФ. 1989. Т. 59. № 6. С. 88.
4. Лыньков Л.М., Махмуд М.Ш., Криштопова Е.А. // Вестн. Полоц. гос. ун-та. Сер. С. Фундаментальные науки. 2012. № 4. С. 103.
5. Bruce P.G., Scrosati B., Tarascon J.M. // Angew. Chem. Int. Ed. 2008. V. 47. P. 2930.
6. Tarascon J.M., Armand M. // Nature. 2001. V. 414. P. 359.
7. Chou N.H., Pierce N., Lei Yu. et al. // Carbon. 2018. V. 130. P. 105.
8. Филиппов М.М. Шунгитоносные породы Онежской структуры. Петрозаводск: Карельский НЦ РАН, 2002.
9. Соколов В.А., Калинин Ю.К., Дюккиев Е.Ф. Шунгиты – новое углеродистое сырье. Петрозаводск: Карелия, 1984.
10. Шека Е.Ф., Голубев Е.А. // ЖТФ. 2016. Т. 86. № 7. С. 74.
11. Голубев Е.А., Уляшев В.В., Велигжанин А.А. // Кристаллография. 2016. Т. 61. № 1. С. 74.
12. Голованов О.А., Макеева Г.С., Ринкевич А.Б. // ЖТФ. 2016. Т. 86. № 2. С. 119.
13. Макеева Г.С., Голованов О.А. Математическое моделирование электронноуправляемых устройств терагерцового диапазона на основе графена и углеродных нанотрубок. Пенза: Изд-во ПГУ, 2018.
14. Ковалевский В.В. Структура углеродного вещества и генезис шунгитовых пород. Дис. ... докт. геол.-мин. наук. Петрозаводск: Инст. геологии Карельского отд. РАН, 2007. 286 с.
15. Гоулдстейн Д., Яковиц Х. Практическая растровая электронная микроскопия. М.: Наука, 1978.
16. Электронная микроскопия в минералогии / Под ред. Г.-Р. Венка. М.: Мир, 1979.
17. Антонец И.В., Котов Л.Н., Некипелов С.В., Голубев Е.А. // ЖТФ. 2004. Т. 74. № 3. С. 24.
18. Антонец И.В., Котов Л.Н., Некипелов С.В. и др. // РЭ. 2004. Т. 49. № 10. С. 1243.
19. Антонец И.В., Котов Л.Н., Шавров В.Г., Щеглов В.И. // РЭ. 2006. Т. 51. № 12. С. 1481.
20. Голубев Е.А. // ФТТ. 2013. Т. 55. № 5. С. 995.
21. Shumilova T.G., Golubev Ye.A., Mayer J. et al. // Carbon. 2017. V. 114. P. 724.
22. Родионов В.В. Механизмы взаимодействия СВЧ-излучения с наноструктурированными углеродсодержащими материалами. Дис. ... канд. физ.-мат. наук. Курск: Курский гос. ун-т, 2014. 140 с.
23. Власов В.С., Котов Л.Н., Шавров В.Г., Щеглов В.И. // РЭ. 2014. Т. 59. № 9. С. 882.
24. Антонец И.В., Котов Л.Н., Калинин Ю.Е. и др. // ПЖТФ. 2014. Т. 40. № 14. С. 1.
25. Антонец И.В., Котов Л.Н., Кирпичева О.А. и др. // РЭ. 2015. Т. 60. № 8. С. 839.
26. Антонец И.В., Власов В.С., Котов Л.Н. и др. // Журн. радиоэлектроники. 2016. № 5. <http://jre.cplire.ru/jre/may16/10/text.pdf>.
27. Шкловский Б.И., Эфрос А.Л. Электронные свойства легированных полупроводников. М.: Наука, 1979.
28. Антонец И.В., Голубев Е.А., Шавров В.Г., Щеглов В.И. // Журн. радиоэлектроники. 2017. № 5. <http://jre.cplire.ru/jre/may17/11/text.pdf>.
29. Антонец И.В., Голубев Е.А., Шавров В.Г., Щеглов В.И. // Журн. радиоэлектроники. 2017. № 9. <http://jre.cplire.ru/jre/sep17/8/text.pdf>.

РАДИОФИЗИЧЕСКИЕ ЯВЛЕНИЯ
В ТВЕРДОМ ТЕЛЕ И ПЛАЗМЕ

УДК 621.382+621.391.822

АНОМАЛЬНЫЕ ТОКИ, ИНДУЦИРУЕМЫЕ ВО ВНЕШНЕЙ ЦЕПИ
ИЗМЕНЕНИЯМИ ПАРАМЕТРОВ ОБРАЗЦА

© 2022 г. С. Г. Дмитриев*

Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова РАН,
пл. Введенского, 1, Фрязино Московской обл., 141190 Российская Федерация

*E-mail: sgd@ms.ire.rssi.ru

Поступила в редакцию 26.08.2020 г.

После доработки 26.08.2020 г.

Принята к публикации 23.11.2020 г.

В развитие идей теоремы Шокли–Рамо проанализирована природа токов во внешней цепи, возникающих при изменении параметров образца. Выведена формула для токов аномальной природы, которые не сводятся к токам, индуцированным движением зарядов в образце, и токам емкостной природы.

DOI: 10.31857/S0033849422020024

ВВЕДЕНИЕ

Теорема Шокли–Рамо (ТШР) [1, 2] и ее обобщения [3–12] связывают движение зарядов между металлическими электродами в вакууме [1–6] или в диэлектрике [7–12] с индуцированными этим движением токами, втекающими в электроды из внешней цепи. Соответствующие формулы полезны при изучении процессов в вакуумных сверхвысокочастотных (СВЧ) приборах [1–6], в датчиках жесткого излучения [7–9, 13], в интегральных схемах и в структурах на основе металл–диэлектрик–полупроводник (МДП) [10–12], а также в других современных приборах. Формула для тока I_α во внешней цепи на отдельный α -й электрод (в системе с N электродами) имеет вид [1–4, 8]

$$I_\alpha = \iiint \vec{E}^{(\alpha)} \cdot \vec{j}_n dV. \quad (1)$$

Здесь $\vec{j}_n(t, \vec{r})$ – плотность полного тока в рассматриваемой системе, определяемая известной формулой

$$\vec{j}_n = \vec{j} + \partial \vec{D} / \partial t, \quad (2)$$

$\vec{j}(t, \vec{r})$ – плотность конвективного тока, $\vec{D}(t, \vec{r})$ – электрическая индукция, равная

$$D_i = P_i + \varepsilon_{ij} E_j, \quad (3)$$

где P_i – плотность поляризации, $\varepsilon_{ij}(t, \vec{r})$ – тензор диэлектрической проницаемости (по одинаковым тензорным индексам здесь и далее предполагается суммирование), $E_j(t, \vec{r})$ – электрическое

поле, $E_i^{(\alpha)}$ – вспомогательное нормированное электрическое поле

$$E_i^\alpha = E_i^{(1\alpha)} / \Phi_0, \quad (4)$$

$\vec{E}^{(1\alpha)} = -\text{grad}\varphi^{(1\alpha)}$ и $\varphi^{(1\alpha)}(t, \vec{r})$ – соответственно электрическое поле и потенциал из вспомогательной задачи для той же системы, но без пространственных зарядов и с потенциалами электродов $\Phi_\beta^{(1\alpha)} = 0$ при $\beta \neq \alpha$ и $\Phi_\alpha^{(1\alpha)} = \Phi_0 = 1$ В.

Интегрирование проводится по пространству без металлических электродов (влияние подводящих ток проводов предполагается незначительным, см. обсуждение вопроса в [12, 14]). Потенциальным должно быть только вспомогательное поле. Ширина области применимости теоремы видна уже из формулы (1). Поляризация в (3) может быть связана не только со структурной поляризацией твердого тела (как, например, в сегнетоэлектриках), но и с отдельными диполями (см., например, оценки проявления генерации дипольных дефектов в полупроводниковых структурах с диэлектриками в [15]) и с другими явлениями.

Обычно ТШР понимают в более узком смысле, удерживая в правой части (1) только конвективные токи:

$$I_{\alpha 1} = \iiint \vec{E}^{(\alpha)} \cdot \vec{j} dV \quad (5)$$

(где $I_{\alpha 1}$ – вклад во внешний ток на α -й электрод от конвективных токов в образце), или даже, как это делалось в первых работах [1, 2], рассматривая движение лишь одного точечного заряда. Отме-

тим, что ту же природу (и тот же вид) имеет также слагаемое в (1), связанное с изменением поляризации ($\partial P/\partial t$), так что и его следует отнести к токам той же природы, что и (5). Тогда

$$I_{\alpha 1} = \iiint \vec{E}^{(\alpha)} \cdot (\vec{j} + \partial \vec{P}/\partial t) dV. \quad (6)$$

Вклад же в (1) от второго слагаемого в (3) имеет другую природу. Во многих случаях токи, связанные с временными изменениями электрической индукции (токи смещения), сводятся к емкостным токам, формулы для которых хорошо известны в теории полупроводниковых приборов [16–18]. Возникает вопрос, не является ли в этих случаях обсуждаемый вклад чисто емкостным. Тогда полный ток в (1) будет складываться из емкостных токов и токов из (5), которые индуцируются конвективными токами. Обычно в полупроводниковых приборах при не слишком высоких частотах так оно и есть.

Отметим в этой связи, что при создании интегральных схем в послевоенные годы независимо рассматривался аналог ТШР для МДП-структур. Речь шла об определении конвективных токов в тонких диэлектрических пленках этих структур путем вычитания емкостных токов из полных измеряемых во внешней цепи токов (“метод вычитания”). Развитие этой высокочувствительной электрофизической методики было необходимо для диагностики рекордно низких концентраций подвижных ионов в пленках окислов структур металл–окисел–полупроводник (МОП) (см. обсуждение истории вопроса в [10, 18], некоторые детали эксперимента отмечены в [19]).

Тем не менее в рассматриваемом в ТШР общем случае при выделении из (1) токов емкостной природы все же остается дополнительное третье слагаемое [11, 12]. В работе [14] приведен поясняющий пример со структурой, в которой неоднородные изменения диэлектрической проницаемости (при отсутствиидвигающихся зарядов в образце) индуцируют во внешней цепи токи только неемкостного характера.

1. ЕМКОСТНЫЕ И АНОМАЛЬНЫЕ ТОКИ

Формула для тока на отдельный электрод после выделения емкостных токов приобретает (при $\epsilon_{ij} = \epsilon_{ji}$) следующий вид [11, 12]:

$$I_{\alpha} = I_{\alpha 1} + I_{\alpha 2} + I_{\alpha 3}. \quad (7)$$

Второе слагаемое здесь описывает емкостные токи:

$$I_{\alpha 2} = \sum_{\beta=1}^N \frac{\partial}{\partial t} (C_{\beta}^{\alpha} \Phi_{\beta}), \quad (8)$$

$$C_{\beta}^{\alpha} = Q_{\beta}^{(\alpha)} / \Phi_0, \quad (9)$$

где C_{β}^{α} – емкостные коэффициенты (в электростатике это коэффициенты емкости (C_{α}^{α}) и коэффициенты электростатической индукции (C_{β}^{α} , $\beta \neq \alpha$) [20]), а $Q_{\beta}^{(\alpha)}$ – заряд на β -м электроде во вспомогательной задаче в рассматриваемом случае. Поле здесь, конечно, должно быть потенциально. Отметим, что изменение потенциала системы на константу не меняет ток $I_{\alpha 2}$, как это видно из следующего равенства (вытекающего из электронейтральности системы во вспомогательной задаче):

$$\sum_{\beta=1}^N \partial C_{\beta}^{\alpha} / \partial t = (1/\Phi_0) \frac{\partial}{\partial t} \sum_{\beta=1}^N Q_{\beta}^{(\alpha)} = 0. \quad (10)$$

Отметим также, что кроме собственно емкостных токов в (8) включены и токи, связанные с изменением емкостных коэффициентов (см. [12]).

После выделения слагаемых указанной природы остается еще третье слагаемое, которое имеет вид

$$I_{\alpha 3} = - \iiint \frac{\partial E_i^{(\alpha)}}{\partial t} \epsilon_{ji} E_j dV. \quad (11)$$

Преобразуем его. Положим $\vec{P} = 0$ и рассмотрим, действуя в духе ТШР (и с теми же оговорками), объемный интеграл

$$J = \iiint \operatorname{div} \left(\frac{\partial \Phi^{(1)}}{\partial t} \vec{D} \right) dV \quad (12)$$

и выразим его, применяя теорему Остроградско–Гаусса, через поверхностный интеграл:

$$J = - \sum_{\beta=1}^N \frac{\partial \Phi_{\beta}^{(1)}}{\partial t} \iint_{S_{\beta}} \vec{D} \cdot \vec{n} dS. \quad (13)$$

Интегрирование здесь проводится по поверхностям S_{β} металлических электродов системы, \vec{n} – внешние нормали к ним, интеграл по бесконечности равен нулю вследствие электронейтральности системы.

Из (13) видно, если потенциалы электродов во вспомогательной задаче постоянны во времени ($\Phi_{\beta}^{(1)} = \text{const}$), то $J = 0$. Вместе с тем, используя уравнение

$$\operatorname{div} \vec{D} = \rho,$$

где ρ – плотность заряда (в основной задаче), можно записать подынтегральное выражение в (12) в виде

$$\operatorname{div} \left(\frac{\partial \Phi^{(1)}}{\partial t} \vec{D} \right) = - \frac{\partial \vec{E}^{(1)}}{\partial t} \vec{D} + \frac{\partial \Phi^{(1)}}{\partial t} \rho. \quad (14)$$

Отсюда для рассматриваемого случая следует

$$\iiint \frac{\partial \vec{E}^{(1\alpha)}}{\partial t} \cdot \vec{D} dV = \iiint \frac{\partial \Phi^{(1\alpha)}}{\partial t} \rho dV. \quad (15)$$

Поскольку левое слагаемое в (15) совпадает (при $\epsilon_{ij} = \epsilon_{ji}$) с интегралом в (11) с точностью до множителя $-(1/\Phi_0)$, то формулу для аномального тока $I_{\alpha 3}$ можно записать в окончательном виде:

$$I_{\alpha 3} = -\frac{1}{\Phi_0} \iiint \frac{\partial \varphi^{(1\alpha)}}{\partial t} \rho dV. \quad (16)$$

Рассмотрим теперь поясняющий пример из [14] с конденсатором, заполненным поровну двумя диэлектриками с относительными диэлектрическими проницаемостями ϵ_1 и ϵ_2 , и найдем ток на второй электрод (обкладка конденсатора, примыкающая ко второму диэлектрику). Емкость такого конденсатора равна

$$C = \frac{2\epsilon_0 S}{(\epsilon_1^{-1} + \epsilon_2^{-1})d}, \quad (17)$$

где S – его площадь, d – толщина, а ϵ_0 – диэлектрическая постоянная вакуума. Если диэлектрические проницаемости изменяются со временем, то условие постоянства емкости имеет вид

$$\partial(\epsilon_1^{-1} + \epsilon_2^{-1})/\partial t = 0. \quad (18)$$

Введем далее поверхностный заряд между диэлектриками с плотностью $-\sigma < 0$. Если этот заряд не движется, потенциалы на электродах не меняются и емкость тоже постоянна, то емкостные и индуцированные токи $I_{\alpha 2}$ и $I_{\alpha 1}$, втекающие в электроды (на обкладки конденсатора) равны нулю. Аномальный же ток из (16) для этой задачи отличен от нуля и описывается выражением

$$I_{\alpha 3} = \frac{\sigma S}{\Phi_\alpha} \frac{\partial \varphi^{(1\alpha)}(d/2)}{\partial t}, \quad (19)$$

где потенциал из вспомогательной задачи (в которой объемного заряда нет) взят в точке $x = d/2$ между диэлектриками (ось Ox направлена от первого электрода с потенциалом 0 ко второму с потенциалом Φ_0). Этот потенциал описывается, очевидно, формулой

$$\varphi^{(1\alpha)}(d/2) = \Phi_0 \epsilon_2 / (\epsilon_1 + \epsilon_2), \quad (20)$$

с помощью которой искомый ток, с учетом (18), можно записать в конечном виде:

$$I_{\alpha 3} = \frac{\sigma S}{\epsilon_1^{-1} + \epsilon_2^{-1}} \frac{\partial(1/\epsilon_1)}{\partial t}. \quad (21)$$

Из (21) видно, что если ϵ_1 растет (а тогда ϵ_2 , как это следует из (18), убывает), то $I_{\alpha 3} < 0$ и заряд на втором электроде убывает (на первом электроде заряд растет). Таким образом, аномальный ток в приведенном примере связан с перераспределением экранирующих зарядов на электродах при неоднородном изменении диэлектрической проницаемости образца (и неподвижных зарядах в образце).

2. ОБЩИЙ СЛУЧАЙ

В общем случае формула для токов, втекающих в электроды, после выделения в ней емкостных токов имеет вид [12]

$$\sum_{\beta=1}^N \Phi_\beta^{(1)} I_\beta = N \sum_{\beta=1}^N \frac{\partial}{\partial t} (Q_\beta^{(1)} \Phi_\beta) + \iiint \{E_i^{(1)}(j_i + \partial P_i \partial t) + E_i^{(1)} \frac{\partial}{\partial t} [(\epsilon_{ij} - \epsilon_{ji})E_j] - \frac{\partial E_i^{(1)}}{\partial t} \epsilon_{ji} E_j\} dV. \quad (22)$$

Заметим, что при наличии магнитного поля \vec{H} симметрия тензора диэлектрической проницаемости определяется равенством

$$\epsilon_{ij}(\omega, \vec{H}) = \epsilon_{ji}(\omega, -\vec{H}) \quad (23)$$

(ω – частота), которое следует из обобщенного принципа симметрии кинетических коэффициентов [20] и допускает асимметрию. Отметим также, что изменение потенциала на константу не изменяет равенство ввиду электронейтральности системы

$$\sum I_\alpha = 0 \quad (24)$$

(см. также (10)).

Для выделения аномальных токов удобно тождественно преобразовать выражение под интегралом в (22) к следующему виду:

$$\sum_{\beta=1}^N \Phi_\beta^{(1)} I_\beta = \sum_{\beta=1}^N \frac{\partial}{\partial t} (Q_\beta^{(1)} \Phi_\beta) + \iiint \{E_i^{(1)}(j_i + \partial P_i \partial t) + \frac{\partial}{\partial t} [E_i^{(1)}(\epsilon_{ij} - \epsilon_{ji})E_j] - \frac{\partial E_i^{(1)}}{\partial t} \epsilon_{ij} E_j\} dV. \quad (25)$$

Отсюда искомая формула, описывающая ток I_α , втекающий в отдельный (α -й) электрод, следует обычным образом (при $\Phi_\beta^{(1\alpha)} = 0$, если $\beta \neq \alpha$ и $\Phi_\alpha^{(1\alpha)} = \Phi_0 = 1$ В):

$$\Phi_0 I_\alpha = \sum_{\beta=1}^N \frac{\partial}{\partial t} (Q_\beta^{(1\alpha)} \Phi_\beta) + \iiint \{E_i^{(1\alpha)}(j_i + \partial P_i \partial t) + \frac{\partial}{\partial t} [E_i^{(1\alpha)}(\epsilon_{ij} - \epsilon_{ji})E_j] - \frac{\partial E_i^{(1\alpha)}}{\partial t} \epsilon_{ij} E_j\} dV \quad (26)$$

или

$$I_\alpha = \sum_{\beta=1}^N \frac{\partial}{\partial t} (C_\beta^\alpha \Phi_\beta) + \iiint \{E_i^{(\alpha)}(j_i + \partial P_i \partial t) + \frac{\partial}{\partial t} [E_i^{(\alpha)}(\epsilon_{ij} - \epsilon_{ji})E_j] - \frac{\partial E_i^{(\alpha)}}{\partial t} \epsilon_{ij} E_j\} dV. \quad (27)$$

От предыдущего выражения (22) формулы (26), (27) отличаются видом последних двух слагаемых

под интегралом. В частности, последнее слагаемое в (27) приняло теперь как раз тот вид

$$\iiint \frac{\partial E_i^{(\alpha)}}{\partial t} \epsilon_{ij} E_j dV = \iiint \frac{\partial E_i^{(\alpha)}}{\partial t} D_i dV, \quad (28)$$

который необходим для вывода (см. формулы (14)–(15)) формулы (16) для аномального тока $I_{\alpha 3}$.

Итак, ток I_{α} , втекающий в отдельный электрод, разбивается, как это видно из (27), на несколько компонент различной природы

$$I_{\alpha} = I_{\alpha 1} + I_{\alpha 2} + I_{\alpha 3} + I_{\alpha 4}, \quad (29)$$

где токи $I_{\alpha 1}$, $I_{\alpha 2}$ и $I_{\alpha 3}$ описываются соответственно формулами (6), (8) и (11), (16), а ток $I_{\alpha 4}$ равен

$$I_{\alpha 4} = \iiint \frac{\partial}{\partial t} [E_i^{(\alpha)} (\epsilon_{ij} - \epsilon_{ji}) E_j] dV \quad (30)$$

и имеет теперь более симметричный, чем в (22), вид и более полно описывает влияние несимметричности тензора диэлектрической проницаемости.

ЗАКЛЮЧЕНИЕ

В развитие идей ТШР показано, что ток из внешней цепи, втекающий в отдельный металлический электрод, включает в себя несколько компонент различной природы, которые описываются формулами (6), (8), (11), (16) и (30). Формула (6) описывает вклад в полный ток, индуцированный движением зарядов в образце (собственно ТШР); формула (8) – вклад от токов емкостной природы; формула (30) – вклад, связанный с асимметрией тензора диэлектрической проницаемости; наконец, формула (16) соответствует еще одному – аномальному – вкладу, который не сводится к предыдущим. Приведен простой поясняющий пример, в котором все вклады, кроме аномального, равны нулю. Отмечено, что ток в этом случае связан с перераспределением экранирующих зарядов на электродах при определенном неоднородном изменении диэлектрической проницаемости

образца и наличии в нем неподвижного заряда. Таким образом, и неподвижный заряд в образце может, в определенных условиях, индуцировать ток во внешней цепи. В этом смысле обсуждаемый механизм дополняет вклад из ТШР.

Работа выполнена по госзаданию.

СПИСОК ЛИТЕРАТУРЫ

1. *Shockley W.* // J. Appl. Phys. 1938. V. 9. № 10. P. 635.
2. *Ramo S.* // Proc. IRE. 1939. V. 27. № 9. P. 584.
3. *Beck A.H.W.* Thermionic Valves: Their Theory and Design. Cambridge: Cambridge Univ. Press, 1953.
4. *Jen C.K.* // Proc. IRE. 1941. V. 29. № 6. P. 345.
5. *Gabor D.* // J. IEE. 1944. V. 91. Pt. 3. № 15. P. 128.
6. *Гвоздовер С., Лопухин В.* // Изв. АН СССР. Сер. физ. 1946. Т. 10. № 1. С. 29.
7. *Cavalleri G., Fabri G., Gatti E., Svelto V.* // Nuclear Instruments and Methods. 1963. V. 21. P. 177.
8. *Cavalleri G., Gatti E., Fabri G., Svelto V.* // Nuclear Instruments and Methods. 1971. V. 92. № 1. P. 137.
9. *He Z.* // Nuclear Instruments and Methods in Physics Research A. 2001. V. 463. № 1–2. P. 250.
10. *Дмитриев С.Г.* // РЭ. 2012. Т. 57. № 11. С. 1229.
11. *Дмитриев С.Г.* // РЭ. 2018. Т. 63. № 10. С. 1115.
12. *Дмитриев С.Г.* // РЭ. 2019. Т. 64. № 9. С. 926.
13. *Tavernier S.* Experimental Techniques in Nuclear and Particle Physics. L: Springer, 2010.
14. *Дмитриев С.Г.* // РЭ. 2020. Т. 65. № 7. С. 725.
15. *Дмитриев С.Г.* // ФТП. 2009. Т. 43. № 6. С. 854.
16. *Бонч-Бруевич В.Л., Калашников С.Г.* Физика полупроводников. М.: Наука, 1990.
17. *Зи С.* Физика полупроводниковых приборов. М.: Мир, 1984.
18. *Nicollian E.R., Brews J.R.* MOS (Metal-Oxide-Semiconductor) Physics and Technology. N.Y.: J. Wiley & Sons, 1982.
19. *Дмитриев С.Г.* // ФТП. 2011. Т. 45. № 2. С. 192.
20. *Ландау Л.Д., Лифшиц Е.М.* Электродинамика сплошных сред. М.: Физматлит, 2005.

ПРИМЕНЕНИЕ РАДИОТЕХНИКИ И ЭЛЕКТРОНИКИ В БИОЛОГИИ И МЕДИЦИНЕ

УДК 57.087.1+004.383.3

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ХАРАКТЕРИСТИК ПУЛЬСОВОЙ ВОЛНЫ МЕТОДАМИ СТРУКТУРНОЙ ДЕКОМПОЗИЦИИ СИГНАЛОВ ФОТОПЛЕТИЗМОГРАФИИ

© 2022 г. Л. В. Лабунец^{a, b, *}, А. Б. Борзов^a, Н. Ю. Макарова^a

^a Московский государственный технический университет им. Н.Э. Баумана,
2-я Бауманская ул., 5, Москва, 105005 Российская Федерация

^b Российский новый университет,
ул. Радио, 22, Москва, 105005 Российская Федерация

*E-mail: labunets@bmstu.ru

Поступила в редакцию 22.06.2021 г.

После доработки 22.06.2021 г.

Принята к публикации 25.08.2021 г.

Предложена методика выделения пульсовой волны методами структурной декомпозиции временных рядов дистанционной фотоплетизмографии (ДФПГ). Структурная декомпозиция сигналов ДФПГ на волновые процессы предусматривает естественный симбиоз методов кратномасштабного, сингулярного и цифрового спектрального анализа нестационарных процессов. С помощью преобразования Гильберта получены адаптивные оценки мгновенных частот сердечных сокращений и их корреляционно-спектральных характеристик. По результатам анализа различных алгоритмов формирования информативных компонент пульсовой волны верифицированы показатели вариационной пульсометрии. Обоснована эффективность структурной декомпозиции временных рядов при исследовании частоты сердечных сокращений методом ДФПГ.

DOI: 10.31857/S0033849422020097

ВВЕДЕНИЕ

Волновые процессы, порождаемые основными системами жизнедеятельности человека, содержат важную информацию о его физиологическом состоянии. В медицинской практике применяют широкий спектр контактных и неконтактных методов диагностики заболеваний пациентов, основанных на различных физических принципах измерения пульсовых, дыхательных, миогенных, нейрогенных и эндотелиальных волн [1]. Эти волновые процессы исследуют с помощью датчиков, функционирующих в различных диапазонах спектра электромагнитных волн [2], методами лазерной доплеровской флоуметрии (ЛДФ) [1], электрокардиографии (ЭКГ) [3], электроэнцефалографии (ЭЭГ) [4], сфигмографии (СФГ) [5]. В этом далеко не полном списке аппаратных методологий исследования биоритмов человека не последнее место занимает дистанционная фотоплетизмография (ДФПГ) [6, 7].

ДФПГ – один из эффективных методов неконтактного мониторинга кровенаполнения сердечно-сосудистой системы человека, например в кардиологии при исследовании variability сердечного ритма (ВСР). Цветное изображение лица испытуемого, регистрируемого RGB-видеокаме-

рой в условиях комнатного или специального освещения в спектральном диапазоне 0.53...0.94 мкм, содержит важную информацию о распространении по телу пульсовой волны. Колебания количества крови в сосудах кожных покровов модулируют поглощение света, проходящего через соответствующий объем ткани испытуемого.

Стандартная методика ДФПГ [6] предусматривает усреднение текущих RGB-видеоизображений по множеству пикселей из так называемой области интереса (Region of Interest – ROI) лица человека. В результате такой предобработки формируют три временных ряда (ВР), соответствующих спектральным диапазонам видеокамеры – “красному” R, “зеленому” G и “синему” B. Динамика этих ВР содержит, как правило, шумы измерений и вычислений, помехи, трендовые и квазициклические компоненты. Помехи в динамике RGB-временных рядов обусловлены условиями освещения и характеристиками отражения – поглощения света кожей испытуемого, а также его движениями или наличием у него физических нагрузок [7]. Соответствующая цифровая постобработка позволяет идентифицировать квазициклы и, в конечном итоге, пульсовую волну.

Современные алгоритмы ДФПГ реализуют сочетание двух подходов к формированию информативных признаков пульсовой волны [8], а именно: фотометрические, индуцированные характеристиками отражения – поглощения света кожей покровами, и априорно известным частотным диапазоном пульсовой волны от 0.667 до 4 Гц, что соответствует 40...240 ударам сердца в минуту. Важно отметить, что границы частотных диапазонов волновых процессов, формируемых системами жизнедеятельности человека, зависят от аппаратных функций этих систем и физических методов измерения волн. В частности, для ЛДФ диагностическим спектральным диапазоном пульсовой волны, порождаемой ритмами сердечно-сосудистой и дыхательной систем, являются частоты от 0.8 до 1.6 Гц [1, с. 16], что соответствует 48...96 ударам сердца в минуту.

Цифровой спектральный анализ сигналов, синхронно измеренных согласно канонам тибетской медицины для шести точек лучевых артерий обеих рук испытуемого [5, 9], убедительно показывает, что 95% энергии пульсовой волны локализованы в частотном диапазоне 0.6...8 Гц. Амплитудно-частотная характеристика фурье-спектра квазипериодического пульсового сигнала, измеренного указанным выше методом, демонстрирует явно выраженную кластерную структуру в виде основной и последующих, как правило, трех локальных мод [9, с. 61]. Частоты мод примерно кратны основной частоте f_1 кардиоритма испытуемого. Для условно здорового человека $f_1 \approx 1$ Гц, а парциальные спектры пульсового сигнала локализованы в диапазонах 0.5...1.5 Гц для основной моды и 1.5...2.5, 2.5...3.5, 3.5...4.5 Гц для последующих мод.

Волновые процессы в системах жизнедеятельности человека демонстрируют нестационарность в широком смысле корреляционно-спектральных характеристик и наличие нелинейных динамических эффектов. Закономерности, скрытые в нестационарности и нелинейности биоритмов человека, обладают несомненной клинической интерпретацией. Фундаментальную методическую основу теоретического анализа такого рода закономерностей, наряду с теорией нелинейных динамических систем [10], составляют методы, модели и алгоритмы структурной декомпозиции сигналов ЛДФ, ЭКГ, ЭЭГ, СФГ.

Частотно-временное исследование локальных особенностей нестационарных ВР биоритмов человека реализуют с помощью непрерывного вейвлет-преобразования [6, 11] или кратномасштабного анализа (КМА) в базе дискретного вейвлет-преобразования [12, 13], а также на основе различных модификаций эмпирической модовой декомпозиции (Empirical Mode Decomposition – EMD) и преобразования Гильберта–Хуанга (Hilbert–Huang Transformation – ННТ) [2, 4, 9, 14, 15]. К этому

списку следует добавить метод “Гусеница” сингулярного спектрального анализа (Singular Spectral Analysis – SSA) структуры закономерностей, скрытых в динамике биомедицинских сигналов [16]. Также отметим рациональную тенденцию к комбинированному применению указанных выше методов структурной декомпозиции нестационарных ВР [17].

Цель данной статьи – рассмотреть содержание последовательных этапов интеллектуальной цифровой обработки сигналов ДФПГ, основанных на сочетании дополняющих друг друга методов кратномасштабного, сингулярного и корреляционно-спектрального анализа для формирования робастных оценок информативных характеристик пульсовой волны.

1. ЭКСПЕРИМЕНТАЛЬНОЕ ИЗМЕРЕНИЕ И ПРЕДОБРАБОТКА СИГНАЛОВ ДФПГ

Измерение сигналов ДФПГ выполнялось в экспериментальной лаборатории научно – образовательного медико-технологического центра МГТУ им. Н.Э. Баумана (МТЦ МГТУ). Информационно-измерительный комплекс содержит панель дневного освещения со спектром от 0.53 до 0.94 мкм, RGB-видеокамеру Logitech HD Webcam C270, сопряженную с ноутбуком через высокоскоростное сертифицированное соединение USB 2.0, и стационарное место расположения испытуемого. Камера формирует последовательность видеоизображений с частотой 30 кадров в секунду размером 640×480 пикселей с глубиной цвета 24 бита. Текущий i, j -й пиксель представляет собой $r_{i,j}(n), g_{i,j}(n), b_{i,j}(n)$ цветовые компоненты с восьмидесятибитовой точностью.

В наших исследованиях дополнительно анализировались результаты измерений ДФПГ базы данных (БД), записанной добровольцами технологического университета Эйденховен (Нидерланды) [18]. БД содержит набор видеозаписей, а также опорные показатели ВСП, измеренные одновременно с фотоплетизмограммой, что является несомненным преимуществом. Имеются видеозаписи области лица трех испытуемых с различным цветом кожи. Для каждого участника эксперимента имеются записи после физической нагрузки, набор записей в спокойном состоянии, но при различной интенсивности освещения, а также записи с поворотами головы.

Область лица испытуемого (так называемую область интереса) ROI формируют методом Виолы–Джонса. Отсчеты $R(n), G(n), B(n)$ ВР рассчитывают усреднением по i, j -м пикселям, принад-

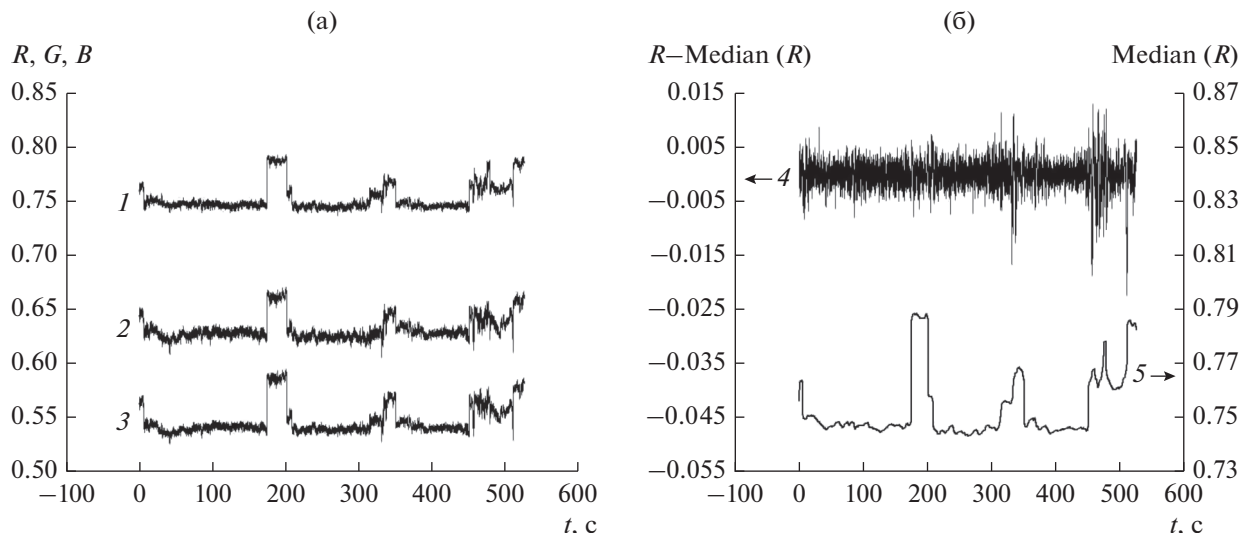


Рис. 1. Сигналы ДФПГ (а) и компоненты сигнала R (б): кривая 1 – R, 2 – G, 3 – B; 4 – очищенный, 5 – оценка выбросов.

лежащим ROI кожного покрова лица в соответствии с методикой, представленной в [6]:

$$\begin{aligned}
 & [R(n)G(n)B(n)]^T = \\
 & = \frac{1}{ROI} \sum_{i,j \in ROI} [r_{i,j}(n)g_{i,j}(n)b_{i,j}(n)]^T. \quad (1)
 \end{aligned}$$

Эффективное по критерию вычислительных затрат удаление выбросов и артефактов RGB-сигналов (рис. 1а), измеренных в МТЦ МГТУ, реализует, по нашему мнению, процедура медианной фильтрации $\text{median}\{R(n)\}$, для R-сигнала (рис. 1б, кривая 5). Апerture фильтра N согласована с интервалом временной дискретизации ВР (1) $\Delta t = 1/30$ с и нижней границей $f_{\text{мин}} = 0.15$ Гц частотного диапазона дыхательной волны [1, с. 16] в соответствии с формулой

$$N = 1 / (\Delta t f_{\text{мин}}) = 201.$$

Второй надежный, но более затратный алгоритм, основан на обнаружении классов типичных значений ВР (1) в двумерном пространстве цветового сигнала и его первой разности (рис. 2а) с помощью алгоритма кластеризации DBSCAN [19], основанного на понятии связности по плотности данных в метрике Евклида. Результаты идентификации аномальных перепадов $g(t)$ текущего уровня R-сигнала представлены на рис. 2б. Параметры DBSCAN представляют собой радиус окружности ближайших соседей $e = 0.002$ и количество ближайших соседей в этой окрестности $\text{minPts} = 10$.

Очищенный сигнал $\tilde{R}(t) = R(t) - g(t)$ формируют в три этапа. На первом шаге DBSCAN маркирует аномальные значения на двумерной диаграмме рассеяния $\{R(t), R(t) - R(t-\Delta t)\}$ (см. рис. 2а). На втором шаге из первой разности $R(t) - R(t - \Delta t)$ исключают

маркированные выбросы с последующей их заменой линейно интерполированными оценками. На третьем шаге интегрируют очищенную первую разность. Отметим, что в отличие от медианной фильтрации динамика цветовых сигналов, очищенных от аномальных перепадов указанным выше способом (см. рис. 2б, кривая 3), сохраняет низкочастотные волновые процессы.

Сигналы цветовых каналов ДФПГ, представленные в БД Технологического университета Эйнховена [18], не содержат артефакты в виде перепадов, что свидетельствует о более высоком качестве методики экспериментальных измерений.

Важным этапом последующей предобработки является удаление шумов, обусловленных ошибками измерений и вычислений. Универсальный подход к идентификации шумов очищенных сигналов $\tilde{R}(t), \tilde{G}(t), \tilde{B}(t)$ основан на применении КМА в базе дискретного вейвлет-преобразования. По определению, шумы – это высокочастотные, практически некоррелированные детализирующие компоненты:

$$\begin{aligned}
 d_l(n, m) &= d_l(n\Delta t, m) = \\
 &= \sqrt{K_m} \sum_{k=0}^{K_m-1} D_l(k, m) \psi(K_m n - k), \quad (2) \\
 K_m &= 2^{M-m},
 \end{aligned}$$

КМА первого $m = 1$ или второго $m = 2$ уровней разложения. Здесь M – количество уровней разложения ВР, согласованное с объемом выборки $\geq 2^M$; n – дискретное время, $\psi(n) = \psi(n\Delta t)$ – вейвлет-функция; $l = 1, 2, 3$ – номер цветового сигнала, соответствующий последовательности R, G, B. В наших вычислительных экспериментах в качестве базиса применялись вейвлеты Добеши по-

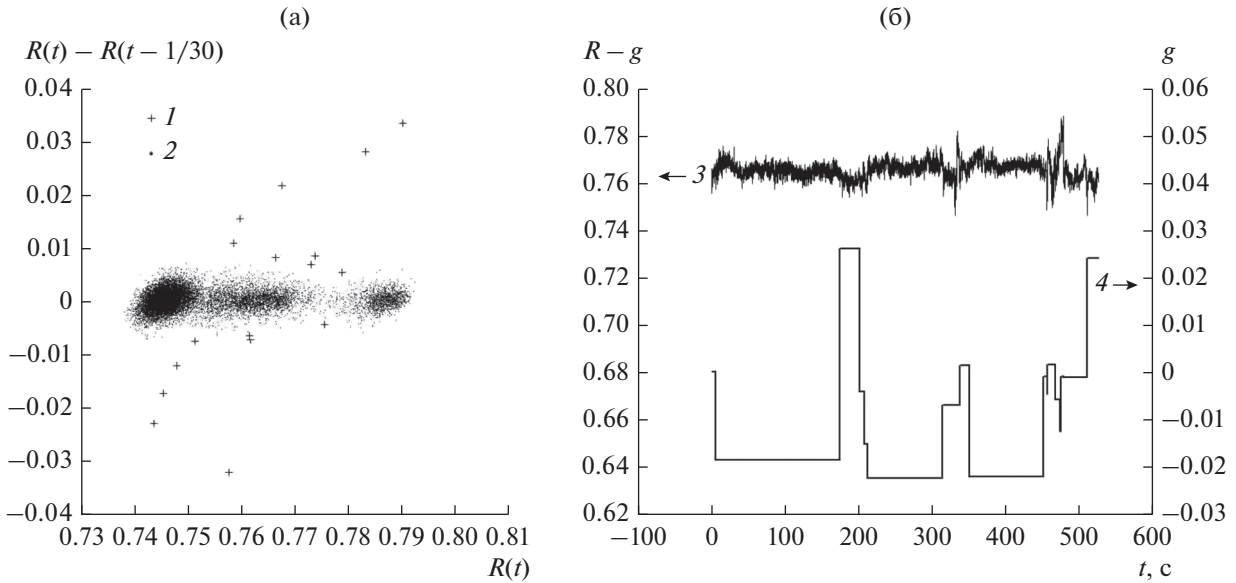


Рис. 2. Кластерная структура R -сигнала ДФПГ (а) и его компоненты (б): 1 – выбросы; 2 – типичные значения; 3 – очищенный, 4 – оценка выбросов.

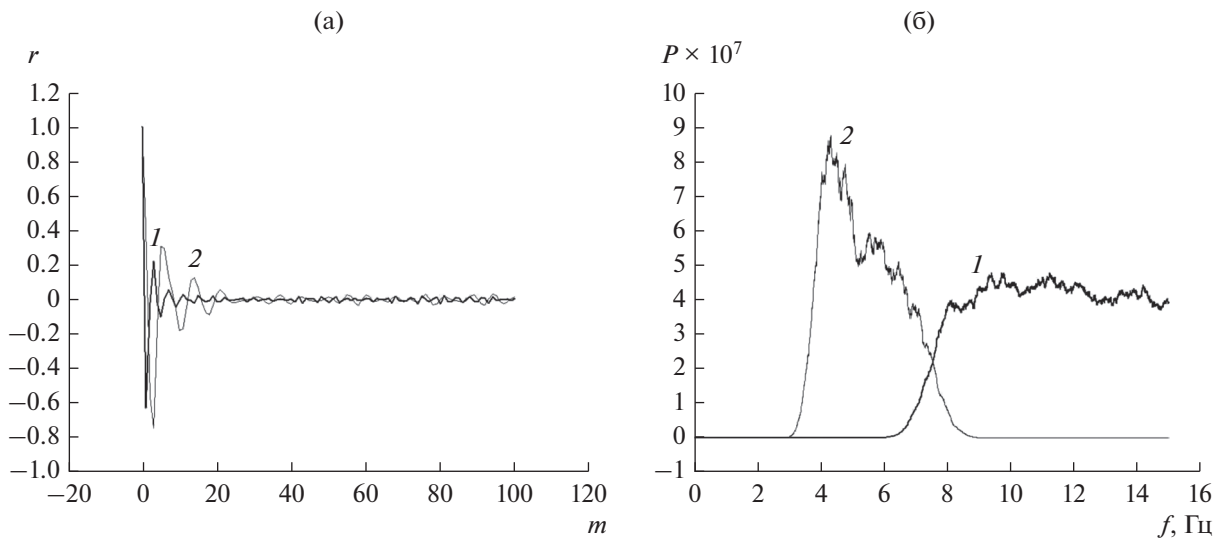


Рис. 3. Корреляционно-спектральные оценки компонент $d_1(n,1)$ (кривая 1) и $d_2(n,1)$ (кривая 2) КМА сигнала $\tilde{R}(n\Delta t)$: АКФ (а); периодограммы (б).

рядка 40 и $M = 10$, т.е. наименьшее количество отсчетов ВР составляет $2^{10} = 1024$, что соответствует длительности записи $1024/30 \approx 34$ с. Такой выбор параметров КМА обеспечил достаточно детальную субполосную фильтрацию и минимизировал эффект перекрытия спектров детализирующих компонент, т.е. обеспечил их приемлемую ортогональность.

На рис. 3а представлены выборочные оценки автокорреляционных функций (АКФ) $r_1(m\Delta t)$ и

$r_2(m\Delta t)$ соответственно первой $d_1(n,1)$ и второй $d_1(n,2)$ детализирующих компонент КМА для сигнала $\tilde{R}(n\Delta t)$. Периодограммы Томсона $P_1(f)$ и $P_2(f)$ с параметрами сглаживания 145 и 89 для указанных компонент приведены на рис. 3б. Корреляционно-спектральные оценки получены для всей реализации ВР, содержащего 15756 отсчетов (≈ 525 с). Первая компонента $d_1(n,1)$ КМА сигнала $\tilde{R}(n\Delta t)$ демонстрирует достаточно низкий уровень корреляции и относительно равномерное распре-

деление мощности гармоник для частот $f \geq 6.5$ Гц. Аналогичный характер поведения корреляционно-спектральных характеристик демонстрируют первые компоненты $d_2(n,1)$ и $d_3(n,1)$ КМА сигналов $\tilde{G}(n\Delta t)$ и $\tilde{B}(n\Delta t)$, измеренных в МТЦ МГТУ, а также ВР из БД [18].

2. ВОЛНОВЫЕ И ТРЕНДОВЫЕ СТРУКТУРНЫЕ КОМПОНЕНТЫ СИГНАЛОВ ДФПГ

Волновые и трендовые, компоненты сигналов цветочных каналов формируют обычно на этапе постобработки [6]. Достаточно большое количество оценок биоритмов нестационарных биомедицинских ВР делят, как правило, на три класса. Первый основан на процедурах синтеза адаптивных цифровых фильтров, таких как, например, НР-фильтр или родственный ему SPA-фильтр. Параметры таких фильтров обеспечивают компромисс между ошибками аппроксимации ВР и критерием простоты модели сглаживания данных. Второй класс оценок, развивающий принцип регуляризации, реализует модели локально взвешенной полиномиальной регрессии Савицкого–Голея, Мешалкина, Клевеланда. Важным достоинством третьего класса оценок, основанных на методологии структурной декомпозиции, является частотно-временной анализ локальных особенностей динамики биоритмов. По этой причине мы отдаем предпочтение методам и моделям структурной декомпозиции.

Цифровой спектральный анализ составляющих КМА обеспечивает идентификацию детализирующих компонент (2), формирующих динамику пульсовой волны. В частности, периодограммы Томсона (рис. 4) третьей $d_l(n,3)$ и четвертой $d_l(n,4)$ ($l = 1, 2, 3$) компонент сигналов цветочных каналов, полученных в МТЦ МГТУ, всецело принадлежат диапазону частот $0.667 \dots 4$ Гц пульсовой волны.

Энергетические спектры $P(f)$ второй $d_l(n,2)$ и пятой $d_l(n,5)$ компонент лишь частично принадлежат указанному выше интервалу частот, а именно “медленная” часть $d_l(n,2)$ и “быстрая” часть $d_l(n,5)$ содержат гармоники с частотами $f \leq 4$ Гц и $f \geq 0.667$ Гц соответственно. Парциальные части реализаций этих компонент, спектральные плотности мощности (СПМ) которых принадлежат указанному диапазону частот, удобно формировать методом аналитической модовой декомпозиции [20]. В соответствии со свойствами дискретного преобразования Гильберта H :

$$h_s(n) = H[d_l(n,m) \sin(2\pi f_b n \Delta t)],$$

$$h_c(n) = H[d_l(n,m) \cos(2\pi f_b n \Delta t)],$$

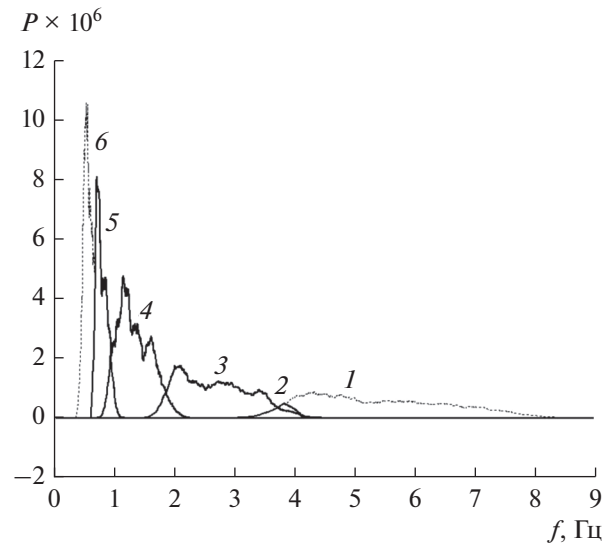


Рис. 4. Периодограммы детализирующих компонент КМА для R-канала: кривая 1 – $d_l(n,2)$; кривая 2 – $d_l^s(n,2)$; кривая 3 – $d_l(n,3)$; кривая 4 – $d_l(n,4)$; кривая 5 – $d_l^f(n,5)$; кривая 6 – $d_l(n,5)$.

для компоненты $d_l(n,m)$ несложно получить “медленную” часть:

$$d_l^s(n,m) = h_c(n) \sin(2\pi f_b n \Delta t) - h_s(n) \cos(2\pi f_b n \Delta t),$$

СПМ которой содержит частоты $f \leq f_b$. Очевидно, что “быстрая” часть $d_l(n,m)$ с частотами СПМ $f \geq f_b$ имеет вид

$$d_l^f(n,m) = d_l(n,m) - d_l^s(n,m).$$

Ясно, что для второй компоненты разделяющая частота равна $f_b = 4$ Гц, а для пятой компоненты $f_b = 0.667$ Гц.

Применяемый нами в данной работе КМА (2), позволил получить адекватные экспертные модели трендовых составляющих цветочных каналов ДФПГ в виде суммы четырех компонент: восьмой $d_l(n,8)$, девятой $d_l(n,9)$, десятой $d_l(n,10)$ детализирующих и десятой $a_l(n,10)$ аппроксимирующей. Реализации нестационарных ВР цветочных каналов, маркированных именем Р1Н1 в БД [18], а также структурные оценки их трендов представлены на рис. 5 (файл данных Р1Н1 содержит описание условий и результаты измерения сигналов ДФПГ испытуемого после физических упражнений на велоэргометре). Аналогичные оценки трендов получены для сигналов цветочных каналов, измеренных в МТЦ МГТУ.

Предварительные центрированные оценки RGB-пульсовых волн формируют в соответствии с предложенной в [21] процедурой масштабиро-

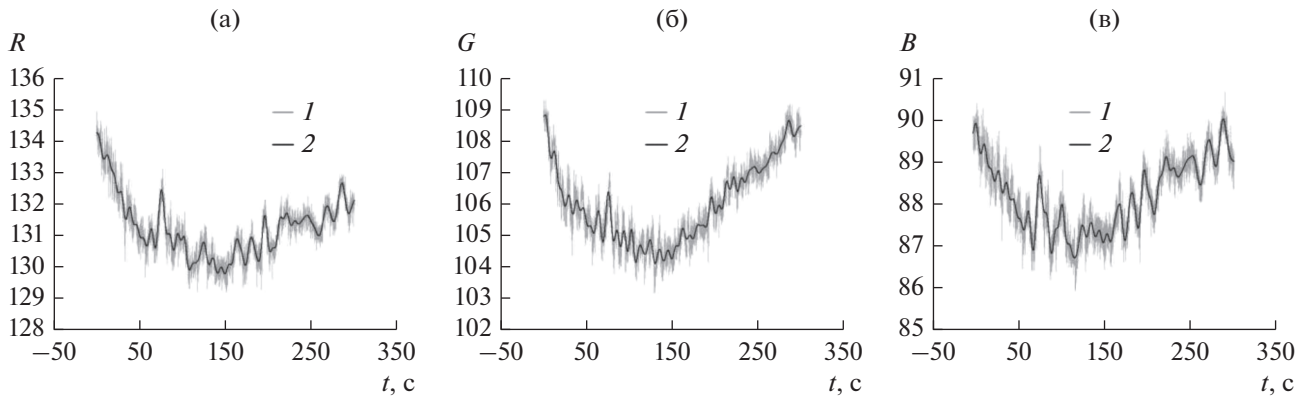


Рис. 5. Временные ряды RGB- каналов (1) и их тренды (2) для сигналов R (а), G (б), B (в).

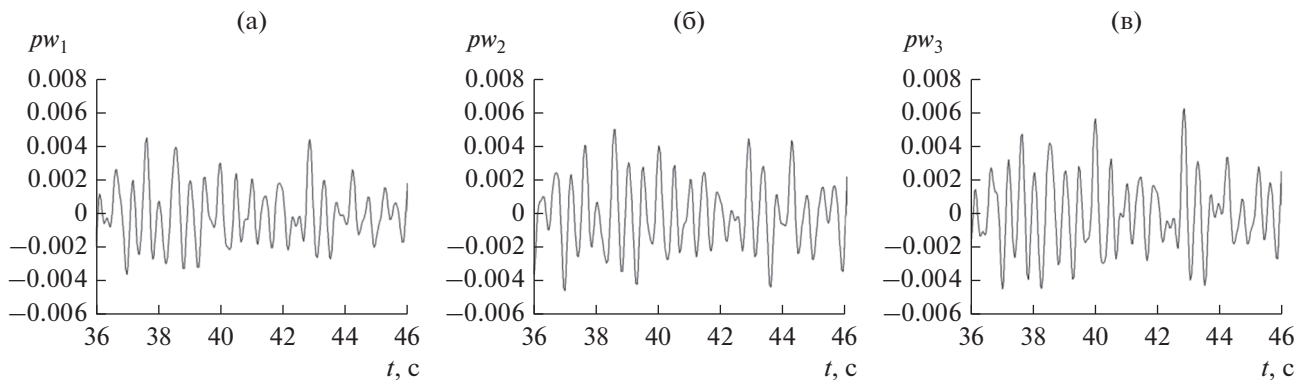


Рис. 6. Фрагменты ВР центрированных пульсовых волн цветочных каналов R (а), G (б), B (в).

вания волновых процессов на величины трендов ВР цветочных каналов:

$$w_l(n) = \frac{d_l^s(n,2) + d_l(n,3) + d_l(n,4) + d_l^f(n,5)}{d_l(n,8) + d_l(n,9) + d_l(n,10) + a_l(n,10)},$$

$$l = 1, 2, 3.$$

Фрагменты реализаций ВР оценок пульсовых волн цветочных каналов для примера Р1Н1 из БД [18] представлены на рис. 6.

3. ИНФОРМАТИВНОЕ ПРОСТРАНСТВО ПУЛЬСОВОЙ ВОЛНЫ

Снижение влияния помеховых факторов, к числу которых относят условия освещения, тон кожного покрова, физические нагрузки, незначительные движения испытуемого в процессе измерения видеоизображений, основано на формировании пространства признаков пульсовой волны и других биоритмов человека в исходном RGB-пространстве ВР цветочных каналов. К настоящему времени предложено и апробировано на практике достаточно большое количество алгоритмов генерации информативных признаков пульсо-

метрии методом ДФПГ [7], а именно: 1) слепое разделение источников (Blind Source Separation – BSS) методами главных и независимых компонент; 2) алгоритмы PVB и CHROM, основанные на модели отражения света кожным покровом; 3) алгоритм POS, формирующий информативный ортонормированный базис в RGB-пространстве на основе ортов тона кожи и импульса объема крови, унаследованных от алгоритма CHROM.

Особое место в этом далеко не полном списке занимает алгоритм 2SR, отслеживающий изменения во времени собственных значений и ориентации собственных векторов взаимной корреляционной матрицы RGB-пикселей видеоизображений. Иными словами, 2SR игнорирует процедуру пространственного усреднения цветочных пикселей в пределах области интереса ROI.

В конечном итоге алгоритмы BSS, PVB, CHROM, POS сводятся к поиску в исходном цветочном пространстве ортонормированного базиса (X, Y, Z), образованного ортом Z и сигнальной плоскостью с ортами (X, Y) (здесь и далее все векторы трактуют как столбцы, если не оговорено иное). Варианты выбора направляющих косину-

Таблица 1. Алгоритмы оценивания пульсовой волны

Алгоритм	X	Y	Z	s_x	s_y
G	$(1, 0, 0)^T$	$(0, 1, 0)^T$	$(0, 0, 1)^T$	0	1
GRD	$(1, 0, 0)^T$	$(0, 1, 0)^T$	$(0, 0, 1)^T$	-1	1
CHROM	$(3, -2, 0)^T/\sqrt{13}$	$(1.5, 1, -1.5)^T/\sqrt{5.5}$	$(1, 1, 1)^T/\sqrt{3}$	1	-1
POS	$(0, 1, -1)^T/\sqrt{2}$	$(-2, 1, 1)^T/\sqrt{6}$	$(1, 1, 1)^T/\sqrt{3}$	1	1

сов ортов (X, Y, Z) для некоторых алгоритмов приведены в табл. 1, где T – обозначает операцию транспонирования.

Проекция отсчетов ВР пульсовых волн цветовых каналов

$$\mathbf{W}_C(n) = \{w_1(n), w_2(n), w_3(n)\}^T$$

из RGB-пространства на плоскость (X, Y) формируют оценку центрированной и нормированной пульсовой волны [6, 7]:

$$\begin{aligned} pw(n) &= w(n)/\sigma_w(n), \\ w(n) &= s_x w_x(n) + s_y w_y(n), \\ w_x(n) &= \mathbf{X}^T \mathbf{W}_C(n)/\sigma_x(n), \\ w_y(n) &= \mathbf{Y}^T \mathbf{W}_C(n)/\sigma_y(n), \end{aligned}$$

где $\sigma_w(n)$ и $\sigma_x(n)$, $\sigma_y(n)$ – робастные оценки средних квадратов отклонений (СКО) ненормированной пульсовой волны $w(n)$ и информативных сигналов $w_x(n)$, $w_y(n)$ соответственно. Наша методика реализует простейший вариант выборочного оценивания СКО с помощью медианной фильтрации квадратов отсчетов ВР $w(n)$ и $w_x(n)$, $w_y(n)$ с окном сглаживания $1/(0.667 \Delta t)$, согласованным с нижней границей 0.667 Гц частот сердечных сокращений.

Наряду с указанными выше алгоритмами в наших вычислительных экспериментах была исследована эффективность модели ненормированной пульсовой волны $w(n) = (112.0, -93.786, -18.214) \mathbf{W}_C(n)/255.0$, основанной на красной цветоразностной компоненте C_r цветового пространства $YC_b C_r$ видеоизображений.

4. СПЕКТРАЛЬНЫЙ И СИНГУЛЯРНЫЙ АНАЛИЗ СЕГМЕНТОВ ПУЛЬСОВОЙ ВОЛНЫ

Предварительные оценки пульсовой волны $pw(n)$, полученные с помощью указанных выше алгоритмов, демонстрируют явно выраженную нестационарную динамику по критериям основных статистик [22]. Для адекватного исследования изменений спектральных характеристик такого рода ВР в нашей методике применяется скользящий интервал анализа. Реализацию пульсовой волны разбивают на последовательность сегментов,

содержащих по 1024 отсчета в каждом. Перекрытие соседних сегментов составляет 98%. Спектральные плотности мощности $P_{ns}(f)$ в пределах каждого сегмента с номером ns оценивают с помощью периодограммы с четырехчленным окном данных Блэкмана–Хэрриса [23] и периодограммы Томсона [24] с параметром сглаживания 3.

Закономерности трансформации периодограмм Томсона по мере увеличения номера сегмента ns пульсовой волны $pw(n)$, полученные для примера P1H1 из БД [18] с помощью алгоритма POS и модели C_r , представлены на рис. 7. Графики наглядно демонстрируют тенденцию стабилизации во времени частоты основного тона сердечных сокращений испытуемого после физических упражнений на велозергометре. Следует также отметить, что альтернативные оценки СПМ по результатам применения алгоритма POS и модели C_r для формирования пульсовой волны практически совпадают.

Рациональной предварительной оценкой изменения частоты основного тона сердечных сокращений (f_{hr}) испытуемого является, на наш взгляд, зависимость абсолютных максимумов периодограммы от номера сегмента (ns) пульсовой волны

$$f_{hr}(ns) = \max_f \{P_{ns}(f)\}. \quad (3)$$

Такие оценки для алгоритма POS и модели C_r пульсовой волны приведены на рис. 8 (кривая 1). Сглаживание подобных оценок с помощью робастной ядерной локально взвешенной квадратичной регрессии Клевеланда (функция `loess` в пакете Matlab) с параметром 0.3 дает более адекватное представление о стабилизации частоты сердечных сокращений после окончания физических упражнений (см. рис. 8, кривая 2).

Результаты скользящего фурье-анализа сегментов *предварительной оценки* пульсовой волны полученной, например, с помощью алгоритма POS или модели C_r , свидетельствуют о наличии обертонов на частотах 2.1 и 2.5 Гц (см. рис. 7). Плодотворный подход к формированию *уточненной оценки* основан на декомпозиции каждого сегмента волны на последовательность квазипериодических циклов, не искаженных фантомными обертонами, в частности, методом сингулярного спектрального анализа [16, 25].

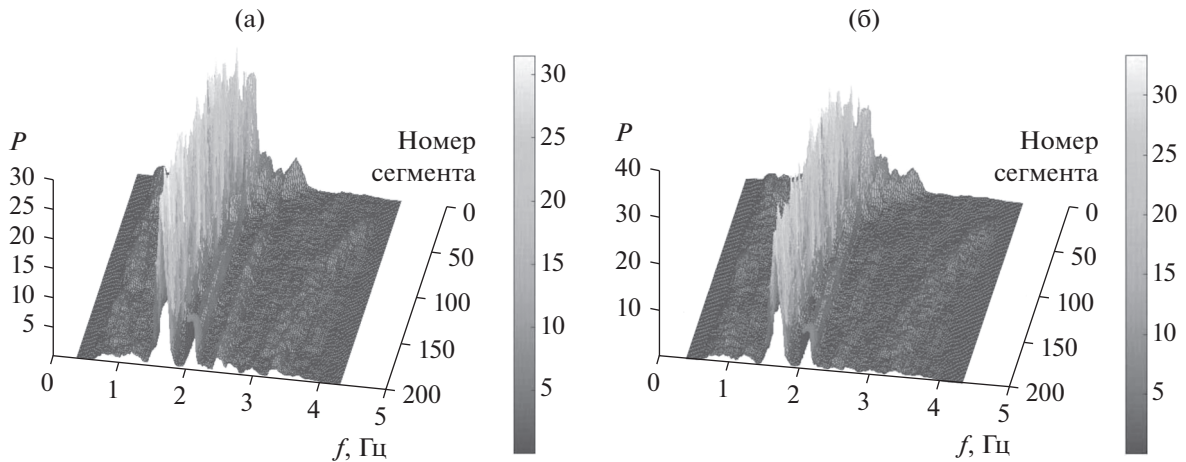


Рис. 7. Зависимость периодограмм Томсона от номера сегмента предварительной оценки пульсовой волны согласно алгоритму POS (а) и модели C_r (б).

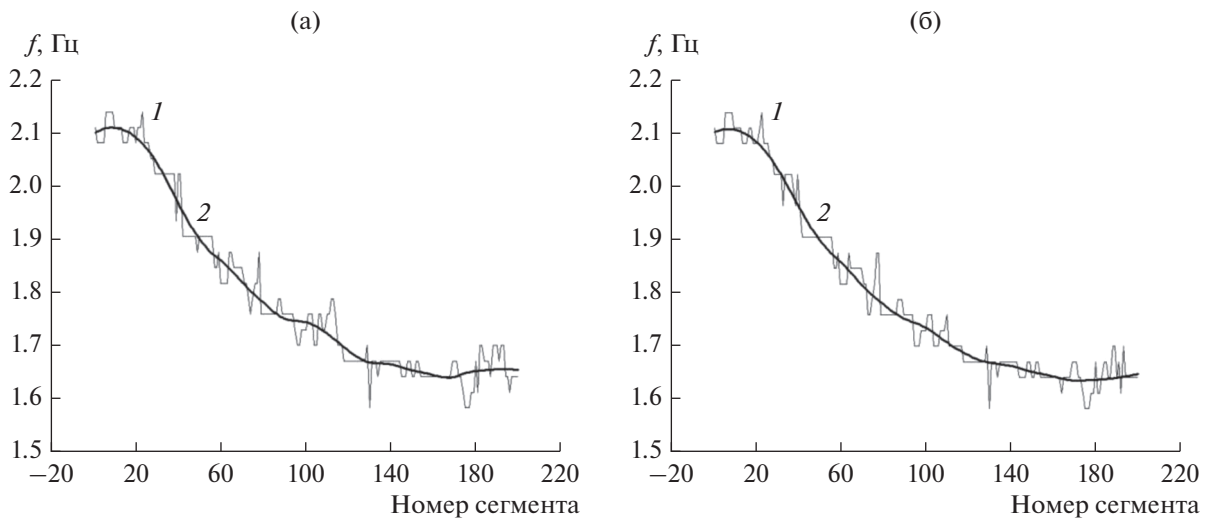


Рис. 8. Зависимость частоты основного тона сердечных сокращений от номера сегмента пульсовой волны согласно алгоритму POS (а) и модели C_r (б): кривая 1 – абсолютные максимумы периодограмм, кривая 2 – сглаженные оценки.

Наши вычислительные эксперименты показали, что результаты SSA-декомпозиции критичны к значению параметра вложения текущего сегмента волны в траекторное пространство. Удачным оказался выбор параметра вложения, равный пяти сглаженным периодам сердечных сокращений $5/\text{gloess}\{f_{hr}(\text{ns}), 0.3\}$. При таком выборе сумма двух первых сингулярных троек формирует для каждого сегмента данных временной профиль основного тона сердечных сокращений (bht), не искаженный фантомными обертонами.

Зависимости периодограмм Томсона суммы двух первых сингулярных троек от номера сегмента ns (рис. 9) иллюстрируют закономерность изменения во времени спектральных характеристик ВР основного тона сердечных сокращений

bht (n) для примера P1N1 из БД [18]. Полученные оценки СПМ подтверждают правильность выбора первой пары сингулярных троек для удаления фантомных обертонов пульсовой волны. Кроме того, сравнение рис. 8б и 9б наглядно демонстрирует состоятельность оценки (3).

5. МГНОВЕННАЯ ЧАСТОТА И ОГИБАЮЩАЯ ОСНОВНОГО ТОНА ПУЛЬСОВОЙ ВОЛНЫ

Естественной процедурой формирования временного профиля bht (n) является, на наш взгляд, агрегирование ансамбля сегментов данных с учетом их фактических временных сдвигов. Зависимость количества перекрывающихся сегментов

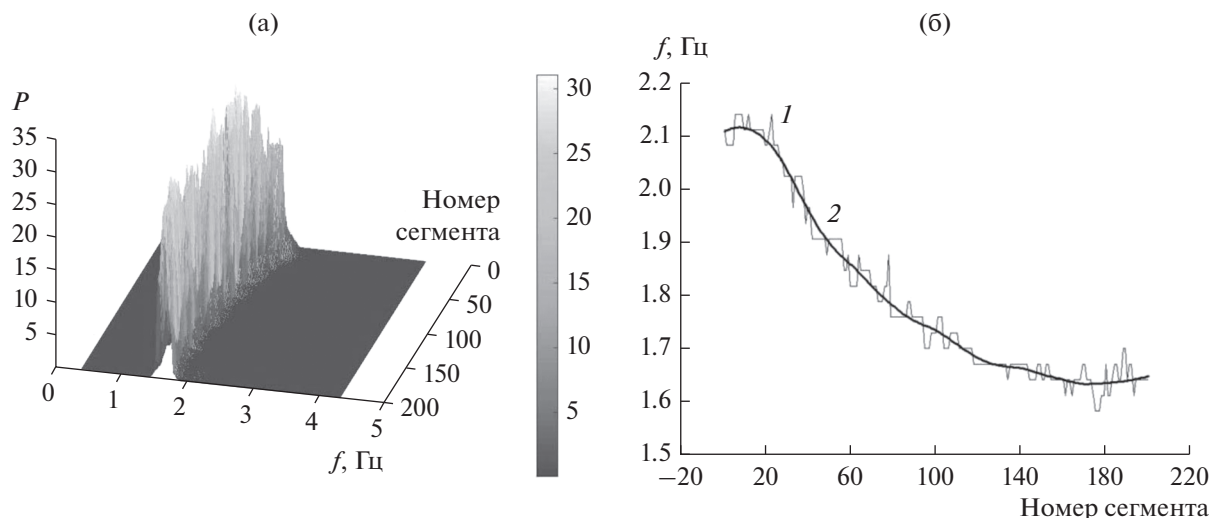


Рис. 9. Зависимость периодограмм Томсона от номера сегмента уточненной оценки пульсовой волны (а) и частот основного тона сердечных сокращений (б) согласно модели C_r ; кривая 1 – абсолютные максимумы периодограмм, кривая 2 – сглаженные оценки.

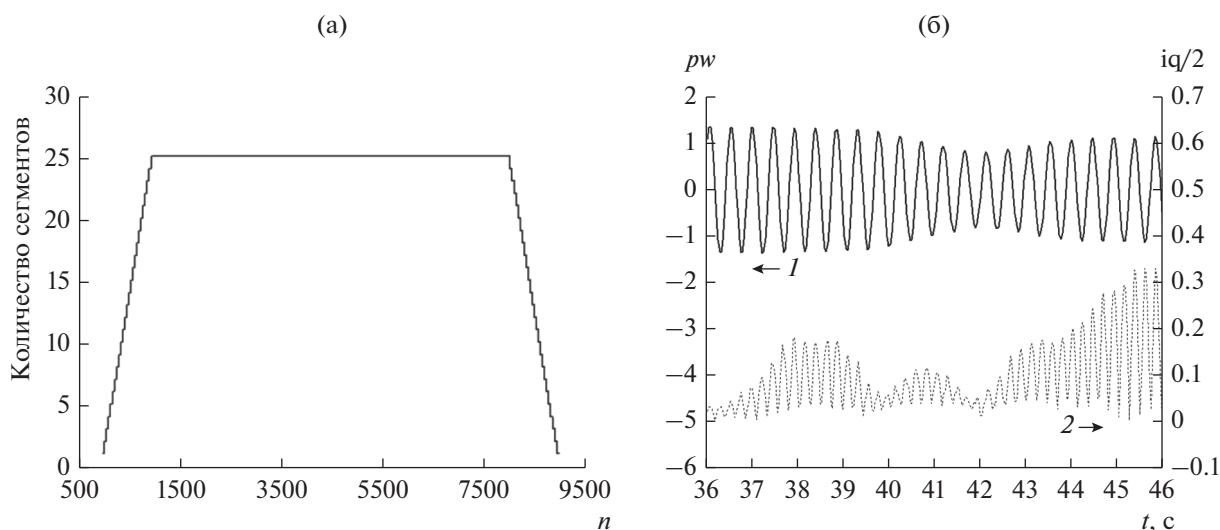


Рис. 10. Усреднение по ансамблю сегментов уточненной оценки пульсовой волны: зависимость количества сегментов от отсчета времени n (а); результаты усреднения (б): кривая 1 – фрагмент реализации медианной оценки агрегированной пульсовой волны, кривая 2 – мера разброса ансамбля сегментов относительно медианной оценки пульсовой волны.

NS от номера n текущего отсчета времени для выбранной нами степени перекрытия 98% представлена на рис. 10а.

Фрагмент ВР $bht(n)$ уточненной оценки основного тона сердечных сокращений, полученной медианным усреднением по ансамблю сдвинутых сегментов, демонстрирует кривая 1 на рис. 10б. Кривая 2 на этом рисунке иллюстрирует степень разброса ансамбля сегментов относительно медианной скользящей средней. Мерой разброса ансамбля является половина его интерквартильного диапазона $iq/2$.

Оценка $bht(n)$ представляет собой ВР с достаточно узкой полосой частот (см. рис. 9а). В этом случае состоятельные оценки мгновенной фазы (ip) и мгновенной огибающей (ie) могут быть получены с помощью дискретного преобразования Гильберта

$$ip_{bht}(n) = \arg\{H[bht(n)]\},$$

$$ie_{bht}(n) = |H[bht(n)]|.$$

В результате с помощью процедуры развертки мгновенной фазы $ip_{bht}(n)$ во времени (функция `unwvar` в пакете Matlab) формируют накопленную

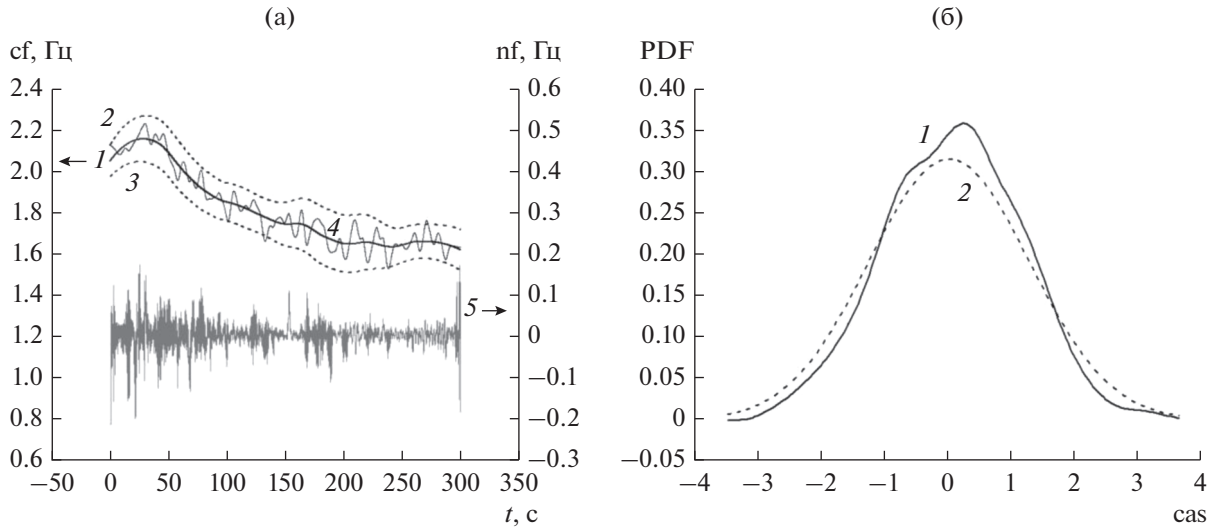


Рис. 11. Статистики мгновенных частот основного тона сердечных сокращений: реализации структурных компонент мгновенных частот (а): кривая 1 – тренд, кривые 2 и 3 – границы полосы Боллинджера, кривая 4 – сумма тренда и квазициклической составляющей, кривая 5 – шумовая компонента; плотность распределения вероятностей центрированных и нормированных мгновенных частот (б): кривая 1 – гистограмма сглаженная сдвигом, кривая 2 – гауссовское приближение гистограммы.

мгновенную фазу $iph_{bht}(n)$ для основного тона сердечных сокращений. Как правило, этот ВР содержит, к сожалению, разрывную не дифференцируемую составляющую, что приводит к появлению отрицательных величин для мгновенной частоты. Для устранения этих физически не интерпретируемых значений в нашей методике предусмотрено формирование двух структурных компонент накопленной мгновенной фазы. Первая – это непрерывная (continuous – cont), неубывающая составляющая $iph_{bht}^{(cont)}(n)$. Вторая разрывная (прерывистая) компонента представлена суммой смещенных во времени ступенчатых функций, аналогичных функции включения Хевисайда.

Физически интерпретируемые оценки мгновенной частоты основного тона сердечных сокращений дает первая производная по времени непрерывной, неубывающей компоненты накопленной фазы $iph_{bht}^{(cont)}(n)$. Численное дифференцирование этого ВР рационально, на наш взгляд, выполнять с помощью сохраняющей форму кусочно-кубической интерполяции полиномами Эрмита [26] (функция `rsph` в пакете Matlab). Для полинома третьей степени

$$iph_{bht}^{(cont)}(t) = a(n)(t - t_n)^3 + b(n)(t - t_n)^2 + c(n)(t - t_n) + d(n), \quad t_n \leq t < t_{n+1}$$

первая производная в момент времени $t = t_n = n\Delta t$ равна коэффициенту $c(n)$. Иными словами, ВР

зашумленной оценки мгновенной частоты (if) основного тона сердечных сокращений имеет вид

$$if_{bht}(n) = c(n)/(2\pi\Delta t).$$

Важно отметить, что динамика этого ВР демонстрирует наличие аномальных значений.

6. СТАТИСТИКИ МГНОВЕННЫХ ЧАСТОТ ОСНОВНОГО ТОНА ПУЛЬСОВОЙ ВОЛНЫ

Оценки статистик ВР $if_{bht}(n)$ рационально формировать с помощью процедуры робастной ядерной локально взвешенной квадратичной регрессии Клевеланда. В частности, тренд $af_{bht}(n)$, СКО $sf_{bht}(n)$ и квазициклическая компонента $cf_{bht}(n)$ представлены следующими моделями сглаживания данных:

$$\begin{aligned} af_{bht}(n) &= rloess\{if_{bht}(n), 0.3\}, \\ sf_{bht}(n) &= \sqrt{rloess\{[if_{bht}(n) - af_{bht}(n)]^2, 0.3\}}, \\ cf_{bht}(n) &= rloess\{if_{bht}(n), 0.03\} - af_{bht}(n). \end{aligned}$$

Реализации тренда (кривая 1), полос Боллинджера шириной ± 3 доли СКО (кривые 2 и 3), а также сумма тренда и квазициклической компоненты, т.е. $rloess\{if_{bht}(n), 0.03\}$ (кривая 4) для примера Р1Н1 из БД [18] и модели пульсовой волны C_r представлены на рис. 11а. Шумовую компоненту мгновенных частот сердечных сокращений

иллюстрирует кривая 5.

Оценку плотности распределения вероятностей (Probability Density Function – PDF) центрированного и нормированного ВР (cas) для квазициклической компоненты мгновенных частот основного тона сердечных сокращений

$$\text{cas}_{\text{bht}}(n) = \text{cf}_{\text{bht}}(n) / \text{sf}_{\text{bht}}(n)$$

в виде гистограммы, сглаженной сдвигом (Average Shifted Histogram – ASH), [27] иллюстрирует рис. 11б (кривая 1). Для сравнения там же представлено нормальное распределение (кривая 2), соответствующее робастным выборочным оценкам математического ожидания и СКО. В качестве нормированного окна данных ASH-оценки распределения применялось трижды взвешенное ядро Епанечникова. Оптимальное количество разрядных интервалов гистограммы выбиралось по правилу Фридмана–Дьякониса. В конечном итоге модель

$$\text{PDF}\{\text{cf}_{\text{bht}}(n)\} = \text{ASH}\left\{\frac{\text{cf}_{\text{bht}}(n)}{\text{sf}_{\text{bht}}(n)}\right\} / \text{sf}_{\text{bht}}(n)$$

описывает зависимость от времени $n\Delta t$ одномерной плотности распределения вероятностей PDF $\{\text{cf}_{\text{bht}}(n)\}$ квазициклической компоненты $\text{cf}_{\text{bht}}(n)$ мгновенных частот основного тона сердечных сокращений.

ЗАКЛЮЧЕНИЕ

В работе последовательно изложены этапы интеллектуального анализа мгновенных частот основного тона пульсовой волны очищенного от обертонов сердечных сокращений. Результаты обработки экспериментальных данных наглядно демонстрируют эффективность алгоритмов выделения пульсовой волны методами структурной декомпозиции нестационарных временных рядов ДФПГ. В частности, адекватность оценок статистических характеристик биоритмов человека обеспечивает естественный симбиоз современной методологии кратномасштабного, сингулярного и корреляционно-спектрального анализа биомедицинских данных.

СПИСОК ЛИТЕРАТУРЫ

1. Лазерная доплеровская флоуметрия микроциркуляции крови: Руководство для врачей / Под ред. А.И. Крупаткина, В.В. Сидорова. М.: Медицина, 2005.
2. Al-Naji A., Gibson K., Sang-Heon Lee, Chahl J. // IEEE Access. 2021. V. 5. № 17. P. 15776.
3. Успенский В.М. Информационная функция сердца. Теория и практика диагностики заболеваний внутренних органов методом информационного

анализа электрокардиосигналов. М.: Экономика и информация, 2008.

4. Bueno-López M., Muñoz-Gutiérrez P.A., Giraldo E., Molinas M. // IAENG Int. J. Computer Sci. 2019. V. 46. № 2. P. 228. http://www.iaeng.org/IJCS/issues_v46/issue_2/index.html.
5. Боронеев В.В. Пульсовая диагностика заболеваний в тибетской медицине: физические и технические аспекты. Улан-Уде: БНЦ СО РАН, 2005.
6. Unakafov A.M. // Biomedical Physics & Engineering Express. 2018. № 4. P. 045001. <https://doi.org/10.1088/2057-1976/aabd09>
7. Wang W., den Brinker A. C., Stuijk S., de Haan G. // IEEE Trans. 2016. V. BE-64. № 7. P. 1479. <https://doi.org/10.1109/TBME.2016.2609282>
8. Wang W., den Brinker A.C., Stuijk, S., de Haan G. // Biomedical Optics Express. 2017. V. 8. № 3. P. 1965. <https://doi.org/10.1364/BOE.8.001965>
9. Омпоков В.Д. Частотно-временной анализ пульсовых сигналов с помощью преобразования Гильберта–Хуанга // Дис. ... канд. физ.-мат. наук. Улан-Уде: Ин-т физ. материаловедения СО РАН, 2019. 106 с.
10. Флейшман А.Н. Вариабельность ритма сердца и медленные колебания гемодинамики: нелинейные феномены в клинической практике. Новосибирск: Изд-во СО РАН, 2009.
11. Боронеев В.В., Гармаев Б.З., Омпоков В.Д. // Вест. Бурят. гос. ун-та. 2012. № 3. С. 221.
12. Martinez J.P., Almeida R., Olmos S. et al. // IEEE Trans. 2004. V. BE-51. № 4. P. 570.
13. Islam M.R., Ahmad M. // Conf. Dig. 2 nd Int. Conf. Electrical, Computer and Communication Engineering (ECCE). Cox's Bazar. 7–9 Feb. 2019. N.Y.: IEEE, 2019. P. 8679156 <https://doi.org/10.1109/ECACE.2019.8679156>
14. Cheng J., Chen X., Xu L., Wang Z.J. // IEEE J. Biomedical Health Informatics. 2017. V. 21. № 5. P. 1422. <https://doi.org/10.1109/JBHI.2016.2615472>
15. Xu L., Cheng J., Chen X. // Electron. Lett. 2017. V. 53. № 4. P. 216. <https://doi.org/10.1049/el.2016.3611>
16. Sanei Saeid, Hassani Hossein. Singular Spectrum Analysis of Biomedical Signals. Boca Raton: CRC Press/Taylor and Francis Group, 2019.
17. Tang S.K.D., Goh Y.Y.S., Wong M.L.D., Lew Y.L.E. // 6th Int. Conf. Intelligent and Advanced Systems (ICIAS). Kuala Lumpur, 15–17 Aug. 2016. N.Y.: IEEE, 2016. P. 7824118 <https://doi.org/10.1109/ICIAS.2016.7824118>
18. Hoffman W.F.C., Lakens D. Public Benchmark Dataset for Testing rPPG Algorithm Performance. Delft: 4TU.Centre for Research Data, 2019. <https://doi.org/10.4121/uuid:2ac74fbd-2276-44ad-aff1-2f68972b7b51>
19. Martin E., Kriegel H.-P., Sander J., Xu Xiaowei // Proc. Second Int. Conf. Knowledge Discovery and Data

- Mining (KDD-96). Portland. 02–04.08.1996, Palo Alto: AAAI Press, 1996. P. 226.
20. *Zhongzhe Ch., Baqiao L., Xiaogang Y., Hongquan Y.* // *Energies*. 2019. V. 12. № 16. P. 3077. <https://doi.org/10.3390/en12163077>
21. *de Haan G., Jeanne V.* // *IEEE Trans.* 2013. V. BE-60. № 10. P. 2878.
22. *Borzov A., Kasikin A., Labunets L., Ryakhina M.* // *Proc. Int. Scientific and Practical Conf. "Information Technologies and Intelligent Decision Making Systems" (ITIDMS 2021) Moscow. 20 Jan. 2021. Aachen: CEUR Workshop Proc., 2021. V. 2843. P. 034. <http://ceur-ws.org/Vol-2843>.*
23. *Harris F.J.* // *Proc. IEEE*. 1978. V. 66. № 1. P. 51.
24. *Thomson D.J.* // *Proc. IEEE*. 1982. V. 70. № 9. P. 1055.
25. *Golyandina N., Nekrutkin V., Zhigljavsky A.* *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton: Chapman and Hall/CRC, 2001.
26. *Fritsch F.N., Carlson R.E.* // *SIAM J. Numerical Analysis*. 1989. V. 26. № 1. P. 230.
27. *Scott D.W.* *Multivariate density estimation: theory, practice, and visualization*. N.Y.: John Wiley & Sons Inc., 1992.

ЭЛЕКТРОННАЯ И ИОННАЯ ЭМИССИЯ

УДК 537.533.2

ПОЛЕВАЯ ЭМИССИЯ ЭЛЕКТРОНОВ МНОГООСТРИЙНЫХ КАТОДНЫХ МАТРИЦ В СИЛЬНЫХ ИМПУЛЬСНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

© 2022 г. Р. К. Яфаров*

*Саратовский филиал Института радиотехники и электроники им. В.А. Котельникова РАН,
ул. Зеленая, 38, Саратов, 410019 Российская Федерация*

**E-mail: pirpc@yandex.ru*

Поступила в редакцию 20.12.2020 г.

После доработки 07.01.2021 г.

Принята к публикации 12.01.2021 г.

Исследованы динамические свойства и рассмотрены особенности механизмов электронного транспорта в приповерхностных слоях кремниевых многоострийных катодных матриц различных типов электропроводности при полевой эмиссии в сильных импульсных электрических полях микросекундной длительности. Показано, что увеличение длительности импульсов повышает прозрачность потенциальных барьеров и уменьшает пороги начала автоэмиссии. Влияние длительностей импульсов на динамические свойства и параметры полевой эмиссии электронов определяются типом электропроводности кристаллов кремния и видом химически активных газовых сред, которые использовались при плазмохимической обработке многоострийных катодных матриц.

DOI: 10.31857/S0033849422020152

ВВЕДЕНИЕ

Развитие вакуумной микроэлектроники в настоящее время во многом определяется возможностями увеличения частотного диапазона работы полевых источников электронов, используемых при создании активной элементной базы, и снижением времени готовности, которое характеризуется длительностью выхода параметров автоэмиссии на стационарный режим функционирования. Одним из наиболее привлекательных материалов для полевых источников электронов вакуумной микроэлектроники остается кристаллический кремний [1, 2]. В первую очередь это обусловлено развитой технологической базой, а также возможностями его использования для создания приборов и устройств наноэлектроники и нанофотоники [3, 4].

Аналитическим выражением, описывающим туннельную эмиссию электронов при приложении сильных электрических полей, является формула Фаулера и Нордгейма [5], характеризующая прозрачность потенциального барьера D на границе твердое тело–вакуум и позволяющая оценить плотность тока автоэлектронной эмиссии. Квантово-механические расчеты показывают (см., например, [6]), что вероятность прохождения частицы сквозь потенциальный барьер очень сильно зависит от ширины барьера и его превышения над величиной энергии туннелирующего электрона.

Несмотря на упрощения, использованные при получении формулы Фаулера–Нордгейма, она хорошо описывает автоэлектронную эмиссию из металлов в вакуум. Однако процессы исследований автоэмиссии из наноструктурированных полупроводниковых и различных нанокompозитных углеродных структур не всегда удается адекватно описать теорией Фаулера–Нордгейма. Основными причинами этого, по мнению некоторых исследователей, являются размерное квантование, специфика проникновения поля в катодную матрицу и туннелирование электронов сквозь потенциальный барьер [7–11]. Поэтому не вызывает сомнений необходимость учета упомянутых факторов при анализе процессов автоэмиссии из неметаллических катодных структур. Признание этого диктует необходимость привлечь при рассмотрении функциональных связей между основными параметрами автоэмиссии новые для данной области исследований физические явления. К ним можно отнести вопросы, связанные с поляризацией, обусловленной электронным смещением, а также с изменениями степени заполнения энергетических состояний в структурах в сильных импульсных электрических полях. Исследование подобных явлений в автоэмиссионных процессах имеет не только определенный фундаментальный, но и большой прикладной интерес, который связан с возможностями применения автоэмиссионных структур в субтерагерцовой вакуумной микроэлектронике.

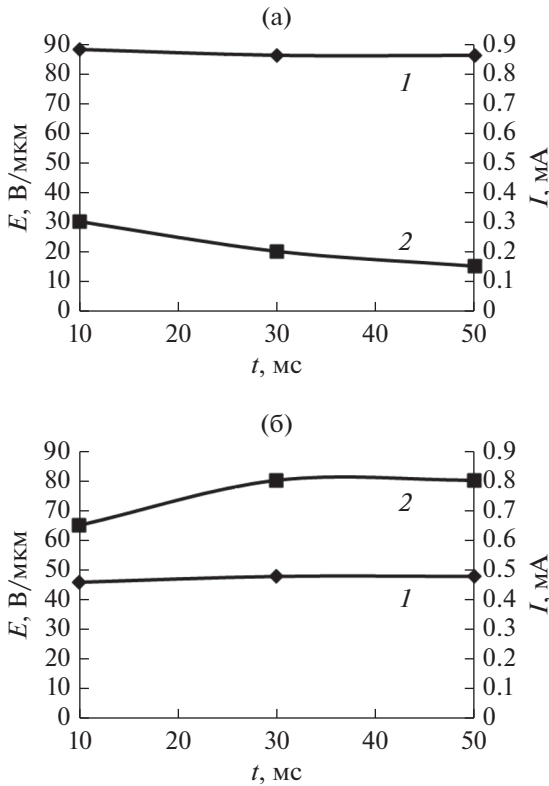


Рис. 1. Зависимости порогов (1) и токов автоэмиссии (2) при напряженности поля, превышающей порог автоэмиссии на 10 В/мкм от длительности импульсов анодных напряжений для кремния *n*-типа с плазменной обработкой в различных химически активных средах Ag (а) и CF₄ (б).

Исследование автоэмиссионных свойств кремниевых катодных структур в сильных импульсных электрических полях микросекундной длительности, в отличие от стационарных полей, представляет собой самостоятельную и более сложную задачу. Это связано, в первую очередь, с различной динамикой передислокации электронной плотности и зависимостью от энергии эффективности взаимодействия туннелирующих электронов с широким спектром зарядовых состояний в приповерхностном слое полупроводника. В отличие от монотонного изменения энергии электронов при увеличении внешнего электрического поля при автоэмиссии в стационарных полях, практически мгновенное увеличение энергии и длины свободного пробега в сильных импульсных полях микросекундной длительности существенным образом изменяет характер взаимодействий электронов с его окружением в твердом теле.

Цель данной работы — исследовать эмиссионную способность в сильных импульсных электрических полях многоострижных катодных матриц, изготовленных на основе кристаллов кремния различных типов электропроводности с исполь-

зованием плазмохимической обработки в различных химически активных средах.

1. МЕТОДИКА И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В экспериментах были использованы кристаллы кремния ориентации (100) различных типов электропроводности с одинаковым удельным сопротивлением, 0.01...0.02 Ом см. Изготовление и исследования автоэмиссионных свойств многоострижных катодных матриц на их основе проводили в условиях высокого вакуума (10^{-6} Па) на диодной структуре, способной изменять расстояние между электродами с точностью до 1 мкм по методикам, изложенным в работах [12–15]. Расстояние между анодом и катодной структурой составляло 40 мкм. Диаметр рабочей поверхности анода, изготовленного из углеродного материала МПГ-6, составлял 3 мм.

На рис. 1–3, по аналогии с экспериментальными результатами [15] исследования в сильных импульсных электрических полях динамических автоэмиссионных свойств катодных матриц на основе кристаллов кремния *p*-типа, приведены зависимости порогов и токов автоэмиссии от длительности импульсов анодных напряжений, а также амплитуд автоэмиссионных токов от количества импульсов анодных напряжений различной длительности для кристаллов кремния *n*-типа с плазменной обработкой в различных химически активных средах. Исследования проводили при напряженности поля, превышающей порог автоэмиссии на 10 В/мкм. Как видим из рис. 1, при плазменной обработке в среде аргона при увеличении длительности импульсов анодных напряжений пороги автоэмиссии для кристаллов кремния *n*-типа, так же как в ранее исследованных кристаллах *p*-типа [15], слабо уменьшаются. Однако токи при этом не увеличиваются, как для кристаллов *p*-типа, а уменьшаются. При долговременных импульсных испытаниях, в течение 5 мин, с длительностями импульсов 10 мкс и скважностью 4×10^4 наблюдается небольшое уменьшение огибающих амплитуд анодных токов (см. рис. 2). При длительностях импульсов 30 и 50 мкс уменьшение огибающих анодных токов не наблюдается. Отметим, что для кристаллов кремния *p*-типа при всех длительностях импульсных воздействий амплитуды автоэмиссионных токов практически мгновенно выходят на стационарный уровень [15].

При плазмохимической обработке кристаллов кремния *n*-типа в среде хладона-14 пороги начала автоэмиссии от длительности импульсов практически не зависят, а автоэмиссионные токи с увеличением длительности импульсов увеличиваются (см. рис. 1б). Пороги начала процесса автоэмиссии значительно меньше, а токи в несколько

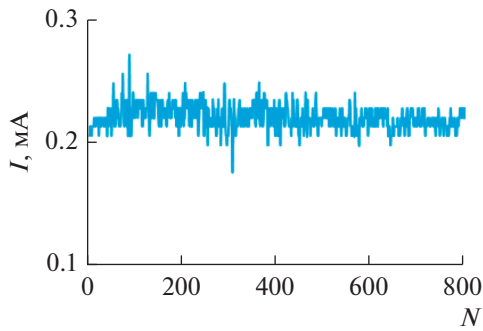


Рис. 2. Зависимость амплитуд анодных токов от количества импульсов с длительностью 10 мкс кристаллов *n*-типа, обработанных в плазме аргона.

раз больше, чем при обработке в плазме аргона. Для кристаллов *p*-типа, обработанных в плазме CF_4 , характеры изменений величин напряжений порогов автоэмиссии и автоэмиссионных токов в зависимости от длительностей импульсов такие же как при их обработке в плазме аргона (см. рис. 1б). Однако наблюдаемые пороги эмиссии значительно меньше, а токи выше [15].

При долговременных испытаниях катодных матриц на основе кристаллов *n*-типа, выполняемых с уменьшением длительностей импульсов напряжений, длительность выхода на стационарный режим автоэмиссионных токов увеличивается. При длительностях импульсов больше 30 мкс выход на стационарный уровень, так же как для кристаллов *p*-типа [15], происходит практически мгновенно (см. рис. 3).

2. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Из полученных результатов следует, что для обоих типов кристаллов кремния и упомянутых видов плазменной обработки общим является тенденция уменьшения порогов автоэмиссии при увеличении длительности импульсов. В сильных электрических полях это связано, возможно, с увеличением длины свободного пробега электронов и их кинетической энергии (как известно, собственная энергия электронов определяется квадратом напряженности внешнего электрического поля и длительностью между двумя последовательными столкновениями, и увеличивается при возрастании энергии), это приводит к увеличению прозрачности потенциальных барьеров. Для кристаллов кремния *p*-типа это выражается в уменьшении порогов начала автоэмиссии и увеличении автоэмиссионных токов для обоих видов плазменной обработки. Для кристаллов *n*-типа тенденция на уменьшение порогов автоэмиссии с увеличением длительностей импульсов выражена более слабо (см. рис. 1). Более того, при их плазменной обработке в среде аргона автоэмиссион-

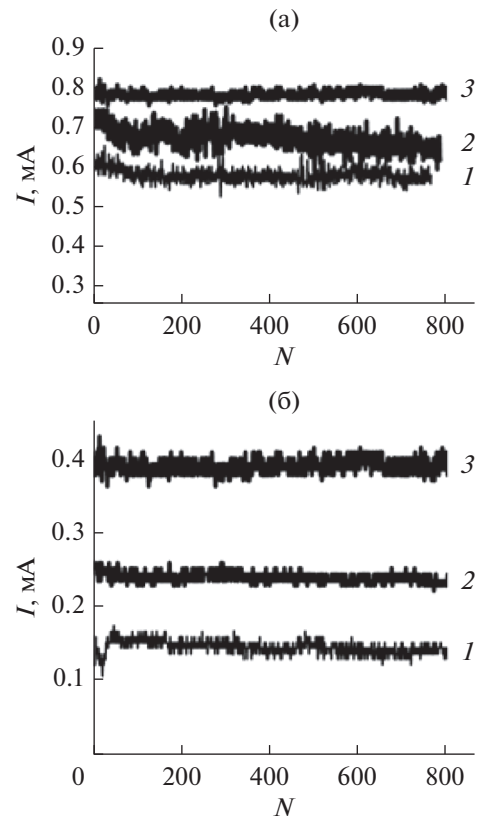


Рис. 3. Зависимости амплитуд автоэмиссионных токов от количества импульсов анодных напряжений различной длительности – 10 (1), 30 (2) и 50 мкс (3) – для кристаллов кремния *n*- (а) и *p*-типа (б) при плазменной обработке в среде CF_4 .

ные токи не увеличиваются, а уменьшаются (см. рис. 1б). Эти различия в параметрах полевой эмиссии для кристаллов различных типов проводимости могут быть интерпретированы влиянием встроенных поверхностных потенциалов и дипольных моментов эмитирующих поверхностей, которые формируются при их плазменном травлении в различных химически активных средах [13, 14].

При использовании низкоэнергетичной плазмы аргона травление кристаллов кремния осуществляется за счет распыления ускоренными ионами поверхностных атомов легирующей примеси и кремния, а также частиц его естественного окисла. Вследствие инертности аргона, какие-либо химические связи с поверхностными атомами кремния не образуются. В результате такой обработки поверхность кремния приобретает структуру и плотность оборванных (ненасыщенных) связей, характерных для атомночистой поверхности кремния заданной кристаллографической ориентации, в частности, к образованию трехвалентного кремния. Таким образом, становится возможным образование поверхностных

состояний либо глубоких энергетических центров донорного типа [16], что приводит к ухудшению проводимости приповерхностного слоя и росту толщины потенциального барьера.

При плазмохимическом травлении кремния в CF_4 химически активными частицами являются ионы C^+ , CF_n^+ ($n = 1...4$), а также радикалы CF_n и нейтральные атомы фтора, которые при хемосорбции образуют адкомплексы $\text{Si}-\text{C}$ и $\text{Si}-\text{F}$ [12]. Хемосорбированные комплексы SiC и SiF имеют энергии химической связи (4.55 и 5.6 эВ соответственно), которые превышают энергию химических связей атомов кремния (3.38 эВ). Вследствие этого поверхностная плотность ненасыщенных связей и встроенный положительный заряд при плазмохимическом травлении кристаллов кремния в хладоне-14 меньше, чем при травлении в плазме аргона. При одинаковых режимах обработки за счет лучшей электрокомпенсированности плотность ненасыщенных электронных связей и величина встроенного приповерхностного электрического поля, направленного в объем полупроводника, на кремнии n -типа меньше, чем на кремнии p -типа.

Как следует из изложенного выше, в термодинамически равновесном состоянии обеднение приповерхностной области носителями заряда сильнее происходит при плазменной обработке в среде аргона. При воздействии положительных импульсов анодных напряжений область пространственного заряда (ОПЗ) экранирует объем полупроводника от проникновения внешнего электрического поля. Вследствие меньшей концентрации ненасыщенных связей ослабление внешнего поля при обработке кристаллов в плазме хладона-14 меньше, чем при плазменной обработке в аргоне. Это уменьшает при обработке в плазме хладона-14 напряженности пороговых полей автоэмиссии и увеличивает автоэмиссионные токи при меньших амплитудах импульсных анодных напряжений.

Следует отметить, что увеличение автоэмиссионных токов при обработке в плазме хладона-14, по сравнению с обработкой в плазме аргона, не является достаточно существенным, как это следовало бы ожидать из снижения порогов автоэмиссии. Причиной этого может быть высокая электроотрицательность атомов фтора, которые увеличивают дипольный момент эмиттирующих центров из-за вероятного присутствия на поверхности кремния $\text{Si}-\text{F}$ диполей [14, 17]. Влияние встроенных полей (дипольного момента) обнаруживается при увеличении длительности импульсов анодных напряжений, что может указывать на уменьшение энергии глубоких энергетических центров из-за влияния внешнего электрического

поля. Для кристаллов n -типа это выражается в увеличении автоэмиссионных токов при увеличении длительностей импульсов (см. рис. 1б).

В процессе полевой эмиссии осуществляется модуляция заряда поверхностных ловушек посредством захвата электронов, дрейфующих к поверхности. Как показали эксперименты, при плазменной обработке кристаллов кремния в среде аргона их быстрое заполнение выражается в практически мгновенном выходе на стационарный режим огибающих амплитуд анодных токов уже при длительности импульсов около 10 мкс [15]. При плазмохимической обработке кристаллов кремния в среде хладона-14 наблюдаемое во времени уменьшение огибающих амплитуд токов анодных импульсов в процессе долговременных испытаний с короткими длительностями импульсов обусловлено, возможно, увеличением суммарного дипольного момента эмиссионной поверхности (рис. 3). Увеличение тока при увеличении длительности импульсов происходит в результате захвата электронов поверхностными ловушками и ослабления кулоновского взаимодействия между зарядами, образующими $\text{Si}-\text{F}$ -диполи. Поэтому при длительностях более 30 мкс повышается прозрачность потенциальных барьеров для туннелирующих электронов. Благодаря этому в процессе долговременных испытаний амплитуды автоэмиссионных токов с увеличением длительностей импульсов увеличиваются, а их выход в стационарный режим происходит практически мгновенно (рис. 3, кривые 2 и 3).

При обработке в плазме аргона и/или хладона-14 кристаллов кремния n -типа вследствие лучшей электрокомпенсированности встроенный поверхностный потенциал и пороги начала автоэмиссии меньше, чем для кристаллов кремния p -типа. При плазмохимическом травлении кристаллов n -типа в плазме хладона-14 вследствие меньшей концентрации ненасыщенных связей и меньшей величины ОПЗ, возникающей при приложении анодных импульсов напряжений, пороговые напряжения автоэмиссии практически в два раза меньше, чем при обработке в плазме аргона (см. рис. 1, кривые 1).

Хемосорбция электроотрицательных атомов фтора при плазмохимическом травлении кристаллов кремния n -типа в хладоне-14, в отличие от травления кристаллов кремния p -типа, приводит к обогащению поверхностного слоя полупроводника электронами. Это существенно увеличивает автоэмиссионные токи при более низких анодных напряжениях по сравнению с обработкой кристалла в плазме аргона (см. рис. 1). Влияние дипольного момента поверхности от предполагаемого присутствия на кремнии $\text{Si}-\text{F}$ -комплексов обнаруживается при увеличении длительностей импульсов

анодных напряжений. При их увеличении в интервале от 10 до 50 мкс автоэмиссионные токи сначала увеличиваются, а затем стабилизируются на некотором постоянном уровне (рис. 3б, кривая 2).

При плазменной обработке кристаллов кремния *n*-типа в среде аргона его ионы разрывают химические связи, как между атомами кремния, так и между атомами сурьмы и кремния. При разрыве последних не связанных с атомом кремния электрон атома сурьмы нейтрализует бомбардирующий его ион аргона (резонансная нейтрализация [16]). Следствием этого является изменение степени заполнения поверхностных и глубоких уровней, что приводит к формированию на поверхности тонкого слоя, обедненного основными носителями. При полевой эмиссии дрейфующие под действием импульсов анодных напряжений электроны формируют ОПЗ. Вследствие возникающей экранировки внешнего поля пороги автоэмиссии при обработке в плазме аргона увеличиваются, а автоэмиссионные токи по сравнению с обработкой в плазме хладона-14 уменьшаются.

В процессе длительных импульсных воздействий динамическая нейтрализация поверхностных центров захвата осуществляется достаточно быстро и выражается в практически мгновенном выходе на стационарный режим огибающих амплитуд анодных токов уже при длительности импульсов около 10 мкс (см. рис. 3). При плазмохимической обработке кристаллов кремния *n*-типа в среде хладона-14 уменьшение огибающих амплитуд токовых импульсов, наблюдаемое в процессе долговременных импульсных испытаний, обусловлено влиянием суммарного дипольного момента эмиссионной поверхности.

ЗАКЛЮЧЕНИЕ

В результате экспериментальных исследований импульсных автоэмиссионных характеристик кремниевых катодных структур в электрических полях микросекундной длительности, были выявлены следующие закономерности.

1. С увеличением длительности импульсов сильных электрических полей наблюдается уменьшение напряженностей внешних электрических полей начала полевой эмиссии в многоострийных катодных матрицах на основе кристаллов кремния различных типов проводимостей. В соответствии с представлениями Фаулера и Нордгейма это обусловлено увеличением прозрачности потенциальных барьеров для туннелирующих электронов за счет уменьшения величины барьера либо из-за увеличения собственной энергии электронов.

2. Характер влияния длительностей импульсов на автоэмиссионные токи определяется типом

электропроводности кристаллов кремния и видом химически активных газовых сред, которые использовались при плазмохимической обработке многоострийных катодных матриц. Для катодных матриц на кристаллах кремния *p*-типа, полученных с использованием плазменной обработки в химически нейтральной (аргон) и химически активной газовых средах (хладон-14), увеличение длительностей импульсов приводит к увеличению амплитуд автоэмиссионных токов. Использование хладона-14 при обработке приводит к большей величине и более быстрому увеличению автоэмиссионных токов при увеличении длительностей импульсов. Последнее, возможно, обусловлено уменьшением эффективности кулоновского взаимодействия туннелирующих электронов с дипольными моментами хемосорбированных на поверхности кремния Si–F-комплексов. Отрицательным эффектом от использования хладона-14 является увеличение по сравнению с плазменной обработкой в среде аргона длительности выхода амплитуд автоэмиссионных токов на стационарный режим.

3. Плазменная обработка кристаллов кремния *n*-типа в среде аргона приводит к изменению степени заполнения поверхностных состояний и глубоких энергетических центров. С увеличением длительности импульсов анодных напряжений это способствует дополнительному увеличению ОПЗ, что уменьшает величину автоэмиссионных токов. Динамическая нейтрализация ловушек осуществляется достаточно быстро и выражается в практически мгновенном выходе на стационарный режим огибающих амплитуд анодных напряжений уже при длительности импульсов около 10 мкс.

При плазмохимической обработке кристаллов кремния *n*-типа в среде хладона-14 наблюдается уменьшение огибающих амплитуд анодных импульсов. При увеличении длительности импульсов амплитуды автоэмиссионных токов увеличиваются, а длительности выхода на стационарные режимы уменьшаются.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках государственного задания.

СПИСОК ЛИТЕРАТУРЫ

1. Marcus R.B., Ravi T.S., Gmitter T. et al. // Appl. Phys. Lett. 1990. V. 56. № 3. P. 236.
2. Velasquez-Garcia L.F., Guerrero S.A., Niu Y., Akinwande A.I. // IEEE Trans. 2011. V. ED-58. № 6. P. 1775.
3. Герасименко Н.Н., Пархоменко Ю.Н. Кремний — материал нанoeлектроники. М.: Техносфера, 2007.

4. Успехи нанотехнологии: электроника, материалы, структуры / Под ред. Дж. Дэвиса, М. Томсона. М.: Техносфера, 2011.
5. *Fowler R.H., Nordheim L.W.* // Proc. Royal Soc. London. A. 1928. V. 119. № 781. P. 173.
6. *Блохинцев Д.И.* Основы квантовой механики. М.: Наука, 1983.
7. *Фурсей Г.Н., Поляков М.А., Кантонистов А.А. и др.* // ЖТФ. 2013. Т. 83. № 6. С. 71.
8. *Давидович М.В., Яфаров Р.К.* // ЖТФ. 2018. Т. 88. № 2. С. 283.
9. *Яфаров Р.К.* // ЖТФ. 2017. Т. 87. № 10. С. 1578.
10. *Лобанов В.М., Шешин Е.П.* // Письма в ЖТФ. 2007. Т. 33. № 9. С. 11.
11. *Forbes R.G., Xanthakis J.P.* // Surf. Interface Anal. 2007. V. 39. P. 139.
12. *Яфаров Р.К.* Физика СВЧ вакуумно-плазменных нанотехнологий. М.: Физматлит, 2009.
13. *Яфаров Р.К., Шаныгин В.Я.* // ФТП. 2017. Т. 51. № 4. С. 558.
14. *Яфаров Р.К.* // ФТП. 2019. Т. 53. № 1. С. 18.
15. *Яфаров Р.К.* // Письма в ЖТФ. 2019. Т. 45. № 9. С. 3.
16. *Оура К., Лифшиц В.Г., Саранин А.А. и др.* Введение в физику поверхности. М.: Наука, 2006.
17. *Яфаров Р.К.* // Письма в ЖТФ. 2018. Т. 44. № 13. С. 68.

ФИЗИЧЕСКИЕ ЯВЛЕНИЯ В ЭЛЕКТРОННЫХ ПРИБОРАХ

УДК 535.215.1,51-73

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ФОТОЗАРЯДОВОГО ЭФФЕКТА В ЭЛЕКТРОНИКЕ

© 2022 г. П. С. Мартьянов^а,*, Д. В. Чуриков^а

^а Научно-технологический центр уникального приборостроения РАН,
ул. Бутлерова, 15, Москва, 117342 Российская Федерация

*E-mail: La3232@mail.ru

Поступила в редакцию 29.01.2021 г.

После доработки 08.08.2021 г.

Принята к публикации 15.09.2021 г.

Рассмотрена возможность получения фотозарядового эффекта на различных материалах. Создана экспериментальная установка для исследования эффекта и получены результаты, подтверждающие наличие этого эффекта на некоторых полупроводниковых образцах. Обнаруженный эффект можно использовать для оценки качества материалов, применяемых для изготовления электронных компонентов.

DOI: 10.31857/S0033849422020127

ВВЕДЕНИЕ

В оптоэлектронике исследование механизмов фоточувствительности представляет практическую значимость. Для понимания этих процессов и реализации новых алгоритмов видения, а также электронных регистрирующих приборов целесообразно провести исследование в данной области оптоэлектроники, поскольку современные электронные и акустооптические устройства позволяют получить желаемые результаты.

Фотозарядовый эффект начали исследовать в конце XX в. Была построена теоретическая модель, которая основана на эффекте перераспределения электронов по поверхности образца при облучении светом. Также была получена зависимость величины фотозарядового эффекта от диэлектрической проницаемости, проводимости и формы поверхности исследуемого образца. Стоит отметить, что при фотозарядном эффекте происходит процесс геометрического усиления, заметный на объектах, которые имеют широкую освещаемую площадку и утоньшение на обратной стороне. В этом случае даже небольшое изменение электронной плотности под воздействием излучения приводит к появлению избытка электронов на утончении образца, который может иметь различную форму. Но более значительный эффект геометрического усиления будет наблюдаться на объектах, которые по форме похожи на каплю: это широкое основание и сужение к вершине. В результате при освещении объектов такой формы возникает значительное электрическое поле [1, 2], которое можно измерить.

При облучении образца светом с определенной модуляцией и интенсивностью на его поверхности образуется небольшой электрический заряд. Это приводит к возникновению электрических потенциалов, которые можно зарегистрировать. Данный эффект имеет место для материалов с различными проводящими свойствами: проводниках, полупроводниках и диэлектриках. Он также наблюдается и на биологических объектах: листьях, ветках, тканях. Физический механизм возникновения напряжения будет различным для каждого материала. Например, при облучении диэлектрических образцов возникают силы, действующие на дипольные моменты и зависящие от градиента диэлектрической проницаемости. Это приводит к перераспределению заряда и появлению напряжения.

1. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Появление фотозарядового эффекта на поверхности объекта возможно только при модулированном облучении. Это позволяет проверить, вызвано ли измеренное напряжение этим или другими эффектами, такими как внешние и внутренние фотоэффекты, тепловое электричество и др. При облучении исследуемого образца немодулированным светом эффект возникновения напряжения не наблюдается [3].

Следует отметить, что данный эффект чувствителен к локальной проводимости освещаемого участка поверхности образца, что позволяет использовать его для бесконтактного определе-

ния свойств поверхности материала и определения его структуры. Поскольку эффект обнаруживается в длинноволновой области спектра, то его также можно использовать при разработке методов регистрации излучения ближнего и дальнего инфракрасного (ИК) и терагерцового диапазонов. Это особенно важно при исследовании различных областей спектра электромагнитных волн [4].

Фотозарядовый эффект – это явление перераспределения заряда, которое возникает под действием излучения. Перераспределение заряда приводит к изменению электрического потенциала двойного слоя вблизи поверхности образца, изготовленного из проводящего материала [5]. Для описания эффекта обычно используют связанную систему уравнений: условия механического равновесия электронов

$$en(x)\varphi'(x) + P'(x) = F(x) \quad (1)$$

и уравнения Пуассона

$$\varphi''(x) + 4\pi en_0(x) = 4\pi en(x), \quad (2)$$

где $P(x)$ – давление электронного газа, $F(x)$ – сила, действующая на электроны со стороны излучения падающего излучения, $\varphi(x)$ – скалярный потенциал двойного слоя, $n(x)$ – концентрация электронов, $n_0(x)$ – концентрация положительных зарядов (ионов), e – заряд электрона [6].

Приведенные выше выражения описывают зависимость изменения потенциала под действием падающего излучения. Эта зависимость может быть представлена в следующем виде:

$$\varphi(x) = \varphi(P(x), E, \varepsilon(x), n_0(x), n(x)), \quad (3)$$

где E – амплитуда электрического поля падающего излучения, $\varepsilon(x)$ – диэлектрическая проницаемость. Следовательно, величина эффекта зависит от мощности излучения, свойств материала и состояния его поверхности, которые описываются функциями $\varepsilon(x)$, $n_0(x)$, $n(x)$ и $P(x)$.

Практическое использование фотозарядового эффекта может быть в следующих областях.

1. Бесконтактное измерение характеристик полупроводниковых материалов при производстве электрорадиоэлементов (ЭРЭ). Характеристики полупроводниковых устройств зависят от чистоты используемого материала, формы и состояния поверхности. Фотозарядовый эффект чувствителен к наличию примесей и дефектов, поэтому его можно применять для изучения электрической структуры поверхности. Такое исследование возможно реализовать не только для полупроводников и проводников, но также и для других перспективных материалов, например феррогранатов, свойства которых существенно зависят от примесей и дефектов, [7].

2. Исследование органических соединений. Для многих сельскохозяйственных культурных

растений и плодов возможно применить оценку их степени созревания на основе предложенного эффекта. Также предложенный метод можно использовать для диагностики органических тканей в биологии и медицине.

3. Реализация оптических систем видения. Это инженерно-практическое применение методики фотозарядового эффекта для разработки электронно-оптических средств, спектр видения у которых будет значительно превосходить восприятие человеческого глаза [6].

2. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Для проведения эксперимента был реализован экспериментальный стенд для исследования фотозарядового эффекта на поверхности различных материалов [1–5, 8, 9]. На рис. 1 представлена схема экспериментальной установки.

Излучение лазера 1 модулируется оптомеханическим модулятором 2 с частотой от 10 до 650 Гц. С помощью поворотных зеркал 3, через фокусирующую линзу 4 луч направляли в окно коробки 5, где находился исследуемый образец. После рассеивания лазерного луча на поверхности образца и электрода возникал электрический сигнал, который поступал на измерительный прибор 6 – синхронный детектор. Затем уровень сигнала увеличивался усилителем 7 и отображался на осциллографе 8. Детектор 6 и модулятор 2 были синхронизированы по частоте. Усилитель сигнала 7 представляет из себя один каскад с NPN-транзистором, включенным по схеме с общим эмиттером (рис. 2).

Эксперимент проводили в несимметричном режиме измерения, поскольку использовался только один электрод в экранирующей коробке (рис. 3).

Экранирующая коробка была сделана из алюминия, ее входное окно покрыто оксидом индия и олова (ИТО). Коробка 6 имела цилиндрическую форму с одним окном для BNC-разъема, расположенного симметрично. К разъему 4 был подключен электрод 3. Лазерное излучение через входное отверстие 5 попадало на исследуемый образец 1, вызывая перераспределение зарядов, расположенных вблизи поверхности, и накопление их на обратной неосвещенной стороне. При помощи конденсаторной связи на электроде 3 образовывался электрический сигнал, который через разъем 4 поступал на измерительный прибор. Внутри коробки также была помещена черная поглощающая бумага 2, расположенная между образцом 1 и электродом 3, она выполняла роль диэлектрика, тем самым исключая контакт между ними. Также она препятствовала попаданию излучения на электрод. Цилиндрическая коробка 6 была заземлена и экранирована от внешних электрических наводок. На входное окно для BNC-

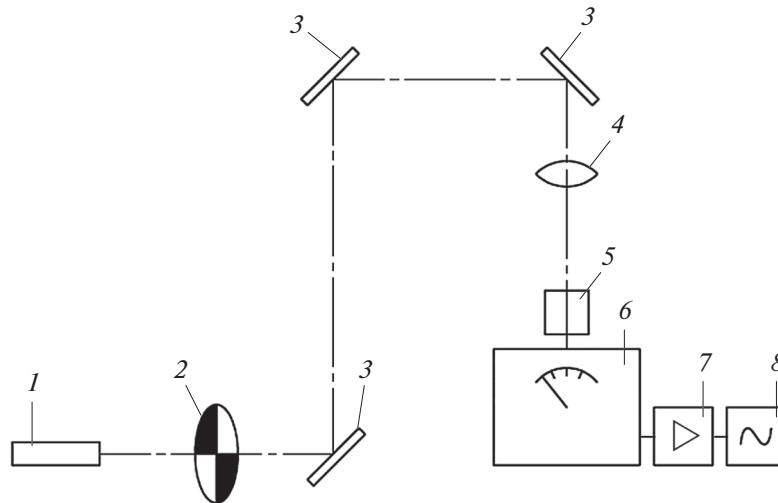


Рис. 1. Схема исследования фотозарядового эффекта: 1 – лазер, 2 – модулятор, 3 – поворотные зеркала, 4 – фокусирующая линза, 5 – коробка с исследуемым образцом, 6 – синхронный детектор, 7 – усилитель, 8 – осциллограф.

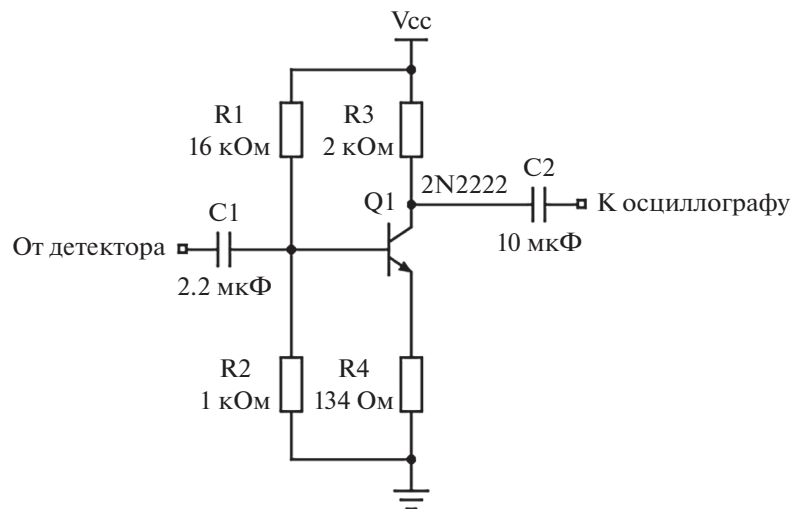


Рис. 2. Схема усилителя сигнала.

разъема было нанесено токопроводящее покрытие, для того чтобы разъем находился в контакте с коробкой.

Было изготовлено два электрода – один из меди и второй из алюминия. Однако в эксперименте использовался только медный электрод (рис. 4) поскольку медь имеет лучшие проводящие свойства, чем алюминий. У электродов с такой формой, как на рис. 4, плотность зарядов зависит от кривизны поверхности: растет с увеличением выпуклости поверхности и убывает с увеличением вогнутости. Заряды скапливаются на тех участках, где больше кривизна, т.е. у острия электрода. При изменении кривизны поверхности плотность заряда тоже будет изменяться.

Фотография экспериментальной установки показана на рис. 5. Источник излучения представлял собой гелий-неоновый лазер с длиной

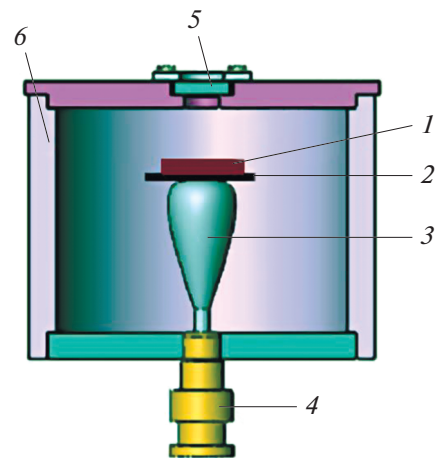


Рис. 3. Схема экранирующей коробки: 1 – исследуемый образец, 2 – черная бумага, 3 – электрод, 4 – BNC-разъем, 5 – входное окно, 6 – металлическая коробка.



Рис. 4. Внешний вид электрода из меди.

волны $\lambda = 632$ нм и мощностью 1 мВт. Модулятор позволял проводить прерывание луча лазера с частотой от 10 до 650 Гц. Сигнал с электрода измеряли при помощи синхронного детектора модели Stanford SR510. Для проведения эксперимента были использованы образцы кремния, арсенида галлия и германия (рис. 6).

На рис. 7а–7в представлены экспериментальные временные зависимости, а также частотные спектры для исследуемых образцов. В процессе эксперимента образцы облучали лазером с частотой модуляции 485 Гц. Можно увидеть, что нормированные по амплитуде спектры исследуемых сигналов с учетом погрешности имеют максимумы на частоте $485 \pm 1.5\%$ Гц, которая соответствует частоте модуляции. Таким образом, амплитуда сигнала на осциллографе свидетельствовала о регистрации фотозарядового эффекта на поверхности исследуемых материалов. Величина амплитуды за-



Рис. 5. Внешний вид экспериментальной установки.

висела от качества поверхности и от физико-химических свойств образцов. Амплитуда сигнала на пластинке кремния была выше, чем у образцов арсенида галлия и германия, поскольку пластинка была лучше технологически обработана (гладкая поверхность, собственная толщина 1 мм). Поэтому для количественного сравнения результатов эксперимента желательно, чтобы исследуемые образцы имели одинаковые геометрические параметры.

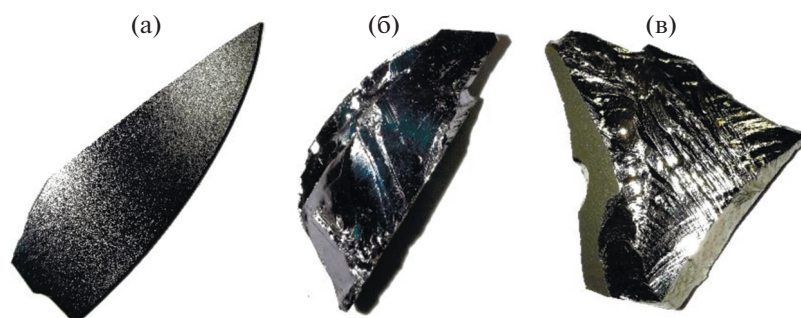


Рис. 6. Исследуемые образцы: кремний (а), арсенид галлия (б), германий (в).

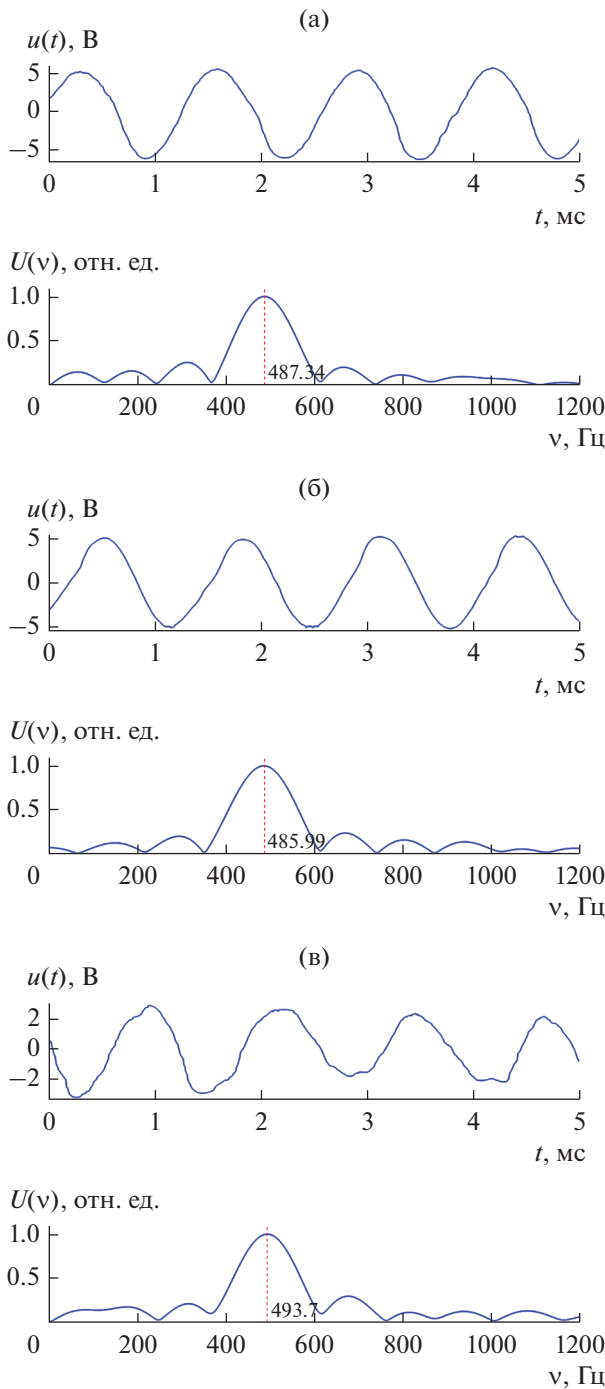


Рис. 7. Временные и частотные характеристики сигнала с электрода на частоте модуляции 485 Гц для образцов: а – кремний, б – арсенид галлия, в – германий.

Для более полного исследования полученных результатов был проведен их вейвлет-анализ [10], что позволило визуализировать частотно-временное распределение, а также локальные свойства. На рис. 8 представлены результаты непрерывного вейвлет-преобразования для исследуемых сигналов.

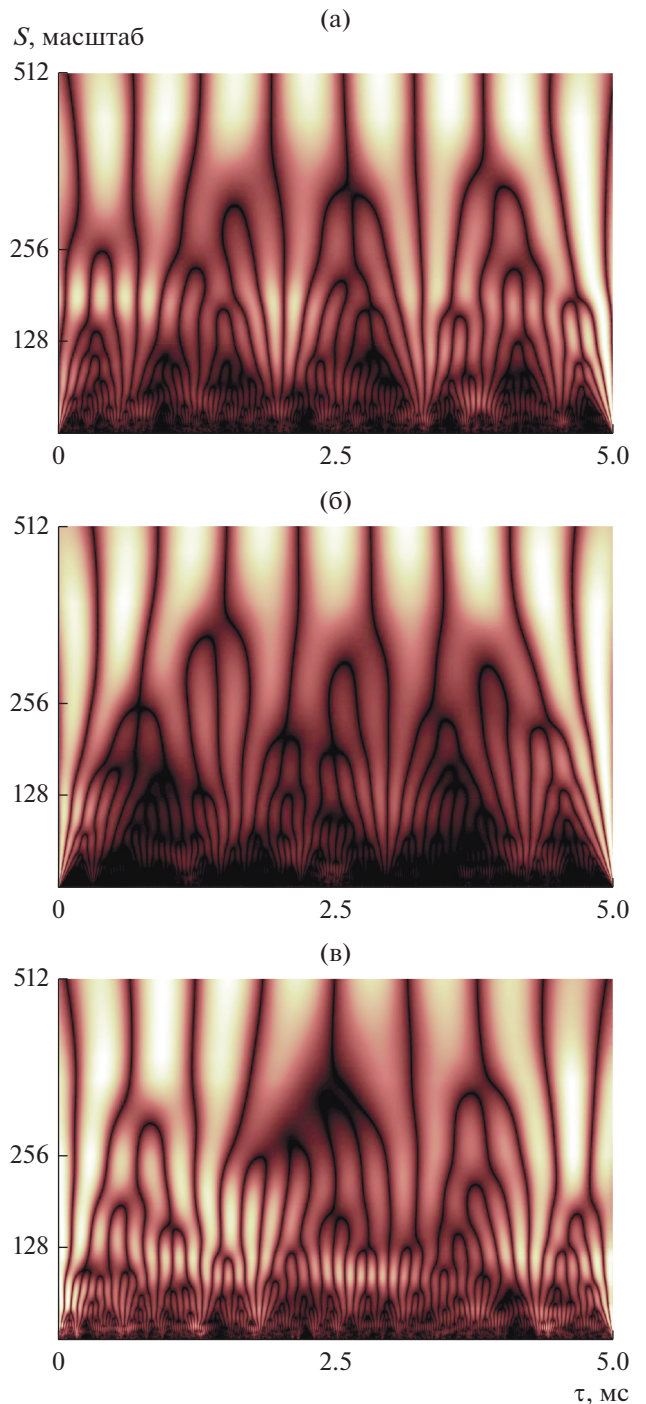


Рис. 8. Результаты вейвлет-анализа сигналов для следующих образцов: а – кремний, б – арсенид галлия, в – германий.

двух сигналов. Преобразование выполняли согласно формуле:

$$\tilde{U}(\tau, s) = \int_a^b U(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right) dt, \quad (4)$$

где $U(t)$ – анализируемый сигнал, s – масштаб вейвлет-функции, $\psi(t)$ – вейвлет-функция. В

данном примере в качестве вейвлет-функции выбран вейвлет Морле, который описывается формулой

$$\psi(t) = \exp\left(-\frac{t^2}{2}\right) \cos(5t). \quad (5)$$

Из результатов вейвлет-анализа для 512 уровней масштабирования определены частотно-временные свойства исследуемых сигналов: устойчивые частоты (основные и кратные), их временное распределение, а также наличие регулярной шумовой составляющей в области высоких частот (см. рис. 8а–8в, нижняя часть каждого рисунка).

ЗАКЛЮЧЕНИЕ

Таким образом, были получены следующие результаты.

1. Доказана возможность получения фотозарядового эффекта на поверхности полупроводниковых материалов с использованием созданной экспериментальной установки, а полученные результаты позволяют разработать методику для оценки качества используемых материалов при изготовлении ЭРЭ.

2. Реализован экспериментальный стенд для возможности исследования фотозарядового эффекта, который в целом позволяет проводить наблюдения на различных материалах, а не только на полупроводниках.

3. Получены зависимости амплитуд напряжений от времени для полупроводниковых матери-

алов при облучении их лазером, что свидетельствует о возникновении исследуемого эффекта на их поверхности.

4. Проанализированы спектральные характеристики зависимостей, а также построены частотно-временные преобразования на основе вейвлет-анализа.

СПИСОК ЛИТЕРАТУРЫ

1. *Pustovoit V.I., Borissov M., Ivanov O.* // Bulgarian J. Phys. 1990. V. 17. № 1. P. 32.
2. *Mam-Chun Park, Won-Gi Min, Kuhn-Il Lee* // Proc. 5th Int. Conf. on Properties and Applications of Dielectric Materials. Seoul. 25–30 May 1997. N.Y.: IEEE, 1997. V. 2. P. 952.
3. *Ivanov O., Mihailov V., Djulgerova R.* // Spectroscopy Lett. 2000. V. 33. № 3. P. 393.
4. *Курчанов А.Ф., Епихин Г.Е., Ефреев З.Л., Фаенов А.Я.* // Квантовая электрон. 1988. Т. 15. № 4. С. 720.
5. *Мартьянов П.С.* // Техника. Технологии. Инженерия. 2017. № 4(6). С. 1.
6. *Pustovoit V.I., Borissov M., Ivanov O.* // Sol. State Commun. 1989. V. 72. № 6. P. 613.
7. *Булатов М.Ф., Чуриков Д.В.* // Поверхность. Рентгеновские синхротронные и нейтронные исследования. 2019. № 3. С. 45. <https://doi.org/10.1134/S0207352819030041>
8. *Мартьянов П.С.* // РЭ. 2018. Т. 63. № 11. С. 1197. <https://doi.org/10.1134/S0033849418110049>
9. *Мартьянов П.С., Чуриков Д.В.* // РЭ. 2020. Т. 65. № 10. С. 1037. <https://doi.org/10.31857/S0033849420100046>
10. *Кравченко В.Ф., Пустовойт В.И., Чуриков Д.В.* // ДАН. 2011. Т. 436. № 5. С. 615.

К 90-ЛЕТИЮ АНАТОЛИЯ ВАСИЛЬЕВИЧА ВАШКОВСКОГО

DOI: 10.31857/S0033849422020164



19 декабря 2021 г. исполнилось 90 лет со дня рождения профессора, доктора физико-математических наук, лауреата Государственной премии СССР, члена редколлегии журнала “Радиотехника и электроника”, Анатолия Васильевича Вашковско­го.

А.В. Вашковский — известный ученый в области спинволновой электроники СВЧ и физики магнитных явлений, один из создателей целого ряда научных направлений, которые связаны с исследованием линейных и нелинейных процессов, возникающих при распространении магнитных волн в ферритовых средах и различных структурах на их основе.

Анатолий Васильевич родился в г. Опочка Псковской области в семье учителей. После окончания средней школы в Москве, он в 1950 г. поступил на радиотехнический факультет Московского энергетического института, по оконча-

нии которого был принят на работу в качестве младшего научного сотрудника в Институт радиотехники и электроники АН СССР (ныне — ИРЭ им. В.А. Котельникова РАН) в лабораторию, которой заведовал чл.-корр. АН СССР А.А. Пистолькорс. Анатолий Васильевич начал свою научную деятельность с измерения электрических параметров ферритов и исследования возможности создания параметрических ферромагнитных усилителей. После окончания в 1963 г. заочной аспирантуры он защитил диссертацию на соискание ученой степени кандидата технических наук, а в 1972 г. — диссертацию “Нестабильность магнитостатических волн в ферритах” на соискание ученой степени доктора физико-математических наук.

А.В. Вашковский является одним из основателей научного направления “Спинволновая электроника СВЧ”, лежащего на стыке радиофизики и физики магнитных явлений. Большинство научных достижений Анатолия Васильевича посвящено исследованию линейных и нелинейных характеристик и свойств дипольных и обменных спиновых волн (в том числе волн с неколлинеарной ориентацией групповой и фазовой скоростей) в различных анизотропных ферритовых средах и структурах. В 1988 г. за разработку научных основ спинволновой электроники А.В. Вашковский с коллегами был удостоен звания лауреата Государственной премии СССР.

А.В. Вашковский автор около 400 опубликованных научных трудов и изобретений. Вместе с Я.А. Моносовым он является автором открытия “Эффект взаимодействия СВЧ и ультразвуковых колебаний в ферритах” (зарегистрированного под № 42 в Государственном реестре открытий СССР с приоритетом от 21 марта 1963 г.). В соавторстве с В.С. Стальмаховым и Ю.П. Шаравским Анатолий Васильевич написал монографию “Магнитостатические волны в электронике сверхвысоких частот” (Изд-во Саратовского университета, 1993).

Большое значение А.В. Вашковский придавал подготовке кадров и научно-общественной дея-

тельности. Под его руководством защитили кандидатские диссертации более 20 человек, а пятеро впоследствии стали докторами наук. В 1982 г. Анатолию Васильевичу было присвоено научное звание профессора. В течение своей научной карьеры А.В. Вашковский руководил десятками НИР и принимал активное участие во многих научных конференциях, проводившихся в СССР, США,

Англии, Германии, Италии, Канаде, Венгрии, Польше и Болгарии, а также основал и возглавил Международную конференцию “Спинволновая электроника СВЧ”.

Коллеги, друзья и все, кто знает А.В. Вашковского, поздравляют его с юбилеем и желают ему крепкого здоровья и счастья.