

НАНО- И МИКРОСИСТЕМНАЯ ТЕХНИКА

ISSN 1813-8586

Том 20. №1. 2018

- Нанотехнологии
- Зондовая микроскопия
- Микромашины и наносистемы
- Молекулярная электроника
- Биоактивные нанотехнологии
- Элементы датчиков и биочипы
- Микроэлектромеханические системы
- Микрооптоэлектромеханические системы
- Биомикроэлектромеханические системы

NANO- and MICROSYSTEMS
TECHNOLOGY

Vol. 20, No.1, 2018

The articles are published
in Russian and English

Рисунки к статье В. С. Кондратенко, В. И. Иванова
«МЕТОД ПРЕЦИЗИОННОЙ ЛАЗЕРНОЙ РЕЗКИ КРЕМНИЕВЫХ ПЛАСТИН НА КРИСТАЛЛЫ С ОРГАНИЧЕСКИМИ СВЕТОИЗЛУЧАЮЩИМИ СТРУКТУРАМИ»

V. S. Kondratenko, V. I. Ivanov
«METHOD OF PRECISION LASER CUTTING OF SILICON WAFERS INTO CRYSTALS WITH ORGANIC LIGHT-EMITTING STRUCTURES »

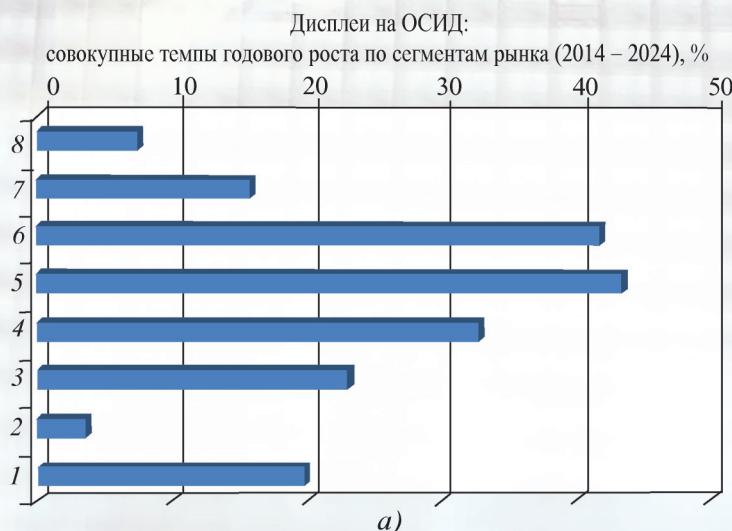


Рис. 1. Рост объемов продаж ОСИД дисплеев за 2014 – 2024 гг.:

a – темпы годового роста по сегментам рынка:

1 – микродисплеи; 2 – другие; 3 – промышленные/профессиональные устройства; 4 – портативная электроника; 5 – автомобильная и аэрокосмическая промышленность; 6 – Телевизоры; 7 – планшеты и ноутбуки; 8 – мобильные телефоны;

b – прогноз рынка устройств на ОСИД

Fig. 1. Growth of sales of OLED displays in 2014–2024:

a – rates of annual growth by market segments:

1 – Microdisplays; 2 – Others; 3 – Industrial/Professional; 4 – Wearable Electronics; 5 – Automotive & Aerospace; 6 – TV; 7 – Tablets & Notebooks; 8 – Mobile phones; *b* – forecast of the market of the devices on OLED

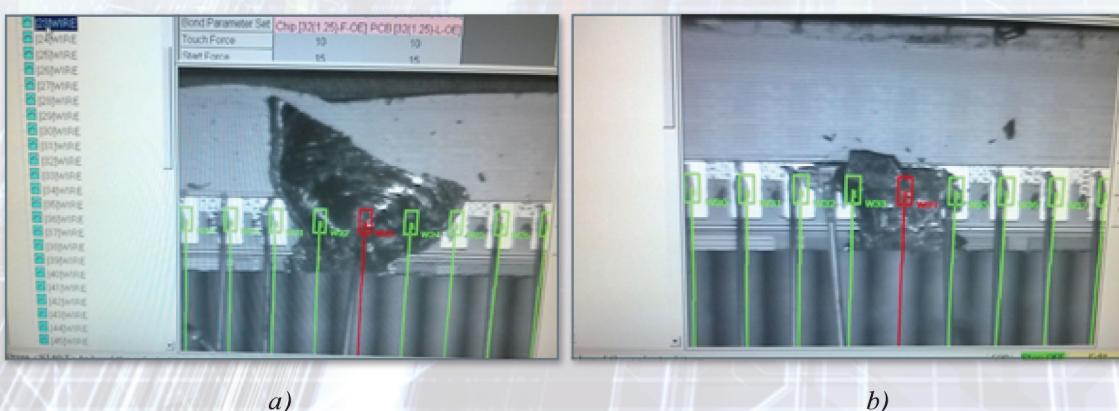


Рис. 3. Фотографии кромок кристаллов со сколами после дисковой алмазной резки

Fig. 3. Photos of the edges of crystals with cleavages after a disk diamond cutting

НАНО- И МИКРОСИСТЕМНАЯ ТЕХНИКА

Том 20. № 7 ◆ 2018

ЕЖЕМЕСЯЧНЫЙ МЕЖДИСЦИПЛИНАРНЫЙ ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Журнал включен в международные базы данных на платформе Web of Science: Chemical Abstracts Service (CAS), которая входит в Medline, и Russian Science Citation Index (RSCI).

Журнал индексируется в системе Российского индекса научного цитирования (РИНЦ) и включен в международную базу INSPEC.

Журнал включен в Перечень международных реферируемых баз данных по научному направлению 02.00.00 химические науки

и в Перечень научных и научно-технических изданий ВАК России по научным направлениям: 01.04.00 физика, 05.27.00 электроника.

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук

Статьи имеют DOI и печатаются в журнале на русском и английском языках

ISSN 1813-8586 DOI: 10.17587/issn1813-8586

Главный редактор

Мальцев П. П., д.т.н., проф.

Зам. гл. редактора

Лучинин В. В., д.т.н., проф.

Шур М., д.ф.-м.н., проф. (США)

Редакционный совет:

Аристов В. В., д.ф.-м.н., проф., чл.-кор. РАН

Асеев А. Л., д.ф.-м.н., проф., акад. РАН

Грибов Б. Г., д.х.н., чл.-кор. РАН

Ковалчук М. В., д.ф.-м.н., проф., чл.-кор. РАН

Кульчин Ю. Н., д.ф.-м.н., проф., акад. РАН

Лабунов В. А., д.т.н., проф., акад. НАНБ (Беларусь)

Рыжий В. И., д.ф.-м.н., проф., чл.-кор. РАН

Сауров А. Н., д.т.н., проф., акад. РАН

Сигов А. С., д.ф.-м.н., проф., акад. РАН

Чаплыгин Ю. А., д.т.н., проф., акад. РАН

Шевченко В. Я., д.х.н., проф., акад. РАН

Редакционная коллегия:

Абрамов И. И., д.ф.-м.н., проф. (Беларусь)

Андреев А., к.ф.-м.н. (Великобритания)

Астахов М. В., д.х.н., проф.

Бакланов М. Р., д.х.н., проф. (Китай)

Басаев А. С., к.ф.-м.н.

Викулин В. В., к.х.н., д.т.н., проф.

Горнев Е. С., д.т.н., проф.

Карякин А. А., д.х.н., проф.

Кузнецов В. И., д.т.н. (Нидерланды)

Леонович Г. И., д.т.н., проф.

Панин Г. Н., к.ф.-м.н., проф. (Южная Корея)

Панич А. Е., д.т.н., проф.

Пожела К., д.ф.-м.н. (Литва)

Рыжий М. В., д.т.н., проф. (Япония)

Сантос Э. Х. П., PhD, Ful. Prof. (Бразилия)

Сингх К., к.т.н. (Индия)

Телец В. А., д.т.н., проф.

Хабибуллин Р. А., к.ф.-м.н.

Шашкин В. И., д.ф.-м.н., проф.

Шубарев В. А., д.т.н., проф.

Редакция:

Антонов Б. И. (директор изд-ва)

Лысенко А. В. (отв. секретарь)

Чугунова А. В.

Фокин В. А., к.х.н. (ред. перевода)

Щетинкин Д. А. (сайт)

СОДЕРЖАНИЕ

Издается с 1999 г.

НАНОТЕХНОЛОГИИ И ЗОНДОВАЯ МИКРОСКОПИЯ

- Ларионов Ю. В., Озерин Ю. В. Разворот изображения выступа при его многократном сканировании в растром электронном микроскопе 387

МОДЕЛИРОВАНИЕ И КОНСТРУИРОВАНИЕ МНСТ

- Мустафаев А. Г., Мустафаев Г. А., Черкесова Н. В. Учет воздействия полной дозы излучения в компактных моделях КМОП-транзисторов 396

- Садков В. Д., Пилькевич А. В., Фомина К. С. Моделирование процесса лазерной подгонки пленочных поглощающих элементов 402

МАТЕРИАЛОВЕДЧЕСКИЕ И ТЕХНОЛОГИЧЕСКИЕ ОСНОВЫ МНСТ

- Кондратенко В. С., Иванов В. И. Метод прецизионной лазерной резки кремниевых пластин на кристаллы с органическими светоизлучающими структурами 409

- Борзов П. А., Филиппов С. Е., Тополов В. Ю., Брыль О. Е., Панич А. Е. Электромеханические свойства и акустические характеристики композита сегнетопьезокерамика — корундовая керамика — воздух 415

ЭЛЕМЕНТЫ МНСТ

- Ильков В. К., Михалев А. О., Майтама М. В. Коммутаторы на арсениде и нитриде галлия 425

МОЛЕКУЛЯРНАЯ ЭЛЕКТРОНИКА И БИОЭЛЕКТРОНИКА

- Горнев Е. С., Теплов Г. С. Математическая модель конечного автомата абстрактного нейрона и сетей на его основе 434

Аннотации и статьи на русском и английском языках доступны на сайте журнала (<http://microsystems.ru>; <http://novtex.ru/nmst/>) в разделе "Архив статей с 1999 г.".

ПОДПИСКА:

по каталогу Роспечати (индекс 79493);

по каталогу "Пресса России" (индекс 27849)

в редакции журнала (тел./факс: (499) 269-55-10)

Адрес для переписки:

107076 Москва,

Стромынский пер., д. 4

e-mail: nmst@novtex.ru

Учредитель:

Издательство "Новые технологии"

NANO- and MICROSYSTEMS TECHNOLOGY

(Title "NANO- I MIKROSISTEMNAYA TEKHNIKA")

ISSN 1813-8586 DOI: 10.17587/issn1813-8586

Vol. 20
No. 7
2018

CHIEF EDITOR

Maltsev P. P., Dr. Sci. (Tech.), Prof.

DEPUTY CHIEF EDITOR

Luchinin V. V., Dr. Sci. (Tech.), Prof.

DEPUTY CHIEF EDITOR

Shur M. S., Dr. Sci. (Phys.-Math.), Prof. (USA)

Editorial council:

Aristov V. V., Dr. Sci. (Phys.-Math.), Prof., Cor.-Mem. RAS
 Aseev A. L., Dr. Sci. (Phys.-Math.), Prof., Acad. RAS
 Chaplygin Ju. A., Dr. Sci. (Tech.), Prof., Acad. RAS
 Gribov B. G., Dr. Sci. (Chem.), Cor.-Mem. RAS
 Kovalchuk M. V., Dr. Sci. (Phys.-Math.), Prof., Cor.-Mem. RAS
 Kuljchin Yu. N., Dr. Sci. (Phys.-Math.), Prof., Acad. RAS
 Labunov V. A. (Belorussia), Sci. (Tech.), Acad. NASB
 Ryzhii V. I., Dr. Sci. (Phys.-Math.), Prof., Cor.-Mem. RAS
 Saurov A. N., Dr. Sci. (Tech.), Prof., Acad. RAS
 Shevchenko V. Ya., Dr. Sci. (Chem.), Prof., Acad. RAS
 Sigov A. S., Dr. Sci. (Phys.-Math.), Prof., Acad. RAS

Editorial board:

Abramov I. I. (Belorussia), Dr. Sci. (Phys.-Math.), Prof.
 Andreev A. (UK), Cand. Sci. (Phys.-Math.), Prof.
 Astahov M. V., Dr. Sci. (Chem.), Prof.
 Baklanov M. R., Dr. Sci. (Chem.), Prof. (China)
 Basaev A. S., Cand. Sci. (Phys.-Math.)
 Gornev E. S., Dr. Sci. (Tech.), Prof.
 Karjakin A. A., Dr. Sci. (Chem.), Prof.
 Khabibullin R. A., Cand. Sci. (Phys.-Math.)
 Kuznetsov V. I., Dr. Sci. (Tech.) (Netherlands)
 Leonovich G. I., Dr. Sci. (Tech.), Prof.
 Panich A. E., Dr. Sci. (Tech.), Prof.
 Panin G. N., PhD, Prof. (South Korea)
 Pozhela K. (Lithuania), Dr. Sci. (Phys.-Math.)
 Ryzhii M. V., (Japan), Dr. Eng., Prof.
 Santos E. J. P., PhD, Prof. (Brasil)
 Shubarev V. A., Dr. Sci. (Tech.), Prof.
 Shashkin V. I., Dr. Sci. (Phys.-Math.), Prof.
 Singh K., PhD (India)
 Telets V. A., Dr. Sci. (Tech.), Prof.
 Vikulin V. V., Cand. Chem. Sci., Dr. Sci. (Tech.), Prof.

Editorial staff:

Antonov B. I. (Director Publ.)
 Lysenko A. V. (Executive secretary)
 Chugunova A. V.
 Fokin V. A., Cand. Sci. (Chem.)
 Shchetinkin D. A. (site)

The Journal is included in the international databases of the chemical sciences — Chemical Abstracts Service (CAS) and of the engineering sciences — INSPEC, and it is also indexed in the Russian Science Citation Index (RSCI) based on the Web of Science platform. The Journal is included in the Russian System of Science Citation Index and the List of Journals of the Higher Attestation Commission of Russia. Its articles have DOI and are printed in the Journal in Russian and English languages. The Journal is published under the scientific-methodical guidance of the Branch of Nanotechnologies and Information Technologies of the Russian Academy of Sciences.

CONTENTS

NANOTECHNOLOGY AND SCANNING PROBE MICROSCOPY

Larionov Yu. V., Ozerin Yu. V. Rotation of an Image of the Protrusion at its Repeated Scanning in a SEM 392

MODELLING AND DESIGNING OF MNST

Mustafaev A. G., Mustafaev G. A., Cherkesova N. V. Consideration of the Effect of the Total Dose of Radiation in the Compact Models of the CMOS Transistors 399

Sadkov V. D., Pilkevich A. V., Fomina K. S. Modeling of the Laser Trimming Process of the Film Absorbing Elements 406

SCIENCE OF MATERIALS AND TECHNOLOGICAL BASICS OF MNST

Kondratenko V. S., Ivanov V. I. Method of Precision Laser Cutting of Silicon Wafers into Crystals with Organic Light-Emitting Structures 413

Borzov P. A., Filippov S. E., Topolov V. Yu., Brill O. E., Panich A. E. Electromechanical Properties and Acoustic Characteristics of a Composite "Ferroelectric Piezo-Ceramic — Corundum Ceramic — Air" 421

MICRO- AND NANOSYSTEM TECHNIQUE ELEMENTS

Il'kov V. K., Mikhalev A. O., Maytama M. V. Arsenide and Nitride Gallium Switches 431

MOLECULAR ELECTRONICS AND BIOELECTRONICS

Gornev E. S., Teplov G. S. Mathematical Model of a Finite Automation of an Abstract Neuron and the Networks Based on It 442

Our:

Web: www.microsystems.ru/eng
 e-mail: nmst@novtex.ru

To subscribe, please contact with:

JSC "MK-Periodica":
 Tel: +7 (495) 672-7012
 Fax: +7 (495) 306-3757
 E-mail: import@periodicals.ru

НАНОТЕХНОЛОГИИ И ЗОНДОВАЯ МИКРОСКОПИЯ

NANOTECHNOLOGY AND SCANNING PROBE MICROSCOPY

УДК 537.533.35

DOI: 10.17587/nmst.20.387-395

Ю. В. Ларионов, ст. науч. сотр., luv@kapella.gpi.ru,
Институт общей физики им. А. М. Прохорова РАН, 119991, Москва,
Ю. В. Озерин, вед. инженер, yozerin@mikron.ru,
ПАО "Микрон", Москва, 124460, Зеленоград

РАЗВОРОТ ИЗОБРАЖЕНИЯ ВЫСТУПА ПРИ ЕГО МНОГОКРАТНОМ СКАНИРОВАНИИ В РАСТРОВОМ ЭЛЕКТРОННОМ МИКРОСКОПЕ

Поступила в редакцию 25.12.2017

Выявлено, что при многократном сканировании выступа в кремниевой пластине в растровом электронном микроскопе (РЭМ) его изображение изменяет свой наклон, как бы поворачиваясь вокруг оси, перпендикулярной плоскости изображения. Сделано предположение, что это явление связано с наведением электрического заряда в подповерхностном слое выступа и его влиянием на формирование изображения. Предложен механизм разворота изображения, оценено влияние обнаруженного явления на значения линейных величин при РЭМ-измерениях.

Ключевые слова: нанометрология, низковольтный растровый электронный микроскоп, разворот изображения, наведенный электрический заряд, ширина структуры

Введение

В работе [1] было выявлено, что при многократном сканировании в растровом электронном микроскопе (РЭМ) выступа, расположенного на кремниевой пластине, происходит искажение его изображения, приводящее к уменьшению измеренного значения его ширины. Это искажение изображения выступа было объяснено наведением электрического заряда в подповерхностном слое выступа сканирующим пучком и влиянием наведенного локализованного заряда на эмиссию из поверхности электронов, формирующих изображение этого выступа. В данной работе выявилось, что искажение изображения выступа проявляется также в развороте изображения выступа вокруг оси, перпендикулярной плоскости изображения. Цель работы — исследовать детали, объяснить причины обнаруженного явления и оценить его влияние на измерение линейных размеров.

Эксперимент

Исследования проводили для двух типов выступов на кремниевой пластине с пологими боковыми стенками: на выступах тест-объекта меры МШПС-2К [2] и на выступах с боковыми стенками, сформированными с помощью химического

травления. Выступы первого типа формировали анизотропным травлением пластины на глубину ~0,41 мкм при периоде их расположения на пластине 2000 нм при угле наклона боковой стенки 54,7°. Изображения выступов были получены с помощью РЭМ S4800 в медленных вторичных электронах (МВЭ) при энергии электронов падающего пучка 20 кэВ. На рис. 1 представлены конечные изображения после непрерывного сканирования в течение времени 20 и 160 с. На рис. 1, b заметен разворот изображения выступа против часовой стрелки по сравнению с его изображением на рис. 1, a.

Мы зафиксировали изменение направления разворота изображения на противоположное в результате смены режима сканирования: с паузами в 2 мин между отдельными кадрами продолжительностью 10 с (рис. 2, a и b).

Для выступов второго типа глубина травления составила ~0,35 мкм, ширина верхнего основания (ВО) одиночного выступа ~290 нм, угол наклона боковой стенки ~65°. Изображения были получены с помощью измерительного РЭМ S-9260A при увеличении ~150.000x, энергии падающих электронов 800 В, токе пучка ~6 пА, числе пикселей на изображении 512 × 512 при размере пикселя 1,75 нм.

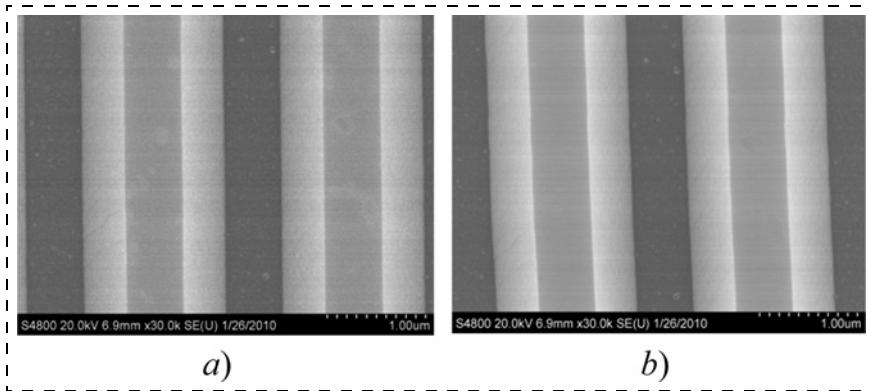


Рис. 1. Вариация наклона на изображении выступа после непрерывного сканирования в течение времени сканирования 20 (a) и 160 с (b). Увеличение — 30 000x
Fig. 1. Variation of the protrusion image after an uninterrupted scanning during time of 20 s (a) and 160 s (b). Magnification — 30 000x

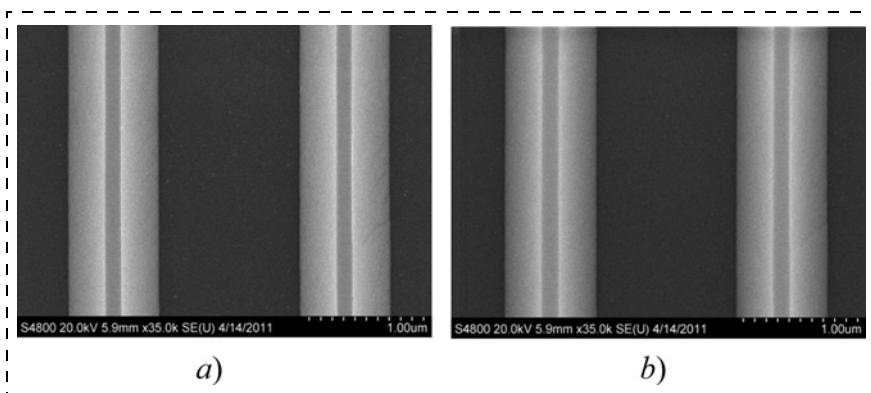


Рис. 2. Вариация наклона на изображении выступа после сканирования с паузами: первый кадр (a), 35-й кадр (b). Увеличение — 35 000x
Fig. 2. Variation of inclination the protrusion image after scanning with pauses: the first frame (a), the 35th frame (b). Magnification — 35 000x

Исследуемый выступ сканировали многократно. Число непрерывных сканирований (кадров) одного выступа составляло 400 или 800. В каждом цикле формировалось всего 25 изображений (сканов) выступа. При 400-кратном сканировании один скан формировался усреднением по 16 кадрам, при 800-кратном сканировании — по 32 кадрам, что увеличивало четкость деталей изображения в последнем случае. Было выполнено несколько циклов 800-кратного сканирования. Выявилось, что значение разворота изображений, полученных в различных циклах сканирования, заметно варьируется. Причем изменялось не только значение разворота, но даже его направление. Поэтому для выявления закономерностей в феномене разворота изображения было выполнено семь циклов непрерывных сканирований с промежутками между ними от нескольких часов до нескольких суток.

Методика измерений наклона изображения выступов

Вариацию наклонов изображений выступов оценивали количественно с помощью программы

обработки изображений [3]. С ее помощью изображение разбивалось на несколько десятков полос — секций, расположенных вдоль горизонтальной оси. В каждой из них фиксировались кривые видеосигнала (ВС) от всех строк развертки внутри нее, после чего формировались усредненные кривые ВС для каждой из секций. На рис. 3 показана кривая ВС, усредненная по всем строкам. Видно, что ее можно разделить на несколько квазилинейных участков (между точками 1—2, 2—3, 4—5, 5—6). Затем каждый из этих участков аппроксимировался прямой линией [4]. И наконец, на пересечениях соседних прямых выделялись координаты опорных точек: начала роста амплитуды (1); точек изменения наклона участков кривой (2 и 5); точек положения максимумов модельной ВС (3 и 4); точек спада амплитуды ВС до фонового значения (6). Согласно [4] расстояния между этими точками позволяют определить все линейные размеры выступа. Ширину выступа по границам верхнего основания следует определять по расстоянию W между точками А и В кривой видеосигнала (ВС) на рис. 3. Точки располагаются на серединах участков спада ВС от ее пиков к участку, соответствующему верхнему основанию выступа.

После выделения опорных точек формировалась зависимость их горизонтальных координат по длине выступа (по номерам секций разбиения). Полученные зависимости аппроксимировались прямыми методом наименьших квадратов с использованием программы Origin 7.0, после чего с ее же помощью вычислялся тангенс наклона этих прямых на изображении. Наклоны линий способны влиять на значение линейных величин, определяющих размеры выступа. Поэтому вместе с тангенсами наклонов вычислялись расстояния между пиками L_p , расстояния между точками 1, 2 и 5, 6, а также ширина верхнего основания выступа W .

Ширина W может быть измерена непосредственно по кривой ВС (см. рис. 3), а также вычислена по данным измерения L_p и значению диаметра пучка (см. далее). Непосредственно измеренная W также включает в себя значение эффективного диаметра в скрытой форме [4]. Было выявлено, что значение этого, другого типа эффективного диа-

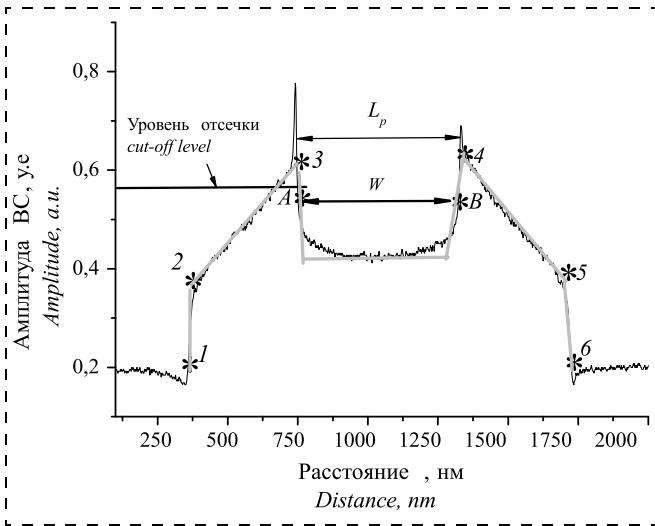


Рис. 3. Кривая ВС от выступа с обозначением точек 1 – 6 и аппроксимирующие ее прямые (серые). Расстояния между точками 1, 2 и 5, 6 вдоль горизонтальной оси определяют диаметр электронного пучка, расстояние W между точками А и В соответствует ширине ВО [4]; L_p – расстояние между пиками кривой

Fig. 3. VS curve from the protrusion with marked points 1 – 6 and straight lines (grey) approximating it. The distances between points 1, 2 and 5, 6 along the horizontal axis define the diameter of the electron beam, distance W between points A and B corresponds to the width of TB [4]; L_p – distance between the curve peaks

метра отличается от значения, получаемого по точкам на рис. 3 [5].

Результаты

В результате расчетов были оценены тангенсы углов наклона прямых, соединяющих наборы точек от 1 до 6 в секциях разбиения. Для разворотов изображений выступов, полученных в высоковольтном РЭМ и показанных на рис. 1, значения тангенсов были усреднены по всем линиям, соединяющим точки 1–6: для рис. 1, *a* усредненное значение тангенса составило 0,0028, для рис. 1, *b* с наибольшим в данной работе наклоном линии – 0,0157 (что соответствует углам разворота $\sim 0,2^\circ$ и $\sim 0,9^\circ$).

На низковольтном РЭМ наиболее заметные развороты выступа были получены при усреднении сканов по 16 кадрам. В табл. 1 представлены тангенсы углов наклона прямых (T_g), соединяющих соответствующие точки от 1 до 6 и случайные погрешности их определения σ . Последняя величина определяет степень близости аппроксимирующих прямых к линиям соответствующих им участков кривых ВС.

Как видно, погрешности оценки тангенсов наклонов сильно отличаются для различных точек. Погрешности определения наклонов для точек 2 и 5 особенно велики, что ведет к увеличенной случайной погрешности значений эффективного диаметра. Наиболее воспроизводимы наклоны для то-

чек 3 и 4, вносящие вклад в погрешность измерения ширины выступа.

Из данных табл. 1 следует важный вывод: наклоны прямых, соединяющих соответствующие точки на изображении, не одинаковы даже в первом скане (после 16 кадров). Это означает, что прямые, соединяющие точки 1, 6, уже в первом скане, не параллельны, а поэтому форма соответствующих участков кривых ВС в разных секциях изображения неодинакова (сжата или растянута). В результате многократного (от первого до 25-го сканов) сканирования наклоны изображений линий, как видно, изменяются. Заметна тенденция к уменьшению наклонов линий в 25-х сканах для большинства наборов точек в табл. 1, что зрительно соответствует развороту изображения выступа по часовой стрелке. Наиболее заметно уменьшаются наклоны линий, объединяющих точки 2 и 6, наименее заметно – линий, объединяющих точки 3 и 4, где значение ВС наибольшее.

Неопределенность измерения линейных величин

Ширина топологического элемента ИС является важнейшей контролируемой линейной величиной. Она определяется как расстояние между парой выделенных точек на кривой ВС (с учетом увеличения РЭМ). Такие точки обычно расположены симметрично относительно центра кривой ВС и задаются значением сигнала, называемым "уровнем отсечки" [4] (см. рис. 3). В работе [4] показано, как для модельной кривой выбрать уровень отсечки так, чтобы выделяемые им на кривой точки точно соответствовали границам верхнего основания модельного трапецидального выступа (а расстояние между ними – его ширине). Однако реально соответствие между положением задаваемых то-

Таблица 1
Table 1
Значения тангенсов наклона T_g и их воспроизводимость σ для прямых, соединяющих точки типа 1, 2 и до 6 на кривых видеосигнала на первом и последнем сканах в непрерывном цикле сканирования

Номер скана Scan number	Величина Quantity	Параметры прямой, объединяющих точки Parameters of the straight line, connecting the points					
		1	2	3	4	5	6
1	T_g	0,0085	0,0128	0,0093	0,0089	0,0058	0,0115
	σ	0,0012	0,0059	0,0006	0,0008	0,0026	0,0014
25	T_g	0,0060	0,0106	0,0097	0,0085	0,0059	0,0089
	σ	0,0014	0,0051	0,0011	0,0009	0,0017	0,0017

чек на кривой и краев основания обычно отсутствует [6].

Важнейшим условием определения ширины ВО выступа в нанометрологии является учет значения диаметра электронного пучка, формирующего изображение [4]. Значение диаметра используется в форме поправки, позволяющей определить ширину ВО, если измерять расстояние между пиками кривой (см. рис. 3) [4]. Измерение диаметра электронного пучка, эмитируемого поверхностью и формирующего изображение, является, по-видимому, самой важной и наиболее сложной задачей РЭМ-метрологии. Погрешность в значении диаметра пучка особенно критична при измерении размеров в диапазоне порядка 10 нм с помощью РЭМ с малым диаметром падающего пучка (т. е. с высоким разрешением).

В нашей стране диаметр пучка, формирующего РЭМ-изображение, предложено определять по кривой ВС, получаемой при сканировании выступа меры ширины МШПС-2К (в рамках используемой модели формирования изображения выступа) [4]. В рамках этой модели диаметр пучка не зависит от каких-либо влияющих на его значение величин и априорно совпадает с параметрами падающего пучка. Размер этого модельного пучка назван эффективным диаметром (d_{ef}) [4]. Использование этой величины позволяет определить ширину ВО выступа относительно просто и точнее, чем по произвольным точкам кривой ВС. В соответствии с [4] значение d_{ef} определяется усредненным расстоянием вдоль оси сканирования между точками 1—2 и 5—6 (см. рис. 3). Значение d_{ef} определяет ширину выступа W :

$$W = L_p - d_{ef}, \quad (1)$$

где L_p — расстояние между пиками на кривой ВС (см. рис. 3) [4]. Поэтому для оценки изменения значения ширины выступа при развороте изображения необходимо определять вариацию d_{ef} и L_p .

Таблица 2
Table 2

Значения Δd_{ef} , ΔL_p и ΔW для семи циклов сканирования
Values of Δd_{ef} , ΔL_p and ΔW for seven cycles of scanning

Цикл Cycle	Δd_{ef} , nm	ΔL_p , nm	ΔW , nm
1	1,7	-0,5	-1,1
2	0,7	-0,5	-0,6
3	2,0	-0,5	-0,7
4	0,9	0,7	-0,4
5	0,5	-0,2	-0,7
6	2,3	-0,1	-0,6
7	0,2	0,0	-0,6
Среднее Average	1,2	-0,16	-0,7

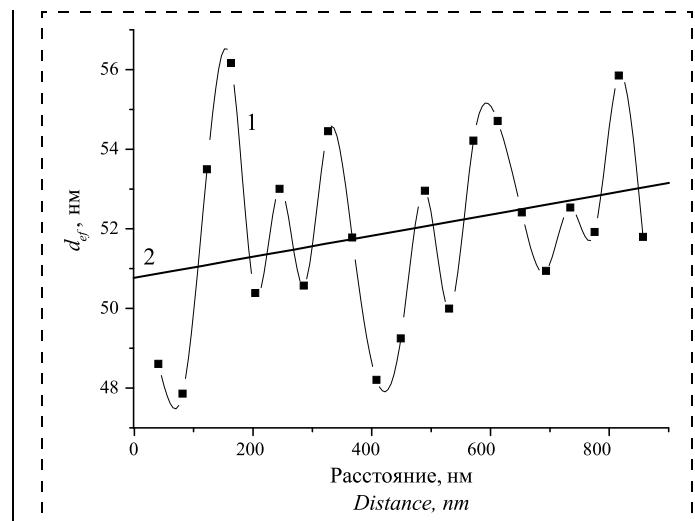


Рис. 4. Вариация d_{ef} вдоль длины выступа (1). Прямая 2 отображает закономерный рост d_{ef} от верха к низу изображения

Fig. 4. Variation d_{ef} along the length of the protrusion (1). Straight line 2 presents a tendency of growth d_{ef} from the top to the bottom of the image

Вследствие плохой воспроизводимости наклонов прямых (см. табл. 1 наборы точек 1—2 и 5—6) мы оценивали изменение d_{ef} по результатам сканирования выступа в семи циклах (цикл состоял из 800 непрерывных сканирований). Результаты оценки вариации величин d_{ef} , L_p в виде разницы значений, полученных по 25-му и первому изображениям (Δd_{ef} , ΔL_p) в семи циклах представлены в табл. 2. Однако результаты определения ΔW получены непосредственно по кривой ВС (не с помощью (1)). Случайная погрешность измерения ширины выступа (2σ) в серии кадров (и соответствующих кривых ВС) составляла 0,2...0,4 нм. При усреднении сканов-изображений по 32 кадрам случайная погрешность ширины в скане не должна превышать 0,1 нм. Поэтому вариацию ΔW в несколько десятых нанометра мы связываем с систематическим изменением ширины в разных сканах и циклах.

Погрешность d_{ef} должна быть подвержена случайным изменениям в наибольшей степени. На рис. 4 показана вариация d_{ef} по полосам одного изображения вдоль длины выступа (при усреднении его по 32 кадрам).

Видно, что стандартное отклонение от среднего d_{ef} по длине выступа близко к ~0,3 нм (2σ). При усреднении его по всем полосам изображения оно должно быть меньше 0,1 нм. Заметно изменение среднего значения d_{ef} в пределах одного изображения (вдоль прямой) на ~2 нм. Однако увеличение значения среднего d_{ef} по 25 изображениям составило всего 1,7 нм (табл. 2, цикл 1). По-видимому, лишь доли наведенных добавок к d_{ef} в каждом изображении суммируются по всем изображениям в пределах одного цикла.

Из данных табл. 2 видно, что многочисленные сканирования выступа внутри каждого цикла приводят к увеличению эффективного диаметра d_{ef} (в среднем на $\sim 1,2$ нм) при уменьшении расстояния между пиками (в среднем на $\sim 0,16$ нм). В соответствии с (1) эти вариации должны приводить к уменьшению ширины выступа почти на $\sim 1,4$ нм. Но по данным табл. 2 уменьшение значений ширины ΔW составляет всего $\sim 0,7$ нм. Разницу мы объясняем отличием значения эффективного диаметра d_{ef} от его значения, которое неявно определяет ширину выступа непосредственно при измерениях по кривой ВС. В работе [6] было указано, что при сканировании ВО сфокусированным пучком в эксперименте методически неверно определять его ширину по значениям эффективного диаметра d_{ef} . Дело в том, что значение d_{ef} в модели [4] определяется по участкам кривой ВС, формируемым деталями профиля выступа около нижнего основания (так как падающий и эмитированные электронные пучки в модели параллельны).

Развороты исходных и конечных изображений в разных циклах, которые использовались для получения результатов в табл. 2, происходили против часовой стрелки с большими разбросами по углу (в одном случае зафиксирован разворот по часовой стрелке). Это свидетельствует о заметном влиянии на значение разворота случайных неучтенных величин. При этом их влияние на вариацию d_{ef} и L_p различно: размах d_{ef} по всем циклам сканирований составил 1,5 нм, в то время как размах L_p (кроме четвертого, где зафиксирован выброс) — 0,5 нм. Эти значения размаха коррелируют с воспроизведимостью тангенсов углов наклона прямых, определяющих значения эффективного диаметра и расстояний между пиками кривой ВС в табл. 1 (для циклов 1, 2, 5, 6 и соответственно 3, 4). Из табл. 2 видно также, что размах ΔW (непосредственно измеряемой ширины ВО) составил только 0,7 нм, т. е. расчеты ширины с использованием выражения (1) менее воспроизводимы.

Причина вариации наклона

Причиной вариации наклона края изображения выступа является, по-видимому, электрический заряд, локально наводимый сканирующим пучком на поверхности выступа. Пучок, сканирующий строку изображения N , наводит заряд предположительно не только в полосе сканирования, но также и в полосе поверхности, где в последующем произойдет сканирование строки $N + 1$. Наведенный заряд не успевает релаксировать за время сканирования строки N и изменяет эмиссию МВЭ в полосе развертки строки $N + 1$. Эмиссионный ток для строки $N + 1$ может либо увеличиваться, либо уменьшаться. Если в последующей строке

развертки он больше, то ВС в соответствующих точках начнет расти раньше и положение этих точек сместится по координате X влево. В этом случае изображение выступа развернется по часовой стрелке. Если эмиссия уменьшится, то точки сдвинутся вправо, а разворот изображения произойдет против часовой стрелки. Разные направления разворота изображения выступа подтверждены экспериментально при разных режимах сканирования одного выступа (см. рис. 1, 2). Факт монотонной вариации эмиссии электронов из поверхности выступа в ходе сканирования с формированием многих кадров изображения представлен в работе [7]. В этой работе сканирование на высоковольтном РЭМ проводили в двух режимах: при непрерывном сканировании и с паузами между отдельными кадрами. Оказалось, что при непрерываемом сканировании амплитуда на кривой ВС на последнем кадре уменьшалась по сравнению с первым кадром, а при сканировании с паузами — наоборот.

Заключение

Длительное сканирование выступа с пологими боковыми стенками в кремнии в РЭМ приводит к различным искажениям его РЭМ-изображения. В данной работе представлен один из видов искажения — разворот изображения выступа (в направлении и по часовой стрелке и против нее) на угол до 1° при числе сканирований несколько сот раз. Развороты изображений происходят как в низковольтном, так и в высоковольтном РЭМ. Причиной разворота изображения является, по видимому, электрический заряд, наводимый сканирующим пучком в подповерхностном слое выступа. Предложен механизм влияния заряда на разворот изображения с учетом направления его вращения.

В связи с разворотом изображения значение ширины верхнего основания выступа может уменьшиться на 2,4 нм. (Отметим, что абсолютная погрешность измерения ширины указанных участков кривой должна быть ± 1 нм [8].) Выявилось, что в ходе разворота происходит неодинаковое смещение опорных точек (между которыми измеряется расстояние на одной строке кривой видеосигнала) в разных строках одного кадра. (Поэтому различные линии, соединяющие однотипные опорные точки в последовательных строках кадра, разворачиваются на разный угол.) В результате размер однотипных участков кривой линейно изменяется в последовательных строках кадра, приводя в итоге к уменьшению среднего значения ширины верхнего основания. Можно выбрать другие опорные точки для измерения ширины верхнего основания (в рамках метода [4]), что приведет к меньшей вариации ее значений при наклонах изображения.

Список литературы

1. **Ларионов Ю. В., Озерин Ю. В.** Вариация изображения выступа в ходе его многократного сканирования в низковольтном РЭМ // Нано и микросистемная техника. 2017. Т. 19, № 12. С. 722—729.
2. **Меры** рельефные нанометрового диапазона из моноцисталического кремния. Национальный стандарт РФ, ГОСТ Р 8.628.2007. М.: Стандартинформ, 2007.
3. **Программа** анализа изображений микроскопов, программная платформа NDPL, <https://yadi.sk/i/YtOL8GEid9nK4>, 2009.
4. **Новиков Ю. А., Озерин Ю. В., Плотников Ю. И., Раков А. В., Тодуя П. А.** Проблема измерения геометрических характеристик электронного зонда растрового электронного микроскопа // Труды ИОФАН, 2006. Т. 62. С. 77.
5. **Ларионов Ю. В.** Оценка диаметра пучка РЭМ по кривым видеосигнала от трапециoidalной структуры: 1-й модельный эксперимент // Поверхность. 2016. № 3. С. 87—97.
6. **Ларионов Ю. В.** Оценка диаметра пучка РЭМ по кривым видеосигнала от трапециoidalной структуры: 2-й эксперимент // Поверхность. 2016. № 12. С. 96—108.
7. **Ларионов Ю. В.** Сканирование поверхности кремния в высоковольтном РЭМ без контаминации // Нано- и микросистемная техника. 2017. Т. 19, № 6. С. 323—330.
8. **Свидетельство** об утверждении типа средств измерений RU.C.27.010.A № 26441. Федеральное агентство по техническому регулированию и метрологии. 2012.

Yu. V. Larionov, Senior Research Scientist, luv@kapella.gpi.ru,
A. M. Prokhorov General Physics Institute, Russian academy of Science, Moscow, 119991, Russian Federation
Yu. V. Ozerin, Leading Engineer, yozerin@mikron.ru
SC "Mikron", Zelenograd, Russian Federation

Corresponding author:

Larionov Yury V. Senior Research Scientist, A. M. Prokhorov General Physics Institute, Russian academy of Science, Moscow, 119991, Russian Federation, e-mail: luv@kapella.gpi.ru

Rotation of an Image of the Protrusion at its Repeated Scanning in a SEM

Received on December 25, 2017

Accepted on January 26, 2018

The image of the protrusion on a silicon plate scanned in a SEM is discovered to change its inclination to screen axes as if it is rotated around the axe which is perpendicular to an image plane. A phenomenon occurs at scanning of a protrusion in a high or a low voltage SEM. It is supposed to link with induction of electric charge in the protrusion undersurface layer which influence at image formation. A mechanism of an image rotation is proposed and the influence of the phenomenon on protrusion linewidth values is estimated. Linewidth values of nanostructures are changed mainly due to variation of effective electron beam diameter values of a SEM measured by using video signal curve from the scanned protrusion. Variation of a beam diameter is estimated to be about 2,5 nm for a low voltage SEM.

Keywords: nanometrology, LV SEM, induced image rotation, induced electric charge, structure linewidth

For citation:

Larionov Yu. V., Ozerin Yu. V. Rotation of an Image of the Protrusion at its Repeated Scanning in a SEM, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 387—395.

DOI: 10.17587/nmst.20.387-395

Introduction

In [1] it was discovered, that during a repeated SEM scanning of a protrusion located on a silicon plate, a distortion of its image occurred leading to reduction of the measured value of its linewidth. This distortion of the protrusion image was explained by the induction of an electric charge in the protrusion subsurface by a scanning beam and influence of the induced localized charge on the emission of electrons from the surface, which formed the image of that protrusion. In the given work it was revealed, that a distortion also occurred in the form of a rotation of the protrusion image around its axis, perpendicular to the plane of the image. The aim of the work is to investigate the details, explain the reasons for the discovered phenomenon and estimate its influence on measurement of the linear sizes.

Experiment

Research was done for two types of protrusions on a silicon plate with gentle slopes: on the protrusions of a test object of measure of MShPS-2K [2] and on the protrusions with the slopes formed by means of chemical etching. The protrusions of the first type were formed by anisotropic etching of a plate to the depth of ~0.41 μm at the angle of inclination of the slopes of 54.7°. The images of the protrusions were obtained by means of SEM S4800 in a regime of slow secondary electrons (SSE) at the energy of the electrons of the incident beam of 20 keV. Fig. 1 presents the final images of the protrusion after its an uninterrupted scanning during the periods of time of 20 s and 160 s. In fig. 1, b one can see a turn of the image of the protrusion in the counter-clockwise direction in comparison with its image in fig. 1, a.

We recorded a change of the turn of rotation direction of the image for the opposite one as a result of the change of the mode of scanning: with pauses of 2 min. between separate frames with duration of 10 s (fig. 2, *a* and *b*).

For the protrusions of the second type the depth of etching was $\sim 0.35 \mu\text{m}$, the width of the top basis (TB) of a protrusion $\sim 290 \text{ nm}$, the angle of inclination of the slope $\sim 65^\circ$. The images were obtained by means of measuring SEM S-9260A at magnification of ~ 150.000 , the energy of the incident electrons of 800 V, the beam current $\sim 6 \text{ pA}$, number of image pixels — 512×512 at the size of a pixel of 1.75 nm. The tested protrusion was scanned repeatedly. The number of uninterrupted scannings (frames) of one protrusion was 400 or 800. In each cycle only 25 images (scans) of the protrusion were formed. In case of 400 frames one scan was formed by averaging of 16 frames, in case of 800 frames — by averaging of 32 frames, which in the latter case improved the image clarity. Several cycles of 800-times scanning were done. It turned out that the values of the image turns recorded in various cycles of scanning varied considerably. At that, not only the value of a turn, but even its direction changed. Therefore, in order to reveal the regularities of the phenomenon of the image rotation seven cycles of uninterrupted scannings were done with the intervals between them from several hours up to several days.

Technique for measurement of turns of the protrusions images

The variation of turns of the protrusion images were estimated quantitatively by means of a program for image processing [3]. With its help an image was divided into several tens of strips—sections located along the horizontal axis. In each of them the video signal (VS) curves from all the scanning lines in it were recorded, after which the average VS curves for each of the sections were formed. Fig. 3 presents a VS curve averaged by all the lines. It is visible, that it can be divided into several quasi-linear sites (between points 1—2, 2—3, 4—5, 5—6). Then, each of those sites was approximated by a straight line [4]. And, finally, on the crossings of the neighboring straight lines the co-ordinates of reference points were defined: beginning of the amplitude growth (1), the points of inclination change of the sites of the curve (2 and 5), the points of positions of the maxima of the model VS (3 and 4), and the points of recession of the VS amplitude down to the background value (6). According to [4], the distances between these points allow us to determine all the linear sizes of the protrusion. The width of the protrusion between the borders of the top basis should be determined by distance W between points *A* and *B* of the video signal (VS) curve in fig. 3. These points are arranged in the middle of the sites of VS recession from its peaks to the site corresponding to the top basis of the protrusion.

After determination of the reference points the dependence of their horizontal co-ordinates along the

length of the protrusion on the number of image strips-section was formed. The recorded dependences were approximated by the straight lines by the method of the least squares with the use of Origin 7.0 program, and then the program was used to calculate the tangent of inclinations of these straight lines on the image. Inclinations of the lines can influence the values of the linear quantities, which determine the dimensions of a protrusion. Therefore, together with the tangents of line inclinations the distances between peaks L_p , the distances between points 1, 2 and 5, 6, and also the width of the top basis of the protrusion W were calculated.

Width W can be measured directly by VS curve (see fig. 3), and it also can be calculated from the value of quantity L_p and the value of the effective beam diameter. Directly measured W also includes the value of the effective diameter in a latent form but this diameter is of another type [4]. As it was revealed, the value of the other type of the effective diameter differed from the value received by the points in fig. 3 [5].

Results

As a result of the calculations, the tangents of the angles of inclinations of the straight lines connecting the sets of reference points from 1 up to 6 in the strips-sections were estimated. For the turns of the images of the protrusions recorded in a high-voltage SEM and shown in fig. 1, the values of the tangents were averaged on all the lines connecting points 1—6: for fig. 1, *a* the average value of the tangent was 0.0028, for fig. 1, *b* with the greatest inclination line in the given work — 0.0157 (which corresponds to the turn angles of ~ 0.2 and $\sim 0.9^\circ$).

For a low-voltage SEM the most noticeable turns of the protrusion image were obtained in case of averaging of the scans by 16 frames. Table 1 presents the tangents of the angles of the straight lines (T_g) connecting the corresponding points from 1 up to 6 and the random errors of their definition σ . The latter value determines the degree of affinity of the approximating straight lines to the lines of the curve VS sites corresponding to them.

Apparently, the errors of estimation of the tangents of inclinations differ greatly for various points. Errors of definition of inclinations for points 2 and 5 are especially great, which leads to an increased casual error of the values of the effective diameter. Inclinations for points 3 and 4, bringing a contribution to the error of measurement of the protrusion width, are mostly reproduced.

From the data presented in table 1 an important conclusion follows: inclinations of the straight lines connecting the corresponding reference points on the image are not identical even in the first scan (after 16 frames). This means that the straight lines connecting points 1—6 are not parallel already in the first scan, and, therefore, the forms of the corresponding sites of VS curves in different sections of the image are not identical (either compressed or stretched). Apparently, as a result of the repeated (from the first up to the

25-th scans) scannings the inclinations of the line images vary. A trend for reduction of the inclinations of lines in 25-th scans for most sets of reference points in table 1 is appreciable, which visually corresponds to a clockwise turn of the protrusion image. The reduction of the inclination of lines connecting set of points 2 and 6 is most visible while for the lines connecting points 3 and 4, where the value of VS is the greatest, the inclination reduction is the least.

Results. Uncertainty of measurement of the linear sizes

The width of the topological element of IC is the most important controllable linear quantity. It is defined as the distance between a pair of the marked points on VS curve (with account of SEM magnification). As a rule, such points are disposed symmetrically in relation to the centre of the VS curve and are set by the value of the signal called "a cut-off level" [4] (fig. 3). In [4] it is shown, how to select a cut-off level for a model curve so that the reference points marked on the curve would correspond precisely to the borders of the top basis of the model trapezoidal protrusion (and the distance between them would correspond to its width). However, a real correspondence between the positions of the reference points set on the curve and the borders of the top basis is usually absent [6].

A major condition for definition of the width of protrusion TB in nanometrology is account of the value of the diameter of the electron beam forming an image [4]. The values of the diameter are used in the form of an amendment, allowing us to determine the width of TB, if the distance between by the peaks of the curve is measured (fig. 3) [4]. Measurement of the diameter of the electron beam, emitted by the surface and forming the image, is, apparently, the most important and most difficult problem of SEM metrology. An error in the value of the beam diameter is especially critical during measurement of the dimensions within the range of about 10 nm by means of SEM with a small diameter of the incident beam (i.e. with a high resolution).

In our country it was suggested to measure the diameter of the beam using VS curve obtained during scanning of the protrusion of the width measure of MShPS-2K (within the framework of the used model for formation of the protrusion image) [4]. Within the framework of this model the diameter of the beam does not depend on any influence quantity and a priori coincides with the parameters of the incident beam. The size of this model beam is called an effective diameter (d_{ef}) [4]. Use of this value allows us to determine the width of TB of the protrusion rather simply and more precisely, than by any points of the VS curve. According to [4], the value of d_{ef} is determined by the average distance along the axis of scanning between points 1–2 and 5–6 (fig. 3). Value of d_{ef} determines the width of protrusion W :

$$W = L_p - d_{ef}, \quad (1)$$

where L_p — distance between the peaks on the VS curve (see fig. 3) [4]. Therefore, for an estimation of the change of the value of the protrusion width during an image turn it is necessary to determine variation of d_{ef} and L_p .

Because of a bad reproducibility of the inclinations of the straight lines (table 1, sets of points 1–2 and 5–6), we estimated the change of d_{ef} by the results of scanning of the protrusion in seven cycles (a cycle consisted of 800 uninterrupted scannings). The results of the estimation of the variation of values of d_{ef} , L_p in the form of the difference of the values, recoded for the 25-th and the first images (Δd_{ef} , ΔL_p) in seven cycles, are presented in table 2. However, the results of determination of ΔW were defined directly from the VS curve (not by means of (1)).

A random error of measurement of the protrusion width (2σ) in the series of frames (and the corresponding VS curves) was 0.2...0.4 nm. In case of averaging of the scans — images by 32 frames a random error of the width in a scan should not exceed 0.1 nm. Therefore, we put in correspondence a variation of ΔW of several tenths of a nanometer with the systematic change of the width in various scans and cycles.

Error d_{ef} is expected to be subjected to random changes in the highest degree. Fig. 4 presents variation d_{ef} for the strips of one image along the length of the protrusion (in case of its averaging by 32 frames).

It is visible that a standard deviation from mean value d_{ef} along the length of the protrusion is close to ~0.3 nm (2σ). If averaged by all the image strips, it should be less than 0.1 nm. A change of the mean value of d_{ef} within one image (along the straight line) by ~2 nm — is noticeable. However, an increase of the value of d_{ef} averaged by 25 images was only 1.7 nm (table 2, cycle 1). Apparently, only fractions of the induced additives to d_{ef} in each image are summed up by all the images within one cycle.

From table 2 it is visible, that numerous scannings of the protrusion in each cycle lead to an increase of the effective diameter d_{ef} (on average by ~1.2 nm) at a reduction of the distance between the peaks (on average, by ~0.16 nm). According to (1), these variations should lead to a reduction of the width of the protrusion almost by ~1.4 nm. But, according to table 2, reduction of the values of width ΔW equals only to ~0.7 nm. We explain the difference by the fact that the value of the effective diameter d_{ef} differs from its value, which implicitly defines the width of the protrusion in case of measurements directly by the VS curve. In [6] it was specified, that during scanning of TB by a focused beam in the experiment it was methodically incorrect to define its width by the values of the effective diameter d_{ef} . The matter is that value d_{ef} in model [4] is determined by the sites of the VS curve formed by parts of the protrusion profile near the bottom basis (because the incident and the emitted electron beams in the model are parallel).

The turns of the initial and final images in different cycles, which were used for obtainment of the results in table 2, occurred in the counter-clockwise direction with a great angle spread (in one case a clockwise turn was recorded). This testifies to an appreciable influence of the random, not considered quantities, on the value of a turn. At that, their influence on variation d_{ef} and L_p was different: the span of d_{ef} in all cycles of scannings was 1.5 nm, while the span of L_p (except the fourth one, where random spike was recorded) was 0.5 nm. These span values correlate with the reproducibility of the tangents of the angles of inclination of the straight lines defining the values of the effective diameter and the distances between the peaks of VS curve in table 1 (for cycles 1, 2, 5, 6 and, accordingly — 3, 4). From table 2 it is also visible, that the span of ΔW (of the directly measured width of TB) was only 0.7 nm, i.e. the calculations of the width with the use of expression (1) were less reproducible.

The reason for variation of the inclination

The reason for variation of the rotation of the edge of the image of the protrusion is, apparently, the electric charge, locally induced by a scanning beam on the surface of the protrusion. The beam scanning a the line of an image induces a charge, presumably, not only in the strip of scanning, but also and in a surface strip, where the line $N + 1$ will subsequently be scanned.

The induced charge has no time for relaxation during scanning of line N and changes emission of SSE in the scanning strip of line $N + 1$. The emission current for line $N + 1$ can either increase, or decrease. If in the next line of scanning it is more, then VS in the corresponding image points will start to grow earlier and the position of these points will be displaced by along axis X to the left. In this case the image of the protrusion will turn clockwise. If the emission lessens, then the image points will move to the right, while the image will turn in the counter-clockwise direction. Different directions of the turn of the image of the protrusion were confirmed experimentally at different modes of scanning of the same protrusion (fig. 1, 2). The fact of a monotonous variation of the emission of electrons from the surface of the protrusion during scanning with formation of many frames of the image is presented in the work [7]. In this work scanning in a high-voltage SEM was done in two modes: during an uninterrupted scanning and with pauses between the separate frames. As it turned out, in case of an uninterrupted scanning the amplitude on the VS curve in last frame decreased in comparison with the first frame, while in case of scanning with pauses it was vice-versa.

Conclusion

A prolonged SEM scanning of the protrusion with low gradient slopes in silicon leads to various distortions of its SEM images.

The given work presents one kind of distortions — a rotation of the image of the protrusion (in the clockwise and the counter clockwise directions) by the angle up to 1° at the number of scannings of several hundreds. The image rotated both in a low-voltage and in a high-voltage SEM. The reason of an image rotation was, apparently, the electric charge induced by the scanning beam in the subsurface layer of the protrusion. A mechanism of influence of the charge on a turn of the image was proposed taking into account the direction of its rotation.

In connection with a turn of the image the value of the width of the top basis of the protrusion can decrease by 2.4 nm. (We should point out, that an absolute error of measurement of the width of the top basis should be ± 1 nm [8].) It was discovered that a turn caused unequal displacements of the reference points (between which the distance was measured at the curve of the video signal) for different scanning lines of one frame. (Therefore, various lines connecting the similar reference points in the consecutive lines of a frame turned to a different angle.) As a result, the size of the similar sites of the curve changed linearly in the consecutive lines of a frame, leading, as a result, to a reduction of the average value of the width of the top basis. It is possible to select other reference points for measurement of the width of the top basis (within the same method [4]), which will lead to a smaller variation of its values in case of the image rotation.

References

1. **Larionov Yu. V., Ozerin Yu. V.** Varitsia izobradzeniya vistupu v hode ego mnogokratnogo skanirovaniya v nizkovoltnom REM, *Nano- i mikrosistemnaya tekhnika*, 2017, vol. 19, no. 12. pp. 722–729 (in Russian).
2. **Meri** rel'efnie nanometrovogo diapazona iz monokristallicheskogo kremniya. Natsional'nii standart RF, GOST R 8.628.2007.M. Standartinform. 2007 (in Russian).
3. **Programma** obrabotki i analiza izobrazhenii mikroskopov, platforma NDPL, available at: <https://yadi.sk/i/YtOL8GEid9nK4>. 2009 (in Russian).
4. **Novikov Yu. A., Ozerin Yu. V., Plotnikov Yu. I., Rakov A. V., Fillipov M. N., Todua.** Problema izmereniya geometricheskikh kharakteristik elektronnogo zonda rastrovogo elektronnogo mikrosropa, *Trudi IOFAN*, 2006, vol. 62, pp. 77. (in Russian).
5. **Larionov Yu. V.** Otsenka diametra puchka REM po krivim videosignalu ot trapezoidalnoi strukturi: 1-modelnii eksperiment, *Poverkhnost*, 2016, no. 3. pp. 87–97. (in Russian).
6. **Larionov Yu. V.** Otsenka diametra puchka REM po krivim videosignalu ot trapezoidalnoi strukturi: 1 — modelnii eksperiment, *Poverkhnost*, 2016, no. 9. pp. 96–108 (in Russian).
7. **Larionov Yu. V.** Skanirovaniye poverkhnosti kremniya v visokovoltnom REM bez kontaminatsii, *Nano- i mikrosistemnaya tekhnika*, 2017. vol. 19, no. 6, pp. 323–330 (in Russian).
8. **Svidetel'stvo** ob utverdzenii tipa sredstv izmerenii RU. C.27.010.A № 26441. Federalnoe agenstvo po tekhnicheskemu regulirovaniyu i metrologii. 2012 (in Russian).

МОДЕЛИРОВАНИЕ И КОНСТРУИРОВАНИЕ МНСТ MODELLING AND DESIGNING OF MNST

УДК 621.382

DOI: 10.17587/nmst.20.396-402

А. Г. Мустафаев, д-р техн. наук, проф., e-mail: arslan_mustafaev@mail.ru,
Дагестанский государственный университет народного хозяйства, г. Махачкала,
Г. А. Мустафаев, д-р техн. наук, проф., Н. В. Черкесова, канд. техн. наук, доц.,
Кабардино-Балкарский государственный университет, г. Нальчик

УЧЕТ ВОЗДЕЙСТВИЯ ПОЛНОЙ ДОЗЫ ИЗЛУЧЕНИЯ В КОМПАКТНЫХ МОДЕЛЯХ КМОП-ТРАНЗИСТОРОВ

Поступила в редакцию 18.01.2018

Успешное решение задач проектирования и разработки КМОП СБИС, а также построения приборов на их основе невозможно без применения САПР как на элементном, так и на схемотехнических уровнях. Для радиационно-стойких СБИС эти задачи существенно усложняются необходимостью дополнительно учитывать воздействие радиационных факторов. Чтобы смягчать риски их воздействия, необходимо предусматривать радиационную стойкость приборов с начала их проектирования.

Ключевая роль в таких САПР принадлежит компактным моделям КМОП-элементов, поскольку от степени учета влияющих факторов и их точности в первую очередь зависит достоверность результатов схемотехнического проектирования. Таким образом, разработка новых и улучшение существующих SPICE-моделей элементов радиационно-стойких КМОП БИС является актуальной задачей.

Ключевые слова: КМОП-транзистор, моделирование, ионизирующее излучение, поверхностный потенциал

Натурные эксперименты по анализу воздействия ионизирующего излучения на СБИС достаточно трудоемки и не всегда информативны относительно протекающих физических процессов [1, 2], поэтому хорошей альтернативой является включение эффектов ионизирующего излучения в компактные модели устройств, используемых в стандартных симуляторах электрических цепей. Включение этих эффектов требует знания физических процессов, которые способствуют возникновению дефектов вследствие ионизирующего излучения, и влияния, которые эти эффекты оказывают на электрические характеристики компонентов в передовых КМОП-технологиях [3, 4].

Приборно-технологическое моделирование МОП-транзисторов КНС- и КНИ-типа, работающих в условиях радиационного воздействия, представляет собой сложную проблему [5]. Наряду с особенностями технологического процесса необходимо учитывать изменение физических параметров, происходящее под действием радиации, и его влияние на выходные электрические характеристики.

Эффективность проектирования зависит от точности аналитических моделей транзисторов, применяемых в моделировании [6].

При моделировании МОП-транзисторов применяют подходы, основанные на понятии порогово-

го напряжения (V_{th}), инверсионного заряда (q_i) и поверхностного потенциала (φ_s) [7].

В моделях, основанных на понятии порогового напряжения, плотность заряда в канале определяется как плотность заряда плоского конденсатора, образованного затвором и каналом, в предположении, что поверхностный потенциал в режиме сильной инверсии не зависит от напряжения на затворе. Плотность подвижного заряда в канале определяется разностью между напряжением на затворе и пороговым напряжением, при котором подвижный заряд в канале отсутствует. Недостатки моделей, основанных на пороговом напряжении, такие как разрывы в разных рабочих областях, отрицательная емкость и отрицательная проводимость, приведены в работе [8].

Модели, основанные на понятии заряда инверсионного слоя [9, 10], описывают поведение транзистора во всех режимах работы и позволяют уменьшить число параметров модели. Недостатком является невозможность описания режима накопления заряда под затвором и в области перекрытия затвором истока и стока. Плотность заряда в инверсионном слое также является неудобной величиной для описания шумовых свойств транзистора, разделения тока затвора в канале между стоком и истоком.

В результате появился значительный интерес к разработке моделей на основе поверхностного потенциала [11–16]. Предложено несколько компактных моделей на основе поверхностного потенциала [6, 17, 18], на основе модели заряда Брюса [19]. Физические эффекты, такие как эффекты квантования, часто рассматриваются как дополнения к таким моделям [20].

Модели, основанные на понятии поверхностного потенциала, учитывают практически все физические эффекты в нанометровых МОП-транзисторах [21]. Их основным достоинством является физически обоснованное моделирование области умеренной инверсии, которая становится преобладающей в современных ИС, работающих при напряжении питания от 1,2 до 1,5 В. Это делает модели пригодными для проектирования цифровых, аналоговых и радиочастотных ИС, созданных по нанометровым технологиям. Такой подход позволяет найти единое выражение, описывающее как подпороговую область, область умеренной и сильной инверсии без применения сглаживающих функций.

На основе соотношений, определяющих закономерности процессов, протекающих в полупроводниковых структурах при воздействии ионизирующего излучения, в работе предложены зависимости деградации параметров элементов схемотехнического уровня, которые дополняют *spice*-модели.

Так как *spice*-параметры взаимосвязаны и характеризованы по температуре, подставив полученные значения подвижности и порогового напряжения, можно получить остальные параметры, характерные для излучения.

Экспериментально установлено, что деградация электрических параметров, а также возникающие функциональные отказы в значительной мере определяются мощностью дозы, т. е. скоростью, с которой был осуществлен набор дозы [22].

Исследования показывают, что ключевыми механизмами деградации, обусловленными поглощенной дозой излучения, являются накопление положительного заряда (N_{ot}) в области затвора [23, 24] и образование поверхностных состояний на границе раздела оксид — полупроводник (интерфейсные ловушки) (N_{it}) [25, 26]. В интегральных КМОП-структуратах, подверженных воздействию высокой дозы ионизирующей радиации, возможно защелкивание [27–29].

В результате масштабирования уменьшение толщины подзатворного оксида (t_{ox}) и увеличение уровня легирования канала и объема полупроводника улучшили радиационную стойкость субмикронной КМОП-технологии. Это обусловлено уменьшением влияния классических радиационных эффектов (например, накопления фиксированного заряда в подзатворном оксиде), который определяется толщиной оксида t_{ox} . Однако накопление N_{ot} и N_{it} вблизи границы раздела полупро-

водник — оксид при глубокой изоляции канавками является потенциальной радиационной угрозой [30–32] и оказывает заметное влияние на основные характеристики интегральных схем [33].

В работе [34] описана методика, позволяющая разделить радиационно-индукционный сдвиг порогового напряжения ΔU_{th} , на его составляющие ΔU_{Nit} — вклад заряда, локализованного на поверхностных ловушках границы Si—SiO₂ и ΔU_{Not} — вклад соответственно объемно-встроенного заряда подзатворного диэлектрика SiO₂ в предположении, что общее изменение U_{th} равно аддитивной сумме составляющих вкладов напряжений $\Delta U_{th} = \Delta U_{Nit} + \Delta U_{Not}$.

В работах [35–38] предложены адаптированные модели транзисторов, учитывающие влияние излучения на электрофизические параметры КМОП-транзисторов.

Имплементация влияния поглощенной дозы излучения в модель на основе поверхностного потенциала осуществляется путем перенормировки уравнения поверхностного потенциала, которое имеет следующий вид [39]:

$$(V_g - V_{FB} - \xi \psi_s)^2 = \gamma^2 \phi_t H(\beta \psi_s), \quad (1)$$

где V_g — напряжение на затворе; V_{FB} — напряжение плоских зон; ψ_s — поверхностный потенциал;

$\gamma = \frac{\sqrt{2q\epsilon_s N_a}}{C_{ox}}$; ϕ_t — температурный потенциал; $\beta = \frac{1}{\phi_t}$; ϵ_s — диэлектрическая проницаемость полупроводника; N_a — уровень легирования; C_{ox} — емкость оксида; H — оператор Гамильтона.

Напряжение плоских зон V_{FB} имеет следующий вид [40]:

$$V_{FB} = \phi_{MS} - \frac{q(N_{ot} + D_{it}\phi_b)}{C_{ox}},$$

где ϕ_{MS} — разность работ выхода затвор-полупроводник; q — заряд электрона; ϕ_b — объемный потенциал; N_{ot} — заряд в объеме оксида; D_{it} — плотность ловушек на поверхности, а параметр ξ равен

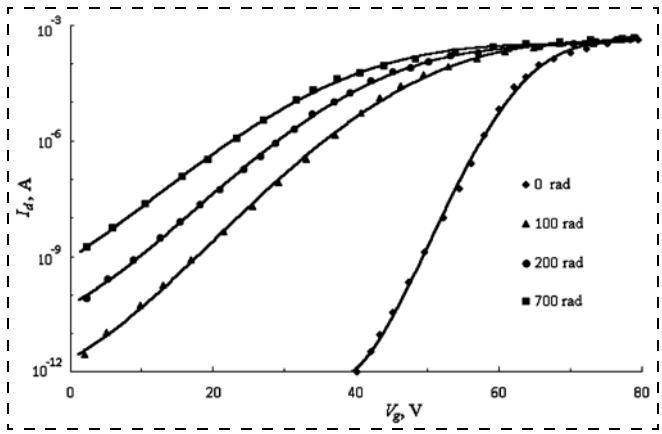
$$\xi = 1 + \frac{qD_{it}}{C_{ox}}.$$

Уравнение поверхностного потенциала [41], использованное в *spice*-модели на основе поверхностного потенциала, имеет вид

$$(V_g - V_{FB} - \psi_s)^2 = \gamma^2 \phi_t H(\beta \psi_s).$$

Таким образом, уравнение (1) расширяет стандартную модель за счет учета параметра ξ , характеризующего плотность поверхностных состояний на границе раздела кремний — оксид.

Чтобы включить эффекты, обусловленные поглощенной дозой излучения, в существующую



Сравнение экспериментальных (точка) и расчетных (линия) сток-затворных характеристик *p*-канального МОП-транзистора при различных значениях поглощенной дозы ионизирующего излучения

Comparison of the experimental (point) and calculated (line) drain-gate characteristics of a p-channel MOS transistor at various values of the absorbed dose of the ionizing radiation

структуре модели на основе поверхностного потенциала, выполняется перенормировка:

$$V_g^* = V_g/\xi, \quad V_{FB}^* = V_{FB}/\xi.$$

Отсюда следует, что

$$(V_g^* - V_{FB}^* - \psi_s)^2 = (\gamma^*)^2 \phi_t H(\beta \psi_s). \quad (2)$$

В отличие от (1), уравнение (2) можно решить относительно ψ_s , используя высокоточные аналитические аппроксимации, которые применяют в *spice*-моделировании [6].

При рассчитанных поверхностных потенциалах истока ψ_{ss} и стока ψ_{sd} , значение тока стока может быть получено методом симметричной линеаризации [42]:

$$I_d = \frac{W}{L} \mu_{eff} C_{ox} (q_{im} + \alpha_m \phi_t) \Delta \psi.$$

Было проведено сравнение значений тока стока *p*-канального МОП-транзистора при различных значениях поглощенной дозы ионизирующего излучения, полученных экспериментальным путем [43] и рассчитанных согласно модели (см. рисунок). Параметры транзистора: длина канала 1,5 мкм; ширина канала 180 мкм; толщина подзатворного оксида 400 нм.

Проверка исследуемой компактной модели показала хорошую корреляцию с экспериментально полученными сток-затворными характеристиками и адекватность сделанных предположений.

Список литературы

1. **Максименко С. Л., Мелехин В. Ф., Филиппов А. С.** Анализ проблемы построения радиационно-стойких информационно-управляющих систем // Информационно-управляющие системы. 2012. № 2. С. 18–25.
2. **Boychenko D., Kalashnikov O., Nikiforov A.** et al. Total ionizing dose effects and radiation testing // Facta Universitatis. Series: Electronics and Energetics. 2015. Vol. 28, N. 1. P. 153–164.
3. **Ma T. P., Dressendorfer P. V.** Ionizing Radiation Effects in MOS Devices and Circuits, New York: Wiley, 1989.
4. **Barnaby H. J., McLain M. L., Esqueda I. S., Xiao Jie V.** Modeling Ionizing Radiation Effects in Solid State Materials and CMOS Devices // IEEE Trans. on Circuits and Systems I. 2009. Vol. 56. P. 1870–1833.
5. **Харитонов И. А., Петросянц К. О., Орехов Е. В.** и др. Приборно-технологическое моделирование элементной базы КМОП КНИ БИС с учетом факторов радиационного воздействия // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС). 2008. № 1. С. 266–271.
6. **Gildenblat G., Li X., Wu W.** et al. PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation // IEEE Trans. Electron Devices. 2006. Vol. 53. P. 1979–1993.
7. **Денисенко В. В.** Компактные модели МОП-транзисторов для SPICE в микро- и наноэлектронике. М.: Физматлит, 2010. 407 с.
8. **Cheng Y., Hu C.** MOSFET Modeling and BSIM3 User's Guide. Boston, MA: Kluwer Academic Publishers, 1999. 462 p.
9. **Enz C. C., Krummenacher F., Vittoz E. A.** An analytical MOS transistor model valid in all regions of operation and dedicated to low voltage and low-current applications // Journ. Analog Integrated Circuit and Signal Processing. 1995. Vol. 8. P. 81–114.
10. **Bucher M., Enz C. C., Krummenacher F.** et al. The EKV compact MOS transistor model: accounting for deep-submicron aspects // Modeling and Simulation of Microsystems 2002. P. 670–673.
11. **Langevelde R. V., Klassen F. M.** An explicit surface-potential-based MOSFET model for circuit simulation // Solid State Electronics. 2000. Vol. 44, N. 3. P. 409–418.
12. **Joardar K., Gullapalli K. K., McAndrew C. C.** et al. An improved MOSFET model for circuit simulation // IEEE Trans. Electron Devices. 1998. Vol. 45, N. 1. P. 134–148.
13. **Boothroyd A. R., Taraswicz S. W., Slaby C.** MISNAN—a physically based continuous MOSFET model for CAD applications // IEEE Trans. Computer Aided Des. Integr. Circuits Syst. 1991. Vol. 10, N. 12. P. 1512–1529.
14. **Miura-Mattausch M., Feldmann U., Rahm A.** et al. Unified complete MOSFET model for analysis of digital and analog circuits // IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 1996. Vol. 15, N. 1. P. 1–7.
15. **Rios R., Mudanai S., Wei-Kai Shih** et al. An efficient surface potential solution algorithm for compact MOSFET models // IEDM Tech. Digest. 2004. P. 755–758.
16. **Chen T. L., Gildenblat G.** Analytical approximation for the MOSFET surface potential // Solid State Electronics. 2001. Vol. 45, N. 2. P. 335–339.
17. **Sadachika N., Kitamaru D., Uetsuji Y.** et al. Completely Surface-Potential-Based Compact Model of the Fully Depleted SOI-MOSFET Including Short-Channel Effects // IEEE Trans. Electron Devices. 2006. Vol. 53, N. 9. P. 2017–2024.
18. **He J., Song Y., Niu X.** et al. Punsim: an advanced surface potential based MOSFET model // Proc. MIXDES, Gdynia, Poland. 2006. P. 111–116.
19. **Brews J. R.** A charge-sheet model of the MOSFET // Solid State Electronics. 1978. Vol. 21, N. 2. P. 345–355.
20. **Pregaldiny F., Lallement C., Mathiot D.** Accounting for quantum mechanical effects from accumulation to inversion, in a fully analytical surface-potential-based MOSFET model // Solid State Electronics. 2004. Vol. 48, N. 5. P. 781–787.
21. **Mattausch H. J., Miyake M., Yoshida T.** et al. HiSIM2 Circuit simulation — Solving the speed versus accuracy crisis // IEEE Circuits and Devices Magazine. 2006. Vol. 22, N. 5. P. 29–38.
22. **Скляр В. А.** Проектирование микросхем с учетом воздействия статического ионизирующего излучения космического пространства: диссертация... кандидата технических наук: 05.13.12. Воронеж, 2014. 174 с.
23. **Kooi E.** Influence of X-ray Irradiations on the Charge Distribution of Metal-Oxide-Silicon Structures // Philips Res. Rept. 1965. Vol. 20. P. 306–314.

24. Szedon J. R., Sandor J. E. The Effect of Low-Energy Electron Irradiation of Metal-Oxide-Semiconductor Structures // Appl. Phys. Letters. 1965. Vol. 6. P. 181–182.
25. Kooi E. Effects of Ionizing Irradiations on the Properties of Oxide-Covered Silicon Surfaces // Philips J. Res. 1965. Vol. 20. P. 595–619.
26. Zaininger K. H. Electron Bombardment of MOS Capacitors // Applied Physics Letters. 1966. Vol. 8. P. 140–142.
27. Li R., Chen W., Lin D. et al. Study of latch-up immunization in bulk CMOS integrated circuits exposed to transient ionizing radiation // Sci. China Technol. Sci. 2012. Vol. 55. P. 3242–3247.
28. Lacoe R. C., Moss S. C., Osborn J. V. et al. Neutron and proton irradiation for latchup suppression in a radiation-tolerant commercial submicron CMOS process // Proceedings of Fifth European Conference on Radiation and Its Effects on Components and Systems (RADECS), 1999. P. 340–345.
29. Plaag R. E., Baze M. P., Johnston A. H. A distributed model for radiation-induced latchup // IEEE Transactions on Nuclear Science. 1988. Vol. 35, N. 6. P. 1563–1569.
30. Barnaby H. J. Total-Ionizing-Dose Effects in Modern CMOS Technologies // IEEE Trans. Nucl. Sci. 2006. Vol. 53. P. 3103–3121.
31. Brady F. T., Maimon J. D., Hurt M. J. A scaleable, radiation hardened shallow trench isolation // IEEE Trans. Nucl. Sci. 1999. Vol. 46, N. 6. P. 1836–1840.
32. Oldham T. R., McLean F. B. Total ionizing dose effects in MOS oxides and devices // IEEE Trans. Nucl. Sci. 2003. Vol. 50, N. 3. P. 483–499.
33. Мустафаев Г. А., Мустафаев А. Г. Разработка процесса формирования глубокой изоляции структур кремний на изолиторе // Нано- и микросистемная техника. 2009. № 1. С. 30–32.
34. McWhorter F. G., Winokur P. S. Simple technique for separating the effects of interface traps and trapped—oxide charge in metal—oxide—semiconductor transistors // Appl. Phys. Lett. 1986. Vol. 48, N. 2. P. 133–135.
35. Esqueda S. I., Barnaby H. J., King M. P. Compact Modeling of Total Ionizing Dose and Aging Effects in MOS Technologies // IEEE Transactions on Nuclear Science. 2015. Vol. 62, N. 4. P. 1501–1515.
36. Дворников О. В., Гришков В. Н. Учет влияния проникающей радиации в "Spice-подобных" программах // Проблемы разработки перспективных микро- и наноэлектронных систем — 2010. Сборник трудов / под общ. ред. акад. А. Л. Степпковского. М.: ИППМ РАН, 2010. С. 301–306.
37. Pershenkov V. S. Conversion model of the radiation-induced interface-trap building and its hardness assurance application // FACTA UNIVERSITATIS Series: Electronics and Engetics. 2015. Vol. 28, N. 4. P. 557–570.
38. Gorbunov M., Danilov I., Zebrev G. et al. Verilog-A Modeling of Radiation-Induced Mismatch Enhancement // IEEE Transactions on Nuclear Science. 2011. Vol. 58. P. 785–792.
39. Esqueda I. S. Modeling of total ionizing dose effects in advanced complementary metal-oxide-semiconductor technologies, PhD thesis, Arizona State Univ., 2011. 179 p.
40. Гуртов В. А. Твердотельная электроника. 3-е изд., доп. М.: Техносфера, 2008. 512 с.
41. Wu W., Ten-Lon C., Gildenblat G. et al. Physics-based mathematical conditioning of the MOSFET surface potential equation // IEEE Trans. Electron Devices. 2004. Vol. 51. P. 1196–1199.
42. Saha S. Compact Models for Integrated Circuit Design. Boca Raton: CRC Press, 2015. 545 p.
43. Мустафаев Г. А., Мустафаев А. Г. Радиационная стойкость КНИ МОП-транзисторов к накопленной дозе ионизирующего излучения // Нано- и микросистемная техника. 2007. № 12. С. 47–49.

A. G. Mustafaev, D. Sc., Professor, arslan_mustafaev@mail.ru,
Dagestan State University of National Economy, Makhachkala, 367008, Russian Federation,
G. A. Mustafaev, D. Sc., Professor, **N. V. Cherkesova**, Ph. D., Associate Professor,
Kabardino-Balkarian State University, Nalchik, Russian Federation

Corresponding author:

Mustafaev Arslan G., D. Sc., Professor, Dagestan State University of National Economy, Makhachkala, 367008, Russian Federation, e-mail: arslan_mustafaev@mail.ru

Consideration of the Effect of the Total Dose of Radiation in the Compact Models of the CMOS Transistors

Received on January 18, 2018
Accepted on February 6, 2018

Successful design solutions, development of CMOS VLSI and construction of the devices based on them are impossible without the use of CAD, both at the elemental and circuit levels. For the radiation-resistant VLSI these tasks are complicated significantly due to the necessity of additional consideration of the effect of the radiation factors. In order to reduce the risks of their impact it is necessary to envisage the radiation resistance of the devices from the stage of their designing.

A key role in such CAD systems belongs to the compact models of CMOS elements, because the reliability of the results of the circuit design depends primarily on the degree of consideration of the influencing factors and their accuracy. Thus, development of new and improvement of the existing SPICE models of the elements of the radiation-resistant CMOS LSI is an urgent task.

Keywords: CMOS transistor, simulation, ionizing radiation, surface potential

For citation:

Mustafaev A. G., Mustafaev G. A., Cherkesova N. V. Consideration of the Effect of the Total Dose of Radiation in the Compact Models of the CMOS Transistors, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 396–402.

DOI: 10.17587/nmst.20.396-402

Natural experiments for analysis of the influence of the ionizing radiation on VLSI are labor-consuming and not always informative enough about the proceeding physical processes [1, 2]. A good alternative for them is inclusion of the effects of the ionizing radiation in the compact models of the devices used in the standard simulators of the electric circuits. Inclusion of the effects demands knowledge of the processes, which promote occurrence of the defects owing to the ionizing radiation, and the influences of the effects on the electric characteristics of the components in the CMOS technologies [3, 4].

The device-technological modeling of the MOS transistors of the SOS and SOI types, operating in the conditions of the radiation influence, is a complex problem [5]. Alongside with the specific features of the technological process it is necessary to take into account the variation of their physical parameters under the influence of radiation, and its influence on the output electric characteristics.

Efficiency of designing depends on the accuracy of the analytical models of the transistors [6].

The approaches applied for modeling of the MOS transistors are based on the concept of the threshold voltage (V_{th}), inversion charge (q_i) and surface potential (φ_s) [7].

In the models based on the threshold voltage concept, the charge density in the channel defined as the density of the charge of the flat condenser formed by the gate and the channel, assuming, that in the mode of a strong inversion the surface potential does not depend on the gate voltage. The density of the mobile charge in the channel defined by the difference between the gate voltage and the threshold voltage, at which the mobile charge in the channel is absent. The drawbacks of the models based on the threshold voltage, such as ruptures in different working areas, negative capacity and negative conductivity, presented in [8].

The models based on the concept of the inversion layer charge [9, 10] describe the behavior of a transistor in all the operating modes and allow us to reduce the number of the model parameters. A drawback is impossibility of a description of the mode of a charge accumulation under the gate and in the area of overlapping of the source and the drain. The charge density in the inversion layer is also an inconvenient value for description of the noise properties of a transistor, separation of the gate current in the channel between the drain and the source.

As a result a considerable interest appeared to development of the models based on surface potential [11–16]. Several compact models based on surface potential [6, 17, 18] and based on Bruce charge model [19] were proposed. The physical effects, such as the effects of quantization, often considered as an addition to such models [20].

The models based on the concept of the surface potential take into account practically all the physical effects in the nanoscale MOS transistors [21]. Their basic advantage is the physically substantiated modeling of the

moderate inversion area, which becomes prevailing in IC, working at the power supply voltage from 1.2 up to 1.5 V. This makes the models suitable for designing of the digital, analogue and radio-frequency IC, created by the nanoscale technologies. Such an approach allows to find an uniform expression describing the subthreshold area and the region of moderate and strong inversion without application of the smoothing functions.

On the basis of the correlations defining the order of the processes in the semiconductor structures under the influence of the ionizing radiation, the dependences of degradation of the parameters of the elements of the circuit level, which expands the *spice* models were proposed.

Since the *spice* parameters are interconnected and characterized by temperature, by substituting the received values of the mobility and the threshold voltage it is possible to obtain the other parameters, characteristic for the radiation.

Experimentally established, that the degradation of the electric parameters and arising functional failures are determined by the dose rate [22].

The researches demonstrate that the key mechanisms of degradation determined by the absorbed dose of radiation are accumulation of a positive charge (N_{ot}) in the area of the gate [23, 24] and formation of the surface states of the oxide-semiconductor interface (interface traps) (N_{it}) [25, 26]. In the integral CMOS structures, exposed to influence of the high-dose ionizing radiation, a latching is possible [27–29].

As a result of scaling, the reduction of the thickness of the gate oxide (t_{ox}) and increase of the channel doping level and the semiconductor volume, the radiation resistance of the submicronic CMOS technology was improved. This was due to the reduction of the influence of the classical radiation effects (for example, accumulation of the fixed charge in the gate oxide), which is determined by thickness t_{ox} of oxide. However, accumulation of N_{ot} and N_{it} near the semiconductor — oxide interface with deep trench isolation appear a potential radiation threat [30–32] and has a significant impact on the basic characteristics of the integrated circuits [33].

In [34] described the method dividing the radiation-induction shift of the threshold voltage ΔU_{th} into its components ΔU_{Nit} — contribution of the charge localized on Si—SiO₂ surface interface traps, and ΔU_{Not} — contribution of the volume-built-in charge of the gate oxide in assumption, that the total change U_{th} is equal to the additive sum of the contributing voltages:

$$\Delta U_{th} = \Delta U_{Nit} + \Delta U_{Not}$$

In [35–38] offered the adapted models of the transistors, which take into account the influence of radiation on the electrophysical parameters of the CMOS transistors.

Implementation of the influence of the absorbed dose of radiation in the model on the basis of the surface potential is carried out by means of a renormalization of the equation of the surface potential, which looks like the following [39]:

$$(V_g - V_{FB} - \xi\psi_s)^2 = \gamma^2\phi_t H(\beta\psi_s), \quad (1)$$

where V_g — gate voltage; V_{FB} — flat band voltage; ψ_s — surface potential; $\gamma = \frac{\sqrt{2q\epsilon_s N_a}}{C_{ox}}$; ϕ_t — temperature potential; $\beta = \frac{1}{\phi_t}$, ϵ_s — dielectric permeability of a semiconductor; N_a — doping level; C_{ox} — oxide capacity; H — Hamilton operator.

Flat band voltage — V_{FB} expressed as [40]

$$V_{FB} = \varphi_{MS} - \frac{q(N_{ot} + D_{it}\phi_b)}{C_{ox}},$$

where φ_{MS} — gate-semiconductor work functions difference; q — electron charge; ϕ_b — volume potential; N_{ot} — oxide trapped charge; D_{it} — interface traps density, while ξ parameter is equal

$$\xi = 1 + \frac{qD_{it}}{C_{ox}}.$$

Equation of the surface potential [41], used in the *spice* model based on the surface potential, looks like the following:

$$(V_g - V_{FB} - \psi_s)^2 = \gamma^2 \phi_t H(\beta \psi_s).$$

Thus, the equation (1) expands the standard model due to account of the parameter characterizing the density of the surface states on the interface of silicon — oxide — ξ .

In order to include the effects caused by the absorbed dose of radiation into the existing structure of the model on the basis of the surface potential, the following renormalization is carried out:

$$V_g^* = V_g/\xi, \quad V_{FB}^* = V_{FB}/\xi.$$

Hence follows

$$(V_g^* - V_{FB}^* - \psi_s)^2 = (\gamma^*)^2 \phi_t H(\beta \psi_s). \quad (2)$$

Unlike (1), equation (2) can be solved in relation to ψ_s by using the high-precision analytical approximations, which are applied in the *spice* modeling [6].

At the calculated surface potentials of source ψ_{ss} and drain ψ_{sd} , the value of the drain current can be received by the method of symmetric linearization [42]

$$I_d = \frac{W}{L} \mu_{eff} C_{ox} (q_{im} + \alpha_m \phi_t) \Delta \psi.$$

A comparison of p-channel MOS transistor drain current, for various values of the absorbed dose of ionizing radiation, obtained experimentally [43] and calculated according to the model, was made (figure). The transistor's parameters: length of the channel — 1.5 μm ; width of the channel — 180 μm ; thickness of the gate oxide — 400 nm.

Testing of the investigated compact model demonstrated a good correlation with the experimentally received *drain*-gate characteristics, and adequacy of the made assumptions.

References

1. Maksimenko S. L., Melehin V. F., Filippov A. S. Analiz problemy postroenija radiacionno-stojkikh informacionno-upravljajushhih sistem, *Informacionno-upravljajushchie sistemy*, 2012, no. 2, pp. 18—25 (in Russian).
2. Boychenko D., Kalashnikov O., Nikiforov A. et al. Total ionizing dose effects and radiation testing, *Facta Universitatis, Series: Electronics and Energetics*, 2015, vol. 28, no. 1, pp. 153—164.
3. Ma T. P., Dressendorfer P. V. *Ionizing Radiation Effects in MOS Devices and Circuits*, New York, Wiley, 1989.
4. Barnaby H. J., McLain M. L., Esqueda I. S., Xiao Jie V. Modeling Ionizing Radiation Effects in Solid State Materials and CMOS Devices, *IEEE Trans. on Circuits and Systems I*, 2009, vol. 56, pp. 1870—1833.
5. Haritonov I. A., Petrosjanc K. O., Orekhov E. V. et al. Pri-borno-tehnologicheskoe modelirovanie jelementnoj bazy KMOP KNI BIS s uchetom faktorov radiacionnogo vozdejstvija, *Problemy razrabotki perspektivnyh mikro- i nanojelektronnyh sistem (MJeS)*, 2008, no. 1, pp. 266—271 (in Russian).
6. Gildenblat G., Li X., Wu W. et al. PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation, *IEEE Trans. Electron Devices*, 2006, vol. 53, pp. 1979—1993.
7. Denisenko V. V. *Kompaktnye modeli MOP-tranzistorov dlja SPICE v mikro- i nanojelektronike*, Moscow, Fizmatlit, 2010, 407 p. (in Russian).
8. Cheng Y., Hu C. *MOSFET Modeling and BSIM3 User's Guide*, Boston, MA, Kluwer Academic Publishers, 1999, 462 p.
9. Enz C. C., Krummenacher F., Vittoz E. A. An analitical MOS transistor model valid in all regions of operation and dedicated to low voltage and low-current applications, *Journ. Analog Integrated Circuit and Signal Processing*, 1995, vol. 8. pp. 81—114.
10. Bucher M., Enz C. C., Krummenacher F. et al. The EKV compact MOS transistor model: accounting for deep-submicron aspects, *Modeling and Simulation of Microsystems*, 2002, pp. 670—673.
11. Langevelde R. V., Klassen F. M. An explicit surface-potential-based MOSFET model for circuit simulation, *Solid State Electronics*, 2000, vol. 44, no. 3, pp. 409—418.
12. Joardar K., Gullapalli K. K., McAndrew C. C. et al. An improved MOSFET model for circuit simulation, *IEEE Trans. Electron Devices*, 1998, vol. 45, no. 1, pp. 134—148.
13. Boothroyd A. R., Taraswicz S. W., Slaby C. MISNAN-a physically based continuous MOSFET model for CAD applications, *IEEE Trans. Computer Aided Des. Integr. Circuits Syst.*, 1991, vol. 10, no. 12, pp. 1512—1529.
14. Miura-Mattausch M., Feldmann U., Rahm A. et al. Unified complete MOSFET model for analysis of digital and analog circuits, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, 1996, vol. 15, no. 1, pp. 1—7.
15. Rios R., Mudanai S., Wei-Kai Shih et al. An efficient surface potential solution algorithm for compact MOSFET models, *IEDM Tech. Digest*, 2004, pp. 755—758.
16. Chen T. L., Gildenblat G. Analytical approximation for the MOSFET surface potential, *Solid State Electronics*, 2001, vol. 45, no. 2, pp. 335—339.
17. Sadachika N., Kitamaru D., Uetsuji Y. et al. Completely Surface-Potential-Based Compact Model of the Fully Depleted SOI-MOSFET Including Short-Channel Effects, *IEEE Trans. Electron Devices*, 2006, vol. 53, no. 9, pp. 2017—2024.
18. He J., Song Y., Niu X. et al. Punsim: an advanced surface potential based MOSFET model, *Proc. MIXDES, Gdynia, Poland*, 2006, pp. 111—116.
19. Brews J. R. A charge-sheet model of the MOSFET, *Solid State Electronics*, 1978, vol. 21, no. 2, pp. 345—355.
20. Pregaldiny F., Lallement C., Mathiot D. Accounting for quantum mechanical effects from accumulation to inversion, in a fully analytical surface-potential-based MOSFET model, *Solid State Electronics*, 2004, vol. 48, no. 5, pp. 781—787.

21. Mattausch H. J., Miyake M., Yoshida T. et al. HiSIM2 Circuit simulation — Solving the speed versus accuracy crisis, *IEEE Circuits and Devices Magazine*, 2006, vol. 22, no. 5, pp. 29—38.
22. Skljar V. A. Proektirovaniye mikrosistem s uchetom vozdeystviya staticheskogo ionizirujushhego izlucheniya kosmicheskogo prostranstva: dissertaciya... kandidata tehnicheskikh nauk: 05.13.12. Voronezh, 2014. 174 p. (in Russian).
23. Kooi E. Influence of X-ray Irradiations on the Charge Distribution of Metal-Oxide-Silicon Structures, *Philips Res. Rept.*, 1965, vol. 20, pp. 306—314.
24. Seddon J. R., Sandor J. E. The Effect of Low-Energy Electron Irradiation of Metal-Oxide-Semiconductor Structures, *Appl. Phys. Letters*, 1965, vol. 6, pp. 181—182.
25. Kooi E. Effects of Ionizing Irradiations on the Properties of Oxide-Covered Silicon Surfaces, *Philips J. Res.*, 1965, vol. 20, pp. 595—619.
26. Zaininger K. H. Electron Bombardment of MOS Capacitors, *Applied Physics Letters*, 1966, vol. 8, pp. 140—142.
27. Li R., Chen W., Lin D. et al. Study of latch-up immunization in bulk CMOS integrated circuits exposed to transient ionizing radiation, *Sci. China Technol. Sci.*, 2012, vol. 55, pp. 3242—3247.
28. Lacoe R. C., Moss S. C., Osborn J. V. et al. Neutron and proton irradiation for latchup suppression in a radiation-tolerant commercial submicron CMOS process, *Proceedings of Fifth European Conference on Radiation and Its Effects on Components and Systems (RADECS)*, 1999, pp. 340—345.
29. Plaag R. E., Baze M. P., Johnston A. H. A distributed model for radiation-induced latchup, *IEEE Transactions on Nuclear Science*, 1988, vol. 35, no. 6, pp. 1563—1569.
30. Barnaby H. J. Total-Ionizing-Dose Effects in Modern CMOS Technologies, *IEEE Trans. Nucl. Sci.*, 2006, vol. 53, pp. 3103—3121.
31. Brady F. T., Maimon J. D., Hurt M. J. A scaleable, radiation hardened shallow trench isolation, *IEEE Trans. Nucl. Sci.*, 1999, vol. 46, no. 6, pp. 1836—1840.
32. Oldham T. R., McLean F. B. Total ionizing dose effects in MOS oxides and devices, *IEEE Trans. Nucl. Sci.*, 2003, vol. 50, no. 3, pp. 483—499.
33. Mustafaev G. A., Mustafaev A. G. Razrabotka processa formirovaniya glubokoj izoljacii struktur kremnij na izoljatore, *Nano- i mikrosistemnaya tekhnika*, 2009, no. 1, pp. 30—32 (in Russian).
34. McWhorter F. G., Winokur P. S. Simple technique for separating the effects of interface traps and trapped—oxide charge in metal—oxide—semiconductor transistors, *Appl. Phys. Lett.*, 1986, vol. 48, no. 2, pp. 133—135.
35. Esqueda S. I., Barnaby H. J., King M. P. Compact Modeling of Total Ionizing Dose and Aging Effects in MOS Technologies, *IEEE Transactions on Nuclear Science*, 2015, vol. 62, no. 4, pp. 1501—1515.
36. Dvornikov O. V., Grishkov V. N. Uchet vlijaniya pronikavushhej radiacii v "Spice-podobnyh" programmah, *Problemy razrabotki perspektivnyh mikro- i nanoelektronnyh sistem — 2010. Sbornik trudov* / ed. A. L. Stempkovskiy, Moscow, IPPM RAN, 2010, pp. 301—306 (in Russian).
37. Pershenkov V. S. Conversion model of the radiation-induced interface-trap building and its hardness assurance application, *FACTA UNIVERSITATIS, Series: Electronics and Energetics*, 2015, vol. 28, no. 4, pp. 557—570.
38. Gorbunov M., Danilov I., Zebrev G. et al. Verilog-A Modeling of Radiation-Induced Mismatch Enhancement, *IEEE Transactions on Nuclear Science*, 2011, vol. 58, pp. 785—792.
39. Esqueda I. S. Modeling of total ionizing dose effects in advanced complementary metal-oxide-semiconductor technologies, PhD thesis, Arizona State Univ., 2011, 179 p.
40. Gurkov V. A. *Tverdotel'naja elektronika*. 3-e izd., dop., Moscow, Tehnosfera, 2008, 512 p. (in Russian).
41. Wu W., Ten-Lon C., Gildenblat G. et al. Physics-based mathematical conditioning of the MOSFET surface potential equation, *IEEE Trans. Electron Devices*, 2004, vol. 51, pp. 1196—1199.
42. Saha S. *Compact Models for Integrated Circuit Design*, Boca Raton: CRC Press, 2015. 545 p.
43. Mustafaev G. A., Mustafaev A. G. Radiacionnaja stojkost' KNI MOP-tranzistorov k nakoplennoj doze ionizirujushhego izlucheniya, *Nano- i mikrosistemnaya tekhnika*, 2007, no. 12, pp. 47—49 (in Russian).

УДК 621.3.049.776

DOI: 10.17587/nmst.20.402-408

В. Д. Садков, канд. техн. наук, доц., **А. В. Пилькевич**, студент, **К. С. Фомина**, студент
Федеральное государственное бюджетное образовательное учреждение высшего образования
"Нижегородский государственный технический университет им. Р. Е. Алексеева" (НГТУ), ktpp@nntu.ru

МОДЕЛИРОВАНИЕ ПРОЦЕССА ЛАЗЕРНОЙ ПОДГОНКИ ПЛЕНОЧНЫХ ПОГЛОЩАЮЩИХ ЭЛЕМЕНТОВ

Поступила в редакцию 23.01.2018

Методом конформных отображений проведено математическое моделирование лазерной подгонки пленочных поглощающих элементов СВЧ аттенюаторов по ослаблению и входному (выходному) сопротивлению с учетом реальных размеров лазерных резов и последовательности их введения. Учет выявленных в предложенной модели зависимостей позволяет повысить точность подгонки, ее производительность и выход годных.

Работоспособность методики проверена моделированием в отечественном программном комплексе *Elcut*, а также подгонкой реальных образцов малогабаритных навесных поглощающих элементов микрополосковых аттенюаторов.

Ключевые слова: пленочный поглощающий элемент, лазерная подгонка, моделирование с учетом толщины лазерного реза

Введение

Пленочные поглощающие элементы (ПЭ) в виде прямоугольных диэлектрических пластин с нанесеной на них резистивной пленкой широко используют в волноводных, коаксиальных и микро-

полосковых аттенюаторах [1—5]. Качество ПЭ и устройств на их основе существенно зависит от точности воспроизведения входного (выходного) сопротивления R_{bx} (Ом) и ослабления g (дБ). Снижение разброса по R_{bx} и g , обусловленного техно-

логическими погрешностями изготовления, обеспечивается лазерной подгонкой [6], заключающейся во введении прорези симметрично вдоль оси удлиненной c и шириной $2e \geq 0$, а также в подрезке заземленных контактов на длину f с каждой стороны [7].

В основе методики подгонки лежит П-образная схема замещения с сопротивлением R_{13} в горизонтальном и $R_{12} = R_{23}$ в вертикальных плечах, причем введение тонкой центральной прорези оказывает влияние на R_{13} , а подрезка заземленных контактов на R_{13} и R_{12} одновременно. Поэтому подрезкой контактов обеспечивают требуемые R_{12} и R_{23} , а введением центральной прорези — R_{13} [7, 8].

В случае конечной ширины лазерного реза, что актуально для современных и перспективных малогабаритных ПЭ, и подрезка заземленных контактов, и центральная прорезь оказывают влияние сразу на R_{13} и R_{12} , что затрудняет подгонку. Использование только экспериментальных данных вследствие ограниченного их объема оказывается недостаточным для полного анализа процесса подгонки. Требуется построение математической модели лазерной подгонки, которому и посвящена настоящая работа.

Методика расчета

Задача решается методом конформных отображений [8, 9]. Топология несимметричного ПЭ представлена на рис. 1, a. Линии M_9M_{10} , M_3M_4 и M_6M_7 — входной, выходной и заземленный контакты соответственно. В большинстве случаев используются симметричные относительно оси x ПЭ, расчет которых сводится к расчету несимметричных ПЭ с уменьшенным вдвое удельным поверхностным сопротивлением ρ резистивной пленки.

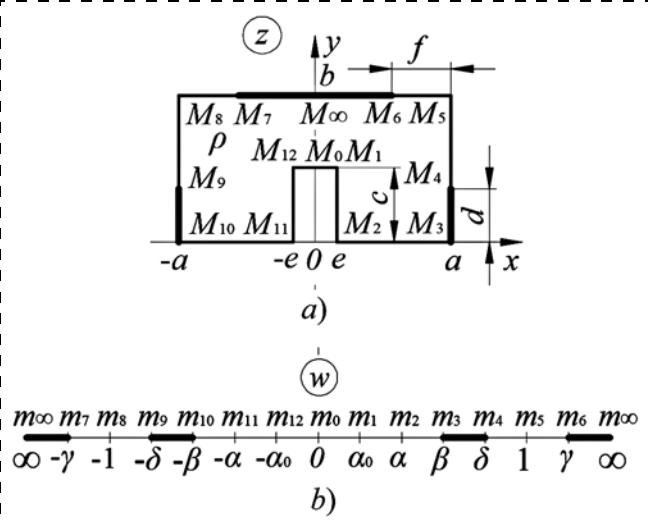


Рис. 1. Пленочный ПЭ с центральной прорезью размерами $2e$, c и длиной f подрезки заземленного контакта в плоскости z (a) и его отображение на верхнюю полуплоскость $Imw \geq 0$ (b)

Fig. 1. Film AE with the central cut of the size $2e$, c and length f of the cutting of the earthed contact in plane z (a) and its mapping on the top semiplane $Imw \geq 0$ (b)

Отображаем многоугольник $M_1M_2M_3M_5M_8M_{10}M_{11}M_{12}$ в плоскости z (рис. 1, a) на верхнюю полуплоскость $Imw \geq 0$ с указанным на рис. 1, b соответствием точек с помощью интеграла Кристоффеля — Шварца [9]:

$$z = A \int_0^w \sqrt{(w^2 - \alpha_0^2)/(w^2 - \alpha^2)(w^2 - \beta^2)(w^2 - 1)} dw. \quad (1)$$

В случае симметричного ПЭ нижняя часть (при $y \leq 0$) отображается на нижнюю полуплоскость $Imw \leq 0$ с тем же соответствием точек.

Константы α_0 , α , β определяются из системы уравнений

$$\left. \begin{aligned} e/c &= I_1/I_2, \quad e/(a-e) = I_1/I_3, \quad c/(a-e) = I_2/I_3, \\ I_1 &= \\ &= A \int_0^{\alpha_0} \sqrt{(\alpha_0^2 - w^2)/(\alpha^2 - w^2)(\beta^2 - w^2)(1 - w^2)} dw, \\ I_2 &= \\ &= A \int_{\alpha_0}^{\alpha} \sqrt{(w^2 - \alpha_0^2)/(w^2 - \alpha^2)(\beta^2 - w^2)(1 - w^2)} dw, \\ I_3 &= \\ &= A \int_{\alpha}^{\beta} \sqrt{(w^2 - \alpha_0^2)/(w^2 - \alpha^2)(\beta^2 - w^2)(1 - w^2)} dw. \end{aligned} \right\} (2)$$

Гиперэллиптические интегралы I_1 — I_3 вычисляют с использованием квадратурных формул наивысшей степени точности после приведения интервалов интегрирования к стандартному $(-1, 1)$ и выделению особенностей в весовую функцию $1/(1 - w^2)^{1/2}$ [10]:

$$\int_{-1}^1 \frac{g(w)dw}{\sqrt{1 - w^2}} \approx \frac{\pi}{N} \sum_{n=1}^N g\left(\cos \frac{2n-1}{2N}\pi\right).$$

После определения констант α_1 — α_3 величина α_4 находится из уравнения

$$\frac{d}{b} = \frac{\int_{\alpha_3}^{\alpha_4} \sqrt{(w^2 - \alpha_1^2)/(w^2 - \alpha_2^2)(w^2 - \alpha_3^2)(1 - w^2)} dw}{\int_{\alpha_3}^1 \sqrt{(w^2 - \alpha_1^2)/(w^2 - \alpha_2^2)(\alpha_3^2 - w^2)(1 - w^2)} dw}.$$

Для параметров схемы замещения симметричной относительно оси x структуры ПЭ (рис. 1, a) получаем [8]

$$\begin{aligned} R_{12} = R_{23} &= \frac{\rho}{2} \frac{K'(m_1)}{K(m_1)}; \quad m_1^2 = \frac{\beta^2}{\gamma^2} \frac{(\gamma^2 - \delta^2)}{(\gamma^2 - \beta^2)}, \\ R_{13} &= \rho / \left[\frac{K'(m)}{K(m)} - \frac{K(m_1)}{K'(m_1)} \right]; \quad m^2 = \frac{\delta^2 - \beta^2}{\gamma^2 - \beta^2}, \end{aligned} \quad (3)$$

где ρ — удельное поверхностное сопротивление резистивной пленки; $K(m)$ и $K(m_1)$ — полные эллиптические интегралы 1-го рода модуля m и m_1 соответственно; $K'(m)$ и $K'(m_1)$ — полные эллиптические интегралы 1-го рода дополнительного модуля $m' = (1 - m^2)^{1/2}$ и $m'_1 = (1 - m_1^2)^{1/2}$.

В частном случае узкой прорези $e = 0$, $\alpha_0 = 0$ интеграл (1) приводится к виду [9]:

$$\begin{aligned} z - (a + ib) &= A \int_1^w \frac{wdw}{\sqrt{(w^2 - \alpha^2)(w^2 - \beta^2)(w^2 - 1)}} = \\ &= \frac{A}{\sqrt{1 - \alpha^2}} F\left(\arcsin \sqrt{\frac{w^2 - 1}{w^2 - \beta^2}}, k\right), \end{aligned} \quad (4)$$

где A — константа отображения; $F(\varphi, k)$ — эллиптический интеграл 1-го рода аргумента φ и модуля $k = [(\beta^2 - \alpha^2)/(1 - \alpha^2)]^{1/2}$.

Из соответствия точек $z = ib$ и $w = \infty$ получаем $A = -a(1 - \alpha^2)^{1/2}/K$, где K — полный эллиптический интеграл 1-го рода модуля k .

Обращая интеграл (4) и используя далее эллиптические функции Якоби, получаем

$$\sqrt{\frac{w^2 - 1}{w^2 - \beta^2}} = \operatorname{sn}\left[\left(1 + \frac{ib - z}{a}\right)K, k\right]. \quad (5)$$

Из соответствия точек $z = a$ и $w = \beta$ следует соотношение для определения K и k :

$$b/a = K'/K, \quad (6)$$

где K и K' — полные эллиптические интегралы 1-го рода модуля k и дополнительного модуля $k' = (1 - k^2)^{1/2}$.

Соотношение (5) с учетом условия (6) перепишем в виде

$$\sqrt{\frac{w^2 - 1}{w^2 - \beta^2}} = \operatorname{dn}(zK/a, k)/k\operatorname{cn}(zK/a, k).$$

Из соответствия точек $z = ic$ и $w = 0$, $z = e$ и $w = \alpha$, $z = (a - f) + ib$ и $w = \gamma$, $z = a + id$ и $w = \delta$, $z = e + ic$ и $w = \alpha_0$ получаем [9]:

$$\begin{aligned} \alpha_0 &= \frac{\sqrt{1 - \beta^2} \operatorname{sn}^2 \{[1 - e/a + i(b - c)/a]K, k\}}{\operatorname{cn} \{[1 - e/a + i(b - c)/a]K, k\}}, \\ \alpha &= \frac{\sqrt{\beta^2 - k^2} \operatorname{sn}^2 [(1 - e/a)K, k]}{\operatorname{dn} [(1 - e/a)K, k]}, \\ \delta &= \frac{\sqrt{\beta^2 - k^2} \operatorname{sn}^2 [(d/b)K', k'] \operatorname{sn}^2 [(c/b)K', k']}{\operatorname{dn} [(d/b)K', k']}, \quad (7) \\ \gamma &= \frac{\sqrt{1 - \beta^2} \operatorname{sn}^2 [(f/a)K, k]}{\operatorname{cn} [(f/a)K, k]}, \\ \beta &= \operatorname{dn} [(1 - c/b)K', k']. \end{aligned}$$

Как видно, параметр β зависит только от длины прорези; α_0 , α , δ — от длины и ширины прорези; γ — от длины как прорези, так и подрезки заземленных контактов. При $e = f = c = 0$, $\beta = k$, $\alpha_0 = \alpha = 0$, $\gamma = 1$, $\delta = k/\operatorname{dn}(dK'/b, k')$ и соотношения (3) приводятся к известным выражениям [8]:

$$\begin{aligned} R_{\text{bx}} &= \frac{\rho}{2} \sqrt{\frac{K'(m)K(m_1)}{K(m)K'(m_1)}}, \\ q &= 20 \lg \left[1 + 2 \left/ \left(\sqrt{\frac{K'(m)K'(m_1)}{K(m)K(m_1)}} - 1 \right) \right. \right], \end{aligned} \quad (8)$$

$$m = \operatorname{cn}[(1 - p)K', k'], \quad m_1 = \operatorname{cn}(pK', k'), \quad p = d/b.$$

При $b/a \geq 3$, $R_{\text{bx}} = (\rho/2)\sqrt{(1-p)/p}$ и $q = 8,68/\sqrt{(1-p)p(b/2a)}$ дБ.

При $(b - c)/a \leq 1/3$, $R_{\text{bx}} = (\rho/2)K'(\sin\pi p/2)/K \times (\sin\pi p/2)$ и $q = 8,68\pi a/b$ дБ.

В другом частном случае $f = 0$, $e = a$ и $c < d$ $\alpha_0 = \alpha = \beta$ и для входного сопротивления и ослабления справедливы соотношения (8) при замене d и b на $d - c$ и $b - c$.

Введение узкой ($e \approx 0$) прорези эквивалентно увеличению длины a исходного ПЭ без прорези в $K'(t)/K(t)$ раз при прежних размерах d и b , $t = \operatorname{dn}[(1 - c/b)K', k']$. Параметры K' , k' находятся из соотношения $K'/K = b/a$.

Результаты моделирования для ПЭ при $b/a = 1$ приведены на рис. 2 и показывают влияние вводимых прорезей. Подгонка на 10...15 % по ослаблению и входному сопротивлению обеспечивается введением прорезей c/b и f/a , не превышающих 7 %.

Задача подгонки имеет единственное решение только при $e = 0$. По заданным значениям R_{bx} , g определяются значения R_{13} и R_{12} , m , m_1 и константы β , γ и δ :

$$R_{13} = R_{\text{bx}}(A^2 - 1)/(2A);$$

$$R_{12} = R_{23} = R_{\text{bx}}(A + 1)/(A - 1);$$

$$A = U_{\text{bx}}/U_{\text{вых}} = 10^{q/20};$$

$$\beta = mm_1/(m'm'_1); \quad \delta = m_1/m'_1, \quad \gamma = 1;$$

$$m'_1 = (1 - m_1^2)^{1/2}; \quad m' = (1 - m^2)^{1/2}.$$

При $e \neq 0$ множество решений находят следующим образом. Для любого выбранного $\beta < 1$ (задает длину прорези c) определяют величины $\delta = \beta m'/m'_1$, $\gamma = \delta m'_1/m = \beta m'm'_1/(mm_1)$, а выбором параметров α_0 и α из условия $0 \leq \alpha_0 < \alpha < \beta$ находится все множество решений задачи подгонки с различными размерами прорези и длинами подрезки заземленных контактов. Зависимости нормированных величин R_{13}/ρ и R_{12}/ρ от размеров прорезей для частного случая $b/a = 1$, $d/b = 0,5$

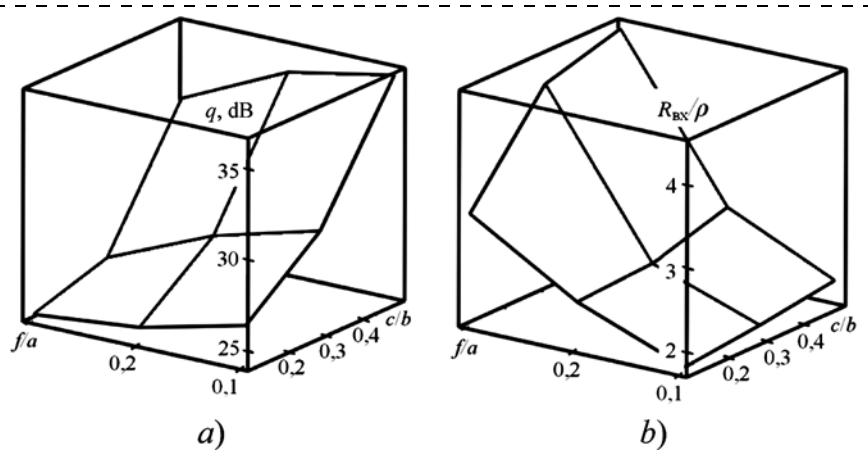


Рис. 2. Зависимость ослабления (а) и нормированного входного сопротивления (б) от f/a и c/b при $b/a = 1$, $d/b = 0,5$ и $e/a = 0,2$

Fig. 2. Dependence of attenuation (a) and normalized input resistance (b) on f/a and c/b at $b/a = 1$, $d/b = 0.5$ and $e/a = 0.2$

ность и выход годных без учета выявленных связей существенно снижаются.

Работоспособность методики проверена моделированием в отечественном программном комплексе *Elcut* и подгонкой образцов навесных ПЭ микрополосковых аттенюаторов с размерами 2×2 и 1×1 мм.

Заключение

Построена математическая модель лазерной подгонки ПЭ для коаксиальных, полосковых и микрополосковых аттенюаторов по ослаблению и входному (выходному) сопротивлению с учетом реальных размеров лазерных резов и последовательности их введения. Приведены зависимости ослабления, нормированного входного сопротивления и параметров П-образной схемы замещения, используемой при подгонке, от размеров лазерных резов. Отмечено, что без учета выявленных зависимостей точность подгонки, ее производительность и выход годных существенно снижаются. Работоспособность предложенной методики проверена моделированием в программном комплексе *Elcut* и экспериментами на реальных образцах малогабаритных ПЭ.

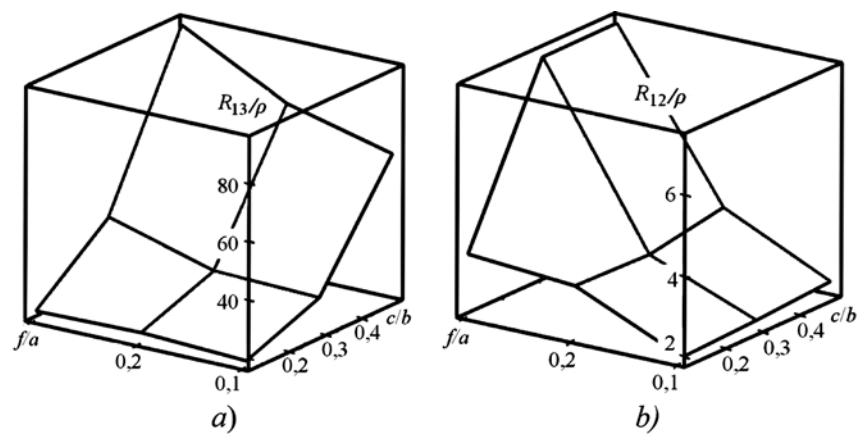


Рис. 3. Зависимость нормированных величин R_{13}/ρ (а) и R_{12}/ρ (б) от f/a и c/b при $b/a = 1$, $d/b = 0,5$ и $e/a = 0,2$

Fig. 3. Dependence of the normalized values R_{13}/ρ (a) and R_{12}/ρ (b) on f/a and c/b at $b/a = 1$, $d/b = 0.5$ and $e/a = 0.2$

приведены на рис. 3. Величины R_{13}/ρ и R_{12}/ρ зависят от всех параметров подгонки (f/a , c/b , e/a), причем для коротких ПЭ ($b/a > 5$) зависимости выражены резче.

Параметры R_{13} и R_{12} в реальном образце ПЭ определяют путем замыкания соответствующих контактов, при этом в симметричном ПЭ необходимо обеспечить симметрию вводимых прорезей и электрическую симметрию. По измеренным значениям R_{13} и R_{12} с помощью предложенной модели определяют параметры подгонки f/a и c/b при значении e/a , не меньшем диаметра лазерного пятна. Подрезку заземленных контактов для обеспечения требуемого значения R_{12} проводят с учетом последующего его увеличения за счет введения центральной прорези для получения требуемого R_{13} . При $e/a > 0$ точность подгонки, ее производитель-

Список литературы

- Львов А. Е., Рудоясова Л. Г., Шишков Г. И., Щербаков В. В. Тонкопленочные полосковые аттенюаторы // Труды НГТУ им. Р. Е. Алексеева. 2014. № 5. С. 40–46.
- Шишков Г. И. Широкополосные фиксированные аттенюаторы // Труды НГТУ им. Р. Е. Алексеева. 2011. № 1. С. 110–114.
- Чижиков С. В. Анализ тенденций развития чип-резистивных СВЧ-аттенюаторов // Успехи современной радиоэлектроники. 2017. № 8. С. 36–42.
- Гошин Г. Г., Екимов С. Ю., Семибраторов В. П., Фатеев А. В. Анализ и моделирование сверхширокополосных фиксированных аттенюаторов СВЧ-сигналов // Доклады ТУСУР. 2011. № 3 (24). С. 223–226.
- Седаков А. Ю., Смолин В. К. Тонкопленочные элементы в микроэлектронике: основы проектирования и изготовления. М.: Радиотехника, 2011. 168 с.
- Кондрашов В. В. Задача автоматизации процесса лазерной подгонки пленочных резистивных элементов // Вест-

ник Саратовского государственного технического университета. 2012. № 1. Вып. 2. С. 207–214.

7. **Садков В. Д.** Расчет и юстировка тонкопленочных звеньев затухания // Известия вузов СССР. Радиоэлектроника. 1977. Т. XX, № 9. С. 115–118.

8. **Ямпурин Н. П., Широков Л. В., Садков В. Д., Потехин В. А.** Современные проблемы радиоэлектроники с по-

зиций теории конформных отображений. Арзамас: АГПИ, 2014. 209 с.

9. **Лаврентьев М. А., Шабат Б. В.** Методы теории функций комплексного переменного. М.: Гл. ред. физ.-мат. лит., 1973. 749 с.

10. **Крылов В. И.** Приближенное вычисление интегралов. М.: Наука, 1967. 500 с.

V. D. Sadkov, Ph. D., Associate Professor, sadvd2016@mail.ru, **A. V. Pilkevich**, Student, **K. S. Fomina**, Student, Institute of Radio Engineering and Information Technologies, Nizhny Novgorod State Technical University named after R. Y. Alekseev (NNSTU), Nizhny Novgorod, 603950, Russian Federation

Corresponding author:

Sadkov Viktor D., Associate Professor, Institute of Radio Engineering and Information Technologies, Nizhny Novgorod State Technical University named after R. Y. Alekseev (NNSTU), Nizhny Novgorod, 603950, Russian Federation, e-mail: sadvd2016@mail.ru,

Modeling of the Laser Trimming Process of the Film Absorbing Elements

Received on January 23, 2018

Accepted on February 20, 2018

By the method of conformal mapping a mathematical modeling was carried out of the laser trimming of the film absorbing microwave elements of attenuators for attenuation and input (output) resistance with account of the real sizes of the laser cuts and the sequence of their introduction. The account of the dependencies identified in the proposed model allows us to improve the accuracy of trimming, its efficiency and product yield.

The operability of the technique was verified by modeling in Elcut domestic program complex and also by trimming of the real samples of the small-sized hinged absorbing elements of the microstrip attenuators.

Keywords: film absorbing element, laser trimming, modeling with account of laser cutting thickness, laser trimming, modeling

For citation:

Sadkov V. D., Pilkevich A. V., Fomina K. S. Modeling of the Laser Trimming Process of the Film Absorbing Elements, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 402–408.

DOI: 10.17587/nmst.20.402-408

Introduction

The film absorbing elements (AE) in the form of the rectangular dielectric plates with a deposited resistive film are widely used in the waveguide, coaxial and microstrip attenuators [1–5]. The quality of the devices on their basis depends essentially on the accuracy of reproduction of the input (output) resistance R_{BX} (Ω) and attenuation g (dB). A decrease of variation by R_{BX} and g , which is due to technological errors, is ensured by the laser trimming [6] consisting in introduction of a cut, symmetric to axis y with length c and width $2e \geq 0$, and also in cutting of the earthed contacts for length f from each side [7].

The method of trimming is based on a U-shaped equivalent circuit with resistance R_{13} in the horizontal arm and $R_{12} = R_{23}$ in the vertical arm, at that, introduction of a thin central cut influences R_{13} , while cutting of the earthed contacts influences R_{13} and R_{12} , simultaneously. Therefore, the cut of the contacts ensures the demanded R_{12} and R_{23} , while introduction of the central cut ensures R_{13} [7, 8].

In case of the final width of the laser cut, which is important for the small-sized AE, and cutting of the

earthed contacts, and the central cut have an influence on R_{13} and R_{12} , which complicates trimming. Use of only experimental data due to their limited volume appears insufficient for a complete analysis of the trimming process. Construction of a mathematical model of laser trimming is required. And this is exactly the aim of the present work.

Design procedure

The task is solved by the method of conformal mapping [8, 9]. The topology of an asymmetrical AE is presented in fig. 1, a. Lines M_9M_{10} , M_3M_4 and M_6M_7 are the input, output and earthed contacts, accordingly. In most cases, the symmetric in relation to axis x AE are used, the calculation of which boils down to calculation of the asymmetrical AE with twice as little specific surface resistance ρ of the resistive film.

We map a polygon of $M_1M_2M_3M_5M_8M_{10}M_{11}M_{12}$ in plane z (fig. 1, a) on the top of semiplane $Imw \geq 0$ with the specified in fig. 1, b conformity of points by means of Schwarz-Christoffel-integral [9]:

$$z = A \int_0^w \sqrt{(w^2 - \alpha_0^2)/(w^2 - \alpha^2)(w^2 - \beta^2)(w^2 - 1)} dw. \quad (1)$$

In case of a symmetrical AE the bottom part (at $y \leq 0$) is mapped on the bottom semiplane $\operatorname{Im} w \leq 0$ with the same correspondence of points.

Constants α_0, α, β are found from the following system of equations:

$$\left. \begin{aligned} e/c &= I_1/I_2, \quad e/(a-e) = I_1/I_3, \quad c/(a-e) = I_2/I_3, \\ I_1 &= \\ &= A \int_0^{\alpha_0} \sqrt{(\alpha_0^2 - w^2)/(\alpha^2 - w^2)(\beta^2 - w^2)(1 - w^2)} dw, \\ I_2 &= \\ &= A \int_{\alpha_0}^{\alpha} \sqrt{(\alpha^2 - w^2)/(\alpha^2 - w^2)(\beta^2 - w^2)(1 - w^2)} dw, \\ I_3 &= \\ &= A \int_{\alpha}^{\beta} \sqrt{(\alpha^2 - w^2)/(\alpha^2 - w^2)(\beta^2 - w^2)(1 - w^2)} dw. \end{aligned} \right\} \quad (2)$$

The hyperelliptic integrals $I_1 - I_3$ are calculated with the use of the quadrature formulas of the highest degree of accuracy after reduction of the integration intervals to the standard $(-1, 1)$ and isolation of the specific features into the weight function $1/(1 - w^2)^{1/2}$ [10]:

$$\int_{-1}^1 \frac{g(w)dw}{\sqrt{1 - w^2}} \approx \frac{\pi}{N} \sum_{n=1}^N g\left(\cos \frac{2n-1}{2N}\pi\right).$$

After determination of constants $\alpha_1 - \alpha_3$ the value of α_4 is found from the following equation:

$$\frac{d}{b} = \frac{\int_{\alpha_3}^{\alpha_4} \sqrt{(\alpha^2 - w^2)/(\alpha_2^2 - w^2)(\alpha_3^2 - w^2)(1 - w^2)} dw}{\int_{\alpha_3}^1 \sqrt{(\alpha^2 - w^2)/(\alpha_2^2 - w^2)(\alpha_3^2 - w^2)(1 - w^2)} dw}.$$

For the parameters of the equivalent circuit, symmetric in relation to axis x of AE structures, fig. 1, a , we get [8]:

$$\begin{aligned} R_{12} = R_{23} &= \frac{\rho}{2} \frac{K'(m_1)}{K(m_1)}; \quad m_1^2 = \frac{\beta^2}{\gamma^2} \frac{(\gamma^2 - \delta^2)}{(\gamma^2 - \beta^2)}; \\ R_{13} &= \rho / \left[\frac{K'(m)}{K(m)} - \frac{K(m_1)}{K'(m_1)} \right]; \quad m^2 = \frac{\delta^2 - \beta^2}{\gamma^2 - \beta^2}, \end{aligned} \quad (3)$$

where ρ — specific surface resistance of the resistive film; $K(m)$ and $K(m_1)$ — full elliptic integrals of the 1st kind of the modulus m and m_1 , accordingly; $K'(m)$ and $K'(m_1)$ — full elliptic integrals of the 1st kind of the additional modulus $m' = (1 - m^2)^{1/2}$ и $m'_1 = (1 - m_1^2)^{1/2}$.

In the particular case of narrow cut $e = 0, \alpha_0 = 0$ and integral (1) it is reduced to the following kind [9]:

$$\begin{aligned} z - (a + ib) &= A \int_1^w \frac{wdw}{\sqrt{(w^2 - \alpha^2)(w^2 - \beta^2)(w^2 - 1)}} = \\ &= \frac{A}{\sqrt{1 - \alpha^2}} F\left(\arcsin \sqrt{\frac{w^2 - 1}{w^2 - \beta^2}}, k\right), \end{aligned} \quad (4)$$

where A — mapping constant; $F(\varphi, k)$ — elliptic integral of the 1st kind of argument φ and modulus $k = [(\beta^2 - \alpha^2)/(1 - \alpha^2)]^{1/2}$.

From the conformity of points $z = ib$ and $w = \infty$ we get $A = -a(1 - \alpha^2)^{1/2}/K$, where K — full elliptic integral of the 1st kind of modulus k .

By turning integral (4) and using further the elliptic Jacobi functions, we receive

$$\sqrt{\frac{w^2 - 1}{w^2 - \beta^2}} = \operatorname{sn}\left[\left(1 + \frac{ib - z}{a}\right)K, k\right]. \quad (5)$$

From the conformity of points $z = a$ and $w = \beta$ the correlation follows for determination of K and k :

$$b/a = K'/K, \quad (6)$$

where K and K' — full elliptic integrals of the 1st kind of modulus k and additional modulus $k' = (1 - k^2)^{1/2}$.

We will rewrite correlation (5) with account of the condition (6) in the following way:

$$\sqrt{\frac{w^2 - 1}{w^2 - \beta^2}} = \operatorname{dn}(zK/a, k)/\operatorname{cn}(zK/a, k).$$

From the conformity of points $z = ic$ and $w = 0, z = e$ and $w = \alpha, z = (a - f) + ib$ and $w = \gamma, z = a + id$ and $w = \delta, z = e + ic$ and $w = \alpha_0$ we get [9]:

$$\begin{aligned} \alpha_0 &= \frac{\sqrt{1 - \beta^2} \operatorname{sn}^2\{[1 - e/a + i(b - c)/a]K, k\}}{\operatorname{cn}\{[1 - e/a + i(b - c)/a]K, k\}}, \\ \alpha &= \frac{\sqrt{\beta^2 - k^2} \operatorname{sn}^2[(1 - e/a)K, k]}{\operatorname{dn}[(1 - e/a)K, k]}, \\ \delta &= \frac{\sqrt{\beta^2 - k^2} \operatorname{sn}^2[(d/b)K', k'] \operatorname{sn}^2[(c/b)K', k']}{\operatorname{dn}[(d/b)K', k']}, \\ \gamma &= \frac{\sqrt{1 - \beta^2} \operatorname{sn}^2[(f/a)K, k]}{\operatorname{cn}[(f/a)K, k]}, \\ \beta &= \operatorname{dn}[(1 - c/b)K', k']. \end{aligned} \quad (7)$$

Apparently, parameter β depends only on the length of the cut; α_0, α, δ — on the length and the width of the cut; γ — on both the length of the cut, and of the cutting of the earthed contacts. At $e = f = c = 0, \beta = k$,

$\alpha_0 = \alpha = 0$, $\gamma = 1$, $\delta = k/\text{dn}(dK'/b, k')$ and correlations (3) are brought to the known expressions [8]:

$$R_{\text{BX}} = \frac{\rho}{2} \sqrt{\frac{K'(m)K(m_1)}{K(m)K'(m_1)}},$$

$$q = 20 \lg \left[1 + 2 / \left(\sqrt{\frac{K'(m)K(m_1)}{K(m)K'(m_1)}} - 1 \right) \right], \quad (8)$$

$$m = \text{cn}[(1-p)K', k'], \quad m_1 = \text{cn}(pK', k'), \quad p = d/b.$$

At $b/a \geq 3$, $R_{\text{BX}} = (\rho/2)\sqrt{(1-p)/p}$ and $q = 8,68/[\sqrt{(1-p)/p}(b/2a)]$ dB.

At $(b-c)/a \leq 1/3$, $R_{\text{BX}} = (\rho/2)K(\sin\pi p/2)/K(\sin\pi p/2)$ and $q = 8,68\pi a/b$ dB.

In another particular case of $f=0$, $e=a$ and $c < d$, $\alpha_0 = \alpha = \beta$ and for the input resistance and attenuation the correlations (8) are fair at replacement of d and b with $d-c$ and $b-c$.

Introduction of a narrow ($e \approx 0$) cut is equivalent to an increase of length a of the initial AE without cutting in $K'(t)/K(t)$ times at the former sizes of d and b , $t = \text{dn}[(1-c/b)K', k']$. Parameters K' , k' are found from correlation $K'/K = b/a$.

The results of modeling for AE at $b/a = 1$ are presented in fig. 2 and demonstrate the influence of the introduced cuts. Trimming by 10–15 % for attenuation and input resistance are ensured by introduction of cuts of c/b and f/a , not exceeding 7 %.

The trimming problem has a solution only at $e=0$. By the set values of R_{BX} , g , the values of R_{13} and R_{12} , m , m_1 and constants β , γ and δ are defined:

$$R_{13} = R_{\text{BX}}(A^2 - 1)/(2A);$$

$$R_{12} = R_{23} = R_{\text{BX}}(A + 1)/(A - 1);$$

$$A = U_{\text{BX}}/U_{\text{вых}} = 10^{q/20};$$

$$\beta = mm_1/(m'm'_1); \quad \delta = m_1/m'_1, \quad \gamma = 1;$$

$$m'_1 = (1 - m_1^2)^{1/2}; \quad m' = (1 - m^2)^{1/2}.$$

At $e \neq 0$ a set of solutions is found as follows. For any selected $\beta < 1$ (which sets the length of cut c) we define $\delta = \beta m'/m'_1$, $\gamma = \delta m'_1/m = \beta m'm'_1/(mm_1)$, while all the set of solutions to the problem of trimming with various sizes of the cuts and the lengths of the cuts of the earthed contacts are found by selection of parameters of α_0 and α from condition $0 \leq \alpha_0 < \alpha < \beta$. The dependences of the normalized values of R_{13}/ρ and R_{12}/ρ on the values of the cuts for the particular case of $b/a = 1$, $d/b = 0.5$ are presented in fig. 3. Values of R_{13}/ρ and R_{12}/ρ depend on all the parameters of trimming (f/a , c/b , e/a), at that, for short AE ($b/a > 5$) the dependences are expressed more sharply.

In a real sample of AE, parameters R_{13} and R_{12} are determined by closing of the corresponding contacts, at that, in a symmetric AE it is necessary to ensure the symmetry of the introduced cuts and the electric sym-

metry. By the measured values of R_{13} and R_{12} by means of the proposed model the trimming parameters of f/a and c/b are defined at the value of e/a not less than the diameter of a laser spot. Cutting of the earthed contacts necessary to ensure the demanded value of R_{12} is carried out with account of its subsequent increase due to introduction of the central cut for obtaining of the demanded R_{13} . At $e/a > 0$ the accuracy of trimming, its efficiency and product yield, without account of the revealed bonds, decrease essentially.

The operability of the technique was verified by modeling in *Elcut* domestic program complex and by trimming of the samples of the hinged AE of the microstrip attenuators with the sizes of 2×2 and 1×1 mm.

Conclusion

A mathematical model was constructed for laser trimming of AE for the coaxial, strip and microstrip attenuators, for attenuation of the input (output) resistance with account of the real sizes of the laser cuts and sequences of their introduction. Dependences were presented for attenuation, normalized input resistance and parameters of the U-shaped equivalent circuit, used for trimming, on the sizes of the laser cuts. It was pointed out, that without the discovered dependences the accuracy of trimming, its efficiency and product yield decrease essentially. The efficiency of the proposed technique was verified by modeling in *Elcut* program complex and by experiments on real samples of the small-sized AE.

References

1. L'vov A. E., Rudojasova L. G., Shishkov G. I., Shherbakov V. V. Tonkoplenochnye poloskovye attenuatory, *Trudy NGTU im. R. E. Alekseeva*, 2014, no. 5, pp. 40–46 (in Russian).
2. Shishkov G. I. Shirokopolosnye fiksirovannye attenuatory, *Trudy NGTU im. R. E. Alekseeva*, 2011, no. 1, pp. 110–114 (in Russian).
3. Chizhikov S. V. Analiz tendencij razvitiya chipe-rezistivnyh SVCh-attenuatorov, *Uspehi sovremennoj radioelektroniki*, 2017, no. 8, pp. 36–42. (in Russian).
4. Goshin G. G., Ekimov S. Ju., Semibratov V. P., Fateev A. V. Analiz i modelirovanie svershirokopolosnyh fiksirovannyh attenuatorov SVCh-signalov, *Doklady TUSUR*, 2011, no. 3 (24), pp. 223–226 (in Russian).
5. Sedakov A. Ju., Smolin V. K. *Tonkoplenochnye jelementy v mikroelektronike: osnovy proektirovaniya i izgotovlenija*, Moscow, Radiotekhnika, 2011, 168 p. (in Russian).
6. Kondrashov V. V. Zadacha avtomatizacii processa lazernoj podgonki pljonochnykh rezistivnykh jelementov, *Vestnik Saratovskogo gosudarstvennogo tehnicheskogo universiteta*, 2012, no. 1. issue. 2, pp. 207–214 (in Russian).
7. Sadkov V. D. Raschet i justirovka tonkoplenochnyh zven'ev zatuhaniya, *Izvestija vuzov SSSR — Radioelektronika*, 1977, vol. XX, no. 9, pp. 115–118 (in Russian).
8. Jampurin N. P., Shirokov L. V., Sadkov V. D., Potekhin V. A. Sovremennye problemy radioelektroniki s pozicij teorii konformnyh otobrazhenij, Arzamas, AGPI, 2014. 209 p. (in Russian).
9. Lavrent'ev M. A., Shabat B. V. *Metody teorii funkciij kompleksnogo peremennogo*, Moscow, Gl. red. fiz.-mat. lit., 1973, 749 p. (in Russian).
10. Krylov V. I. *Priblizhennoe vychislenie integralov*, Moscow, Nauka, 1967, 500 p. (in Russian).

МАТЕРИАЛОВЕДЧЕСКИЕ И ТЕХНОЛОГИЧЕСКИЕ ОСНОВЫ МНСТ SCIENCE OF MATERIALS AND TECHNOLOGICAL BASICS OF MNST

УДК 621.373.8

DOI: 10.17587/nmst.20.409-415

В. С. Кондратенко, д-р техн. наук, проф., e-mail: vsk1950@mail.ru,
Физико-технологический институт, Московский технологический университет, г. Москва
В. И. Иванов, зам. гл. технолога, e-mail: ivi061@gmail.com,
ООО "Технологии органической и печатной электроники", г. Москва

МЕТОД ПРЕЦИЗИОННОЙ ЛАЗЕРНОЙ РЕЗКИ КРЕМНИЕВЫХ ПЛАСТИН НА КРИСТАЛЛЫ С ОРГАНИЧЕСКИМИ СВЕТОИЗЛУЧАЮЩИМИ СТРУКТУРАМИ

Поступила в редакцию 25.01.2018

Описывается метод прецизионной резки кремниевых пластин на основе лазерного управляемого термораскальвания с нерабочей стороны, который позволяет обрабатывать сложные устройства, чувствительные к высоким температурам, а именно органические светоизлучающие диоды OLED (organic light emitting diode) или ОСИД, изготовленные по "сэндвич"-технологии.

Ключевые слова: лазерное управляемое термораскальвание (ЛУТ), органические светоизлучающие диоды (OLED или ОСИД), резка кремниевых подложек

Введение

Развитие органических полупроводников стимулирует интерес к производству дисплеев на основе органических светоизлучающих диодов (ОСИД) (рис. 1, см. вторую сторону обложки) [1], поскольку органические структуры могут менять физико-химические свойства (работу выхода электронов, оптоэлектронные свойства, электропроводность, адсорбцию, плотность и электромагнитное поглощение) и позволяют изготавливать уникальные микроэлектронные устройства, в частности микродисплеи с высоким разрешением.

На сегодняшний день микродисплеи находят широкое применение в самых различных областях техники. Областями применения микродисплеев являются, прежде всего, высокоэффективные индивидуальные средства отображения информации: наголовные или нашлемные видеомодули; устройства визуализации и видоискатели фото- и видеокамер, тепловизоров; видеомодули, входящие в экипировку военнослужащих, различного рода ремонтных бригад, работающих в сложных условиях; технические средства информационно-развлекательной сферы, особенно в рамках развития 3D-технологий отображения графической информации. Индивидуальные средства отображения информации должны иметь малую массу, небольшие размеры и низкое энергопотребление, обеспечивая при этом необходимые светотехнические параметры.

Развитие микроэлектроники ведет к уменьшению размеров микродисплеев при одновременном увеличении разрешающей способности.

По мере развития микродисплейной техники будут появляться все новые области их применения, поскольку их использование существенно облегчает выполнение различных задач практически во всех областях человеческой жизнедеятельности.

Методы резки кремниевых пластин с ОСИД

Обработка кремниевых пластин с органическими структурами имеет ряд существенных ограничений, поскольку температуры выше +90 °C, воздействия воды, кислорода и вибрации, возникающие в процессе резки кремниевых пластин с органическими структурами, вызывают деградацию ОСИД [2].

Передовые производители микродисплеев на ОСИД применяют наиболее распространенный метод дисковой алмазной резки кремниевых пластин, который имеет следующие качественные недостатки [2–4]: дефекты вдоль линии реза, которые влияют на качество кристаллов; долговременные температурные воздействия и большое количество воды, приводящие к деградации и разрушению органических структур.

В производстве микродисплеев на ОСИД МДО 01 [5] на операцию резки поступают светоизлучающие матрицы, изготовленные на кремни-

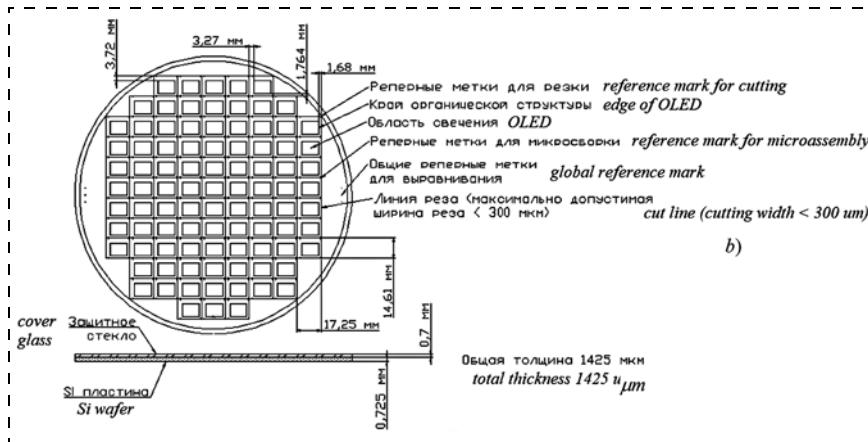


Рис. 2. Чертеж подложки с размерами кристаллов микродисплеев и размерами, имеющими значение при резке

Fig. 2. Drawing of a substrate with the dimensions of the microdisplay crystals and the dimensions important for cutting

евой подложке $\varnothing 200$ мм, состоящие из кристаллов, которые представляют собой кремниевые интегральные схемы. На поверхности этих схем сформированы светоизлучающие органические структуры с защитной стеклянной крышкой $\varnothing 200$ мм, так называемые "сэндвичи" (рис. 2). Требования к качеству и точности обработки кремниевых подложек со структурами ОСИД высокие, ввиду их высокой себестоимости.

Дисковая резка не позволяет управлять качеством кремниевых кристаллов из-за сколов, ввиду износа инструмента нужно отслеживать каждый рез и при необходимости менять технологические режимы или сам режущий диск. Кромки кремниевых кристаллов со сколами после дисковой алмазной резки представлены на рис. 3 (см. вторую сторону обложки). Скорость дисковой резки составляет всего 1...4 мм/с (низкая производительность), а средний выход годных кристаллов с ОСИД 53,9 %.

Применение прецизионных лазерных методов при резке в производстве изделий на базе ОСИД ограничено минимальным расстоянием от линии реза до органических структур на кристаллах, данное расстояние для МДО 01 равно 1,68 мм. Поэтому на базе метода ЛУТ [6, 7] был разработан и внедрен новый метод прецизионной лазерной резки кремниевых приборных пластин с нерабочей стороны в сборе с защитным стеклом [8–10], благодаря ИК-камере с чувствительностью в спектре 1460...1625 нм, позволяющей получать изображение пластины, расположенной лицевой стороной вниз. Разработанный метод позволяет управлять качеством

за счет следующих возможностей: резки кремниевых пластин в сборе с защитным стеклом; снижения нагрева органических структур; уменьшения ширины реза с рабочей стороны практически до 0; устранения дефектов по рабочим кромкам кристаллов за счет отсутствия сколов.

Теплота, выделяемая лазерными источниками нагрева при нанесении предварительных дефектов и последующем ЛУТ кремниевых пластин, распространяется на прилегающие к линиям реза участки. При нагреве и последующем остывании в этих участках изменяются структура и свойства кремния. Участок пластины, подвергающийся в процессе резки нагреву до температуры, при которой происходят видимые или невидимые изменения, называют зоной термического влияния (ЗТВ). Наряду с тепловым воздействием кремний в области ЗТВ подвергается пластической деформации.

Нагрев разных участков ЗТВ происходит в пределах от высокой температуры (по линии реза) до начальной температуры кремниевой пластины. Размеры и строение ЗТВ зависят от теплофизических характеристик, химического состава материала и термического цикла лазерной обработки.

Под влиянием термического воздействия и вызванных им деформаций при процессе ЛУТ может наблюдаться деградация ОСИД. Технологический процесс лазерной резки должен субсидировать оптимальные условия формирования сквозной трещины и получения надлежащего качества кремния и отсутствия деградации органических структур, от этого зависит работоспособность получаемых кристаллов с ОСИД.

Математическая модель лазерной резки

Технологические параметры оптимизируют путем математического моделирования [11, 12]. Расчитывают температуры, возникающие при лазерной резке кремниевой пластины с обеих ее сторон в зоне обработки. На рис. 4 представлено

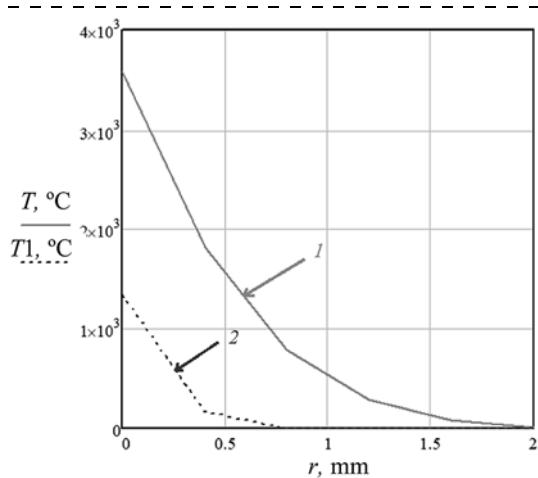


Рис. 4. Распределения температуры при нанесении первоначальных дефектов импульсным УФ лазером со стороны реза (1) и с обратной стороны (2)

Fig. 4. Temperature distribution during deposition of the initial defects by a pulse UV laser from the cut side (1) and from the reverse side (2)

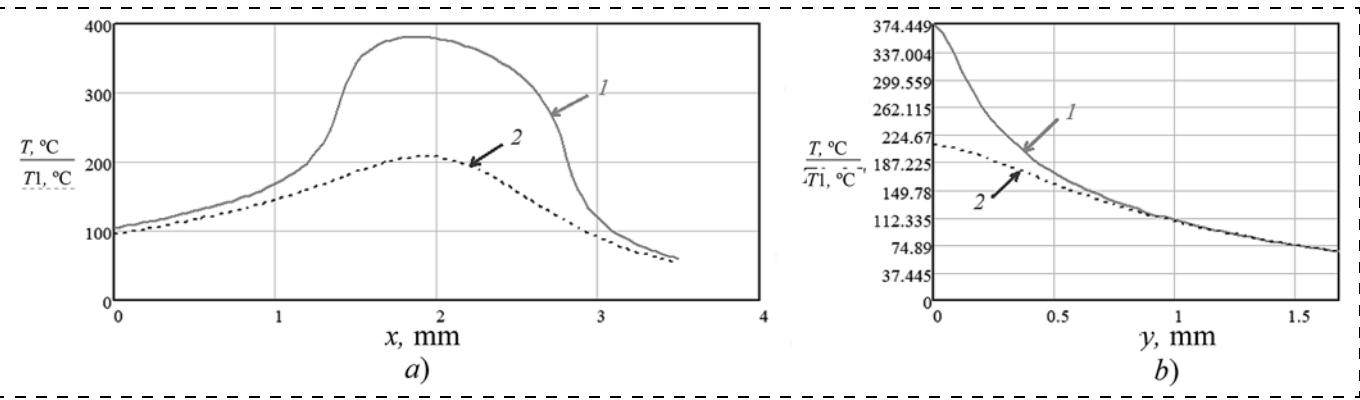


Рис. 5. Распределение температуры при ЛУТ кремниевых пластин по поверхностям со стороны реза (1) и с обратной стороны (2) вдоль линии реза (а) и в зоне максимального нагрева поперек линии реза (б)

Fig. 5. Temperature distribution during LCT of the silicon plates on the surfaces from the cut side (1) and from the reverse side (2) along the cut line (a) and in the zone of the maximal heating across the cut line (b)

распределение температуры при нанесении первоначальных дефектов импульсным УФ лазером.

Таким образом определены оптимальные параметры нанесения первоначальных дефектов УФ лазером ($\lambda = 356$ нм): мощность $\sim 4,5$ Вт, частота модуляции 100 кГц, размеры пучка $0,05 \times 0,05$ мм, скорость 30 мм/с.

Распределения температуры при ЛУТ представлены на рис. 5.

Для ЛУТ оптимальные параметры следующие: мощность излучения 200 Вт; размеры эллиптического лазерного пучка $2a \times 2b \approx 1400 \times 200$ мкм; скорость 80 мм/с [11].

Математическое моделирование позволило определить температуру нагрева рабочей стороны пластины, исключающую деградацию структур ОСИД, на расстоянии 1,68 мм от линии реза, при оптимальных параметрах температура ниже критической ($T_{\text{кр}} = 90$ °С).

Оценка качества резки кремниевых пластин с ОСИД на кристаллы

В табл. 1 приведены параметры качества кристаллов с ОСИД для производства микрорадиодисплеев МДО 01.

В соответствии с параметрами качества разработана и внедрена система технического

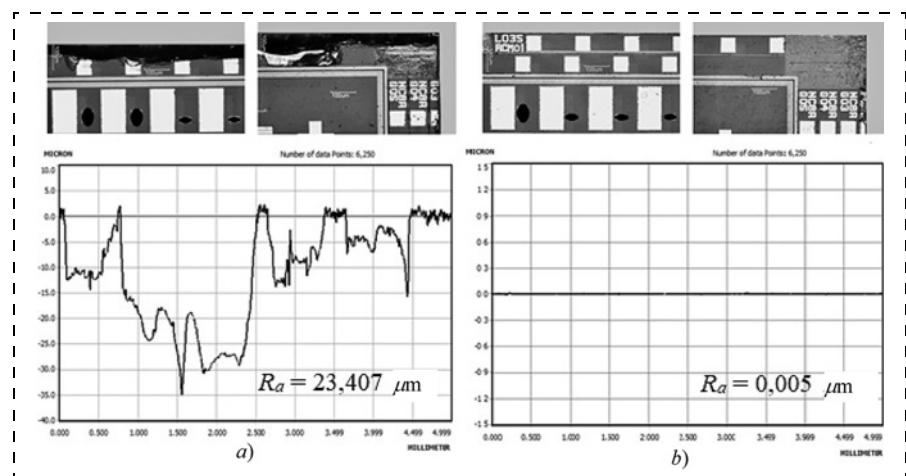


Рис. 6. Фотографии кромок кристаллов и профилограммы поверхностей их кромок после дисковой резки (а) и после ЛУТ (б)

Fig. 6. Photos of the edges of crystals and surface profilograms of their edges after a disk cutting (a) and after LCT (b)

Таблица 1

Table 1

Параметры качества кремниевых кристаллов с ОСИД для микродисплеев МДО 01
Quality parameters of the silicon crystals with OLED for microdisplays MDO 01

Наименование параметра, единица измерения <i>Parameters, units of measurement</i>	Значения параметра <i>Parameters</i>		
	Наименьшее <i>Lowest</i>	Номинальное <i>Nominal</i>	Наибольшее <i>Highest</i>
Размеры кристалла $W \times L$, мм <i>Crystal dimensions $W \times L$, mm</i>	—	$14,508 \times 17,15$	$14,61 \times 17,25$
Шероховатость торцов, R_a , мкм <i>Roughness of the end faces, R_a, μm</i>	$\leq 0,008$	$\leq 12,5$	25
Дефекты на рабочей поверхности кристалла вдоль линий реза*, мм <i>Defects on the working surface of a crystal along the cut lines*, mm</i>	—	—	$\leq 0,05$
Яркость ОСИД**, кд/м ² <i>Brightness of OLED**, cd/m²</i>	150	540	800
Неравномерность яркости, % <i>Non-uniformity of brightness, %</i>	—	—	15

* Повреждение контактных площадок не допускается

** Damage of the pads is not admissible

** Деградация структуры ОСИД не допускается

** Degradation of OLED structure is not admissible

Таблица 2
Table 2

Соответствие кристаллов с ОСИД для микрорадиодисплеев МДО 01 критериям качества
Conformity of the crystals with OLED for MDO 01 microdisplays to the quality criteria

Контрольный образец № <i>Control sample no.</i>	Размеры кристалла Ш × Д, мм <i>Crystal dimensions W × L, mm</i>	R _a , мкм Ra, μm	Дефекты на рабочей поверхности кристалла, мм <i>Defects on the working surface of a crystal, mm</i>	V _{an} , В V _{an} , V	I, мА I, mA	Яркость, кд/м ² Brightness, cd/m ²	Неравномерность яркости, % <i>Non-uniformity of brightness, %</i>
1	14,6 × 17,23					543	
2	14,6 × 17,25					579	
3	14,61 × 17,24	0,005	Отсутствуют <i>Absent</i>	4	5	702	
4	14,61 × 17,25					606	
5	14,6 × 17,24					565	

межоперационного контроля качества кристаллов с ОСИД для обеспечения надежности и эффективности производства [12, 13], которая позволяет в полной мере оценивать качество кристаллов и влияние температурного нагрева и источников лазерного излучения на органические структуры.

Проведено сравнение и количественная оценка качества торцов кристаллов на лицевой стороне; после дисковой алмазной резки и после ЛУТ была измерена шероховатость поверхности кристаллов по рабочим кромкам (рис. 6) с помощью профилометра модели XP-200 фирмы Ambios (США).

Результаты измерений параметров качества опытных образцов МДО 01, полученных при производстве с применением лазерной резки (табл. 2), соответствуют критериям качества.

Заключение

Новый прецизионный метод лазерной резки кремниевых подложек с ОСИД имеет следующие преимущества по сравнению с дисковой алмазной резкой: бездефектная рабочая кромка; отсутствует деградация структур ОСИД; высокая скорость 80 мм/с на порядок увеличивает производительность; выход годных кристаллов с ОСИД выше на 45,7 %.

Список литературы

1. **IDTechEx.** URL: <https://www.idtechex.com/research/reports/oled-display-forecasts-2016-2026-the-rise-of-plastic-and-flexible-displays-000477.asp>.
2. **Иванов В. И.** Методы резки кремниевых приборных пластин на чипы в производстве органических микрорадиодисплеев // Интернет-журнал Науковедение. 2014. № 4 (23). URL: <http://naukovedenie.ru/PDF/87TVN414.pdf>.
3. **Кондратенко В. С., Иванов В. И.** Влияние методов резки кремниевых подложек на качество органических светоизлучающих диодов // Прикладная физика. 2017. № 1. С. 36–40.
4. **Кондратенко В. С., Борисовский В. Е., Иванов В. И.** Разделение органических светоизлучающих диодов на кремниевой подложке методом лазерного управляемого термораскалывания // Оборонный комплекс — научно-техническому прогрессу России. 2014. № 2. С. 76–81.
5. **Усов Н. Н., Грачев О. А., Кондратенко Б. А., Нуриев А. В.** и др. Микрорадиодисплеи на основе органических светодиодов МДО 01 // Современная электроника. 2016. № 1. С. 34–37.
6. **Кондратенко В. С.** Способ резки неметаллических материалов. Патент РФ № 2024441 по заявке № 5030537/33 от 02.04.1992. Опубл. 15.12.1994. Бюл. 23.
7. **Кондратенко В. С., Наумов А. С.** Способ резки пластин из хрупких материалов. Патент РФ № 2404931 по заявке № 2009132338/08 от 28.08.2009. Опубл. 27.11.2010. Бюл. 33.
8. **Кондратенко В. С., Борисовский В. Е., Иванов В. И., Зобов А. К.** Повышение эффективности процесса лазерной резки кремниевых приборных пластин на кристаллы ОСИД // Приборы. 2015. № 9. С. 49–55.
9. **Кондратенко В. С., Иванов В. И.** Технология лазерной резки кремниевых пластин на кристаллы органических светоизлучающих диодов // Российский технологический журнал. 2016. Т. 4, № 3. С. 11–17.
10. **Кондратенко В. С., Иванов В. И., Зобов А. К.** Применение метода лазерного управляемого термораскалывания при разделении кремниевых приборных пластин с ОСИД в сборе с защитным стеклом // Сб. науч. трудов: материалы Всероссийской научно-технической конференции "Оптические технологии, материалы и системы" ("ОПТОТЕХ—2016"): Московский технологический университет, Физико-технологический институт. Вып. I. (электронный ресурс) / Под ред. Кондратенко В. С. М.: МТУ, 2016. С. 53–57.
11. **Кондратенко В. С., Борисовский В. Е., Иванов В. И.** Оптимизация лазерного термораскалывания кремниевых пластин на кристаллы с органическими светоизлучающими структурами // Инновационные технологии в науке и образовании: материалы II Междунар. науч.-практ. конф., Чебоксары, 07 авг. 2015 г. Чебоксары: ЦНС Интерактив плюс, 2015. С. 175–183.
12. **Кондратенко В. С., Иванов В. И.** Управление качеством кристаллов ОСИД за счет применения нового метода лазерной резки кремниевых подложек // Сб. науч. трудов: материалы Международной научно-технической конференции "Информатика и технологии. Инновационные технологии в промышленности и информатике". Москва: Московский технологический университет, Физико-технологический институт. Вып. 23 (XXIII) / Под ред. Булатова М. Ф. 2017. С. 48–52.
13. **Иванов В. И.** Методы контроля качества кремниевых кристаллов с ОСИД после лазерной резки // Результаты научных исследований: сб. статей VII Междунар. науч.-практ. конф., 5 октября 2015 г., г. Екатеринбург. Уфа: АЭТЕРНА, 2015. С. 43–48.

V. S. Kondratenko, D. Sc., Professor, vsk1950@mail.ru,
Moscow Technological University, Moscow, 119454, Russian Federation,
V. I. Ivanov, Deputy Chief Technologist, ivi061@gmail.com,
"Technologies of organic and printed electronics" Ltd., Moscow, 107207, Russian Federation

Corresponding author:

Ivanov Vladimir I., Deputy Chief Technologist, "Technologies of organic and printed electronics" Ltd., Moscow, 107207, Russian Federation, e-mail: ivi061@gmail.com

Method of Precision Laser Cutting of Silicon Wafers into Crystals with Organic Light-Emitting Structures

Received on January 25, 2018
Accepted on February 13, 2018

Requirements for high precision and improved surface finish and demand for production of microdisplays based on organic light emitting diodes (OLED) explain the necessity for effective techniques for cutting of the silicon device wafers. This paper deals with the quality of cutting of the silicon device wafers into chips with the organic light-emitting diodes. This work is aimed to introduce a high efficiency technological precision laser cutting process of the silicon wafers into production of the microdisplays based on OLED. The article describes the precision cutting method of the silicon wafers based on the laser controlled thermocracking (LCT) from the non-working side, which allows processing of high-end technology devices sensitive to high temperatures, namely, the organic light-emitting diodes made by the sandwich technology. This technological process is controlled and has less aggressive modes unlike the other mechanical and laser methods, such as mechanical dicing, scribing, laser scribing and laser cutting. The article presents a qualitative analysis of the crystal edges after cutting. The article also discusses modeling of a moving laser source during the laser cutting of the silicon wafers into crystals based on the laser controlled thermocracking. The authors prove the efficiency of the mathematical modeling for assessment of the heating effects on the organic light-emitting structures. This paper substantiates and experimentally proves possibility and efficiency of the laser controlled thermocracking method in OLED technology.

Keywords: laser controlled thermocracking (LCT), organic light-emitting diodes (OLED), cutting of silicon wafers

For citation:

Kondratenko V. S., Ivanov V. I. Method of Precision Laser Cutting of Silicon Wafers into Crystals with Organic Light-Emitting Structures, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 409–415.

DOI: 10.17587/nmst.20.409-415

Introduction

Development of organic semiconductors stimulates interest to manufacture of the displays based on the organic light emitting diodes (OLED) (fig. 1, see the 2nd side of cover) [1], because the organic structures can change the physical and chemical properties (work function of electrons, optoelectronic properties, electroconductivity, adsorption, density and electromagnetic absorption) and allow us to manufacture unique microelectronic devices, in particular, microdisplays with high resolution.

The microdisplays find wide application in most varied technologies. Their spheres are, first of all, highly efficient individual means for displaying of information: head or helmet video modules; visualization devices and view-finders of the photo- and videocameras, thermal imagers; the video modules as part of equipment for military men, various kinds of the repair brigades working in difficult conditions; hardware for the information-entertaining sphere, especially within development of 3D technologies for displaying of graphic information.

The individual information display devices should have small weight, small dimensions and low energy consumption, ensuring the necessary lighting parameters. Development of microelectronics leads to reduction of the dimensions of the microdisplays at a simultaneous increase of their resolution.

In the process of development of the microdisplay technologies, new areas of their application appear, because their use facilitates solving of problems practically in all the areas of human life.

Methods for cutting of the silicon plates with OLED

Processing of the silicon plates with the organic structures has a number of essential restrictions, because temperatures over +90°C, influence of water, oxygen and vibrations arising in the process of cutting of the silicon plates with organic structures, cause degradation of OLED [2].

The leading producers of OLED microdisplays apply the most wide-spread method of the disk diamond cutting of the silicon plates, which has the following drawbacks [2–4]: defects along the cut line influence the quality of the crystals; long-term temperature influences and a considerable quantity of water lead to a degradation and destruction of the organic structures.

Production of OLED MDO 01 microdisplays envisages cutting of the light-emitting matrixes made on Ø200 mm silicon substrates and consisting of crystals, which are silicon integral circuits. On the surface of the circuits the light-emitting organic structures are formed with a protective Ø200 mm glass cover, sandwiches (fig. 2). The requirements to the quality and accuracy of processing of the silicon substrates with OLED structures are high in view of their high cost price.

The disk cutting does not allow us to control the quality of the silicon crystals, because of the cleavages and because of the tear and wear of the tools it is necessary to check every cut and, if necessary, to change the technological modes or the cutting disk. The edges of the silicon crystals with cleavages after a disk diamond cutting are presented in fig. 3, see the 2nd side of cover. The speed of the disk cutting is only 1...4 mm/s (low productivity), while an average product yield of the crystals with OLED is 53.9 %.

Application of the precision laser methods for cutting in manufacture of products on the basis of OLED is limited by the minimal distance from a cut line to the organic structures on crystals, and the given distance for MDO 01 is equal to 1.68 mm. Therefore, on the basis of LCT [6, 7] a method was developed and introduced for precision laser cutting of the silicon instrument plates from the non-working side in assembly with a protective glass [8—10], due to an IR chamber with sensitivity in the spectrum of 1460...1625 nm, allowing us to obtain an image of the plate located with its face downwards. The developed method ensures control of the quality because it allows us to do the following: to cut the silicon plates in assembly with a protective glass; to lower heating of the organic structures; to decrease the width of a cut from the working side practically down to 0; to eliminate the defects on the working edges of the crystals due to absence of cleavages.

The warmth emitted by the laser sources of heating during deposition of the preliminary defects and subsequent LCT of the silicon plates, extends to the sites adjoining the cut lines. Heating and cooling change the structure and properties of the silicon in these sites. The site of a plate, which during cutting is exposed to heating up to the temperature which makes visible or invisible changes, is called the zone of thermal influence (ZTI). Alongside with the thermal influence, the silicon in the area of ZTI is subjected to a plastic deformation.

Heating of different sites of ZTI occurs from a high temperature (on the cut line) up to the initial temperature of a silicon plate. The dimensions and structure of ZTI depend on the thermal characteristics, chemical compound of a material and the thermal cycle of the laser processing.

Under the thermal influence and due to the deformations caused by it during the LCT the process of degradation of OLED may be observed. The technological process of the laser cutting should subsidize the optimal conditions for formation of a through crack, obtaining of the appropriate quality of the silicon and absence of degradation of the organic structures, the operability of the received crystals depends on this.

Mathematical model of laser cutting

The technological parameters are optimized by mathematical modeling [11, 12]. The temperatures, arising during the laser cutting of a silicon plate from both sides, are calculated. Fig. 4 presents distribution of temperatures during deposition of the initial defects by a pulse UV laser.

Thus the optimal parameters were determined for deposition of the initial defects by an UV laser ($\lambda = 356$ nm): power ~4.5 W, frequency modulation — 100 kHz, dimensions of the beam — 0.05×0.05 mm, speed — 30 mm/s.

Distributions of temperatures during LCT are presented in fig. 5.

For LCT the optimal parameters are the following: radiation power — 200 W, dimensions of the elliptic laser beam — $2a \times 2b \approx 1400 \times 200$ μm , speed — 80 mm/s [11].

The mathematical modeling allowed us to determine the temperature of heating of the working side of the plate, excluding degradation of the OLED structures, at the distance of 1.68 mm from the cut line at the optimal parameters of temperature below the critical level $T_{kp} = 90$ °C.

Estimation of quality of cutting of the silicon plates with OLED into crystals

Table 1 presents the parameters of quality of the crystals with OLED for manufacture of microdisplays MDO 01.

In accordance with the quality parameters, the system of technical interoperational quality control of crystals with OLED was developed and introduced in order to ensure reliability and production efficiency [12, 13], which allows us to estimate fully the quality of the crystals and influence of the temperature heating and sources of laser radiation on the organic structures.

A comparison and quantitative estimation of the quality of the end faces of the crystals on the face side were done; after the disk diamond cutting and after LCT the roughness of the surface of the crystals on the working edges (fig. 6) was measured by means of XP-200 profilometer from Ambios Co. (USA).

The results of measurements of the quality parameters of the MDO 01 prototypes manufactured with application of laser cutting (table 2) correspond to the quality criteria.

Conclusion

The new method for laser cutting of the silicon substrates with OLED has the following advantages in comparison with the disk diamond cutting: faultless working edges; no degradation of OLED structures; high speed of 80 mm/s ensures a 10-fold productivity increase; product yield of crystals with OLED is 45.7 % higher.

References

1. **IDTechEx.** Available at: <https://www.idtechex.com/research/reports/oled-display-forecasts-2016-2026-the-rise-of-plastic-and-flexible-displays-000477.asp>.
2. **Ivanov V. I.** Techniques for cutting silicon device wafers into chips in the production of OLED microdisplays, *Internet-journal Naukovedenie*, 2014, no. 4 (23). Available at: <http://naukovedenie.ru/PDF/87TVN414.pdf>.
3. **Kondratenko V. S., Ivanov V. I.** Vlijanie metodov rezki kremnievyh podlozhek na kachestvo organicheskikh svetoizluchajushhih diodov, *Prikladnaja fizika*, 2017, no. 1, pp. 36—40 (in Russian).
4. **Kondratenko V. S., Borisovskij V. E., Ivanov V. I.** Razdelenie organicheskikh svetoizluchajushhih diodov na kremnievoj podlozhke metodom lazernogo upravlyaemogo termoraskalyvaniya, *Oboronnyj kompleks — nauchno-tehnicheskому progressu Rossii*, 2014, no. 2, pp. 76—81 (in Russian).
5. **Usov N. N., Grachev O. A., Kondrackij B. A., Nuriev A. V.** et al. Mikrodisplei na osnove organicheskikh svetodiodev MDO 01, *Sovremennaja elektronika*, 2016, no. 1, pp. 34—37 (in Russian).
6. **Patent RF № 2024441**, MKI5 SO3 V 33/09. Sposob rezki nevettallitsheskikh materialov, V. S. Kondratenko, Zajavl. № 5020537/33 ot 02.04.1992; Opubl. 15.12.1994, Bjul. no. 23 (in Russian).

7. Patent RF № 2404931, MKI5 SO3 V 33/09. Sposob rezki plastin iz hrupkikh materialov, V. S. Kondratenko, A. S. Naumov, Zajavl. № 2009132338/03 ot 28.08.2009; Opubl. 27.11.2010, Bjul. no. 33 (in Russian).
8. Kondratenko V. S., Borisovskij V. E., Ivanov V. I., Zobov A. K. Povyshenie effektivnosti processa lazernoj rezki kremnievyh pribornyh plastin na kristally OLED, *Pribory*, 2015, no. 9, pp. 49–55 (in Russian).
9. Kondratenko V. S., Ivanov V. I. Tehnologija lazernoj rezki kremnievyh plastin na kristally organicheskikh svetoizluchajushhih diodov, *Rossijskij tehnologicheskij zhurnal*, 2016, vol. 4, no. 3, pp. 11–17 (in Russian).
10. Kondratenko V. S., Ivanov V. I., Zobov A. K. Primeneenie metoda lazernogo upravljajemogo termoraskalyvanija pri razdele-nii kremnievyh pribornyh plastin s OLED v sbore s zashhitnym steklom, *Sbornik nauchnyh trudov: materialy Vserossijskoj nauchno-tehnicheskoy konferencii "Opticheskie tehnologii, materialy i sistemy" ("OPTOTEH—2016")*: Moskovskij tehnologicheskij universitet, Fiziko-tehnologicheskij institut. Part I (electronic source) / Ed. prof. Kondratenko V. S. Moscow, MTU, 2016, pp. 53–57 (in Russian).
11. Kondratenko V. S., Borisovskij V. E., Ivanov V. I. Optimizacija lazernogo termoraskalyvanija kremnievyh plastin na kristally s organicheskimi svetoizluchajushchihi strukturami, *Innovacionnye tehnologii v nauke i obrazovanii: materialy II Mezhdunar. nauch.-prakt. konf.*, Cheboksary, 07 avg., 2015. Cheboksary, CNS "Interaktiv pljus", 2015, pp. 175–183 (in Russian).
12. Kondratenko V. S., Ivanov V. I. Upravlenie kachestvom kristallov OLED za schet primeneija novogo metoda lazernoj rezki kremnievyh podlozhek, *Sbornik nauchnyh trudov: materialy Mezhdunarodnoj nauchno-tehnicheskoy konferencii "Informatika i tehnologii. Innovacionnye tehnologii v promyshlennosti i informatiche"*: Moscow technological university, Physico-technological institute. N.23 (XXIII), Ed. prof. Bulatov M. F., Moscow, 2017, pp. 48–52 (in Russian).
13. Ivanov V. I. Metody kontrolja kachestva kremnievyh kristallov s OLED posle lazernoj rezki, *Rezul'taty nauchnyh issledovanij: sbornik statej VII Mezhdunar. nauch.-prakt. konf.*, 5 oktjabrja 2015, Ekaterinburg. Ufa, AJeTERNA, 2015, pp. 43–48 (in Russian).

УДК 537.228.1 + 539.219.1

DOI: 10.17587/nmst.20.415-424

П. А. Борзов, аспирант, **С. Е. Филиппов**, канд. техн. наук, технолог,
В. Ю. Тополов, д-р физ.-мат. наук, проф.,
Южный федеральный университет, Ростов-на-Дону,
О. Е. Брыль, технолог, **А. Е. Панич**, д-р техн. наук, проф., директор,
Институт высоких технологий и пьезотехники,
Южный федеральный университет, Ростов-на-Дону, 344090, Россия
E-mail: vutopolov@sfedu.ru

ЭЛЕКТРОМЕХАНИЧЕСКИЕ СВОЙСТВА И АКУСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ КОМПОЗИТА СЕГНЕТОПЬЕЗОКЕРАМИКА — КОРУНДОВАЯ КЕРАМИКА — ВОЗДУХ

Поступила в редакцию 30.01.2018

Исследованы эффективные электромеханические свойства композита типа 3—0, в состав которого входят сегнетопьезокерамика ЦТС-19, плотная корундовая керамика (объемная концентрация $0,05 < m_c < 0,18$) и воздух. Установлены практически равные объемные концентрации корундовых включений и воздушных пор, что существенно влияет на свойства композита, его коэффициенты электромеханической связи и скорости распространения звука в нем. Эффективные параметры интерпретированы в рамках модели "композит в композите" с элементами связности 3—0.

Ключевые слова: пьезоактивный композит, сегнетопьезокерамика, корундовая керамика, связность 3—0, коэффициент электромеханической связи, скорость звука, объемная концентрация, модель композита, включение

Введение

Композиты на основе сегнетопьезокерамик (СПК) являются многочисленной группой активных диэлектриков, эффективные свойства которых существенно зависят от свойств и объемных концентраций компонентов, микрогоиметрии, технологических и других факторов [1–3]. Среди двухкомпонентных композитов со связностями α — β (где α и β принимают одно из следующих значений: 0, 1, 2 или 3 [4]) 3— β -композиты [5, 6] остаются до настоящего времени менее исследованными по сравнению с 1—3-, 2—2- и 0—3-компо-

зитами [1–3]. Примеры 3— β -композитов "СПК — полимер" показывают, что их пьезоэлектрические свойства, гидростатические и другие параметры [5, 6] изменяются в достаточно широких пределах и представляют интерес для пьезотехнических применений. Улучшение отдельных параметров композитов достигается при внедрении вместо полимеров неорганических пьезопассивных компонентов, таких как глина, песок, цемент, $\alpha\text{-Al}_2\text{O}_3$ (корунд) и др. [7–11]. В частности, в работе [10] приведены экспериментальные данные по электромеханическим свойствам керамического композита, содер-

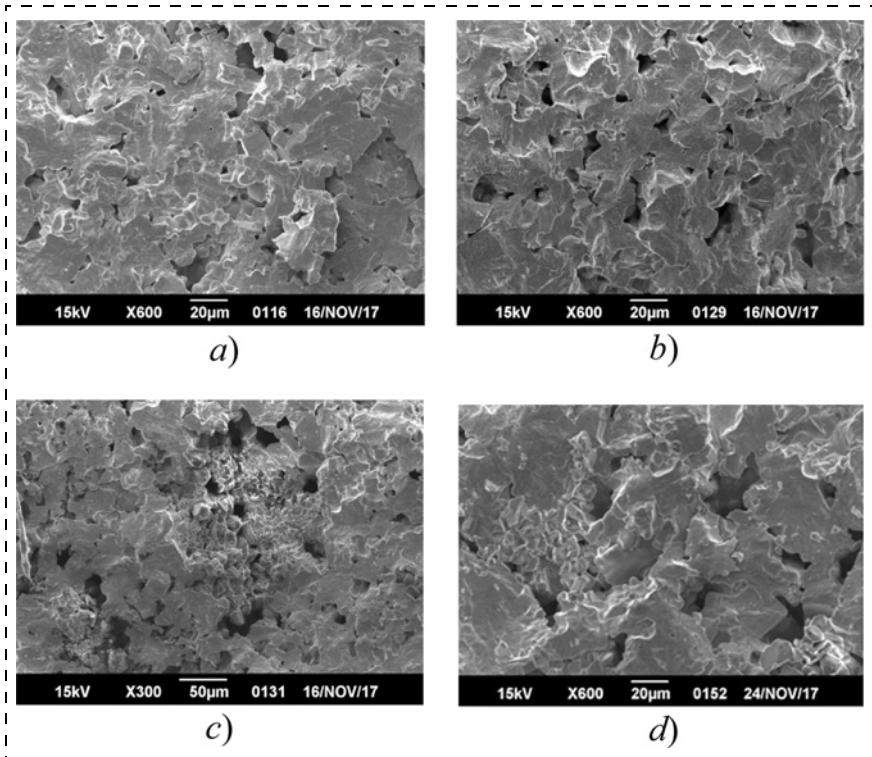


Рис. 1. Микрофотографии сколов неполяризованных образцов композита "керамика ЦТС-19 — корундовая керамика — воздух" при концентрации корундовой керамики $f_{m,c} = 2,5$ мас. % (а); $f_{m,c} = 5,0$ мас. % (б); $f_{m,c} = 7,5$ мас. % (в); $f_{m,c} = 10,0$ мас. % (г). Микроструктуру образцов исследовали с помощью сканирующего электронного микроскопа JSM-6390L

Fig. 1. Micrographs of sections of non-poled ZTS-19 ceramic/corundum ceramic/air composite samples at the volume fraction of corundum ceramic $f_{m,c} = 2.5$ wt. % (a); $f_{m,c} = 5.0$ wt. % (b); $f_{m,c} = 7.5$ wt. % (c); $f_{m,c} = 10.0$ wt. % (d). Microstructure of the samples was studied using the scanning electron microscope JSM -6390 L

жащего СПК типа $\text{Pb}(\text{Zr}, \text{Ti})\text{O}_3$ и $\alpha\text{-Al}_2\text{O}_3$. Последний характеризуется [10, 12] модулями упругости, на порядок превышающими модули упругости СПК, и диэлектрической проницаемостью, которая на два порядка ниже диэлектрической проницаемости СПК, что существенно отличает $\alpha\text{-Al}_2\text{O}_3$ по свойствам от таких компонентов, как полимеры или глина. Однако при достаточно малых объемных концентрациях $\alpha\text{-Al}_2\text{O}_3$ (около 10 % и ниже) детально не исследовались эффективные свойства, коэффициенты электромеханической связи (КЭС) и факторы анизотропии композита, а также не рассматривались особенности его связности. Цель настоящей работы — исследование эффективных свойств, КЭС и скоростей звука в пьезоактивном композите типа 3—0 при объемных концентрациях корундовой керамики менее 20 %.

Экспериментальная часть

Образцы исследуемого композита получены в Институте высоких технологий и пьезотехники Южного федерального университета (НКТБ "Пьезо-прибор"). В качестве пьезоактивного компонента композита была использована сегнетокерамика

ЦТС-19 [3, 13], в качестве пьезопассивного компонента — вакуум-плотная корундовая керамика ВК-94-1 (также известная как 22ХС) с содержанием 94,39 % Al_2O_3 [14, 15]. Прессование и спекание образцов из ВК-94-1 проводили по обычной керамической технологии, а после спекания разрушали образцы под давлением и просеивали через сито для получения спеченной керамической крошки. Керамическую крошку ВК-94-1 тщательно смешивали с порошком ЦТС-19 при фиксированных концентрациях ВК-94-1 $2,5 \text{ мас. \%} \leq f_{m,c} \leq 10,0 \text{ мас. \%}$, в качестве связующего вещества использовали раствор полиvinилового спирта. После этого проводили перемешивание, погружение и просеивание смеси в целях получения гранулированного состава пресс-порошка при $f_{m,c} = \text{const}$. Кроме того, для сравнения параметров композита с параметрами монолитной СПК изготавливали пресс-порошок чистого ЦТС-19.

Обжиг спрессованных композитных образцов проводили в свинецодержащей атмосфере при скорости нагрева 50 К/ч и температуре $T = 1513$ К в течение 2 ч. Процесс охлаждения проходил естественным образом, без фиксированной скорости охлаждения. Образцы, использовавшиеся для дальнейших исследований, имели форму диска диаметром $(10,00 \pm 0,06)$ мм и высотой $(1,00 \pm 0,07)$ мм.

Примеры микрофотографий сколов композитных образцов (до их поляризации) показаны на рис. 1. Микрорельеф полученных образцов усложняется с увеличением концентрации $f_{m,c}$. Наблюдаются сосуществование областей, обогащенных корундом, и воздушных пор, а пористость образцов m_p возрастает с увеличением $f_{m,c}$. Зависимости объемных концентраций корундовых включений m_c и воздушных пор m_p в композитных образцах и их плотности от $f_{m,c}$ графически представлены на рис. 2, а. Отметим "зеркальное" возрастание объемных концентраций m_c и воздушных пор m_p в интервале $2,5 \text{ мас. \%} \leq f_{m,c} \leq 10,0 \text{ мас. \%}$ (рис. 2, а), и подобное поведение объемных концентраций ранее не отмечалось в литературе по композитам на основе СПК. Основываясь на данных рис. 1 и других экспериментальных результатах, мы рассматриваем данный материал как композит типа 3—0.

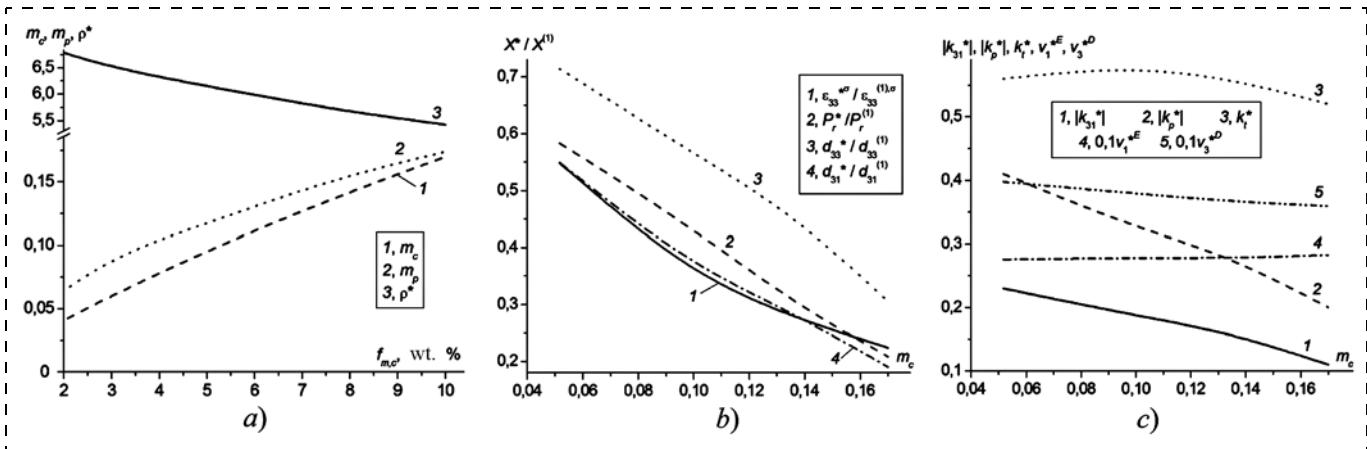


Рис. 2. Экспериментальные данные по параметрам композита "СПК ЦТС-19 — корундовая керамика — воздух" при комнатной температуре: а — объемные концентрации корундовых включений m_c и воздушных пор m_p и плотность образцов ρ^* ($10^3 \text{ кг}/\text{м}^3$) как функции массовой доли корундовой керамики $f_{m,c}$ в образцах; б — нормированные параметры $X^*/X^{(1)}$, характеризующие диэлектрические (кривые 1 и 2) и пьезоэлектрические (кривые 3 и 4) свойства композита при различных m_c ; в — КЭС k_{31}^* , k_p^* , k_t^* и скорости распространения звуковых волн v_1^{*E} , v_3^{*D} (в км/с) в композите при различных m_c

Fig. 2. Experimental data on parameters of the ZTS-19 FPC/corundum ceramic/air composite at room temperature: a — volume fractions of corundum inclusions m_c and air pores m_p , and density of samples ρ^* ($10^3 \text{ kg}/\text{m}^3$) as a function of the mass fraction of corundum ceramic $f_{m,c}$ in the samples; b — normalized parameters $X^*/X^{(1)}$ characterizing dielectric (curves 1 and 2) and piezoelectric (curves 3 and 4) and properties of composite at different m_c values; c — ECFs k_{31}^* , k_p^* , k_t^* and propagation velocity of sound waves v_1^{*E} , v_3^{*D} (km/s) in the composite at different m_c values

Связность композита в исследуемом интервале m_c может быть представлена как 3—0—0, где последний нуль соответствует изолированным воздушным порам, а средний нуль — изолированным корундовым включениям.

Поляризацию композитных образцов проводили в воздухе при напряженности электрического поля $E = 1,25 \text{ кВ}/\text{мм}$, температуре $T = 638 \text{ K}$, а общее время поляризации образца составляло 75 с. Спустя семь дней после поляризации измерения электрофизических параметров поляризованных образцов проводили, используя измерительную аппаратуру "Цензурка-МА" [16] по методу, описанному в ОСТ [17]. Исследуемый композит характеризуется симметрией ∞mm по аналогии с СПК [3, 13, 18], и в дальнейшем остаточная поляризация образца представляется как $P_r^* \parallel OX_3$ в прямоугольной системе координат ($X_1X_2X_3$). По результатам измерений построены экспериментальные кривые эффективных свойств и других параметров композита (рис. 2). При этом интерес представляют концентрационные зависимости КЭС

$$k_{3j}^* = d_{3j}^* / (s_{jj}^{*\sigma} \varepsilon_{33}^{*\sigma})^{1/2}; \quad k_t^* = [(c_{33}^{*D} - c_{33}^{*E}) / c_{33}^{*D}]^{1/2}; \\ k_p^* = k_{31}^* [2s_{11}^{*E} / (s_{11}^{*E} + s_{12}^{*E})]^{1/2} \quad (1)$$

и фазовых скоростей распространения звуковых волн

$$v_1^{*E} = (\rho^* s_{11}^{*E})^{-1/2}, \quad v_3^{*D} = (c_{33}^{*D} / \rho^*)^{1/2} \quad (2)$$

при планарной (v_1^{*E}) и толщинной (v_3^{*D}) модах колебаний. В формулах (1) $j = 1; 3$, d_{3j}^* — пьезомодули, $s_{jj}^{*\sigma}$ — упругие податливости при $E = \text{const}$, $\varepsilon_{33}^{*\sigma}$ — диэлектрическая проницаемость при механическом напряжении $\sigma = \text{const}$, c_{33}^{*D} — модуль упругости при электрическом смещении $D = \text{const}$ и c_{33}^{*E} — модуль упругости при $E = \text{const}$. В ряде источников (например, [10, 18–20]) используются обозначения v_t^{*D} и v_t^* , которые соответствуют скорости звука при толщинной моде колебаний образца.

Наблюдается корреляция между диэлектрическими (кривые 1 и 2 на рис. 2, б) и пьезоэлектрическими (кривые 3 и 4 на рис. 2, б) свойствами. Введение пьезопассивного корунда с большими модулями упругости по сравнению с СПК матрицей и возрастание пористости m_p композита приводят к ослаблению его пьезоэлектрического отклика, однако относительный пьезомодуль $d_{33}^* / d_{33}^{(1)}$ снижается в меньшей степени по сравнению с $d_{31}^* / d_{31}^{(1)}$ (ср. кривые 3 и 4 на рис. 2, б). Это можно объяснить заметным ослаблением электромеханических связей между отдельными областями образца вдоль неполярных осей OX_1 и OX_2 . Последнее способствует усилению пьезоэлектрической анизотропии композита, что проявляется в возрастании отношений $d_{33}^* / |d_{31}^*|$, $k_{33}^* / |k_{31}^*|$ и $k_t^* / |k_p^*|$. При заметном уменьшении диэлектрической проницаемости $\varepsilon_{33}^{*\sigma}$ и остаточной поляризации P_r^* (см. кривые 1 и 2 на рис. 2, б) напряженность коэрцитивного поля E_c^* композита слабо уменьшается:

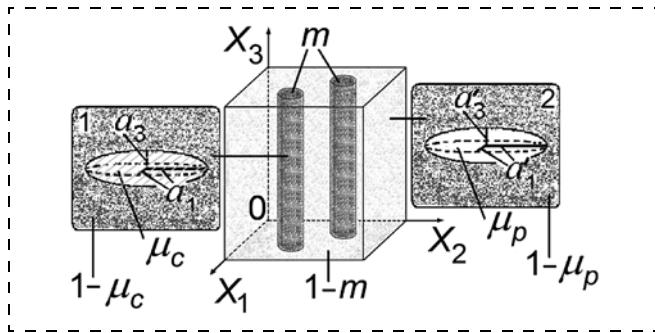


Рис. 3. Схематическое изображение композита типа 3–0. ($X_1X_2X_3$) — прямоугольная система координат, m и $1 - m$ — объемные концентрации стержней и окружающей матрицы соответственно. На вставке 1: μ_c — объемная концентрация корундовых включений в СПК среде, $1 - \mu_c$ — объемная концентрация СПК, a_1 и a_3 — длины полуосей корундового включения. На вставке 2: μ_p — пористость в матрице, $1 - \mu_p$ — объемная концентрация СПК, окружающей поры, a'_1 и a'_3 — длины полуосей поры

Fig. 3. Schematic representation of the 3–0-type composite. ($X_1X_2X_3$) is the rectangular coordinate system, m and $1 - m$ are volume fractions of the rods and surrounding matrix, respectively. In inset 1, μ_c is the volume fraction of corundum inclusions in FPC, $1 - \mu_c$ is the volume fraction of FPC, a_1 and a_3 are semi-axes of the corundum inclusion. In inset 2, μ_p is porosity in the matrix, $1 - \mu_p$ is the volume fraction of FPC that surrounds the pores, a'_1 and a'_3 are lengths of the semi-axes of the pore

например, $E_c^* = 0,860$ кВ/мм при $m_c = 0$ и $E_c^* = 0,815$ кВ/мм при $m_c = 0,17$.

Значительное уменьшение КЭС $|k_{31}^*|$ и $|k_p^*|$ при увеличении m_c (см. кривые 1 и 2 на рис. 2, с) согласуется с соотношениями $|k_{31}^*| \sim |d_{31}^*|$ и $|k_p^*| \sim |d_{31}^*|$, вытекающими из зависимости (1). Вместе с тем КЭС $k_t^* \approx \text{const}$ (кривая 3 на рис. 2, с) благодаря существенной корреляции между зависимостями $c_{33}^{*D}(m_c)$ и $c_{33}^{*E}(m_c)$ в исследуемом композите. Поведение кривых 4 и 5 на рис. 2, с указывает на сравнительно небольшие изменения скоростей звука $v_1^{*E}(m_c)$ и $v_3^{*D}(m_c)$ из формулы (2). Это связано с тем, что изменения плотности образцов $\rho^*(m_c)$ существенно компенсируются изменениями их упругих свойств в продольном ($s_{11}^{*E}(m_c)$) или поперечном ($c_{33}^{*D}(m_c)$) направлении в исследуемом интервале m_c . Примечательно, что удельный акустический импеданс $Z^* = \rho^* v_3^{*D}$ уменьшается монотонно от $35,5 \cdot 10^6$ кг/(м²·с) при $m_c = 0$ до $19,3 \cdot 10^6$ кг/(м²·с) при $m_c = 0,17$, что обусловлено в основном уменьшением плотности $\rho^*(m_c)$ (см. кривую 3 на рис. 2, а).

Модельные представления и интерпретация экспериментальных результатов

Для интерпретации экспериментальных концентрационных зависимостей эффективных параметров (рис. 2, с) предложена модель "композит в композите" (рис. 3). Предполагается, что система корундовых включений регулярно распределена в СПК матрице, и соответствующую область можно

рассматривать как композит типа 3–0 (см. вставку 1 на рис. 3). Стержни этого композита характеризуются связностью 3–0, являются поляризованными вдоль оси OX_3 и регулярно расположены внутри пористой керамической матрицы (см. вставку 2 на рис. 3), связность которой тоже 3–0. Форма корундового включения описывается уравнением сфероида $(x_1/a_1)^2 + (x_2/a_2)^2 + (x_3/a_3)^2 = 1$. При этом $\rho_c = a_1/a_3$ — отношение длин полуосей сфероида, $a_2 = a_1$. Форма воздушной поры удовлетворяет уравнению $(x_1/a'_1)^2 + (x_2/a'_2)^2 + (x_3/a'_3)^2 = 1$, причем $\rho_p = a'_1/a'_3$ — отношение длин полуосей поры, $a'_2 = a'_1$.

Эффективные электромеханические свойства композита определяют в три этапа. На первом этапе проводят определение свойств стержней "корундовая керамика — СПК" $X^{(c)}(\mu_c, \rho_c)$ по методу [3, 21] без учета взаимодействия между пьезопассивными включениями. На втором этапе определяют свойства пористой среды $X^{(m)}(\mu_p, \rho_p)$ с помощью метода [21], применимого к средам с невзаимодействующими порами. При этом среда, окружающая стержни, может быть либо пьезопассивной, либо пьезоактивной (поляризованной вдоль OX_3). На третьем этапе проводят усреднение свойств в рамках метода эффективного поля [3] с учетом электромеханического взаимодействия между пьезоактивными стержнями. В результате эффективное свойство композита представляется в виде $X^*(m, \mu_c, \rho_c, \mu_p, \rho_p)$. В соответствии с рис. 3 объемная концентрация корунда в композите $m_c = m\mu_c$, а объемная концентрация пор в композите $m_p = (1 - m)\mu_p$.

Результаты расчетов эффективных параметров (1), (2) представлены на рис. 4. Скорости звука v_1^{*E} и v_3^{*D} из эксперимента (см. кривые 4 и 5 на рис. 2, с) могут быть аппроксимированы кривыми 1–3 на рис. 4, а, б. Эти кривые рассчитаны для пористой СПК без учета корундовых включений внутри этой среды. При рассмотрении системы "корунд — монолитная СПК" (см. вставку 1 на рис. 3) значения v_1^{*E} и v_3^{*D} (см. кривые 4 и 5 на рис. 4, а, б) выше экспериментальных. Приведенные выше факты указывают на доминирующую роль пористой матрицы в формировании упругих свойств композита и связанных с этими свойствами скоростей звука из формулы (2).

При объемной концентрации корунда $0 < m_c < 0,1$ эффективные параметры из (1) и (2) определяются в рамках предложенной нами модели (см. рис. 3), где пористая матрица содержит сферические включения ($\rho_p = 1$, $m_p = 0,1 \dots 0,2$) и является поляризованной. Присутствие сильно сплющенных корундовых включений с $\rho_c = 100$ в стержнях способствует большой пьезоэлектрической анизотропии. Расчетные кривые для этого случая представлены

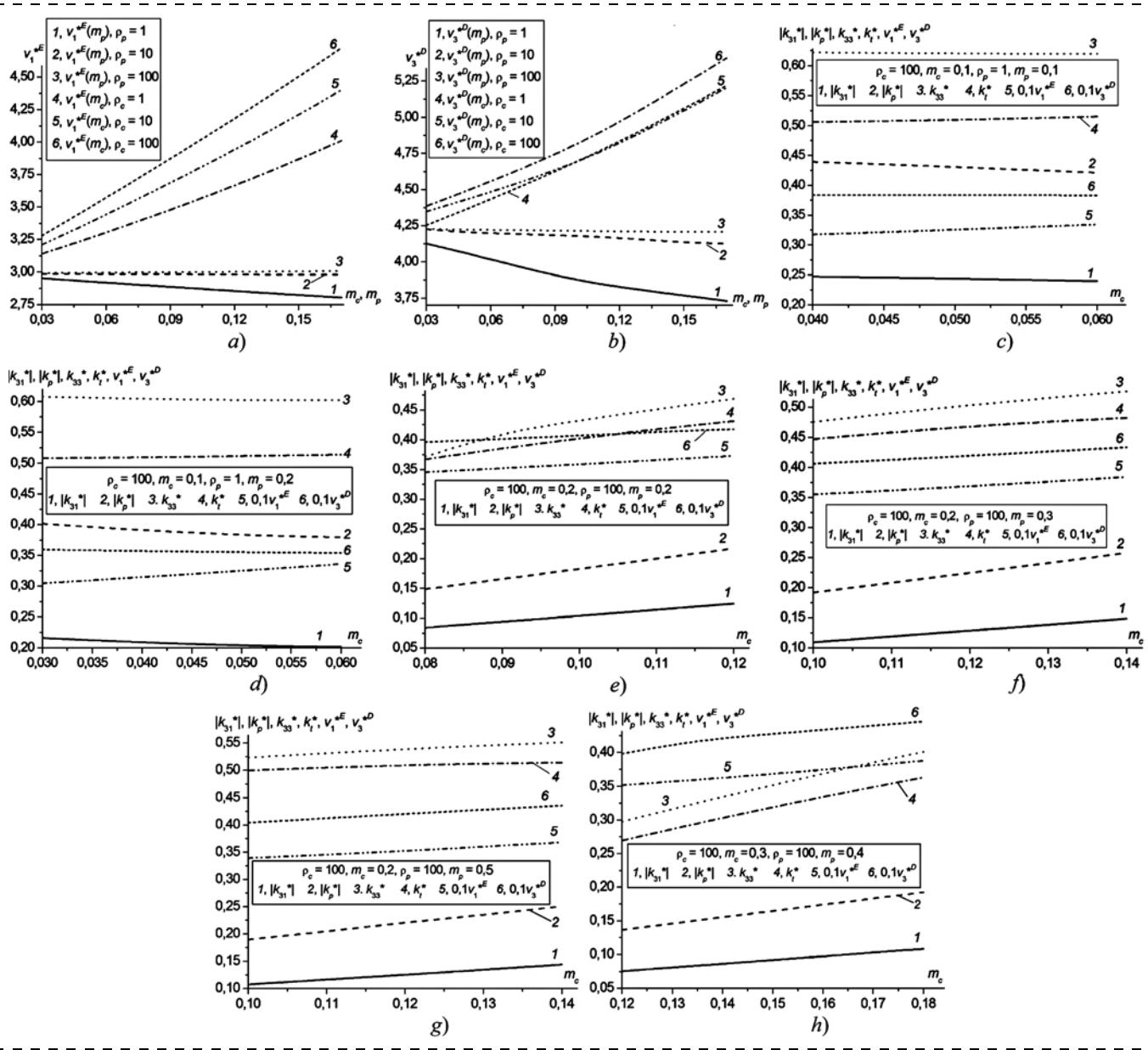


Рис. 4. Расчетные концентрационные зависимости КЭС $|k_{31}^*|$, $|k_p^*|$, k_{33}^* , k_i^* , композита типа 3—0 "СПК ЦТС-19 — корундовая керамика — воздух" и скоростей распространения звуковых волн v_1^{*E} , v_3^{*D} (в км/с) в данном композите. Композит схематически представлен на рис. 3, расчеты проведены с использованием экспериментальных констант компонентов из работ [3, 12, 14, 15]

Fig. 4. Calculated volume-fraction dependence of ECFs $|k_{31}^|$, $|k_p^*|$, k_{33}^* , k_i^* , v_1^{*E} , v_3^{*D} of the 3—0-type ZTS-19 FPC/corundum ceramic/air composite and sound-wave velocities v_1^{*E} , v_3^{*D} (km/s) in this composite. The composite is schematically represented in fig. 3, and calculations are carried out using experimental constants of components from works [3, 12, 14, 15]*

на рис. 4, *c*, *d* и в целом хорошо согласуются с экспериментальными данными из рис. 2, *c*.

При увеличении объемной концентрации m_c большее согласие между расчетом и экспериментом достигается в случае неполяризованной матрицы с сильно сплющенными воздушными порами (например, при $p_p = 100$). На присутствие значительных неполяризованных областей внутри композита указывает кривая 2 на рис. 2, *b*: например, при $m_c = 0,1$ остаточная поляризация компо-

зита P_r^* составляет всего около 40 % от $P_r^{(1)}$ СПК. Расчетные результаты для композита "поляризованные стержни — неполяризованный матрица" представлены на рис. 4, *e*—*h*. Разность между расчетными и экспериментальными значениями параметров связана с упрощениями, заложенными в модели композита (см. рис. 3). Во-первых, пористая матрица в реальном композитном образце частично поляризована, и ее свойства дают вклады в параметры (1) и (2), подобно вкладам от второго

пьезоэлектрического компонента в 1—3-композите [3]. Во-вторых, вместо идеальных цилиндрических стержней "корунд — СПК" в реальных образцах присутствуют некоторые области, протяженные вдоль оси OX_3 , и степень поляризации композита может варьироваться в некотором интервале. В-третьих, отношения длин полуосей включений ρ_c и пор ρ_p могут варьироваться в пределах одного образца, а сами неоднородности могут иметь неправильную форму. В-четвертых, в реальном образце расположение включений, пор и протяженных вдоль OX_3 областей "корунд — СПК" только приближенно можно считать регулярным. Тем не менее данные на рис. 4, e—h целесообразно учитывать при интерпретации анизотропии КЭС, а также при прогнозировании поведения скоростей звука и других характеристик композита.

Проводя сравнение параметров исследуемого композита с известными экспериментальными данными, отметим, что значения d_{3j}^* (см. кривые 3, 4 на рис. 2, b) по модулю больше $|d_{3j}^*|$ композита "СПК ЦТС-19 — фторопласт Ф-2МЭ" [22] при объемных концентрациях фторопласта менее 0,2, однако анизотропия d_{3j}^* несколько меньше установленной в работе [22]. Значения КЭС k_t^* композита (см. кривую 3 на рис. 2, c) близки к экспериментальным значениям $\max k_t^*$ 1—3-композита "СПК типа PZT-7A — аральдит" [1] и k_t^* пористой СПК на основе ПКР-8 [20]. Значения удельного акустического импеданса Z^* композита типа 3—0 близки к Z^* 1—3-композита при объемной концентрации СПК PZT-7A [1] примерно от 0,55 до 0,95. При этом отношение КЭС $k_t^*/|k_p^*|$ композита типа 3—0 больше, чем отношение $k_t^*/|k_p^*|$, измеренное на образцах 1—3-композита [1].

Заключение

Проведено исследование эффективных электромеханических свойств и КЭС композита типа 3—0 "СПК ЦТС-19 — корундовая керамика — воздух", а также фазовых скоростей звуковых волн в этом композите при объемной концентрации корунда $m_c < 0,18$. Установлено "зеркальное" возрастание объемных концентраций m_c и m_p при увеличении массовой доли корундовой керамики $f_{m,c}$ в образцах (см. рис. 2, a), что оказывает влияние на пьезоэлектрический отклик, КЭС и другие параметры композита. Экспериментальные данные (см. рис. 2, c) объяснены в рамках модели "композит в композите" (см. рис. 3). Расчетные кривые КЭС и скоростей звука (рис. 4) показывают, что пористая матрица композита играет важную роль в формировании анизотропных пьезоэлектрических свойств данного композита, а также в значительной степени определяет концентрационные зависимости скоростей звука из (2). Наличие сплю-

щенных корундовых включений и воздушных пор способствует значительной анизотропии КЭС композита даже при малых m_c . При этом пористая матрица, окружающая пьезоэлектрические области "корунд — СПК", может быть частично поляризованной. Исследованный композит типа 3—0 представляет интерес как активный элемент пьезоэлектрических преобразователей и других устройств, в которых ведущую роль играет толщинная мода колебаний.

Авторы благодарны д-ру техн. наук, проф. Нестрову А. А., канд. техн. наук Доле В. К. (ЮФУ, г. Ростов-на-Дону) и проф., д-ру К. Р. Баузуну (Университет Бата, Соединенное Королевство) за интерес к современным пьезоактивным композитам.

Исследования проведены при финансовой поддержке Министерства образования и науки РФ в рамках реализации комплексного проекта "Разработка и создание высокотехнологичного производства мобильного гидроакустического комплекса освещения обстановки в различных акваториях Мирового океана на основе современных пьезоэлектрических средств нового поколения" (Договор № 03.G25.31.0276 от 29.05.2017) с использованием оборудования ЦКП "Высокие технологии" ЮФУ.

Список литературы

1. Chan H. L. W., Unsworth J. Simple model for piezoelectric ceramic/polymer 1—3 composites used in ultrasonic transducer applications // IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. 1989. Vol. 36, N. 4. P. 434—441.
2. Akdogan E. K., Allahverdi M., Safari A. Piezoelectric composites for sensor and actuator applications // IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control. 2005. Vol. 52, N. 5. P. 746—775.
3. Topolov V. Yu., Bowen C. R. Electromechanical properties in composites based on ferroelectrics. London: Springer, 2009. 202 p.
4. Newnham R. E., Skinner D. P., Cross L. E. Connectivity and piezoelectric — pyroelectric composites // Materials Research Bulletin. 1978. Vol. 13. N. 5. P. 525—536.
5. Smay J. E., Ceserano J., Tuttle A. B. et al. Piezoelectric properties of 3—X periodic $Pb(Zr_xTi_{1-x})O_3$ — polymer composites // Journal of Applied Physics. 2002. Vol. 92, N. 10. P. 6119—6127.
6. Smay J. E., Tuttle A. B., Ceserano III J. Robocasting of three-dimensional piezoelectric structures // Piezoelectric and acoustic materials for transducer applications / Eds. A. Safari, E. K. Akdogan. New York: Springer, 2008. P. 305—318.
7. Chaipanich A. Dielectric and piezoelectric properties of PZT—cement composites // Current Applied Physics. 2007. Vol. 7, N. 5. P. 537—539.
8. Filippov S. E., Vorontsov A. A., Topolov V. Yu. et al. Features of the piezoelectric effect in a novel PZT-type ceramic / clay composite. Ferroelectrics Letters Section. 2014. Vol. 41, N. 4—6. P. 82—88.
9. Dongyu X., Xin C., Shifeng H. Investigation of inorganic fillers on properties of 2—2 connectivity cement/polymer based piezoelectric composites // Construction and Building Materials. 2015. Vol. 94. P. 678—683.

10. Rybaynets A. N., Konstantinov G. M., Naumenko A. A. et al. Elastic, dielectric, and piezoelectric properties of ceramic lead zirconate titanate/ α -Al₂O₃ composites // Physics of the Solid State. 2015. Vol. 57, N. 3. P. 527–530.
11. Zhao P., Kim S., Hinderliter B. Investigation of cement–sand-based piezoelectric composites // Journal of Intelligent Material Systems and Structures. 2016. Vol. 27, N. 12. P. 1666–1672.
12. Физические величины: Справочник / Под ред. И. С. Григорьева, Е. З. Мейлихова. М.: Энергоатомиздат, 1991. 1232 с.
13. Гориш А. В., Дудкевич В. П., Куприянов М. Ф. и др. Пьезоэлектрическое приборостроение / Под ред. А. В. Гориша. Т. 1. Физика сегнетоэлектрической керамики. М.: Издат. предпр. ред. жур. "Радиотехника", 1999. 368 с.
14. Корундовая керамика и ее свойства. URL: <http://all-ceramic.ru/materials/korund.html>.
15. Вакуумная керамика. URL: <http://www.vaccer.ru/mater>.
16. Иванов Н. М., Кондаков Е. В., Милюславский Ю. К. Разработка, серийный выпуск и внедрение аппаратуры ряда "Цензурка" для контроля параметров пьезоэлементов, пьезопреобразователей и гидроакустических антенных решеток // Сборник трудов молодежной международной конференции "Актуальные проблемы пьезоэлектрического приборостроения". 23–27 сентября 2013 г. г. Анапа. — Ростов н/Д: Изд-во Юж. фед. ун-та, 2013.— С. 164–167.
17. ОСТ 11 044–87. Отраслевой стандарт. Материалы пьезокерамические. Технические условия. Введ. 1988-01-01. М.: Изд-во стандартов, 1987. 141 с.
18. Berlincourt D. A., Cerran D. R., Jaffe H. Piezoelectric and piezomagnetic materials and their function in transducers // Physical acoustics. Principles and methods / Ed. W. Mason. V. 1. Methods and devices. Pt A. New York: Academic Press, 1964. P. 169–270.
19. Рыбянец А. Н. Физические свойства и применения сегнетоэластичных пористых пьезокерамик на основе ЦТС // Международная научно-практическая конференция "Фундаментальные проблемы пьезоэлектрического приборостроения" — Пьезотехника—2002. 17–21 сентября 2002 г. Сборник докладов. Тверь, 2002. С. 65–72.
20. Ситало Е. И., Смотраков В. Г., Еремкин В. В. и др. Технология изготовления и свойства пористой пьезокерамики // Там же. С. 305–308.
21. Dunn M. L., Taya M. Electromechanical properties of porous piezoelectric ceramics // Journal of the American Ceramic Society. 1993. Vol. 76, N. 7. P. 1697–1706.
22. Борзов П. А., Тополов В. Ю., Воронцов А. А., Брыль О. Е. Пьезоэлектрический отклик и особенности микрогеометрии нового композита на основе сегнетопьезокерамики ЦТС-19 // Нано- и микросистемная техника. 2015. N. 6. С. 22–26.

P. A. Borzov, Postgraduate, **S. E. Filippov**, Ph. D., Technologist, **V. Yu. Topolov**, D. Sc., Professor, **O. E. Brill**, Technologist, **A. E. Panich**, Professor, D. Sc., Director of the Institute, Southern Federal University, Rostov-on-Don, 344006, Russian Federation,

Corresponding author:

Topolov Vitaly Yu., D. Sc., Professor, Southern Federal University, Rostov-on-Don, 344006, Russian Federation, e-mail: vutopolov@sedu.ru

Electromechanical Properties and Acoustic Characteristics of a Composite "Ferroelectric Piezo-Ceramic — Corundum Ceramic — Air"

Received on January 30, 2018
Accepted on February 20, 2018

The effective electromechanical properties have been studied in the composite wherein the ZTS-19 ferroelectric ceramic, dense corundum ceramic and air are components, and the volume fraction of the corundum ceramic m_c is from the range of $0.05 < m_c < 0.18$. Features of microgeometry and connectivity of the composite suggest that in the aforementioned m_c range, this material is of the 3–0 type. Almost equal volume fractions of corundum inclusions and air pores are stated for the composite, and this equality considerably influences the effective properties of the composite, its electromechanical coupling factors and sound velocities therein. The studied effective parameters are interpreted within the framework of the 'composite in composite' model by taking into account elements of 3–0 connectivity. The studied composite is of interest as an active element of piezoelectric transducers, sensors and other devices where the thickness oscillation mode plays the leading role.

Keywords: piezo-active composite, ferroelectric piezo-ceramic, corundum ceramic, 3–0 connectivity, electromechanical coupling factor, sound velocity, volume fraction, composite model, inclusion

For citation:

Borzov P. A., Filippov S. E., Topolov V. Yu., Brill O. E., Panich A. E. Electromechanical Properties and Acoustic Characteristics of a Composite "Ferroelectric Piezo-Ceramic — Corundum Ceramic — Air", *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 415–424.

DOI: 10.17587/nmst.20.415-424

Introduction

Composites based on ferroelectric piezoceramics (FPCs) are the numerous group of active dielectrics, the effective properties of which depend essentially on the properties and volume fractions of the components, microgeometry, and

other technological factors [1–3]. Among the two-component composites with α – β connectivity patterns (where α and β take one of the following values: 0, 1, 2, or 3 [4]), 3– β composites [5, 6] remain less studied than the 1–3, 2–2, and 0–3 composites [1–3]. Examples of the 3– β FPC/polymer

composites show that their piezoelectric properties, hydrostatic and other parameters [5, 6] vary within fairly wide ranges and are of interest for piezotechnical applications. Improvement of the individual parameters of composites is achieved by introducing of piezopassive inorganic components polymers such as clay, sand, concrete, α -Al₂O₃ (corundum), etc. [7–11]. In particular, the experimental data for electromechanical properties of the composite containing the Pb(Zr, Ti)O₃-type FPC and α -Al₂O₃ ceramic are shown in work [10]. The latter is characterized [10, 12] by elastic moduli that exceed by an order-of-magnitude the elastic moduli of FPC, and the dielectric permittivity is two orders-of-magnitude lower than the dielectric permittivity of the FPC, which essentially distinguishes α -Al₂O₃ in properties from other components such as polymers or clay. However, at sufficiently small volume fractions of α -Al₂O₃ (about 10 % or less), the effective properties and electromechanical coupling factor (ECF) and the anisotropy factors of the composite were not studied in details, and also features of its connectivity were not considered. The aim of this work is to study the effective properties, ECF, and sound velocities in a piezoactive composite of 3–0 type at volume fractions of corundum ceramics of less than 20 %.

Experimental part

The samples of the composite being studied were prepared at the Institute of High Technology and Piezo-technics of the South Federal University (SCTB "Piezopribor"). The ferroelectric ceramics ZTS-19 [3, 13] was used as a piezoactive component of the composite, and the vacuum-tight alumina ceramic VK-94-1 (also known as 22XC) with content 94.39 % of Al₂O₃ [14, 15] was used as a piezopassive component. Pressing and sintering of the VK-94-1 samples was carried out by conventional ceramic technology and followed by a destruction of these samples under the pressure after sintering and sifting through a sieve to obtain sintered ceramic crumbs. The ceramic VK-94-1 crumb was thoroughly mixed with the ZTS-19 powder at fixed VK-94-1 mass fractions (2.5 wt. % $\leq f_{m,c} \leq$ 10.0 wt. %); a solution of polyvinyl alcohol was used as a binder. Mixing, grinding and sieving of the mixture was carried out after this stage to produce a granular press-powder composition at $f_{m,c} = \text{const}$. In addition, to compare the parameters of the composite with the parameters of a monolithic FPC, a press-powder of pure ZTS-19 was manufactured.

The burning of the pressed composite samples was carried out in a lead-containing atmosphere at a heating rate of 50 K/h and temperature $T = 1513$ K for 2 hours. The cooling process went naturally without a fixed cooling rate. The samples used for further studies had the shape of a disk with a diameter (10.00 \pm 0.06) mm and height (1.00 \pm 0.07) mm.

The examples of photo-micrographs of cleaved composite samples (before their poling) are shown in fig. 1. Microgeometry of the obtained samples is complicated with increasing of the fraction $f_{m,c}$. The coexistence of regions enriched in corundum and air pores is observed, and the porosity of the samples m_p increases with increasing $f_{m,c}$. Dependences of the volume fractions of corundum inclusions m_c and air pores m_p in the composite samples and their densities on $f_{m,c}$ are graphically presented in fig. 2, a. Let's note the "mirror" increase in the volume fractions of corundum m_c

and air pores m_p in the range of 2.5 wt. % $\leq f_{m,c} \leq$ 10.0 wt. % (fig. 2, a), and the similar behavior of the volume fractions was not previously noted in the literature on composites based on FPC. Based on the data in fig. 1 and the other experimental results, we consider this material as a 3–0-type composite. Connectivity of the composite in the studied m_c range can be represented as 3–0–0, where the last zero corresponds to isolated air pores, and the central zero is related to isolated corundum inclusions.

Poling of the composite samples was carried out in air at an electric field $E = 1.25$ kV/mm, $T = 638$ K, and the total poling time of the sample was 75 s. Seven days after the poling, the measurements of electrophysical parameters of the samples were performed using measuring instruments "Tsenzurka-MA" [16] by a method described in the industrial standard [17]. The investigated composite is characterized by ∞mm symmetry by analogy with FPC [3, 13, 18], and hereafter, the remanent polarization of the sample is represented as $P_r^* \| OX_3$ in a rectangular coordinate system ($X_1X_2X_3$). Based on the measurement results, the experimental curves of the effective properties and other parameters of the composite are built (fig. 2). At the same time, of interest are the volume-fraction dependences of the ECFs

$$k_{3j}^* = d_{3j}^*/(s_{jj}^{*E} \varepsilon_{33}^{*\sigma})^{1/2}, \quad k_t^* = [(c_{33}^{*D} - c_{33}^{*E})/c_{33}^{*D}]^{1/2}; \\ k_p^* = k_{31}^* [2s_{11}^{*E}/(s_{11}^{*E} + s_{12}^{*E})]^{1/2} \quad (1)$$

and phase velocities of sound waves

$$v_1^{*E} = (\rho^* s_{11}^{*E})^{-1/2}, \quad v_3^{*D} = (c_{33}^{*D}/\rho^*)^{1/2} \quad (2)$$

for planar (v_1^{*E}) and thick (v_3^{*D}) oscillation modes. In the formulas (1) $j = 1; 3$, d_{3j}^* — are piezoelectric moduli, s_{jj}^{*E} are elastic compliances at $E = \text{const}$, $\varepsilon_{33}^{*\sigma}$ is the dielectric permittivity at mechanical stress $\sigma = \text{const}$, c_{33}^{*D} is the elastic modulus at electric displacement $D = \text{const}$, and c_{33}^{*E} is the elastic modulus at $E = \text{const}$. A number of sources (see for example, works [10, 18–20]) use the v_t^{*D} and v_t^* notations that correspond to the sound velocity in a case of a thickness oscillation mode of the sample.

There is a correlation between the dielectric (curves 1 and 2 in fig. 2, b) and piezoelectric (curves 3 and 4 in fig. 2, b) properties. The introduction of piezo-passive corundum with large elastic moduli in comparison with the SPC matrix and an increase in porosity m_p of the composite result in a weakening of its piezoelectric response, but the relative piezoelectric modulus $d_{33}^*/d_{33}^{(1)}$ decreases to a lesser extent in comparison with $d_{31}^*/d_{31}^{(1)}$ (compare curves 3 and 4 in fig. 2, b). This can be explained by the noticeable weakening of the electromechanical coupling between the individual regions of the sample along the non-polar OX_1 and OX_2 axes. The latter contributes to the enhancement of the piezoelectric anisotropy of the composite, which is manifested in the increase in the $|d_{33}^*/d_{31}^*|$, $|k_{33}^*/|k_{31}^*|$ and $|k_t^*/|k_p^*|$ ratios. At the marked decrease in the dielectric constant $\varepsilon_{33}^{*\sigma}$ and remanent polarization P_r^* (see curves 1 and 2 in fig. 2, b), the intensity of the coercive field E_c^* of the composite is weakly reduced: for example, $E_c^* = 0.860$ kV/mm at $m_c = 0$, and $E_c^* = 0.815$ kV/mm for $m_c = 0.17$.

A significant decrease in the ECFs $|k_{31}^*|$ and $|k_p^*|$ with an increase in m_c (curves 1 and 2 in fig. 2, c) agrees with the re-

lations $|k_{31}^*| \sim |d_{31}^*|$ and $|k_p^*| \sim |d_{31}^*|$, which follow from the formulas (1). At the same time, the ECF $k_t^* \approx \text{const}$ (curve 3 in fig. 2, c) is due to the significant correlation between the $c_{33}^{*D}(m_c)$ and $c_{33}^{*E}(m_c)$ dependences in the studied composite. The behavior of curves 4 and 5 in fig. 2, c indicates a relatively small change in sound velocities $v_1^{*E}(m_c)$ and $v_3^{*D}(m_c)$ from the formula (2). This is due to the fact that changes in the density of the samples $\rho^*(m_c)$ are substantially compensated by changes in their elastic properties in the longitudinal ($s_{11}^{*E}(m_c)$) or transverse ($c_{33}^{*D}(m_c)$) direction and in the studied m_c range. It is noteworthy that the specific acoustic impedance $Z^* = \rho^* v_3^{*D}$ decreases monotonically from $35.5 \cdot 10^6 \text{ kg}/(\text{m}^2 \cdot \text{s})$ at $m_c = 0$ to $19.3 \cdot 10^6 \text{ kg}/(\text{m}^2 \cdot \text{s})$ at $m_c = 0.17$, that is mainly due to the decreasing density $\rho^*(m_c)$ (curve 3 in fig. 2, a).

Model representations and interpretation of experimental results

To interpret the experimental volume-fraction dependences of the effective parameters (fig. 2, c), a "composite in composite" model is proposed (fig. 3). It is assumed that the system of the corundum inclusions are regularly distributed in the FPC matrix, and the corresponding region may be regarded as a 3–0-type composite (see inset 1 in fig. 3). The rods of this composite are characterized by 3–0 connectivity, poled along the OX_3 axis and regularly distributed in the porous ceramic matrix (see inset 2 in fig. 3), the connectivity of which is also 3–0. The shape of the corundum inclusion is described by the spheroid equation $(x_1/a_1)^2 + (x_2/a_2)^2 + (x_3/a_3)^2 = 1$. Moreover, $\rho_c = a_1/a_3$ is the ratio of the lengths of the semi-axes of the spheroid, and $a_2 = a_1$. The shape of the air pore obeys the equation $(x_1/a'_1)^2 + (x_2/a'_2)^2 + (x_3/a'_3)^2 = 1$, $\rho_p = a'_1/a'_3$ is the ratio of the lengths of the pore semi-axes, and $a'_2 = a'_1$.

The effective electromechanical properties of the composite are determined in three steps. In the first stage the properties of the corundum ceramic/FPC $X^{(c)}(\mu_c, \rho_c)$ are determined by means of the method [3, 21] without taking into account the interaction between piezo-passive inclusions. In the second stage the properties of the porous media $X^{(m)}(\mu_p, \rho_p)$ are determined by using the method [21] being applicable to media with non-interacting pores. The medium surrounding the rods may be either piezopassive or piezoactive (poled along OX_3). In the third stage averaging the properties within the effective field method [3] is carried out by taking into account the electromechanical interaction between the piezoactive rods. As a result, the effective property of the composite is represented in the form $X^*(m, \mu_c, \rho_c, \mu_p, \rho_p)$. In accordance with fig. 3, the volume fraction of corundum in the composite is $m_c = m\mu_c$, and the volume concentration pores in the composite is $m_p = (1 - m)\mu_p$.

The results of calculations of the effective parameters (1), (2) are shown in fig. 4. Sound velocities v_1^{*E} and v_3^{*D} from the experiment (curves 4 and 5 in fig. 2, c) can be approximated by curves 1–3 shown in fig. 4, a, b. These curves are calculated for a porous FPC without allowance for the corundum inclusions within this medium. When considering the corundum/monolithic FPC system (see inset 1 in fig. 3), the v_1^{*E} and v_3^{*D} values (curves 4 and 5 in fig. 4, a, b) are higher than the experimental ones. The aforementioned facts indicate the dominant role of the porous matrix in the

formation of the elastic properties and sound velocities linked with the elastic properties of the composite in accordance with formulas (2).

At a volume fraction of corundum $0 < m_c < 0.1$ the effective parameters from (1) and (2) are determined within the framework of the model proposed by us (see fig. 3), where the porous matrix contains spherical inclusions ($\rho_p = 1$, $m_p = 0.1 \dots 0.2$) and is poled. The presence of highly oblate corundum inclusions with $\rho_c = 100$ in the rods promotes a large piezoelectric anisotropy. The calculated curves for this case are shown in fig. 4, c, d and on the whole are in good agreement with the experimental data from fig. 2, c.

With an increase in the volume fraction m_c a better agreement between calculation and experiment results is achieved in the case of the non-poled matrix with the highly oblate air inclusions (e.g., at $\rho_p = 100$). The presence of the significant non-poled regions inside the composite is indicated by curve 2 in fig. 2, b: for example, for $m_c = 0.1$ the remanent polarization of the composite P_r^* is only about 40 % of $P_r^{(1)}$ related to FPC. Calculated results for the poled rods/non-poled matrix composite are shown in fig. 4, e–h. The difference between the calculated and experimental values of the parameters is related to simplifications embedded in the composite model (see fig. 3). Firstly, the porous matrix in a real composite sample is partially poled, and its properties contribute to the parameters (1) and (2), like the contribution from the second component in 1–3 piezoelectric composites [3]. Second, instead of ideal cylindrical corundum/FPC rods in real samples, there are some regions extending along the OX_3 axis, and the poling degree of the composite can vary within a certain range. Third, the ratio of the lengths of the semi-axes of the inclusions ρ_c and pores ρ_p can vary within one sample, and the inhomogeneities themselves may have an irregular shape. Fourthly, in a real sample, the location of inclusions, pores, and extended corundum/FPC regions along the OX_3 axis only approximately can be considered as a regular location. Nevertheless, the data in fig. 4, e–h are advisable to take into account for an interpretation of the anisotropy of ECFs, as well as in predicting the behavior of sound velocities and other characteristics of the composite.

Comparing the studied composite parameters to known experimental data, we note that the absolute d_{3j}^* values (curves 3, 4 in fig. 2, b) are larger than $|d_{3j}^*|$ of the ZTS-19 FPC / fluoroplast F-2ME composite [22] at the volume fraction of fluoroplast less than 0.2, but the anisotropy of d_{3j}^* is somewhat less than that stated in work [22]. The ECF k_t^* values of the composite (see curve 3 in fig. 2, c) are close to the experimental values of $\max k_t^*$ of the 1–3 PZT-7A FPC/araldite composite [1] and k_t^* of porous FPC based on PCR-8 [20]. The values of specific acoustic impedance Z^* of the 3–0-type composite are close to Z^* of the 1–3 composites at a volume concentration of PZT-7A FPC [1] from the range of ca. 0.55...0.95. At the same time, the ECF ratio $k_t^*/|k_p^*|$ of the 3–0-type composite is larger than the $k_t^*/|k_p^*|$ ratio measured on the 1–3 composite samples [1].

Conclusion

The study of the effective electromechanical properties and ECFs of the 3–0-type ZTS-19 FPC corundum ceramic/air composite and phase velocities of sound waves in this composite is carried out at volume fractions of corundum

$m_c < 0.18$. A "mirror" increase in the volume fractions m_c and m_p was observed with an increase in the mass fraction of corundum ceramics $f_{m,c}$ in the samples (fig. 2, a), which affects the piezoelectric response, ECFs and other parameters of the composite. The experimental data (fig. 2, c) are interpreted within the framework of the "composite in composite" model (fig. 3). The calculated curves of the ECFs and sound velocities (fig. 4) show that the porous matrix of the composite plays an important role in the formation of the anisotropic piezoelectric properties of this composite, and also largely determines the volume-fraction dependence of the sound velocities from formulas (2). The oblate corundum inclusions and air pores promote a significant ECF anisotropy even at low m_c values. In this case, the porous matrix surrounding the piezoelectric corundum/FPC regions can be partially poled. The studied 3–0-type composite is of interest as an active element of piezoelectric transducers and other devices in which the thickness oscillation mode plays a leading role.

The authors are grateful to Dr. Tech. Sc., Prof. A. A. Nesterov, Cand. Tech. Sc. V. K. Dolya (SFedU, Rostov-on-Don) and Prof., Dr. C. R. Bowen (University of Bath, United Kingdom) for their interest in modern piezoactive composites.

The studies were carried out with the financial support of the Ministry of Education and Science of the Russian Federation within the framework of the comprehensive project "Development and creation of high-tech production of a mobile sonar complex for lighting the situation in various areas of the World Ocean based on modern piezoelectric devices of a new generation" (Contract No. 03.G25.31.0276 dated 29 May 2017) using the equipment of the Center for High Technologies, SFedU.

References

1. Chan H. L. W., Unsworth J. Simple model for piezoelectric ceramic/polymer 1–3 composites used in ultrasonic transducer applications, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 1989, vol. 36, no. 4, pp. 434–441.
2. Akdogan E. K., Allahverdi M., Safari A. Piezoelectric composites for sensor and actuator applications, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2005, vol. 52, no. 5, pp. 746–775.
3. Topolov V. Yu., Bowen C. R. *Electromechanical properties in composites based on ferroelectrics*, London, Springer, 2009, 202 p.
4. Newnham R. E., Skinner D. P., Cross L. E. Connectivity and piezoelectric – pyroelectric composites, *Materials Research Bulletin*, 1978, vol. 13, no. 5, pp. 525–536.
5. Smay J. E., Cesarano J., Tuttle A. B. et al. Piezoelectric properties of 3–X periodic $\text{Pb}(\text{Zr}_x\text{Ti}_{1-x})\text{O}_3$ – polymer composites, *Journal of Applied Physics*, 2002, vol. 92, no. 10, pp. 6119–6127.
6. Smay J. E., Tuttle A. B., Cesarano J. Robocasting of three-dimensional piezoelectric structures, *Piezoelectric and acoustic materials for transducer applications* / Eds. A. Safari, E. K. Akdogan, New York, Springer, 2008, pp. 305–318.
7. Chaipanich A. Dielectric and piezoelectric properties of PZT–cement composites, *Current Applied Physics*, 2007, vol. 7, no. 5, pp. 537–539.
8. Filippov S. E., Vorontsov A. A., Topolov V. Yu. et al. Features of the piezoelectric effect in a novel PZT-type ceramic/clay composite, *Ferroelectrics Letters Section*, 2014, vol. 41, no. 4–6, pp. 82–88.
9. Dongyu X., Xin C., Shifeng H. Investigation of inorganic fillers on properties of 2–2 connectivity cement/polymer based piezoelectric composites, *Construction and Building Materials*, 2015, vol. 94, pp. 678–683.
10. Rybaynets A. N., Konstantinov G. M., Naumenko A. A. et al. Elastic, dielectric, and piezoelectric properties of ceramic lead zirconate titanate/ α - Al_2O_3 composites, *Physics of the Solid State*, 2015, vol. 57, no. 3, pp. 527–530.
11. Zhao P., Kim S., Hinderliter B. Investigation of cement–sand-based piezoelectric composites, *Journal of Intelligent Material Systems and Structures*, 2016, vol. 27, no. 12, pp. 1666–1672.
12. Physical values: a handbook / Eds. I. S. Grigoryev, E. Z. Meylikhov, Moscow, Energoatomizdat, 1991, 1232 p. (in Russian).
13. Gorish A. V., Dudkevich V. P., Kupriyanov M. F. et al. *P'ezohlektricheskoe priborostroenie* (Piezoelectric instrument-making) / Ed. A. V. Gorish. Vol. 1. Fizika segnetoehlektricheskoy keramiki., Moscow, Radiotekhnika Publ. Co., 1999. 368 p. (in Russian).
14. Corundum ceramics and properties, available at: <http://allceramic.ru/materials/korund.html> (in Russian).
15. Vacuum ceramics, available at: <http://www.vacker.ru/mater> (in Russian).
16. Ivanov N. M., Kondakov E. V., Miloslavsky Yu. K. Razrabotka, serijnyj vypusk i vnedrenie apparatury ryada "Tsenzurka" dlya kontrolya parametrov p'ezohlementov, p'ezopreobrazovatelej i gidroakusticheskikh antennnyj reshetok (Working out, serial output and introduction of the equipment of the "Tsenzurka" type for controlling the parameters of piezoelectric elements, transducers and hydroacoustic arrays), *Sbornik trudov molodezhnoj mezhdunarodnoj konferencii "Aktual'nye problemy p'ezohlektricheskogo priborostroeniya"*, 23–27 September 2013, Anapa, Rostov-on-Don, South. Fed. Univ. Publ., 2013, pp. 164–167 (in Russian).
17. OST 11 044–87. Branch standard. Piezo-ceramic materials. Technical conditions. Introd. 01/01/1988. Moscow, Standards Press, 1987, 141 p. (in Russian).
18. Berlinecourt D. A., Cerran D. R., Jaffe H. Piezoelectric and piezomagnetic materials and their function in transducers, *Physical acoustics. Principles and methods* / Ed. W. Mason. V. 1. Methods and devices. Pt A. New York, Academic Press, 1964, pp. 169–270.
19. Rybyanets A. N. Fizicheskie svojstva i primeneniya segnetozhestkih poristyh p'ezokeramik na osnove CTS (Physical properties and applications of ferroelectric hard porous piezo-ceramics based on ZTS), *Mezhdunarodnaya nauchno-prakticheskaya konferenciya "Fundamental'nye problemy p'ezohlektricheskogo priborostroeniya" — P'ezotekhnika—2002*. 17–21 September 2002, Collected reports, Tver, 2002, pp. 65–72 (in Russian).
20. Sitalo E. I., Smotrakov V. G., Eremkin V. V. et al. Tekhnologiya izgotovleniya i svojstva poristoj p'ezokeramiki (Technology of manufacturing and properties of porous piezo-ceramics), *Mezhdunarodnaya nauchno-prakticheskaya konferenciya "Fundamental'nye problemy p'ezohlektricheskogo priborostroeniya" — P'ezotekhnika—2002*, 17–21 September 2002, Collected reports, Tver, 2002. P. 305–308 (in Russian).
21. Dunn M. L., Taya M. Electromechanical properties of porous piezoelectric ceramics, *Journal of the American Ceramic Society*, 1993, vol. 76, no. 7, pp. 1697–1706.
22. Borzov P. A., Topolov V. Yu., Vorontsov A. A., Brill O. E. P'ezohlektricheskij otklik i osobennosti mikrogeometrii novogo kompozita na osnove segnetop'ezokeramiki CTS-19 (Piezoelectric response and features of microgeometry of a novel composite based on ferroelectric ZTS-19 ceramic), *Nano- i Mikrosistemnaya Tekhnika*, 2015, no. 6, pp. 22–26 (in Russian).

ЭЛЕМЕНТЫ МНСТ

MICRO- AND NANOSYSTEM TECHNIQUE ELEMENTS

УДК 621.382

DOI: 10.17587/nmst.20.425-433

В. К. Ильков, канд. техн. наук, зав. каф.,
Федеральное государственное бюджетное образовательное учреждение высшего образования
"Московский технологический университет",

А. О. Михалев, аспирант, e-mail: m2lkeny@yandex.ru, **М. В. Майтама**, аспирант,
Федеральное государственное бюджетное учреждение науки Институт сверхвысокочастотной
полупроводниковой электроники Российской академии наук, Москва

КОММУТАТОРЫ НА АРСЕНИДЕ И НИТРИДЕ ГАЛЛИЯ

Поступила в редакцию 23.01.2018

Рассмотрены основные типы МЭМС-коммутаторов. Разработан и изготовлен МЭМС-переключатель для частотного диапазона 10...12 ГГц на гетероструктуре AlGaAs/InGaAs/GaAs. Приведены результаты измерений параметров МЭМС-переключателя на арсениде галлия. Спроектирован и изготовлен переключатель на основе монолитной интегральной схемы (МИС) СВЧ сигнала 1×2 (1 вход – 2 выхода) диапазона частот 0,1...15 ГГц на базе нитридных наногетероструктур. Проведен анализ современного состояния разработок в области применения широкозонных наногетероструктур AlGaN/AlN/GaN в СВЧ диапазоне.

Ключевые слова: X-диапазон, монолитная интегральная схема, МЭМС, коммутатор, арсенид галлия, нитрид галлия

Введение

Необходимость в снижении размеров и массы сверхвысокочастотных (СВЧ) устройств, расширении их динамического диапазона частот, уменьшении потребляемой мощности и стоимости, возможности использования беспроводных систем для увеличения мобильности и доступности для источников потребления информации [1], в наращивании их интеграции и функциональных возможностей при росте рабочих частот требует от разработчиков максимального использования опыта и технологических приемов создания интегральных схем (ИС) на подложках Si, GaAs и GaN.

Одной из проблем, возникающей при разработке СВЧ устройств, является создание переключателей сигналов.

В настоящее время наиболее распространены СВЧ переключатели двух типов:

- электромеханические;
- электронные (на дискретных приборах).

Для переключения СВЧ сигналов наиболее перспективными являются переключатели на основе микроэлектромеханических систем (МЭМС) [2], поскольку они позволяют:

- получать переключатели СВЧ сигналов с нулевым энергопотреблением в состоянии покоя;
- снизить размеры и массу СВЧ устройств;
- снизить энергопотребление СВЧ устройств (мощность, необходимая на переключение коммутатора, порядка 1 нВт);

— значительно уменьшить время переключения по сравнению с механическими СВЧ переключателями;

— получить линейные характеристики переключаемого сигнала.

Неоспоримые достоинства МЭМС-компонентов определяют оправданность и необходимость их применения в современной высоконадежной аппаратуре, в том числе космической и специального назначения [3].

СВЧ коммутаторы

В настоящее время широко применяют СВЧ переключатели трех типов: электромеханические, твердотельные (полупроводниковые) и МЭМС-переключатели. Электромеханические переключатели превосходят все остальные по коммутируемой мощности — до нескольких киловатт на частоте 1 ГГц, имеют вносимые ключом потери (собственные потери) 0,1...0,3 дБ и изоляцию сигнала 60...80 дБ. Но при этом электромеханические коммутаторы имеют высокую потребляемую мощность — до 10 Вт (для мощных переключателей), время переключения 10...20 мс и относительно низкий срок службы — порядка 1 млн циклов. В табл. 1 приведены сравнительные характеристики полупроводниковых коммутаторов на основе PIN-диодов, полевых транзисторов и МЭМС-переключателей СВЧ сигнала.

Из табл. 1 видно, что МЭМС-переключатели имеют преимущества перед твердотельными СВЧ

Таблица 1
Table 1

**Характеристики переключателей на основе МЭМС,
PIN-диодах, полевых транзисторах**
*Characteristics of switches on the basis of MEMS,
PIN-diodes and field transistors*

Параметры <i>Parameters</i>	СВЧ МЭМС <i>Microwave MEMS</i>	PIN-диод <i>PIN diode</i>	Полевой транзистор <i>Field transistor</i>
Напряжение, В <i>Voltagee, V</i>	20...80	±3...5	3...5
Ток, мА <i>Current, mA</i>	0	3...20	0
Энергопотребление, мВт <i>Power consumption, mW</i>	0,05...0,1	3...20	0,05...0,1
Время переключения <i>Switching time</i>	1...300 μ s	1...100 ns	1...100 ns
Границчная частота, ТГц <i>Limiting frequency, THz</i>	20...80	1...4	0,5...2
Изоляция (1...10 ГГц) <i>Isolation (1...10 GHz)</i>	Очень высокая <i>Very high</i>	Высокая <i>High</i>	Средняя <i>Average</i>
Изоляция (10...40 ГГц) <i>Isolation (10...40 GHz)</i>	Очень высокая <i>Very high</i>	Средняя <i>Average</i>	Низкая <i>Low</i>
Изоляция (40...100 ГГц) <i>Isolation (40...100 GHz)</i>	Высокая <i>High</i>	Средняя <i>Average</i>	Нет <i>No</i>
Потери (1...100 ГГц), дБ <i>Losses (1...100 GHz), dB</i>	0,05...0,2	0,3...1,2	0,4...2,5
Переключаемая мощность, Вт <i>Switching power, W</i>	> 10	<10	<10

переключателями: высокое соотношение потерь к изоляции в разомкнутом и замкнутом состояниях, практически нулевое потребление мощности в замкнутом состоянии. Но при этом имеют и недостатки: низкое быстродействие по сравнению с полупроводниковыми переключателями. Для управления необходимо формировать импульс переключающего напряжения от 6 до 80 В. Следует отметить, что срок службы МЭМС-переключателей составляет порядка 10^{10} циклов. Следовательно, можно сказать, что МЭМС-переключатели объединили в себе лучшие черты твердотельных и электромеханических переключателей — малые га-

баритные размеры, высокое быстродействие при низком собственном энергопотреблении, большой срок службы (присуще твердотельным переключателям), при этом появляется возможность коммутации СВЧ сигнала высокой мощности (до 10 Вт) (характерно для электромеханических переключателей). Следует отметить, что МЭМС-переключатели требуют вакуумного корпусирования, что, в свою очередь, позволяет увеличить быстродействие и значительно продлить срок их службы.

Типы МЭМС-переключателей

При изготовлении переключателей средней и малой мощности (до 100 мВт) переключатель с емкостным типом контакта имеет важное преимущество — возможность изготовления переключателя с низким напряжением срабатывания, что позволяет использовать единую цепь питания как для полупроводниковых приборов, так и для управления переключателями.

В табл. 2 приведены различные по типу действия МЭМС-устройства.

Как видно из табл. 2, электростатические реле наиболее пригодны для использования в МЭМС-переключателях [4] ввиду того, что:

- потребляемая мощность близка к нулю;
- МЭМС-переключатель имеет малый размер;
- при изготовлении используется хорошо разработанная планарная технология;
- наиболее быстрое время переключения;
- для срабатывания реле достаточно 50...200 мкН;
- при использовании LC-развязки возможна подача управляющего напряжения через ВЧ волновод.

Преимущества электростатического привода МЭМС-ключа состоят в крайне малой потребляемой мощности (ток потребляется только при переключении, типовое значение потребляемой мощности на один цикл переключения составляет не более 1 нВт) [5, 6]. Единственным серьезным недостатком является необходимость использовать управляющее напряжение 20...80 В, в то время как большинство цепей управляются с помощью напряжения 3...5 В.

Таблица 2
Table 2

Сравнение МЭМС-устройств с различным механизмом работы
Comparison of MEMS devices with various mechanisms of operation

МЭМС <i>MEMS</i>	$V_{\text{упр}}, \text{ В}$ $V_{\text{upr}}, \text{ V}$	$I_{\text{упр}}, \text{ мА}$ $I_{\text{upr}}, \text{ mA}$	$P, \text{ мВт}$ $P, \text{ mW}$	Размер, мкм <i>Size, μm</i>	Время переключения, мкс <i>Switching time, μs</i>	Контактная сила, мкН <i>Contact power, μN</i>
Электростатические <i>Electrostatic</i>	20...80	0	0,002	100 × 440	1...200	50...1000
Термомеханические <i>Thermomechanical</i>	3...5	5...100	15...250	200 × 220	300...10 000	500...4000
Электромагнитные <i>Electromagnetic</i>	3...5	20...150	60...100	1000 × 500	300...1000	50...200
Пьезоэлектрические <i>Piezoelectric</i>	3...20	0	0,002	200 × 340	50...500	50...200

Следует отметить, что в зависимости от условий эксплуатации у каждого типа МЭМС-переключателя может быть своя область применения, но у электростатических она шире, не только ввиду вышеупомянутых преимуществ, но также из-за возможности использования при изготовлении переключателя хорошо известной технологии производства полупроводниковых приборов [7].

Проектирование и изготовление МЭМС-переключателя для частотного диапазона 10...12 ГГц на арсениде галлия

При разработке модели узкополосного МЭМС-переключателя были выставлены следующие требования:

- вносимые потери не более 0,5 дБ;
- развязка принимаемого и передающего сигнала не хуже 30 дБ;
- напряжение питания не более 30 В;
- рабочий диапазон частот 10...12 ГГц;
- развязка по постоянному напряжению питания и СВЧ сигналу.

Для моделирования МЭМС-переключателей был выбран шунтирующий тип переключателя, и использовалась программа AWR Design Environment 2008. Шунтирующий переключатель размещается между СВЧ линией и заземляющим электродом. При подаче постоянного напряжения на СВЧ линию за счет сил электростатического притяжения мембрана опускается — емкость между заземленным электродом и СВЧ электродом резко возрастает и СВЧ сигнал стекает на землю. На рис. 1 представлена геометрическая схема шунтирующего переключателя и его эквивалентная схема.

В результате расчетов были найдены параметры принципиальной схемы СВЧ МЭМС-переключателя, которые приведены в табл. 3.

Результаты моделирования прохождения СВЧ сигнала через МЭМС-переключатель представлены на рис. 2 и 3. Резонансная частота для данного переключателя составила 10 ГГц. В принципиальной схеме рассчитывали идеальные элементы, не

Таблица 3
Table 3
Расчетные значения параметров принципиальной схемы
МЭМС-переключателя
*Calculated values of the parameters of the basic circuit
of the MEMS switch*

Элементы переключателя <i>Switch elements</i>	Параметры элементов переключателя <i>Parameters of the switch elements</i>	
Индуктивность <i>Inductance</i>	0,11 нН	
Сопротивление <i>Resistance</i>	0,15 Ом	
Емкость мембранны <i>Membrane capacity</i>	Вкл. <i>On position</i>	5,531 пФ
	Выкл. <i>Off position</i>	0,0439 пФ

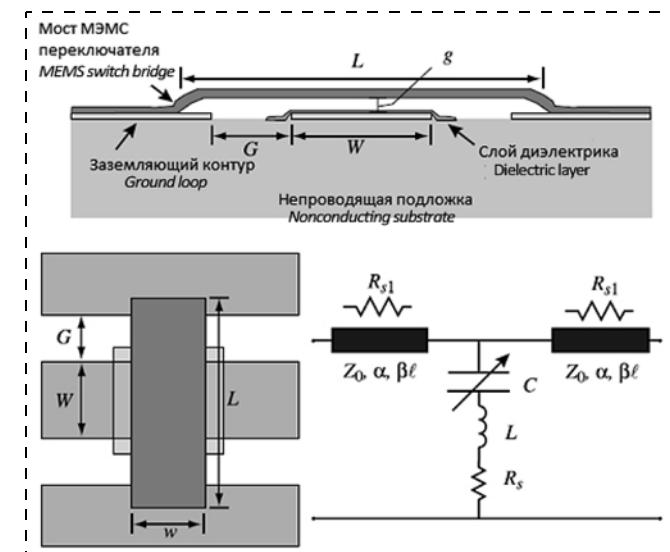


Рис. 1. Геометрическая и эквивалентная электрическая схема шунтирующего МЭМС-переключателя

Fig. 1. Geometrical and equivalent circuitry of the shunting MEMS switch

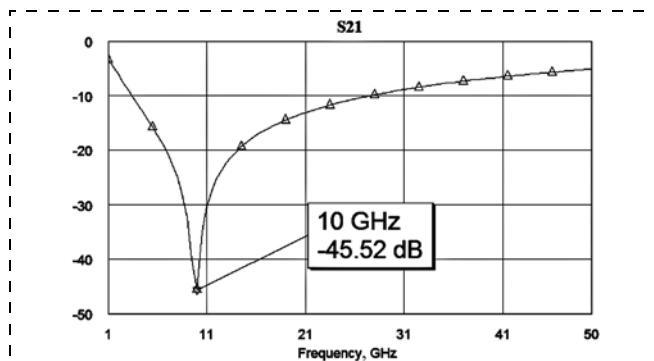


Рис. 2. Прохождение СВЧ сигнала через переключатель с опущенной мембраной

Fig. 2. Microwave signal transmission through the switch with the membrane in down position

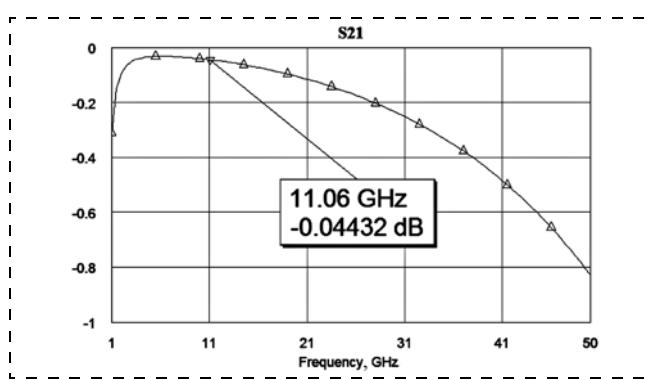


Рис. 3. Прохождение СВЧ сигнала через переключатель с поднятой мембраной

Fig. 3. Microwave signal transmission through the switch with the membrane in position up

отражающие поведение схемы на высоких частотах вследствие возникновения паразитных емкостей и индуктивностей в распределенных элементах. Для учета этих эффектов было проведено электрофи-

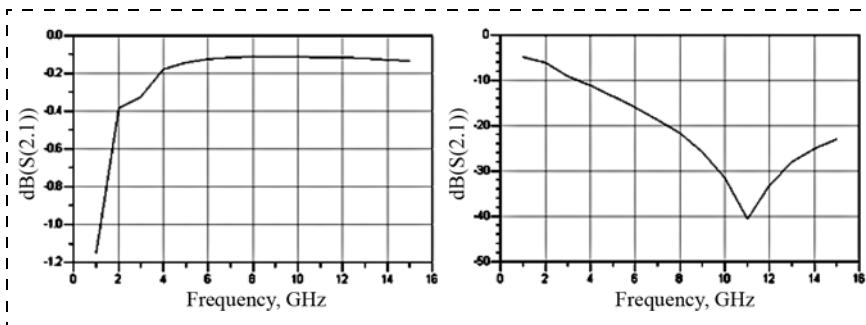


Рис. 5. Электрофизический расчет шунтирующего МЭМС-переключателя с креплением мембранны в виде меандра

Fig. 5. Electrophysical analysis of a shunt MEMS switch with a meander membrane

зическое моделирование данной схемы, которое заключалось в том, что расчет СВЧ характеристик топологии выполняли в широком частотном диапазоне, при этом непосредственно учитывали влияние топологии на прохождение СВЧ сигнала. По результатам электрофизического моделирования проводили конечную оптимизацию топологии и перерасчет скорректированной топологии.

Работу по оптимизации конструкции МЭМС-переключателя проводили в программе ADS (Advanced Design Systems). Топология получившегося

переключателя представлена на рис. 4 (см. третью сторону обложки), расчет переключателя представлен на рис. 5.

Результаты моделирования: вносимые потери менее 0,4 дБ, а изоляция сигнала не менее 32 дБ в диапазоне 10...12 ГГц, что удовлетворяет поставленной задаче.

Проведен расчет напряжения срабатывания мембранны по формуле, позволяющей точно рассчитать коэффициент упругости крепления мембранны в виде меандра произвольной формы.

Напряжение срабатывания мембранны получилось равным 17,2 В [8, 9].

МЭМС-переключатель был изготовлен с применением базовой технологии ИС на основе дискретных полупроводниковых приборов на базе гетероструктур AlGaAs/InGaAs/GaAs [10, 11]. Основные этапы изготовления МЭМС-переключателя и ИС МШУ (малошумящего усилителя) представлены в табл. 4.

Основным важным отличием процесса изготовления СВЧ МЭМС-устройств от стандартной

Таблица 4
Table 4

Сравнение технологических маршрутов изготовления МШУ на основе гетероструктур AlGaAs/InGaAs/GaAs и технологического маршрута изготовления МЭМС-переключателя

Comparison of the technological routes of manufacturing of LNA on the basis of AlGaAs/InGaAs/GaAs heterostructures and the technological route for manufacturing of a MEMS switch

Последовательность технологических операций при изготовлении МЭМС-переключателя <i>Sequence of the technological operations during manufacturing of a MEMS switch</i>	Последовательность технологических операций при изготовлении макетных образцов СнК со встроенными МШУ на гетероструктурах AlGaAs/InGaAs/GaAs <i>Sequence of the technological operations during manufacturing of model SoC samples with built in LNA on AlGaAs/InGaAs/GaAs heterostructures</i>
Подготовка поверхности пластины <i>Preparation of the plate surface</i>	Приборная изоляция <i>Instrument isolation</i>
Формирование нижнего электрода, контактных площадок и нижних обкладок развязывающих конденсаторов <i>Formation of the bottom electrode, contact platforms and bottom facings of the decoupling capacitors</i>	Омические контакты <i>Ohmic contacts</i>
	Формирование затвора длиной <0,15 мкм <i>Formation of a gate with the length of <0.15 μm</i>
Формирование опор моста <i>Formation of the bridge supports</i>	
Формирование пассивации СВЧ линии и нижних электродов <i>Formation of passivation of the microwave line and the bottom electrodes</i>	Пассивация активных поверхностей <i>Passivation of the active surfaces</i>
Формирование верхних обкладок развязывающих конденсаторов и утолщения опор моста <i>Formation of the top facings of the decoupling capacitors and the thickening of the bridge supports</i>	Формирование первого уровня электрических межсоединений <i>Formation of the first level of the electric interconnections</i>
Формирование "жертвенного слоя" Гальваническое утолщение элементов конструкции МЭМС-устройств Обтрав лишнего металла и формирование отверстий в мемbrane <i>Formation of "a sacrificial layer" Galvanic thickening of the design elements of the MEMS devices Removal of the superfluous metal and formation of apertures in the membrane</i>	Металлизация контактных площадок с "воздушными" мостами <i>Metallization of the contact platforms with "the air bridges"</i>

технологии изготовления дискретных полупроводниковых приборов на GaAs является удаление "жертвенного слоя" из-под мембранны МЭМС-переключателя. Данная проблема решалась путем выжигания остатков фоторезиста из-под мембранны в кислородной плазме, процессом жидкостного удаления полимера в растворителе и последующей сублимационной сушке в циклогексане.

Сравнение базовых операций изготовления МШУ на AlGaAs/InGaAs/GaAs и МЭМС-переключателей показали возможность изготовления МЭМС-переключателя и полупроводниковых приборов на базе арсенида галлия в едином технологическом цикле.

Исследование характеристик МЭМС-переключателя для диапазона частот 10...12 ГГц на арсениде галлия

Все изготовленные СВЧ переключатели изготовлены с площадками под зонды Infinity I67-A-GSG-150, используемые при измерениях *S*-параметров. Ниже представлены измеренные *S*-параметры СВЧ МЭМС-переключателей. Измеренная зависимость изоляции узкополосных СВЧ коммутаторов от частоты для двух СВЧ коммутаторов сигналов показана на рис. 6 (см. третью сторону обложки).

После изготовления СВЧ МЭМС-коммутаторов методом сублимационной сушки измеренные резонансные частоты лежат в диапазоне 13,9...14,9 ГГц. На данных частотах коэффициент изоляции составил 45...50 дБ, собственные потери составили 0,49 дБ.

Переключатель СВЧ сигнала 1×2 на нитриде галлия

Полевые транзисторы на нитридных гетероструктурах имеют ряд преимуществ по сравнению с кремниевыми и арсенидными: высокие пробивные напряжения, большой удельный ток, высокая стойкость к тепловому и радиационному воздействию. Это позволяет создавать малогабаритные коммутаторы мощных СВЧ сигналов для экстремальных условий эксплуатации.

В ИСВЧПЭ РАН был спроектирован и изготовлен на основе МИС переключателя СВЧ сигнала 1×2 (1 вход — 2 выхода) диапазона частот 0,1...15 ГГц на базе нитридных наногетероструктур [12]. При разработке конструкции МИС была применена копланарная топология без использования сквозных отверстий с учетом площадок под *flip-chip*-монтаж. Для уменьшения сопротивления сток-исток (R_{ON}) транзисторов до менее 0,2 Ом · мм была применена технология невплавных омических контактов, что позволило уменьшить вносимые потери в открытом канале (IL — *insertion loss*) и улучшить развязку (I_{so}) между каналами. Изготовление невплавных омических контактов было выполнено ИСВЧПЭ РАН совместно с НИЦ Курчатовский институт по разработанной ранее технологии. При

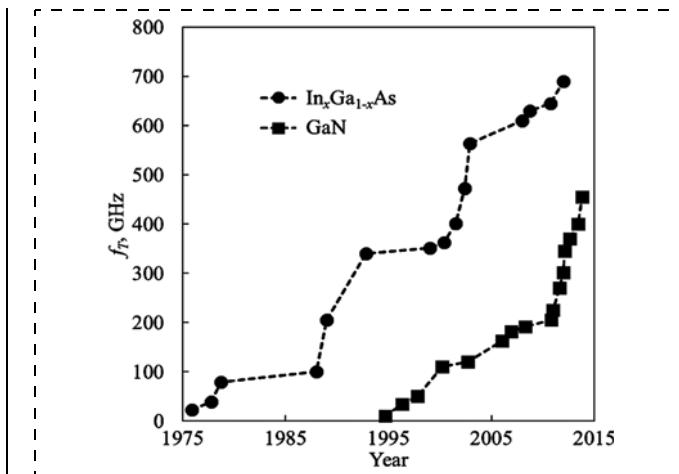


Рис. 7. Динамика увеличения предельных частот арсенидных и нитридных НЕМТ

Fig. 7. Dynamics of growth of the limiting frequencies of the arsenide and nitride HEMT

изготовлении были использованы нитридные гетероструктуры отечественного производства (ЗАО "Элма-Малахит") на сапфировых и SiC подложках. Размер основного варианта МИС-переключателя составил 1,2 × 1,1 мм. Расчеты показали, что можно уменьшить размер кристалла до 0,7 × 1,1 мм с сохранением характеристик.

Изготовленные образцы МИС-переключателя устанавливали на промежуточную подложку GaAs для улучшения теплоотвода и соединения с управляющими сигналами от GaN драйвера (МИЭТ) методом *flip-chip* (МИЭТ, "Исток"). Измерение характеристик переключателя в собранном виде было выполнено на оборудовании НПП "Исток" им. Шокина. Потери в открытом состоянии 1-го канала менее 1,7 дБ, развязка более 18 дБ на частоте 15 ГГц, потери 2-го канала менее 1,13 дБ, развязка 18,4 дБ на частоте 15 ГГц. Максимальная коммутируемая мощность МИС-переключателя составила 8,5 Вт (на сапфировых подложках).

По результатам работы получены свидетельства о государственной регистрации топологии интегрального СВЧ переключателя 1×2 для диапазона 4...18 ГГц в копланарном исполнении [13].

Освоение терагерцевого диапазона в последнее десятилетие является одним из самых приоритетных направлений развития полупроводниковых СВЧ приборов. Частотные параметры полевых транзисторов с двумерным электронным газом на базе арсенидных наногетероструктур (In, Ga, Al, P)As на подложках GaAs и InP и на базе нитридных наногетероструктур (In, Ga, Al)N на подложках Al₂O₃, SiC и Si (далее арсенидные и нитридные НЕМТ соответственно) в эти годы возрастили наиболее высокими темпами (рис. 7). Это стало возможным в результате развития технологии изготовления транзисторов, а именно снижения сопротивления омических контактов путем освоения технологии повторного выращивания высоколегированного

контактного слоя $n+$ -InGaAs или $n+$ -GaN и создания как для арсенидных, так и для нитридных НЕМТ самосовмещенной технологии изготовления затворов и контактов, уменьшающей сопротивление канала транзистора до физического минимума, а также уменьшения длины затворов до 20 нм.

У GaN отличные электрическая плотность и подвижность электронов, что дает в итоге более высокие показатели, чем у SiC. Это справедливо для большинства применений, требующих преобразования электроэнергии, за исключением режимов работы при высоких температурах, где SiC превосходит все известные материалы [14].

На рис. 7 показано, что рост предельной частоты усиления по току f_T арсенидных НЕМТ остановился на отметке $f_T = 688$ ГГц уже в 2011 г., а нитридных НЕМТ после достижения $f_T = 454$ ГГц в 2013 г.

На основе анализа экспериментальных и теоретических исследований радиационной стойкости, проведенных в работе [15], можно сделать вывод, что GaN является достаточно перспективным полупроводниковым материалом с точки зрения радиационно-стойкой СВЧ элементной базы для использования в атомной и космической промышленности [16, 17].

Заключение

На основании проведенного анализа современного состояния разработок в области применения широкозонных наногетероструктур AlGaN/AlN/GaN в СВЧ диапазоне за рубежом и опыта работ ИСВЧПЭ РАН с гетероструктурами AlGaN/GaN, полученного в ходе выполнения ряда НИР и ОКР в предыдущие годы [18], сделан вывод о возможности и необходимости перенесения акцента исследований на создание технологии проектирования и изготовления широкой номенклатуры радиационно-стойких МИС для приемо-передающих модулей сантиметрового и миллиметрового диапазонов на базе широкозонных НЕМТ-гетероструктур отечественных производителей (ЗАО "Элма-Малахит", ЗАО "Светлана-Рост", НИЦ "Курчатовский институт").

Работа выполнена при финансовой поддержке Министерства образования и науки РФ (соглашение о предоставлении субсидии № 14.607.21.0011, уникальный идентификатор проекта RFMEF160714X0011).

Список литературы

1. Вернер В. Д., Мальцев П. П., Сауров А. Н. МЭМС и третья индустриальная революция // Нано- и микросистемная техника. 2012. № 11. С. 2—5.
2. Аравин В. В., Вернер В. Д., Сауров А. Н., Мальцев П. П. МЭМС высокого уровня — возможный путь развития МЭМС в России // Нано- и микросистемная техника. 2011. № 6. С. 28—31.
3. Аскерко А. Н., Бахов О. С., Лучинин В. В. Испытание и тестирование микроэлектромеханических компонентов и систем на их основе // Нано- и микросистемная техника. 2013. № 2. С. 49—54.
4. Амеличев В. В., Генералов С. С., Сывороткин П. А., Шаманаев С. В., Платонов В. В. Устройство управления микроэлектромеханическими реле на основе массива силовых МЭМС-ключей // Нано- и микросистемная техника. 2014. № 2. С. 43—45.
5. DeNatale J., Mihailovich R., Waldrop J. Techniques for reliability analysis of MEMS RF switches // Proc. IEEE Int. Reliability Physics Symposium, Dallas, TX, USA. Apr. 7—11. 2002. P. 116—117.
6. Амеличев В. В., Генералов С. С., Якунина А. В., Шаманаев С. В., Платонов В. В. Конструкция и технология изготовления матрицы силовых микроэлектромеханических ключей // Нано- и микросистемная техника. 2014. № 3. С. 28—32.
7. Платонов В. В., Генералов С. С., Смехова М. И., Амеличев В. В., Поломошинов С. А. Электростатический МЭМС-ключ на структуре кремний-стекло // Нано- и микросистемная техника. 2013. № 7. С. 43—47.
8. Мальцев П. П., Майтама М. В., Павлов А. Ю., Шаврук Н. В. Расчет и изготовление узкополосного СВЧ микроэлектромеханического переключателя для частотного диапазона 10...12 ГГц на подложках арсенида галлия // Нано- и микросистемная техника. 2014. № 6. С. 30—33.
9. Jun Z., Chang L., Schutt-Aine' J. E. Development of a Wide-Tuning-Range Two-Parallel-Plate Tunable Capacitor for Integrated Wireless Communication Systems // J. RF and Microwave CAE. 2001. N. 11. P. 322—329.
10. Jinghong C., Jun Z., Chang L., Schutt-Aine' J. E., Kang S.-M. K. Design and Modeling of a Micromachined High-Q Tunable Capacitor With Large Tuning Range and a Vertical Planar Spiral Inductor // IEEE Transactions on electron devices. 2003. Vol. 50, N. 3. P. 730—739.
11. Мальцев П. П., Лисицкий А. П., Павлов А. Ю., Шаврук Н. В., Побойкина Н. В., Хачатрян В. Д. Возможности формирования МЭМС-варакторов с электростатическим управлением в GaAs-технологии // Нано- и микросистемная техника. 2012. № 9. С. 28—33.
12. Федоров Ю. В., Богданов Ю. М., Егоркин В. И., Заваевский М. Л., Павлов А. Ю., Майтама М. В. Переключатель СВЧ-сигнала 1×2 диапазона частот до 15 ГГц на базе AlGaN/GaN НЕМТ с невжигаемыми омическими контактами, смонтированный методом flip-chip // Труды Международного форума "Микроэлектроника-2016". 2-я научная конференция "Интегральные схемы и микроэлектронные модули". Республика Крым, г. Алушта, 26—30 сентября 2016 г. С. 349—353.
13. Федоров Ю. В., Майтама М. В. Топология ИМС "Интегральный СВЧ переключатель 1×2 для диапазона частот 4—18 ГГц в копланарном исполнении". Свидетельство о государственной регистрации № 2015630112 от 11 января 2016 г.
14. Ботвел Б., Драммонд Д., Пилла М., Зинг У. Г., Джина Д., Кон Г. GaN-транзисторы с вертикальной структурой: эффективные решения для силовой электроники // Электроника. 2017. № 2. С. 92—96.
15. Громов Д. В., Матвеев Ю. А., Федоров Ю. В. Влияние радиации на характеристики элементов на нитриде галлия // Нано- и микросистемная техника. 2011. № 5. С. 39—48.
16. Xinwen Hu, Karmarkar A. P., Bongim Junet et al. Proton — Irradiation Effects on AlGaN/AlN/GaN High Electron — Mobility Transistors // IEEE Transactions on Nuclear Science. 2003. Vol. 50, N. 6. P. 1791—1796.
17. Sonia G., Brunner F., Denker A. et al. Proton and Heavy Ion Irradiation Effects on AlGaN/GaN HFET Devices // IEEE Transactions on Nuclear Science. 2006. Vol. 53, N. 6. P. 3661—3666.
18. Федоров Ю. В., Михайлович С. В. Перспективы замены арсенидных МИС на нитридные // Нано- и микросистемная техника. 2016. Т. 18, № 4. С. 217—226.

V. K. Il'kov, Ph. D., Head of Department, e-mail: m2lkeny@yandex.ru,
Federal State Budget Educational Institution of Higher Education "Moscow Technological University",
Moscow, Russian Federation,

A. O. Mikhalev, Postgraduate Student, e-mail: m2lkeny@yandex.ru, **M. V. Maytama**, Postgraduate Student,
Federal State Budgetary Institution of Science Institute of Ultra-High Frequency Semiconductor Electronics,
RAS, Moscow, 117105, Russian Federation

Corresponding author:

Mikhalev Artem O., Postgraduate Student, Federal State Budgetary Institution of Science Institute of Ultra-High
Frequency Semiconductor Electronics, RAS, Moscow, 117105, Russian Federation, e-mail: m2lkeny@yandex.ru

Arsenide and Nitride Gallium Switches

Received on January 23, 2018
Accepted on February 20, 2018

The article presents the main types of MEMS switches. A MEMS switch on AlGaAs/InGaAs/GaAs heterostructure was developed and manufactured for the 10...12 GHz frequency range. The article presents the results of measurements of the MEMS switch on gallium arsenide.

A nitride nanoheterostructure 1×2 MIS microwave switch for the 0.1...15 GHz frequency range was designed and manufactured. An analysis of the current state of developments in the field of application of the wide-band AlGaN/AlN/GaN nanoheterostructures in the microwave range was carried out.

Keywords: X-band, monolithic integrated circuit, MEMS, switches, gallium arsenide, gallium nitride

For citation:

Il'kov V. K., Mikhalev A. O., Maytama M. V. Arsenide and Nitride Gallium Switches, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 425—433.

DOI: 10.17587/nmst.20.425-433

Introduction

Decrease of the dimensions and weight of the microwave devices, expansion of their dynamic range of frequencies, reduction of their power consumption and costs, possibility of the use of the wireless systems for increasing the mobility and availability to the sources of information consumption [1], for escalation of their integration and functionalities at a growth of the working frequencies, demand from the developers a maximal use of the experience and technological methods for development of integrated circuits (IC) on Si, GaAs and GaN substrates.

One of the problems arising in the process of development of the microwave devices is creation of the signal switches.

Now, the most wide-spread microwave switches are of two types:

- Electromechanical;
- Electronic (on discrete devices).

For switching of the microwave signals the most promising switches are the ones based on microelectromechanical systems (MEMS) [2], because they allow us to do the following:

— To obtain switches of the microwave signals with a zero energy consumption in the state of rest;

— To lower the dimensions and weight of the microwave devices;

— To lower the energy consumption of the microwave devices (the power necessary for switching of a switch is about 1 nW);

— To reduce considerably the switching time in comparison with the mechanical microwave switches;

— To obtain the characteristics of a switched signal as linear ones.

The unquestionable advantages of the MEMS components determine justify their application in highly reliable equipment, including equipment for space and special purposes [3].

Microwave switches

There are three types of widely applied microwave switches: electromechanical, solid-state (semi-conductor) and MEMS

switches. The electromechanical switches surpass all the other ones by the switching power — up to several kilowatts on frequency of 1 GHz, they have key insertion losses (own losses) of 0.1...0.3 dB and signal isolation of 60...80 dB.

At that, the electromechanical switches have higher power consumption — up to 10 W (for powerful switches), switching time of 10...20 ms and relatively short service life — about 1 million cycles. Table 1 presents the comparison characteristics of the semiconductor switches on the basis of PIN diodes, field transistors and MEMS switches of the microwave signals.

It is visible, that MEMS switches have advantages in comparison with the solid-state microwave switches: high ratio of losses to isolation in the open and closed positions, and practically zero power consumption in the closed position. At that, they have drawbacks: low speed in comparison with the semiconductor switches. For control it is necessary to form a pulse of the switching voltage from 6 up to 80 V. We should point out, that the service life of the MEMS switches is about 10^{10} cycles. Hence, it is possible to say, that MEMS switches combine in themselves the best features of the solid-state and the electromechanical switches — small overall dimensions, high speed at a low own power consumption, and long service life (typical for the solid-state switches), at that, a possibility appears for switching of a microwave signal of high power (up to 10 W) (typical for the electromechanical switches). It is necessary to point out, that the MEMS switches demand a vacuum packaging, which, in turn, allows us to increase their speed and to prolong considerably their service life.

Types of MEMS switches

During manufacturing of switches of average and low power (up to 100 mW) a switch with a capacitor type of contact has an important advantage — possibility of manufacturing of the switch with an low response voltage, which allows us to use a single power supply circuit both for the semi-conductor devices and for control of the switches.

Table 2 presents MEMS devices of various operating principles.

It is visible that the electrostatic relays are most suitable for use in the MEMS switches [4] because:

- Their power consumption is close to zero;
- A MEMS switch has a small size;
- During manufacturing the planar technology is used;
- The switching time is the quickest;
- Actuation of the relay requires only 50... 200 μ N;
- When LC decoupling is used, the control voltage can be supplied via a high frequency waveguide.

The advantages of an electrostatic drive of the MEMS key include a very small power consumption (the current is consumed only during switching, a typical value of the power consumption per one cycle of switching does not exceed 1 nW) [5, 6]. Its only serious drawback is the necessity to use the control voltage of 20...80 V, while most of the circuits are controlled by voltage of 3...5 V.

It is necessary to point out, that depending on the conditions of operation, each type of the MEMS switch can have its sphere of application, but that of the electrostatic one is wider, not only because of the above advantages, but also because of the possibility to use the well-known technology of the semiconductor devices for manufacturing of the switch [7].

Designing and manufacturing of the MEMS switches for the frequency range of 10...12 GHz on gallium arsenide

During development of a model of the narrow-band MEMS switch, the following requirements were taken into consideration:

- Insertion loss is not more than 0.5 dB;
- Decoupling of the received and transmitted signal is not worse than 30 dB;
- Voltage of the power supply is not more than 30 V;
- Working range of frequencies — 10...12 GHz;
- Decoupling by constant voltage of power supply and microwave signal.

For modeling of MEMS switches a shunting type of a switch was selected and AWR Design Environment 2008 program was used. The shunting switch was placed between the microwave line and the earthing electrode. When a constant voltage was supplied to the microwave line, due to the forces of electrostatic attraction the membrane fell down — the capacity between the earthed electrode and the microwave electrode increased sharply and the microwave signal came down to earth. Fig. 1 presents the geometrical circuit of the shunting switch and its equivalent circuit.

As a result of calculations, the parameters of the basic circuit of the microwave MEMS switch (presented in table 3) were found.

The results of modeling of transmission of a microwave signal through the MEMS switch are presented in fig. 2 and 3. The resonant frequency for the given switch was 10 GHz. In the basic circuit the ideal elements were calculated, not reflecting behavior of the circuit at high frequencies owing to occurrence of the parasitic capacities and inductances in the distributed elements. For account of these effects an electrophysical modeling of the given circuit was done, which boiled down to calculation of the microwave characteristics of the topology in a wide frequency range, at that, the influence of the topology on transmission of a microwave signal was considered directly. By the results of the electrophysical modeling, the final optimization of the topology and recalculation of the corrected topology were done.

The optimization of the design of the MEMS switch was done in ADS program (Advanced Design Systems). The topology of the obtained switch is presented in fig. 4 (see the 3rd side of cover), the switch calculation is presented in fig. 5.

The results of modeling: the insertion loss was less than 0.4 dB, and signal isolation was not less than 32 dB in the range of 10...12 GHz, which met the aims of the set task.

Calculation of the response voltage of the membrane was carried out according to the formula, allowing us to calculate precisely the elasticity factor of fastening of the membrane in the form of a meander of any form.

The response voltage of the membrane turned out to be equal to 17.2 V [8, 9].

The MEMS switch was made with application of the basic IC technology of the discrete semi-conductor devices based on AlGaAs/InGaAs/GaAs heterostructures [10, 11]. The main stages of manufacturing of the MEMS switch and IC LNA (low-noise amplifier) are presented in table 4.

The basic difference of the process for manufacturing of the microwave MEMS devices from the standard manufacturing techniques of the discrete semi-conductor devices on GaAs is removal of "the sacrificial layer" from under the MEMS switch membrane. The given problem was solved by burning out of the remains of the photoresist from under the membrane in the oxygen plasma by the process of liquid removal of the polymer in solvent and the subsequent sublimation drying in cyclohexane.

Comparison of the basic manufacturing operations of LNA on AlGaAs/InGaAs/GaAs and MEMS switches demonstrated a possibility of manufacturing of the MEMS switches and semi-conductor devices on the basis of gallium arsenide in a single technological cycle.

Research of characteristics of the MEMS switches on gallium arsenide for the range of frequencies of 10...12 GHz

All the manufactured microwave switches were made with the platforms for Infinity I67-A-GSG-150 probes used during the measurements of *S*-parameters. Below the measured *S*-parameters of the microwave MEMS switches are presented. The measured dependence of isolation of the narrow-band microwave switch on frequency for two microwave switches of signals is shown in fig. 6, see the 3rd side of cover.

After manufacturing of the microwave MEMS switches by the method of sublimation drying the measured resonant frequencies were within the range of 13.9...14.9 GHz. On the given frequencies the isolation coefficient was 45...50 dB, the inherent loss was 0.49 dB.

Microwave 1×2 signal switch on gallium nitride

The field transistors on the nitride heterostructures have a number of advantages in comparison with the silicon and arsenide ones: high breakdown voltages, big specific current, high resistance to the thermal and radiation influences. This allows us to create small-sized switches of powerful microwave signals for the extreme operating conditions.

ISVChPE of RAS designed and manufactured a MIC switch of a 1×2 microwave signal (1 input — 2 outputs) in the 0.1...15 GHz frequency range on the basis of nitride heterostructures [12]. During development of the MIC design the coplanar topology was applied without the use of the through apertures and with account of the platforms for a *flip-chip* assembly. In order to reduce the drain-source (R_{ON}) resistance of transistors less than 0.2 $\Omega \cdot \text{mm}$ the technology of non-melting ohmic contacts was applied which allowed us to reduce the insertion loss in the open channel and improve decoupling (I_{so}) between the channels. Manufacturing of the non-melting ohmic contacts was done by ISVChPE of RAS jointly with Kurchatov Institute in accordance with the developed technology. In the process of manufacturing the nitride heterostructures on sapphire and SiC substrates were used of

domestic production (Elma-Malachite Co.). The size of the basic version of the MIC switch was 1.2×1.1 mm. Calculations demonstrated that it was possible to reduce the size of a crystal down to 0.7×1.1 mm with preservation of its characteristics.

The manufactured samples of the MIC switch were placed on GaAs intermediate substrate for improvement of heat-conducting and connection with the control signals from GaN driver (MIET) the *flip-chip* method (MIET, Istok Co.). Measurement of the characteristics of the switch in the assembled form was taken on equipment of Shokin Istok Co. Losses in the open position of the 1st channel were less than 1.7 dB, decoupling was more than 18 dB on frequency of 15 GHz, losses of the 2nd channel were less than 1.13 dB, decoupling — 18.4 dB on frequency of 15 GHz. The maximal switching power of the MIC switch was 8.5 W (on sapphire substrates).

By the results of the work the certificates were obtained on the state registration of the topology of the integrated 1×2 microwave switch for the range of 4...18 GHz in a coplanar version [13].

Mastering of the terahertz range is one of the priority directions in development of the semi-conductor microwave devices. These years witnessed the highest rates of increase of the frequency parameters of the field transistors with the two-dimensional electron gas on the basis of arsenide nanoheterostructures (In, Ga, Al, P)As on GaAs and InP substrates and on the basis of nitride nanoheterostructures (In, Ga, Al)N on Al_2O_3 , SiC and Si substrates (hereinafter, the arsenide and nitride HEMT, accordingly) (fig. 7). This became possible as a result of development of the manufacturing technologies of the transistors, namely, a decrease of resistance of the ohmic contacts due to mastering of the technology of a repeated growth of a high-alloy contact layer of $n+$ -InGaAs or $n+$ -GaN, and development for both arsenide and nitride HEMT of the self-combined manufacturing technology of the gates and contacts, reducing the resistance of the transistor channel to a physical minimum, and also reduction of the length of the gates down to 20 nm.

GaN has excellent electric density and mobility of electrons, which, as a result, ensures higher indicators, than SiC. This is fair for most applications demanding transformation of the electric power, except for the operating modes at high temperatures, where SiC surpasses all the known materials [14].

Fig. 7 shows, that the growth of the limiting frequency of strengthening by current f_T of the arsenide HEMT stopped at the mark $f_T = 688$ GHz already in 2011, while that of the nitride HEMT — after achievement of the level of $f_T = 454$ GHz in 2013.

On the basis of the analysis of the research of the radiation resistance done in [15] it is possible to draw a conclusion, that GaN is a rather promising semi-conductor material from the point of view of the radiation resistance of the microwave element base for use in the nuclear and space industries [16, 17].

Conclusion

On the basis of the analysis of the state of developments in the sphere of application of the wide-band nanoheterostructures of AlGaN/AlN/GaN in the microwave frequency range done abroad and on the basis of the experience of ISVChPE of RAS with heterostructures of AlGaN/GaN, received during performance of R & D in previous years [18], a conclusion was drawn on the possibility and necessity of a transfer of the accent of research on development of the technology for designing and manufacturing of a wide nomenclature of the radiation-resistant MIC for the receiving-transmitting modules of the centimeter and millimeter ranges on the basis of the wide-band NEMT heterostructures from domestic manufacturers (Elma-Malachite Co., Svetlana-Rost Co., and Kurchatov Institute).

The work was done with the financial support of the Ministry of Education and Science of the Russian Federation (grant agreement № 14.607.21.0011, unique identifier of the project — RFMEF160714X0011).

References

1. Verner V. D., Mal'cev P. P., Saurov A. N. MJeMS i tret'ja industrial'naja revoljucija, *Nano- i mikrosistemnaya tekhnika*, 2012, no. 11, pp. 2—5 (in Russian).
2. Aravin V. V., Verner V. D., Saurov A. N., Mal'cev P. P. MJeMS vysokogo urovnya — vozmozhnyj put' razvitiya MJeMS v Rossii, *Nano- i mikrosistemnaya tekhnika*, 2011, no. 6, pp. 28—31 (in Russian).
3. Askero A. N., Bohov O. S., Luchinin V. V. Ispytanie i testirovanie mikroelektromehanicheskikh komponentov i sistem na ih osnove, *Nano- i mikrosistemnaya tekhnika*, 2013, no. 2, pp. 49—54 (in Russian).
4. Amelichev V. V., Generalov S. S., Syvorotkin P. A., Shamanaev S. V., Platonov V. V. Ustrojstvo upravlenija mikroelektromehanicheskikh rele na osnove massiva silovyh MJeMS-kljuchej, *Nano- i mikrosistemnaya tekhnika*, 2014, no. 2 pp. 43—45 (in Russian).
5. DeNatale J., Mihailovich R., Waldrop J. Techniques for reliability analysis of MEMS RF switches, *Proc. IEEE Int. Reliability Physics Symposium*, Dallas, TX, USA, Apr. 7—11, 2002, pp. 116—117.
6. Amelichev V. V., Generalov S. S., Jakunina A. V., Shamanaev S. V., Platonov V. V. Konstrukcija i tehnologija izgotovlenija matricy silovyh mikroelektromehanicheskikh kljuchej, *Nano- i mikrosistemnaya tekhnika*, 2014, no. 3, pp. 28—32 (in Russian).
7. Platonov V. V., Generalov S. S., Smehova M. I., Amelichev V. V., Polomoshnov S. A. Jelektrostaticheskij MJeMS-kljuch na strukture kremnij-steklo, *Nano- i mikrosistemnaya tekhnika*, 2013, no. 7, pp. 43—47 (in Russian).
8. Mal'cev P. P., Majtama M. V., Pavlov A. Ju., Shavruk N. V. Raschet i izgotovlenie uzkopolosnogo SVCh mikroelektromehanicheskogo pereklyuchatelya dlja chastotnogo diapazona 10...12 GGc na podlozhkah arsenida gallija, *Nano- i mikrosistemnaya tekhnika*, 2014, no. 6, pp. 30—33 (in Russian).
9. Jun. Z., Chang L., Schutt-Aine' J. E. Development of a Wide-Tuning-Range Two-Parallel-Plate Tunable Capacitor for Integrated Wireless Communication Systems, *J. RF and Microwave CAE*, 2001, no. 11, pp. 322—329.
10. Jinghong C., Jun Z., Chang L., Schutt-Aine' J. E., Kang S.-M. K. Design and Modeling of a Micromachined High-Q Tunable Capacitor With Large Tuning Range and a Vertical Planar Spiral Inductor, *IEEE Transactions on electron devices*, 2003, vol. 50, no. 3, pp. 730—739.
11. Mal'cev P. P., Lisickij A. P., Pavlov A. Ju., Shavruk N. V., Pobojkina N. V., Hachatryan V. D. Vozmozhnosti formirovaniija MJeMS-varaktorov s elektrostaticeskim upravleniem v GaAs-tehnologii, *Nano- i mikrosistemnaya tekhnika*, 2012, no. 9, pp. 28—33 (in Russian).
12. Fedorov Ju. V., Bogdanov Ju. M., Egorkin V. I., Zanavskin M. L., Pavlov A. Ju., Majtama M. V. Pereklyuchatel' SVCh-signal'a 1h2 diapazona chastot do 15 GGc na baze AlGaN/GaN HEMT s nevzhigayimi omicheskimi kontaktami, smontirovannyj metod flip-chip, *Trudy Mezhdunarodnogo foruma "Mikroelektronika-2016". 2-ja nauchnaja konferencija "Integral'nye shemy i mikroelektronnye moduli"*. Respublika Krym, Alushta, 26—30 September 2016, pp. 349—353 (in Russian).
13. Fedorov Ju. V., Majtama M. V. Topologija IMS "Integral'nyj SVCh pereklyuchatel' 1h2 dlja diapazona chastot 4—18 GGc v koplanarnom ispolnenii". Svidetel'stvo o gosudarstvennoj registraciї № 2015630112 ot 11 janvarja 2016 g. (in Russian).
14. Botvel B., Draymond D., Pilla M., Zing U. G., Dzhina D., Kon G. GaN-tranzistory s vertikal'noj strukturoj: effektivnye reshenija dlja silovoj elektroniki, *Elektronika*, 2017, no. 2, pp. 92—96 (in Russian).
15. Gromov D. V., Matveev Ju. A., Fedorov Ju. V. Vlijanie radiacii na harakteristiki jelementov na nitride gallija, *Nano- i mikrosistemnaya tekhnika*, 2011, no. 5, pp. 39—48 (in Russian).
16. Xinwen Hu., Karmarkar A. P., Bongim Junet et al. Proton — Irradiation Effects on AlGaN/AlN/GaN High Electron — Mobility Transistors, *IEEE Transactions on Nuclear Science*, 2003, vol. 50, no. 6, pp. 1791—1796.
17. Sonia G., Brunner F., Denker A. et al. Proton and Heavy Ion Irradiation Effects on AlGaN/GaN HFET Devices, *IEEE Transactions on Nuclear Science*. 2006, vol. 53, no. 6, pp. 3661—3666.
18. Fedorov Ju. V., Mihajlovich S. V. Perspektivy zameny arsenidnyh MIS na nitridnye, *Nano- i mikrosistemnaya tekhnika*, no. 4, 2016, pp. 217—226 (in Russian).

МОЛЕКУЛЯРНАЯ ЭЛЕКТРОНИКА И БИОЭЛЕКТРОНИКА

MOLECULAR ELECTRONICS AND BIOELECTRONICS

УДК 004.032.26

DOI: 10.17587/nmst.20.434-448

Е. С. Горнев, д-р техн. наук, проф., e-mail: egornev@niime.ru,
Г. С. Теплов, аспирант, мл. науч. сотр., e-mail: gteplov@niime.ru,
АО "НИИМЭ", Москва, Московский физико-технический институт

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ КОНЕЧНОГО АВТОМАТА АБСТРАКТНОГО НЕЙРОНА И СЕТЕЙ НА ЕГО ОСНОВЕ

Поступила в редакцию 12.03.2018

На основе математических моделей мультипликативного и суммирующего искусственных нейронов предлагается теоретико-мноожественная модель конечного автомата абстрактного нейрона, сформулированная в терминах конечных множеств. В отличие от известных моделей искусственных нейронов предлагаемый подход учитывает возможность изменения функции активации в процессе работы или обучения нейрона. Построение модели на основе автомата Мура позволяет использовать ее при аппаратной реализации большинства моделей искусственных нейронов.

Ключевые слова: математическая модель абстрактного нейрона, функция активации искусственного нейрона, конечный автомат абстрактного нейрона, дискретная модель искусственной нейронной сети

Введение

Развитие микро- и наноэлектроники последнего десятилетия, выражющееся в расширении номенклатуры электронной компонентной базы [1], уменьшении минимальных топологических размеров и разработке схемотехнических решений на основе этих компонентов, вновь пробудило интерес к аппаратным реализациям нейронных сетей [2, 3]. Существующие модели искусственных нейронов по способу учета входных сигналов могут быть разделены на два основных типа. Первый тип нейронов — это суммирующие нейроны, впервые предложенные в работе МакКалюка и Питса [4], использующие для учета взвешенных входных сигналов операцию сложения. Второй тип нейронов — это мультипликативные нейроны [5], использующие для учета взвешенных входных сигналов операцию умножения. Пример использования одной из разновидностей данного класса моделей показан в работе [6]. В работе проведено обобщение обеих математических моделей в единую модель, названную абстрактным нейроном и ориентированную на аппаратную реализацию. Более сложные модели, такие как $\Sigma\Pi$ -нейроны, агрегирующие нейроны [7] и нейроны со сложными синапсами, могут быть сведены к мультипликативным и суммирующим моделям нейронов. Более подробный обзор существующих моделей можно найти в работе [8].

Все существующие основные подходы к аппаратной реализации вычислительной техники вне зависимости от схемотехнических решений могут быть проанализированы с применением дискретных конечных множеств. Дискретность цифровой схемотехники не вызывает сомнений. Дискретность гибридной и аналоговой схемотехники определяется соотношением сигнал—шум, в результате которого все сигналы ниже определенного порога являются неразличимыми между собой, что в свою очередь влечет дискретизацию множества их значений. Конечность множества обрабатываемых сигналов обусловливается ограниченностью любой аппаратной реализации вычислительной системы.

Таким образом, описание аппаратной реализации абстрактного нейрона как вычислительного элемента удобно проводить в терминах конечных автоматов. С учетом математического описания рассматриваемых моделей, а именно зависимости результата функции активации от результата операции учета взвешенных сигналов (сложение или умножение), наиболее подходящим является использование автомата Мура [9].

1. Анализ суммирующего и мультипликативного нейронов

Искусственный нейрон описывается набором вида $(n, Y, X, W, C, \delta, \Delta, \lambda)$, где:

$n \in \mathbb{N}$ — число входов (синапсов);

$Y \subseteq \mathbb{R}$ — множество выходных сигналов нейрона;

$X = \bigcup_{i=0}^n X_i, X \subseteq \mathbb{R}$ — множество входных сигналов, подаваемых на синапсы искусственного нейрона;

$W = \bigcup_{i=0}^n W_i, W \subseteq \mathbb{R}$ — множество всех возможных значений весовых коэффициентов синапса;

$C \subseteq \mathbb{R}$ — множество подгоночных коэффициентов функции активации;

δ — операция учета весового коэффициента входного сигнала (традиционно применяется операция умножения $\delta(w_i \cdot x_i) = w_i \cdot x_i; w_i \in W; x_i \in X$);

Δ — операция учета взвешенных входных сигналов $\sum_{i=1}^n (w_i \cdot x_i)$ или $\prod_{i=1}^n (w_i \cdot x_i + c_i), c_i \in C; w_i \in W, x_i \in X$;

λ — функция активации искусственного нейрона.

Здесь и далее $\mathbb{N} \cap 0 = \emptyset$.

2. Обобщение моделей суммирующего и мультипликативного нейронов

Зависимость сигнала на выходе для суммирующего нейрона имеет вид:

$$y = \lambda \left(\sum_{i=1}^n w_i \cdot x_i + c \right), \quad c \in C, \quad (1)$$

а для мультипликативного нейрона имеет вид:

$$y = \lambda \left(\prod_{i=1}^n (w_i \cdot x_i + c_i) \right), \quad c_i \in C. \quad (2)$$

С учетом введенного в разд. 1, имеем следующие описания для абстрактного, суммирующего и мультипликативного нейронов соответственно:

$$\begin{aligned} y &= \lambda_a \left(\sum_{i=1}^n \delta(w_i \cdot x_i) \right), \\ \delta: \mathbb{R}^2 &\rightarrow \mathbb{R}, \Delta: \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_a: \mathbb{R} \rightarrow Y, \end{aligned} \quad (3)$$

$$\begin{aligned} y &= \lambda_s \left(\prod_{i=1}^n w_i \cdot x_i + c \right), \\ \delta: \mathbb{R}^2 &\rightarrow \mathbb{R}, \Sigma: \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_s: \mathbb{R}^2 \rightarrow Y, \end{aligned} \quad (4)$$

$$\begin{aligned} y &= \lambda_m \left(\prod_{i=1}^n (w_i \cdot x_i + c_i) \right), \\ \delta: \mathbb{R}^3 &\rightarrow \mathbb{R}, \Pi: \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_m: \mathbb{R} \rightarrow Y. \end{aligned} \quad (5)$$

Первым условием эквивалентности автоматов M_1, M_2 является равенство алфавитов входных $X_1 = X_2$ и выходных $Y_1 = Y_2$ символов. Вторым условием эквивалентности является получение оди-

наковых выходных символов $y_1 = y_2, y_1 \in Y_1, y_2 \in Y_2$ при задании на входе автомата одинаковых входных символов $x_1 = x_2, x_1 \in X_1, x_2 \in X_2$ для всех элементов множества входных символов.

Аналогично будем считать эквивалентными модели искусственных нейронов, для которых выполняется следующее: множества входных символов (сигналов, элементов) равны; множества выходных символов (сигналов, элементов) равны; при подаче на два эквивалентных нейрона любого входного символа их выходные символы равны.

Приведенное уравнение (2) не совпадает с описанием абстрактного нейрона (3) в силу наличия в нем констант, приведем его к требуемому виду.

Поскольку \mathbb{R} непрерывно (аксиома о полноте \mathbb{R}), из формулы (2) для случая $\forall x_i \neq 0$ справедливо:

$$\begin{aligned} \forall x_i \neq 0, \exists \tilde{w}_i \in \mathbb{R} (\tilde{w}_i \cdot x_i = w_i \cdot x_i + c_i) \Rightarrow \prod_{i=1}^n \tilde{w}_i \cdot x_i = \\ = \prod_{i=1}^n (w_i \cdot x_i + c_i) = y_k, \forall y_k \in Y_k, Y_k \subseteq Y. \end{aligned}$$

Условиями равенства функций f_1, f_2 являются:

$$\begin{aligned} \text{Dom}f_1 = \text{Dom}f_2, \quad \text{ran}f_1 = \text{ran}f_2, \\ \forall x_i \in \text{Dom}f_1 ((f_1(x_i) = f_2(x_i)) \wedge ((\bigcup_i x_i) \cap \text{Dom}f_1 = \emptyset)). \end{aligned}$$

Тогда для выполнения условий равенства функций активации при $\exists x_i = 0$

$$\lambda_{m1} \left(\prod_{i=1}^n \tilde{w}_i \cdot x_i \right) = \lambda_{m1} \left(\prod_{i=1}^n w_i \cdot x_i + c \right),$$

перезапишем функции λ_{m1} в ином виде:

$$\begin{aligned} y = \lambda_{m1} \left(\prod_{i=1}^n \tilde{w}_i \cdot x_i \right) = \lambda_{m1} \left(\prod_{i=1}^n \tilde{w}_i \cdot \prod_{i=1}^n x_i \right) = \\ = \lambda_{m1}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle), \end{aligned}$$

где $\tilde{\mathbf{w}}_{in}, \mathbf{x}_{in}$ — векторы, $\tilde{\mathbf{w}}_{in} = \tilde{w}_1 \times \tilde{w}_2 \times \dots \times \tilde{w}_n$ и $\mathbf{x}_{in} = x_1 \times x_2 \times \dots \times x_n$. Поставив в соответствие значению функции $\lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle)$ соответствующее значение функции λ_{m1} для каждого \mathbf{x}_{in} при $\exists x_i = 0$ получим

$$\begin{aligned} \forall \mathbf{x}_{in}, \exists y_j \left(y_j = \lambda_{m1} \left(\prod_{i=1}^n w_i \cdot x_i + c_i \right) \right) \wedge \\ \wedge (\bigcup_j y_j \subseteq Y) \wedge (\exists x_i = 0) \Rightarrow \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) = \\ = \lambda_{m1} \left(\prod_{i=1}^n w_i \cdot x_i + c_i \right) = y_j, \bigcup_j y_j = Y_j, Y_j \subseteq Y. \end{aligned}$$

Поскольку функция активации не имеет специальных ограничений к своему виду (гладкость, непрерывность и т. д.), можно ее задать в другом виде λ_{m2} . Поставив в соответствие значению функци-

ции $\lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle)$ соответствующее значение функции λ_{m1} для каждого \mathbf{x}_{in} при $\forall x_i \neq 0$ получим

$$y = \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) = \begin{cases} y_j, \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k, \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y \end{cases} \\ (Y_k \cup Y_j) \cap Y = \emptyset,$$

сведем таким образом описание (5) к (3). Сама функция примет вид

$$y = \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) = \lambda_{m2}\left(\prod_{i=1}^n \tilde{w}_i \cdot x_i\right).$$

Приведенное уравнение (1) не совпадает с описанием абстрактного нейрона (3) в силу наличия в нем константы, приведем его к требуемому виду.

Теорема 1. Для любой модели суммирующего искусственного нейрона с определенными на \mathbb{R} множествами входных X , выходных Y и внутренних сигналов, может быть построена эквивалентная ей модель суммирующего нейрона без использования константы смещения.

Доказательство. По условию $n \geq 1$, следовательно

$$y = \lambda_{s1}\left(\sum_{i=1}^n w_i \cdot x_i + c\right) = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right).$$

Поскольку \mathbb{R} непрерывно (аксиома о полноте \mathbb{R}), из формулы (2) для случая $\forall x_i \neq 0$ справедливо

$$\forall x_i \neq 0, \exists \tilde{w}_i \in \mathbb{R} \left(\tilde{w}_i \cdot x_i = w_i \cdot x_i + \frac{c}{n} \right) \Rightarrow \\ \Rightarrow \lambda_{s1}\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right) = y_k, \\ \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y.$$

Задав функцию λ_{s2} , поставим в соответствие значения функции λ_{s1} при $\exists x_i = 0$ и $\forall x_i \neq 0$, аналогично рассмотренному выше способу, получим исходную модель и сведем описание (4) к описанию (3):

$$y = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right) = \lambda_{s2}\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) = \\ = \begin{cases} y_j, \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k, \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y \end{cases} \\ (Y_k \cup Y_j) \cap Y = \emptyset.$$

Теорема доказана.

Докажем эквивалентность мультипликативного и суммирующего нейронов.

Теорема 2. Для любой модели мультипликативного искусственного нейрона с определенными на \mathbb{R} множествами входных X , выходных Y и весо-

вых коэффициентов может быть построена эквивалентная ей модель суммирующего нейрона.

Доказательство. По определению эквивалентности множества входных $X_s = X_m$ и выходных $Y_s = Y_m$ сигналов искусственных нейронов равны (индексы s, m — индексы множеств суммирующего и мультипликативного нейронов соответственно).

Опираясь на рассмотренное выше, уравнения моделей искусственных нейронов можно переписать в следующем виде:

$$y = \lambda_s\left(\sum_{i=1}^n (w_i \cdot x_i)\right), \quad (6)$$

$$y = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right). \quad (7)$$

Из плотности \mathbb{R} при $\forall x_i \neq 0$ выполняется

$$\exists \tilde{w}_i \in \mathbb{R}, \left(\lambda_s\left(\sum_{i=1}^n \tilde{w}_i \cdot x_i\right) = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right) \right).$$

Задав функцию λ_s и поставив в соответствие значения функции активации λ_m при $\exists x_i = 0$ получим исходную модель

$$y = \lambda_s\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right) = \\ = \begin{cases} y_j, \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k, \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y, \\ (Y_k \cup Y_j) \cap Y = \emptyset. \end{cases}$$

Теорема доказана.

Обратное построение эквивалентной модели мультипликативного нейрона для модели суммирующего также возможно ввиду отсутствия ограничений на тип функции и равенства множеств входных и выходных сигналов.

Приведенные уравнения (6) и (7) могут рассматриваться как частные случаи модели абстрактного нейрона (3).

Эквивалентность моделей искусственных нейронов, включающих переменные определенные на \mathbb{R} , не гарантирует их эквивалентности при определении всех их переменных на \mathbb{N} , ввиду отсутствия у \mathbb{N} непрерывности. Учитывая условие ограниченности любых аппаратных реализаций, докажем эквивалентность моделей определенных на конечных множествах элементов из \mathbb{N} .

Теорема 3. Для любой модели мультипликативного искусственного нейрона с определенными на \mathbb{N} конечными множествами входных X , выходных Y и внутренних переменных (W, C), может быть построена эквивалентная ей модель суммирующего нейрона.

Доказательство. Множества $X_s = X_m = X$, $Y_s = Y_m = Y$ равны по определению (s, m — индексы суммирующего и мультиплексивного нейронов). Результатом функции учета веса синапса является конечное множество ввиду конечности множеств W_m , X_m , C_m . Результатом конечной n -мерной операции умножения для учета взвешенных входных сигналов является конечное линейно упорядоченное множество $\text{Dom}\lambda_m$, поскольку все элементы операции принадлежат \mathbb{N} . Для эквивалентности моделей требуется равенство выходных сигналов (символов) обеих моделей в зависимости от одинаковых входных сигналов (символов) на всем множестве возможных входных сигналов.

Поскольку отсутствуют ограничения на значение константы c в модели суммирующего нейрона, модель может быть перезаписана в следующем виде:

$$y = \lambda_s \left(\sum_{i=1}^n w_{is} \cdot x_i + c \right) = \lambda_s \left(\sum_{i=1}^n (w_{is} \cdot x_i + c_{is}) \right),$$

$$c = \sum_{i=1}^n c_{is}.$$

Поскольку множества W_m , C_m конечны, могут быть подобраны такие элементы w_{is} , c_{is} , которые обеспечат линейно упорядоченное множество $\text{Dom}\lambda_s$ как результат n -мерной операции сложения для модели суммирующего нейрона, имеющей ровно тот же порядок и мощность что и множество $\text{Dom}\lambda_m$ в зависимости от входных сигналов X .

Поскольку $|\text{Dom}\lambda_s| = |\text{Dom}\lambda_m|$ и их элементы имеют одинаковый порядок в зависимости от входных сигналов, то можно задать функцию λ_s , поставив в соответствие выходные элементы Y элементам множества $\text{Dom}\lambda_s$ ровно таким же способом, как и функция λ_m . То есть возьмем наименьший элемент $\min(\text{Dom}\lambda_s)$ из $\text{Dom}\lambda_s$ и поставим ему в соответствие элемент из Y , соответствующий элементу $\min(\text{Dom}\lambda_m)$ по функции λ_m . Возьмем следующий за ним наименьший элемент $\min(\text{Dom}\lambda_s \setminus \min(\text{Dom}\lambda_s))$ из $\text{Dom}\lambda_s$ и поставим ему в соответствие элемент из Y , соответствующий элементу $\min(\text{Dom}\lambda_m \setminus \min(\text{Dom}\lambda_m))$ по функции λ_m . Будем повторять указанные действия для каждого последующего наименьшего, пока не поставим в соответствие всем элементам из $\text{Dom}\lambda_s$ все элементы из Y . Полученная модель эквивалентна модели мультиплексивного нейрона. Теорема доказана.

В случае необходимости более сложные модели нейронов, такие как $\Sigma\Pi$ -нейроны или же агрегирующие нейроны, а также многие другие могут быть сведены к абстрактному. Рассмотрение данных нейронов не проводится по причине необходимости использовать большее число ячеек памяти при аппаратной или программной реализации этих моделей (память для хранения дополнительных переменных), что явным образом увеличивает затраты

на реализацию. Введение дополнительных операций (агрегирующие и $\Sigma\Pi$ -нейроны) явно (в общем случае) увеличивает число реализуемых операций для вычисления выходного сигнала, что ставит вопрос о возможности применения данных моделей для широкого круга задач и является темой отдельной статьи.

3. Модель конечного автомата абстрактного нейрона

В данном разделе приводится формализм для описания абстрактного нейрона. Ввиду конечности любой аппаратной реализации описание проводится в терминах теории конечных множеств. Поскольку основной абстракцией при построении вычислительных систем является автомат, формализм строится на основе конечного автомата Мура.

Автомат Мура описывается упорядоченным множеством $A = (S, S_0, X, Y, \delta, \lambda)$, где S — множество внутренних состояний автомата; $S_0 \in S$ — начальное состояние автомата; X и Y — множества входных символов и множество выходных символов соответственно; $\Delta: S \times X \rightarrow S$ — функция перехода в новое состояние; $\lambda: S \rightarrow Y$ — функция вывода символа. Поведение автомата определяется как $s^{t+1} = \Delta(s^t, x^t)$ и $y^t = \lambda(s^t)$.

Введем формальное описание абстрактного нейрона в терминах конечных множеств. Конечный автомат абстрактного нейрона (КААН) может быть описан набором конечных множеств (N , E , W , Q , Δ , Λ , T), где:

N — множество индексов входов (управляющие входы и синапсы), $N \in \mathbb{N}$;

E — алфавит, включающий алфавит всех возможных входных символов E_{input} и алфавит всех возможных выходных символов E_{output} , $E = E_{\text{input}} \cup E_{\text{output}}$, $E \subset \mathbb{N}$;

W — линейно упорядоченное конечное множество всех возможных значений весовых коэффициентов информационных входов (синапсов), $W \subset \mathbb{N}$;

Q — область (множество) определения функции активации, $\text{Dom}\lambda = Q$, $Q \subset \mathbb{N}$;

Δ — операция учета взвешенных входных сигналов $\sum_{i=1}^n (\delta_i(w_i, E_j))$, $i \in N$, зависящая от операции учета весового коэффициента сигнала каждого входа (синапса) $\delta_i(w_i, E_j)$, $E_j \in E$;

Λ — множество функций активации, реализуемых на данном нейроне $|\Lambda| \in \mathbb{N}$;

T — множество функций изменения параметров функций Δ , Λ и $\tau_i \in T$, $|T| \in \mathbb{N}$ искусственного нейрона.

На основе E_{input} формируются подмножества входных информационных сигналов E_{in} и множества управляющих сигналов E_r , такие что $E_{\text{input}} = E_{\text{in}} \cup E_r$. В общем случае при аппаратной реали-

зации не всегда имеется возможность разделить входящие сигналы на управляющие и информационные. С учетом того что сигнал выхода искусственного нейрона может подаваться на его вход или содержать управляющие символы, подаваемые на другие нейроны, имеем следующие соотношения: $E_{input} \cap E_{output} \neq \emptyset$, $E_{input} \cap E_{output} = \emptyset$ и $E_{input} = E_{output}$. Множество входных символов каждого i -го входа может быть описано так: $I_i = \{(E_{in}, E_r) | ((E_{in} = \emptyset) \wedge (E_r \neq \emptyset)) \vee ((E_{in} \neq \emptyset) \wedge (E_r = \emptyset)) \vee ((E_{in} = \emptyset) \wedge (E_r = \emptyset))\}$, $i \in N$. Например, подача сигнала выше определенного порога может не только переводить мемристор в новое состояние проводимости, но и распознаваться как информационный сигнал.

Множества символов на различных входах образуют алфавит входных векторов нейрона $I_{in} = \{(i_j, \dots, i_k) | i_j \in I_1, \dots, i_k \in I_n, \cup j_i = I_1, \dots, \cup k_i = I_n\}$, $n = N$. Алфавит выходных символов имеет вид $O = \{o_i | o_i = (e_{out}, e_r), e_{out} \in E_{out}, e_r \in E_r, E_{out} \cup E_r = E_{output}\}$, где E_{out} — множество символов информационных выходных сигналов. Резюмируя все перечисленное выше: множествам входных и выходных символов X и Y автомата Мура могут быть поставлены в соответствие алфавиты I_{in} и O .

С учетом возможности реализации нескольких функций на одном нейроне, функция активации КААН $\lambda_i \in \Lambda$, $i \in \mathbb{N}$ и определяется зависимостями

$$\lambda_i: Q \rightarrow O, \quad (8)$$

$$o^t = \lambda_i(q^t)^t, \quad (9)$$

что позволяет поставить в соответствие данную функцию функции вывода автомата Мура.

Все возможные комбинации весовых коэффициентов на всех выходах КААН могут быть представлены в виде вектора

$$V = \{v_i | v_i = (w_j, \dots, w_k), w_j \in W_1, \dots, w_k \in W_n, \cup j w_j = W_1, \dots, \cup k w_k = W_n\}, n \in N,$$

где W_n — подмножество весовых коэффициентов каждого i -го синапса.

Внутренние состояния абстрактного нейрона описываются множеством $V \times \Delta \times \Lambda \times T$.

Функция учета взвешенных информационных входных сигналов Δ является заданной на множестве $I_{in} \times V$ и имеет вид

$$\Delta: I_{in} \times V \rightarrow Q, \quad (10)$$

$$q^{t+1} = \sum_{i=1}^n (\delta(v_i^t, i_i^t))^t. \quad (11)$$

Функция изменения параметров КААН $\tau_i \in T$, $i \in \mathbb{N}$ в наиболее общем виде описывается как

$$\tau_i: E_r \rightarrow V \times \Delta \times \Lambda \times T. \quad (12)$$

Динамика изменения параметров имеет вид

$$(v_i^{t+1}, \Delta^{t+1}, \lambda_i^{t+1}, \tau_i^{t+1}) = \tau_j(i_i)^t, \\ \tau_i^{t+1} \in T, \quad \tau_j \in T. \quad (13)$$

Функция перехода КААН в новое состояние является результатом двух независимых друг от друга функций (11) и (12), что может быть записано следующим образом:

$$(q^{t+1}, v_i^{t+1}, \Delta^{t+1}, \lambda_i^{t+1}, \tau_i^{t+1}) = \\ = \left(\sum_{i=1}^n (\delta(v_i^t, i_i^t))^t, \tau_j(i_i)^t \right), \quad (14)$$

Соответствие между автоматами Мура и КААН
Conformity between Moore automaton and FAAN

Объект соответствия <i>Object of conformity</i>	Автомат Мура $A = (S, S_0, X, Y, \delta, \lambda)$ <i>Moore automaton</i> $A = (S, S_0, X, Y, \delta, \lambda)$	КААН $(N, E, W, Q, \Delta, \Lambda, T)$ <i>FAAN</i> $(N, E, W, Q, \Delta, \Lambda, T)$
Множество внутренних состояний <i>Set of internal states</i>	S	$V \times \Delta \times \Lambda \times T$
Алфавит входных символов <i>Alphabet of the input symbols</i>	X	I_{in}
Алфавит выходных символов <i>Alphabet of the output symbols</i>	Y	O
Начальное состояние <i>Initial state</i>	S_0	$\Delta_{i=1}^n (\delta(v_i^{t=0}, \emptyset))^{t=0}$ $\tau_j(\emptyset)^{t=0}$
Функция переходов в новое состояние <i>Function of transition into a new state</i>	$\delta: S \times X \rightarrow S$	$\Delta: I_{in} \times V \rightarrow Q$ $\tau_i: E_r \rightarrow V \times \Delta \times \Lambda \times T$
Функция вывода символов <i>Function of the output of symbols</i>	$\lambda: S \rightarrow Y$	$\lambda_i: Q \rightarrow O$
Динамика переходов <i>Dynamics of transitions</i>	$s^{t+1} = \delta(s^t, x^t)$	$q^{t+1} = \sum_{i=1}^n (\delta(v_i^t, i_i^t))^t$ $(v_i^{t+1}, \Delta^{t+1}, \lambda_i^{t+1}, \tau_i^{t+1}) = \tau_j(i_i)^t$
Динамика вывода символов <i>Dynamics of the output of symbols</i>	$y^t = \lambda(s^t)$	$o^t = \lambda_i(q^t)^t$

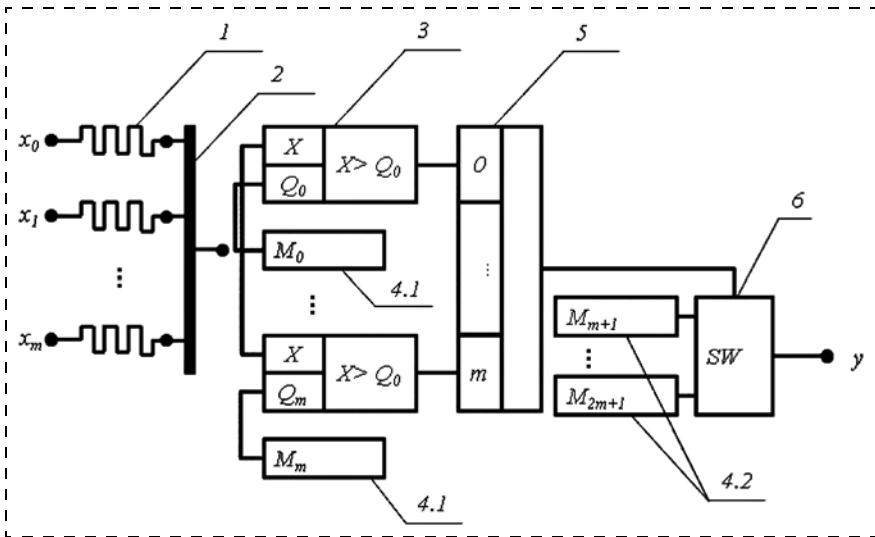


Рис. 1. Минимальная реализация КААН: 1 — мемристор; 2 — аналоговый сумматор токов; 3 — схема сравнения со значениями q ; 4.1 — блок регистров для хранения q ; 4.2 — блок регистров для хранения y ; 5 — схема определения наибольшего q ; 6 — функция λ

Fig. 1. Minimal realization of FAAN: 1 — memristor; 2 — analog adder of currents; 3 — scheme of comparison with the values of q ; 4.1 — unit of registers for storage of q ; 4.2 — unit of registers for storage of y ; scheme of definition of the highest relevant q ; 6 — function λ

что позволяет поставить в соответствие функцию перехода автомата Мура.

Начальное состояние КААН выражается через (11) и (13) с пустым множеством сигналов на входе:

$$\left(\bigcup_{i=1}^n (\delta(v_i^{t=0}, \emptyset)^{t=0})^{t=0}, \tau_j(\emptyset)^{t=0} \right), \text{ где } t = 0 \text{ — значения в начальный момент времени.}$$

Исходя из всего перечисленного выше автомата Мура может быть поставлен в соответствие КААН (см. таблицу).

Минимальная реализация КААН показана на рис. 1.

4. Свойства синапсов КААН

Возможны следующие ситуации при задании множества E_{in} и его подмножеств значений входных

сигналов i -го синапса $E_{in} = \bigcup_{i=1}^n E_i$, $E_i = \{e_i | e_i \in E_{in}\}$, n — число связей, для которых справедливо $w_i \neq 0$ при $\delta(w_i, e_i) = w_i \cdot e_i$.

Для информационных сигналов i -х входов (синапсов) возможны различные ситуации при их задании. Элементы различных множеств E_{in} равны между собой

$$E_i \cup E_j = E_{in}, \forall e_i \in E_i, \forall e_j \in E_j ((e_i > e_j) \vee (e_i < e_j)) \Rightarrow \forall e_i \in E_i, \forall e_j \in E_j (e_i = e_j).$$

Например, на 1-й вход подается информация о цвете (красный, зеленый и т. д.), на 2-й вход подается информация о форме (шар, параллелепипед и т. д.).

Элементы не всех подмножеств сравнимы между собой: $E_i \cup E_j = E_{in}, \exists e_i \in E_i, e_j \in E_j ((e_i \not> e_j) \vee (e_i \neq e_j) \vee (e_i \not< e_j))$. На один из входов подаются сигналы, неравнимые с другими.

На элементах подмножеств может быть задан нестрогий порядок $\forall e_i \in E_{in}, \exists e_j \in E_{in} ((e_i > e_j) \vee (e_i < e_j)) \Rightarrow (E_{in} \cap (U_j e_j) = \emptyset)$. Все подаваемые на входы сигналы сравнимы друг с другом. Выход хотя бы одного нейрона (входной нейрон) связан с входом хотя бы еще одного нейрона (для минимальной сети необходимо более одного нейрона). В общем случае множество выходных информационных сигналов нейрона E_{out} может иметь мощность, отличную от мощности входных сигналов последующего нейрона E_{in} , например, диапазон напряжений на выходе 1...3 В первого нейрона подается на вход второго нейрона, принимающий сигналы в диапазоне 2...3 В. Тогда очевидно, что входные сигналы в диапазоне [1, 2) не должны рассматриваться как информационные.

Данный случай формально описывается применением характеристической функции к множеству входных информационных сигналов каждого синапса: $\chi E_i(e_i)$, $e_i \in E_i$, $\bigcup_{i=1}^n E_i = E_{in}$, где значения функции $\chi E_i(e_i) = 0$ не учитываются на данном нейроне.

На множествах информационных сигналов синапсов $\chi E_i(e_i) \neq 0$, учитываемых принимающим их

нейроном $\bigcup_{i=1}^n E_i = E_{in}$, $|E_i| > 1$, может быть задан линейный порядок. В случае если порядок не задан, все сигналы (или их часть, в случае задания нестрогого порядка) будут неразличимы между собой.

Резюмируя все описанное выше, можно заключить следующее — входные информационные сигналы на разных синапсах, с одной стороны, могут быть несравнимы между собой, с другой стороны, множество сигналов каждого синапса описывается характеристической функцией. Результатом чего является определение оптимальной мощности множеств информационных сигналов для общего случая:

$$\bigcup_{i=1}^n E_i = E_{in}, |E_i| = |E_{out}| = 1.$$

Отсутствие связи может быть задано через значение весовых коэффициентов синапсов $w_i = 0$ при $\delta_i(w_i, i_i) = w_i \cdot i_i$. Поскольку множество весовых ко-

эффективов W линейно упорядочено и для всех несуществующих связей $w_i = 0$, то элементы множества принимают следующие значения:

$$\chi W_i(\emptyset) = 0 \text{ для отсутствующих связей,}$$

$$\min(W_i) = k, \text{ где } k \in \mathbb{N} \text{ некоторая константа.}$$

$$\min(W_i \setminus \min(W_i)) = k_w + h, \text{ где } h \geq k, h \in \mathbb{N} \text{ некоторая константа и } W_i = \{w_j | w_{j+1} - w_j = h\}.$$

Традиционным подходом к реализации учета веса входных информационных сигналов каждого i -го входа является умножение на его весовой коэффициент

$$\delta : w_i \cdot e_i \rightarrow \mathbb{N}, e_i \in E_i, \bigcup_{i=1}^n E_i = E_{in}, w_i \in W_i,$$

$$\bigcup_{i=1}^n W_i = W.$$

При соотношении элементов множества $e_i \in E_{in}$, $w_j \in W$, задаваемых константой, одинаковой для обоих множеств, максимальная возможная мощность множеств в предельном случае определяется следующим образом:

$$\forall e_{i+1} \in E_{in}, \forall e_i \in E_{in}, \forall w_{j+1} \in W,$$

$$\forall w_j \in W((e_{i+1} - e_i = h) \wedge (w_{i+1} - w_i = h)) \Rightarrow$$

$$\Rightarrow |\text{Dom}\Delta| = |E_{in}| \cdot |W|.$$

Увеличение мощности множества области определения $\text{Dom}\Delta$ КААН при сохранении мощности множеств E_{in} , W возможно за счет задания нелинейной зависимости между элементами $w_j \in W$. Первый из возможных способов поставить в соответствие элементам множества W элементы линейно упорядоченного множества с нелинейной зависимостью: $W = \{\emptyset = 0, w_1 = 1, w_2 = n + 1, w_3 = n \cdot w_2 + 1, \dots, w_j = n \cdot w_{j-1} + 1\}$, где n — число связей, для которых справедливо $w_j \neq 0$. Указанный метод максимально увеличивает мощность области определения $\text{Dom}\Delta$ КААН при аппаратной реализации. К недостаткам подхода можно отнести полное отсутствие возможности задания линейной зависимости входными сигналами на синапсах, что ограничивает его применение при аппаратных реализациях, предназначенных для построения сетей на основе КААН.

Второй метод может быть определен как внесение нелинейности между подмножествами множества W , он имеет следующую схему задания:

$$W = \{\emptyset = 0, w_1 = 1, w_2 = 2 \cdot w_1, \dots,$$

$$w_i = i \cdot w_1, w_{i+1} = n \cdot w_i + 1, w_{i+2} = 2 \cdot w_{i+1}, \dots,$$

$$w_{i+j} = j \cdot w_{i+1}, \dots, w_{k+1} = n \cdot w_k + 1, \dots,$$

$$w_{k+l} = l \cdot w_{k+1}, \bigcup_{n=1}^i w_i = W_i, \bigcup_{j=1}^l w_j = W_j, \dots,$$

$$\bigcup_{k+1}^l w_l = W_b, \bigcup_p W_p = W,$$

где n — число связей, для которых справедливо $w_i \neq 0$.

Подход является компромиссом между возможностью задавать линейные зависимости при учете входных сигналов на синапсах и увеличением мощности множества $\text{Dom}\Delta$ КААН. Нелинейная зависимость между входными сигналами E_{in} также приводит к увеличению множества $\text{Dom}\Delta$ и может быть описана одним из двух вариантов, рассмотренных выше для множества W .

5. Формализм искусственных нейронных сетей на основе КААН

Формализм описания сети на основе КААН определяется кортежем $(E, W, Q, \Delta, \Lambda, T)$ без использования N . Покажем это. Используемые понятия входных нейронов могут быть распространены не только на информационные входы (синапсы), но и на управляющие входы. Используя данное условие, получим единое описание управляющих и информационных входов, на случай аппаратной реализации гарвардской архитектуры построения вычислительных систем. Определив, что выходной сигнал каждого КААН подается на вход всех остальных КААН сети, получим полно связную сеть. Задав для несуществующих связей значение весовых коэффициентов синапсов $w_i \neq 0$, получим описание искусственной нейронной сети (ИНС) без использования множества N .

Классы реализуемых на КААН функций напрямую зависят от мощности множества Q . Множество Q является линейно упорядоченным множеством $\text{Dom}\Delta$, с одной стороны, и областью значений Δ , с другой стороны. Для случая $|E_{in}| = 1$ возможны следующие ситуации. Поскольку $\text{Dom}\Delta$ линейно упорядочено и $\forall a \in \text{Dom}\Delta (\Delta(a_{i+1}) > \Delta(a_i))$, то при $|Q| = 1$ возможна реализация только пороговой функции.

При $|Q| = 2$ возможны реализации пороговой, линейной и RBF-функций. Задание пороговой функции возможно четырьмя способами, первые два:

$$\forall \lambda \in \Lambda((\lambda(\max Q) = 1) \wedge (\lambda(q_i) = 0, q_i \neq \max Q)),$$

$$\forall \lambda \in \Lambda((\lambda(\min Q) = 0) \wedge (\lambda(q_i) = 1, q_i \neq \min Q)).$$

Инвертирование значений функции λ с 0 на 1 дает два других способа.

Линейная функция задается следующим образом:

$$\forall q \in Q, \forall \lambda \in \Lambda(((\lambda(q_{i+1}) > \lambda(q_i)) \wedge (q_{i+1} > q_i)) \vee ((\lambda(q_{i+1}) < \lambda(q_i)) \wedge (q_{i+1} > q_i))).$$

RBF-функция задается следующим образом:

$$\forall \lambda \in \Lambda(((\lambda(q_1) > \lambda(q_2)) \wedge (\lambda(q_1) > \lambda(\emptyset))) \vee ((\lambda(q_1) < \lambda(q_2)) \wedge (\lambda(q_1) < \lambda(\emptyset)))).$$

Минимальная необходимая мощность для задания периодических функций $|Q| = 3$.

Аппаратная реализация задания периодических функций может быть реализована путем модификации функции учета взвешенных входных сигналов с применением операции взятия по модулю над суммой или произведением взвешенных входных сигналов $\Delta_i(\delta(w_i, i_j)) = \text{mod}_d(\sum_i \delta(w_i, i_j))$, где d — размер периода на $\text{Dom}\Delta$. Результатом применения модифицированной функции будет значение от начала текущего периода. Путем задания соответствующих значений E_{out} функция активации может принимать вид RBF-, линейной или пороговой функции на текущем периоде.

В предельном случае $|\text{Dom}\Delta| = |Q|$, что позволяет задавать не только линейные, пороговые, нелинейные, RBF- или периодические функции, но и почти периодические функции. Включение в модель КААН множества функций изменения параметров T позволяет учитывать в алгоритмах обучения сети КААН изменения функции активации в процессе обучения сети. Например, осуществлять переход от пороговых функций к линейным функциям, переход от линейных функций к нелинейным (сигмоидальная функция, заданная на дискретном множестве) и RBF-функциям; обеспечивать переход от RBF- и нелинейных функций к периодическим, а также от периодических функций к почти периодическим функциям во время тренировки ИНС.

Имеющееся описание КААН применимо для описания работы синхронных искусственных нейронных сетей. Сетей, в которых все нейроны срабатывают (либо не срабатывают) в один момент времени, после чего данные переменных обнуляются и начинается новый цикл работы сети. Попытка асинхронного описания работы одиночного КААН невозможна ввиду отсутствия объекта сравнения, что позволяет сделать вывод о ее практической бесполезности. Аппаратная реализация предложенной модели КААН (гибридное схемотехническое решение) в минимальном составе требует для реализации набора из двух блоков регистров (рис. 2).

В отличие от единичных КААН сети на их основе могут иметь асинхронный механизм работы. Обобщение сетей КААН на случай асинхронной работы сети требует внесения изменений в выражения (9), (11). Вносимые изменения должны учи-

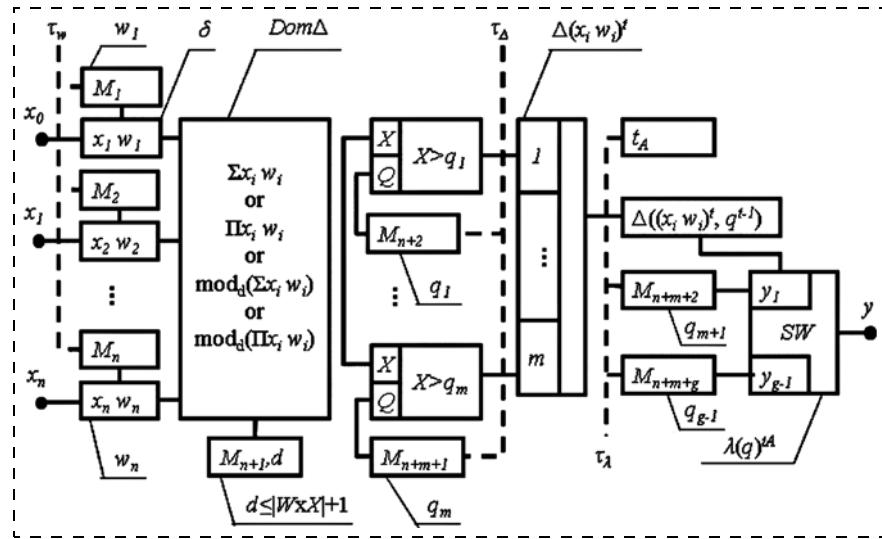


Рис. 2. Обобщенная схема КААН: M_x — блок хранения данных; δ — блок формирования взвешенного входного сигнала Δ ; $\text{Dom}\Delta$ — блок учета всех взвешенных входных сигналов; $X > q_x$ — блок сравнения со значениями Q ; $\Delta(x_i, w_i)$ — блок формирования q от текущих значений на синапсах; $\Delta((x_i, w_i)^t, q^{t-1})$ — блок учета предыдущих q ; SW — блок вывода выходного сигнала; t_A — блок таймера

Fig. 2. Generalized scheme of FAAN: M_x — data storage unit; δ — unit for formation of the weighted input signal Δ ; $\text{Dom}\Delta$ — unit of account of all the weighted input signals; $X > q_x$ — unit of comparison with the values of Q ; $\Delta(x_i, w_i)$ — unit of formation of q from the current values on the synapses; $\Delta((x_i, w_i)^t, q^{t-1})$ — unit of account of the previous q ; SW — unit of output of the output signal; t_A — timer unit

тывать возможность настройки времени срабатывания функции активации КААН t_A и/или возможность учета предыдущего состояния функцией Δ . Уравнение изменения состояний КААН, описываемых уравнением (11), и уравнение динамики вывода символов (9) для асинхронных сетей КААН принимают следующий вид:

$$o^{t_A} = \lambda_i(q^{t_A}); \quad (15)$$

$$q^{t+1} = \prod_{i=1}^n (\delta(v_i^t, i_i^t)^t, q^t)^t. \quad (16)$$

Поблочный состав аппаратной реализации КААН для асинхронных сетей приведен на рис. 2.

Заключение

Обобщенная модель КААН позволяет моделировать широкий набор искусственных нейронов с различными функциями активации, что при аппаратной реализации КААН гарантирует относительную универсальность технического решения и широкий набор возможностей при проектировании сети.

Построение модели на основе конечных множеств позволяет просчитать необходимое число нейронов (для известного алгоритма работы сети), исходя из данных о множествах, на основе которых реализуются сети КААН, либо провести оптимизацию имеющихся сетей.

На основании предложенных моделей могут быть построены сети, реализующие помимо распознавания изображений математические операции сложения, вычитания, умножения и др. Реализация математических операций позволяет применить часть ресурсов сети (реализовав на подсети алгоритм обучения) для обучения остальной части сети без применения классических (неймановских и гарвардских) архитектур.

Предлагаемая модель КААН позволяет учесть в алгоритмах обучения изменение активационной функции КААН и построить алгоритм, автоматически генерирующий сеть с оптимальными активационными функциями нейронов. Помимо генерации сетей КААН (путем применения алгоритмов обучения) появляется возможность задания алгоритма работы сети с динамически изменяющимися функциями активации.

Автор статьи выражает благодарность и глубокую признательность сотруднику отдела функциональной электроники АО НИИМЭ Игорю Валерьевичу Матюшкину и сотруднику ИППМ РАН Дмитрию Владимировичу Тельпухову за ценные замечания и обсуждение материала статьи.

Работа выполнена при содействии РФФИ, проект № 17-07-00570 A.

Список литературы

1. Strukov D. B., Snider G. S., Stewart D. R., Williams R. S. The missing memristor found // Nature LETTERS. 2008. Vol. 453. P. 80–83. DOI: 10.1038/nature06932.
2. Merolla P. A., Arthur J. V., Alvarez-Icaza R. et al. A million spiking-neuron integrated circuit with a scalable communication network and interface // Science. 2014. Vol. 354, Issue 6197. P. 668–673. DOI: 10.1126/science/1254642.
3. Degans H. Imec demonstrates self-learning neuromorphic chip that composes music. URL: <http://www.imec-int.com>
4. McCulloch W. S., Pitts W. A. A logical Calculus of Ideas Immanent in Nervous Activity // Bull. Mathematical Biophysics. 1943. Vol. 5, Issue 4. P. 115–133.
5. Koch C., Poggio T., McKenna T., Davis J., Zornetzer S. F. (Eds.) Multiplying with Synapses and Neurons in Single Neuron Computations. Academic Press Inc., San Diego, CA, USA, 1992. Chapter 12. P. 315–319.
6. Kandpal P. K., Mehta A. Comparison and Analysis of Multiplicative neuron and Multilayer Perceptrons using three different datasets // International Journal of Engineering Research and General Science. 2017. Vol. 5, Issue 3. P. 58–70.
7. Шибзухов З. М., Чередников Д. Ю. О моделях нейронов агрегирующего типа // Машинное обучение и анализ данных. 2015, Т. 1. С. 1706–1716. DOI: 10.21469/22233792.1.12.06.
8. Ходашинский И. А., Малютин Н. Д. Модели статических искусственных нейронов. Обзор // Доклады ТУСУРа. 2015. № 2 (36). С. 100–107.
9. Moore E. P. Gedanken-experiments on sequential machine // Automata Studies, Annals of Mathematical Studies. 1956. Vol. 34. P. 129–153.

E. S. Gornev^{1,2}, D. Sc., Professor, egornev@niime.ru,

G. S. Teplov^{1,2}, Postgraduate Student, Junior Researcher, gteplov@niime.ru,

¹ Molecular Electronics Research Institute, Zelenograd, Moscow, 124460, Russian Federation,

² Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141701, Russian Federation,

Corresponding author:

Teplov Georgiy S., Postgraduate Student, Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141701, Russian Federation, e-mail: gteplov@niime.ru

Mathematical Model of a Finite Automaton of an Abstract Neuron and the Networks Based on It

Received on March 12, 2018

Accepted on March 20, 2018

On the basis of the mathematical models of the multiplicative and summing artificial neurons, a set-theoretic model of a finite automaton of an abstract neuron is proposed, formulated in terms of the finite sets. Unlike the known models of the artificial neurons, the proposed approach takes into account a possibility of variation of the activation function in the process of the neuron work or training. Design of the model based on the Moore automaton allows us to use it in the hardware realization of most models of the artificial neurons.

Keywords: mathematical model of an abstract neuron, activation function of an artificial neuron, finite-state automaton of an abstract neuron, discrete model of an artificial neural network

For citation:

Gornev E. S., Teplov G. S. Mathematical Model of a Finite Automaton of an Abstract Neuron and the Networks Based on It, *Nano- i Mikrosistemnaya Tekhnika*, 2018, vol. 20, no. 7, pp. 434–448.

DOI: 10.17587/nmst.20.434-448

Introduction

Development of micro- and nanoelectronics expressed in widening of the nomenclature of the electronic component base [1], minimization of the topological sizes and development of the circuit solutions on their basis again provokes an interest to the hardware realizations of the neural networks [2, 3].

By the method of account of the input signals the existing models of the artificial neurons can be divided into two basic types. The first type of the neurons, proposed for the first time in [4], is the type of the summing neurons, which use operation of addition for the account of the weighted input signals.

The second type of the neurons is the type of the multiplicative neurons [5], which use operation of multiplication for the account of the weighted input signals. An example of one of the kinds of the given class of models is presented in [6]. In this work both mathematical models were generalized into a single model, called an abstract neuron and intended for a hardware realization.

More complex models, such as $\Sigma\Pi$ neurons, aggregating neurons [7] and the neurons with complex synapses can be reduced to the multiplicative and summing models of neurons. A more detailed review of the existing models is presented in [8].

All the basic approaches to the realization of the computer hardware, regardless of the circuitry solutions, can be analyzed with application of the discrete finite sets. There are no doubts about the discreteness of the digital circuitry.

The discreteness of the hybrid and analogue circuitry is defined by the signal-to-noise ratio, as a result of which all the signals below a certain threshold are indiscernible between themselves, which in turn, leads to a digitization of the set of their values. Finiteness of the set of the processed signals is determined by the limits of any hardware realization of the computing system.

Thus, a description of the hardware realization of the abstract neuron as a computing element is convenient to implement in terms of the finite automata. Taking into account the mathematical description of the considered models, namely, the dependence of the result of the activation function on the result of operation of the account of the weighed signals (addition or multiplication), the most suitable is the use of Moore automaton [9].

1. Analysis of the summing and multiplicative neurons

An artificial neuron is described by the following kind of a set $(n, Y, X, W, C, \delta, \Delta, \lambda)$, where:

$n \in \mathbb{N}$ — number of inputs (synapses);

$Y \subseteq \mathbb{R}$ — a set of the output signals of a neuron;

$X = \bigcup_{i=0}^n X_i, X \subseteq \mathbb{R}$ — a set of the input signals sent to the synapses of the artificial neuron;

$W = \bigcup_{i=0}^n W_i, W \subseteq \mathbb{R}$ — a set of all possible values of the weight factors of the synapse;

$C \subseteq \mathbb{R}$ — a set of the adjustable factors of the activation function;

δ — operation of the account of the weight factor of the input signal (multiplication operation is traditionally applied $\delta(w_i, x_i) = w_i \cdot x_i; w_i \in W; x_i \in X$);

Δ — operation of the account of input signals $\sum_{i=1}^n (w_i \cdot x_i)$
or $\prod_{i=1}^n (w_i \cdot x_i + c_i), c_i \in C; w_i \in W, x_i \in X$;
 λ — function of activation of the artificial neuron.
Here in after $\mathbb{N} \cap 0 = \emptyset$.

2. Generalization of the models of the summing and multiplicative neurons

The dependence of the signal on the output for the summing neuron looks like the following:

$$y = \lambda \left(\sum_{i=1}^n w_i \cdot x_i + c \right), \quad c_i \in C, \quad (1)$$

while for the multiplicative neuron it looks like the following:

$$y = \lambda \left(\prod_{i=1}^n (w_i \cdot x_i + c_i) \right), \quad c_i \in C. \quad (2)$$

Taking into account what was introduced in section 1, we have the following descriptions for the abstract, summing and multiplicative neurons, accordingly:

$$\begin{aligned} y &= \lambda_a \left(\sum_{i=1}^n (\delta(w_i, x_i)) \right), \\ \delta : \mathbb{R}^2 &\rightarrow \mathbb{R}, \Delta : \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_a : \mathbb{R} \rightarrow Y; \end{aligned} \quad (3)$$

$$\begin{aligned} y &= \lambda_s \left(\sum_{i=1}^n w_i \cdot x_i + c \right), \\ \delta : \mathbb{R}^2 &\rightarrow \mathbb{R}, \Sigma : \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_s : \mathbb{R}^2 \rightarrow Y; \end{aligned} \quad (4)$$

$$\begin{aligned} y &= \lambda_m \left(\prod_{i=1}^n (w_i \cdot x_i + c_i) \right), \\ \delta : \mathbb{R}^3 &\rightarrow \mathbb{R}, \Pi : \mathbb{R}^n \rightarrow \mathbb{R}, \lambda_m : \mathbb{R} \rightarrow Y. \end{aligned} \quad (5)$$

The first condition of the equivalence of automata M_1, M_2 is equality of the alphabets of the input $X_1 = X_2$ and output $Y_1 = Y_2$ symbols. The second condition of the equivalence is obtaining of the identical output signals $y_1 = y_2, y_1 \in Y_1, y_2 \in Y_2$ when the identical input signals $x_1 = x_2, x_1 \in X_1, x_2 \in X_2$ are set at the automaton input for all the elements of the set of the input signals.

Similarly, we will consider as equivalent the models of the artificial neurons for which the following conditions are met: the sets of the input symbols (signals, elements) are equal; the sets of the output symbols (signals, elements) are equal; when any input symbol is submitted to two equivalent neurons, their output symbols are equal.

The presented equation (2) does not coincide with the description of the abstract neuron (3) due to the presence of constants in it, we will reduce it to the demanded form.

Since \mathbb{R} is continuous (axiom of completeness of \mathbb{R}), from formula (2) for the case of $\forall x_i \neq 0$ the following is true:

$$\begin{aligned} \forall x_i \neq 0, \exists \tilde{w}_i \in \mathbb{R} (\tilde{w}_i \cdot x_i = w_i \cdot x_i + c_i) \Rightarrow \prod_{i=1}^n \tilde{w}_i \cdot x_i = \\ = \prod_{i=1}^n (w_i \cdot x_i + c_i) = y_k, \forall y_k \in Y_k, Y_k \subseteq Y. \end{aligned}$$

The conditions of equality of the functions f_1, f_2 are:

$$\text{Dom}f_1 = \text{Dom}f_2, \quad \text{ran}f_1 = \text{ran}f_2,$$

$$\forall x_i \in \text{Dom}f_1 ((f_1(x_i) = f_2(x_i)) \wedge ((\bigcup_i x_i) \cap \text{Dom}f_1 = \emptyset)).$$

Then for implementation of the conditions of equality of the activation functions at $\exists x_i = 0$

$$\lambda_{m1}\left(\prod_{i=1}^n \tilde{w}_i \cdot x_i\right) = \lambda_{m1}\left(\prod_{i=1}^n w_i \cdot x_i + c_i\right),$$

we will rewrite the functions λ_{m1} in a different way:

$$y = \lambda_{m1}\left(\prod_{i=1}^n \tilde{w}_i \cdot x_i\right) = \lambda_{m1}\left(\prod_{i=1}^n \tilde{w}_i \cdot \prod_{i=1}^n x_i\right) =$$

$$= \lambda_{m1}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle),$$

where $\tilde{\mathbf{w}}_{in}, \mathbf{x}_{in}$ — vectors; $\tilde{\mathbf{w}}_{in} = \tilde{W}_1 \times \tilde{W}_2 \times \dots \times \tilde{W}_n$ and $\mathbf{x}_{in} = X_1 \times X_2 \times \dots \times X_n$.

By setting the corresponding value of function λ_{m1} for each \mathbf{x}_{in} at $\exists x_i = 0$ in conformity with function $\lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle)$ we will get:

$$\forall \mathbf{x}_{in}, \exists y_j \left(y_j = \lambda_{m1}\left(\prod_{i=1}^n w_i \cdot x_i + c_i\right) \right) \wedge$$

$$\wedge (\bigcup_j y_j \subseteq Y) \wedge (\exists x_i = 0) \Rightarrow \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) =$$

$$= \lambda_{m1}\left(\prod_{i=1}^n w_i \cdot x_i + c_i\right) = y_j, \bigcup_j y_j = Y_j, Y_j \subseteq Y.$$

Since the activation function has no special restrictions to its form (smoothness, continuity, etc.), it is possible to set it in a different way λ_{m2} .

By setting the corresponding value of function λ_{m1} for each \mathbf{x}_{in} at $\forall x_i \neq 0 \quad \forall x_i \neq 0$ in conformity with function $\lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle)$ we will get:

$$y = \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) = \begin{cases} y_j \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y \end{cases},$$

$$(Y_k \cup Y_j) \cap Y = \emptyset,$$

thus we will reduce description (5) to (3). The function itself will look like the following:

$$y = \lambda_{m2}(\langle \tilde{\mathbf{w}}_{in}, \mathbf{x}_{in} \rangle) = \lambda_{m2}\left(\prod_{i=1}^n \tilde{w}_i \cdot x_i\right).$$

The presented equation (1) does not coincide with the description of the abstract neuron (3) because of the presence of a constant in it, we will reduce it to the required form.

Theorem 1. For any model of the summing artificial neuron with the sets of the input X , output Y and inner signals defined on \mathbb{R} , a model of the summing neuron, equivalent to it, can be constructed without the use of the displacement constant.

Proving. According to condition $n \geq 1$, hence:

$$y = \lambda_{s1}\left(\sum_{i=1}^n w_i \cdot x_i + c\right) = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right).$$

Since \mathbb{R} is continuous (axiom of completeness of \mathbb{R}), from formula (2) for $\forall x_i \neq 0$ the following is fair:

$$\forall x_i \neq 0, \exists \tilde{w}_i \in \mathbb{R} \left(\tilde{w}_i \cdot x_i = w_i \cdot x_i + \frac{c}{n} \right) \Rightarrow$$

$$\Rightarrow \lambda_{s2}\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right) = y_k,$$

$$\forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y.$$

Setting function of λ_{s2} we will bring in conformity the values of function λ_{s1} at $\exists x_i = 0$ and $\forall x_i \neq 0$, similarly to the above method, we will receive the necessary model and reduce description (4) to description (3):

$$y = \lambda_{s1}\left(\sum_{i=1}^n \left(w_i \cdot x_i + \frac{c}{n}\right)\right) = \lambda_{s2}\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) =$$

$$= \begin{cases} y_j, \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k, \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y, \end{cases}$$

$$(Y_k \cup Y_j) \cap Y = \emptyset.$$

The theorem is proved.

Let us prove equivalence of the multiplicative and summing neurons.

Theorem 2. For any model of the multiplicative artificial neuron with a set of input X , output Y and weight factors defined on \mathbb{R} , a model of the summing neuron, equivalent to it, can be constructed.

Proving. According to the definition of equivalency, the sets of the input $X_s = X_m$ and output $Y_s = Y_m$ signals of the artificial neurons are equal (indexes s, m are indexes of the sets of the summing and multiplicative neurons, accordingly).

Leaning on what was considered above, the equations of the models of artificial neurons can be rewritten in the following way:

$$y = \lambda_s\left(\sum_{i=1}^n (w_i \cdot x_i)\right), \quad (6)$$

$$y = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right). \quad (7)$$

From density \mathbb{R} at $\forall x_i \neq 0$ the following is done:

$$\exists \tilde{w}_i \in \mathbb{R}, \left(\lambda_s\left(\sum_{i=1}^n \tilde{w}_i \cdot x_i\right) = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right) \right).$$

By setting function λ_s and brining the values of the function at $\exists x_i = 0$ in conformity with the activation function λ_m at $\exists x_i = 0$ we get the sought-for model.

$$y = \lambda_s\left(\sum_{i=1}^n (\tilde{w}_i \cdot x_i)\right) = \lambda_m\left(\prod_{i=1}^n w_i \cdot x_i\right) =$$

$$= \begin{cases} y_j, \exists x_i = 0, \forall y_j \in Y_j, Y_j \subseteq Y \\ y_k, \forall x_i \neq 0, \forall y_k \in Y_k, Y_k \subseteq Y, \end{cases}$$

$$(Y_k \cup Y_j) \cap Y = \emptyset.$$

The theorem is proved.

A reverse construction of an equivalent model of the multiplicative neuron for the model of the summing one is also possible, in view of absence of the restrictions concerning the types of the function and equality of the sets of the input and output signals.

The presented equations (6) and (7) can be considered as special cases of the model of the abstract neuron (3).

Equivalence of the models of the artificial neurons, including the variables defined on \mathbb{R} , do not guarantee their equivalence during definition of all the variables on \mathbb{N} , because \mathbb{N} has no continuity.

Considering the condition of finiteness of any hardware realizations, we will prove the equivalence of the models of the elements defined on the finite sets from \mathbb{N} .

Theorem 3. For any model of the multiplicative artificial neuron with finite sets of input X , output Y and internal variables (W , C), defined on \mathbb{N} , a model of the summing neuron, equivalent to it, can be constructed.

Proving. Sets of $X_s = X_m = X$, $Y_s = Y_m = Y$ are equal by definition (s, m are indexes of the summing and multiplicative neurons). A result of the function of account of the weight of synapse is a finite set because of the finiteness of sets of W_m , X_m , C_m .

A result of the finite n-D operation of multiplication for account of the weighted input signals is a finite linear ordered set $\text{Dom}\lambda_m$, because all the elements of the operation belong to \mathbb{N} .

For equivalence of the models, an equality of the output signals (symbols) of both models depending on the identical input signals (symbols) on all the set of the possible input signals, is required.

Since there are no restrictions on the value of constant c in the model of the summing neuron, the model can be re-written in the following way:

$$y = \lambda_s \left(\sum_{i=1}^n w_{is} \cdot x_i + c \right) = \lambda_s \left(\sum_{i=1}^n (w_{is} \cdot x_i + c_{is}) \right),$$

$$c = \sum_{i=1}^n c_{is}.$$

Since the sets of W_m , C_m are finite, we can select such elements of w_{is} , c_{is} , which will ensure a linear ordered set of $\text{Dom}\lambda_s$ as a result of n-D operation of addition for the model of the summing neuron, having the same order and power as $\text{Dom}\lambda_m$ set depending on the input signals of X .

Since $|\text{Dom}\lambda_s| = |\text{Dom}\lambda_m|$ and their elements have the same order, depending on the input signals, we can set function λ_s (comparing) the output elements Y and elements of $\text{Dom}\lambda_s$ set, exactly in the same way as function λ_m . That is, let us take the smallest element $\min(\text{Dom}\lambda_s)$ from $\text{Dom}\lambda_s$, and put in conformity with it the element from Y corresponding to element $\min(\text{Dom}\lambda_m)$ by function λ_m .

Let us take the next smallest element $\min(\text{Dom}\lambda_s \setminus \min(\text{Dom}\lambda_s))$ from $\text{Dom}\lambda_s$ and put in conformity with it the element from Y , corresponding to element $\min(\text{Dom}\lambda_m \setminus \min(\text{Dom}\lambda_m))$ by function λ_m .

We will repeat the above actions for each next smallest element, till we put all the elements from Y in conformity with all the elements from $\text{Dom}\lambda_s$. The received model will be equivalent to the model of the multiplicative neuron. The theorem is proved.

If necessary, more complex models of neurons, such as $\Sigma\Pi$ -neurons or aggregating neurons, and many other ones, can be reduced to the abstract one.

We do not consider the given neurons, because of the necessity to use a bigger number of the memory cells in case of a hardware or program realization of these models (memory for storage of the additional variables), which evidently increases the costs of their realization.

Introduction of additional operations (aggregating and $\Sigma\Pi$ -neurons) obviously (in a general case) increases the number of the realized operations for calculation of the output signal, which sets the question of a possibility of application of the given models for a wide range of problems and is a topic of a separate article.

3. Model of the finite automaton of an abstract neuron

The given section presents formalism for description of the abstract neuron. In view of the finiteness of any hardware realization, the description is presented in terms of the theory of the finite sets. Since the basic abstraction during construction of the computing systems is the automaton, the formalism is constructed on the basis of Moore finite automaton.

Moore automaton is described by the ordered set of $A = (S, S_0, X, Y, \delta, \lambda)$, where S — set of the internal states of the automaton, $S_0 \in S$ — initial state of the automaton, X and Y — sets of the input symbols and a set of the output symbols, accordingly; $\Delta: S \times X \rightarrow S$ — function of transition into a new state; $\lambda: S \rightarrow Y$ — function of output of a symbol. Behavior of the automaton is defined as $s^{t+1} = \Delta(s^t, x^t)$ and $y^t = \lambda(s^t)$.

Let us introduce a formal description the abstract neuron in terms of the finite sets. The finite automaton of the abstract neuron (FAAN) can be described by a set of finite sets ($N, E, W, Q, \Delta, \Lambda, T$),

where: N — set of indexes of inputs (control inputs and synapses), $N \in \mathbb{N}$; E — alphabet including the alphabet of all possible input symbols E_{input} and alphabet of all possible output symbols E_{output} и E_{output} , $E = E_{\text{input}} \cup E_{\text{output}}$, $E \subset \mathbb{N}$; W — linearly ordered finite set of all possible values of the weight factors of the information inputs (synapses), $W \subset \mathbb{N}$; Q — area (set) for definition of the activation function, $\text{Dom}\Lambda = Q$, $Q \subset \mathbb{N}$; Δ — operation of the account of the weighed input signals $\Delta_{i=1}^n (\delta_i(w_i, E_j))$, $i \in N$ depending on operation of account of the weight signal factor of each input (synapse $\delta_i(w_i, E_j)$, $E_j \in E$); Λ — set of the activation functions realized on a given neuron $|\Lambda| \in \mathbb{N}$; T — set of functions of variation of the functions' parameters Δ, Λ and $\tau_i \in T$, $|T| \in \mathbb{N}$ of the artificial neuron.

On the basis of E_{input} the subsets of the input information signals E_{in} are formed as well as the sets of the control signals E_r such as $E_{\text{input}} = E_{in} \cup E_r$. In a general case, at a hardware realization, it is not always possible to divide the input signals into the control and information ones.

Taking into account the fact that the signal of output of the artificial neuron can move to its input or contain the control symbols sent to the other neurons, we have the following correlations: $E_{\text{input}} \cap E_{\text{output}} \neq \emptyset$, $E_{\text{input}} \cap E_{\text{output}} = \emptyset$, and $E_{\text{input}} = E_{\text{output}}$. The set of the input symbols of each i -input can be described as $I_i = \{(E_{in}, E_r) | ((E_{in} = \emptyset) \wedge (E_r \neq \emptyset)) \vee ((E_{in} = \emptyset) \wedge (E_r = \emptyset)) \vee ((E_{in} \neq \emptyset) \wedge (E_r = \emptyset))\}$,

$i \in N$. For example: supply of a signal above a certain threshold can not only transform the memristor in a new state of conductivity, but it can also be recognized as an information signal.

Sets of symbols on various inputs form the alphabet of the input vectors of a neuron $I_{in} = \{(i_j, \dots, i_k) | i_j \in I_1, \dots, i_k \in I_n, \cup j_i = I_1, \dots, \cup k_i = I_n\}, n = N$.

The alphabet of the input symbols looks like the following $O = \{o_i | o_i = (e_{out}, e_r), e_{out} \in E_{out}, e_r \in E_r, E_{out} \cup E_r = E_{output}\}$, where E_{out} — is a set of symbols of the information output signals. Summing up all the above, the alphabets of I_{in} and O can be put in conformity with the sets of the input and output symbols of X and Y and Moore automaton.

Taking into account a possibility of realization of several functions on one neuron, the activation function of FAAN $\lambda_i \in \Lambda, i \in \mathbb{N}$ is defined by the following dependencies:

$$\lambda_i : Q \rightarrow O, \quad (8)$$

$$o^t = \lambda_i(q^t)^t, \quad (9)$$

which allows us to put in conformity the given function of the output function of Moore automaton.

All the possible combinations of the weight factors on all inputs of FAAN can be presented in the form of the vector

$$V = \{v_i | v_i = (w_j, \dots, w_k), w_j \in W_1, w_k \in W_n, \cup j w_j = W_1, \dots, \cup k w_k = W_n\}, n \in N,$$

where W_n — subset of the weight factors of every i synapse. The internal state of the abstract neuron is described by the following set $V \times \Delta \times \Lambda \times T$.

The function of account of the weighted input signals Δ is a given one on the set of $I_{in} \times V$ and looks like the following:

$$\Delta : I_{in} \times V \rightarrow Q, \quad (10)$$

$$q^{t+1} = \Delta \left(\sum_{i=1}^n (\delta(v_i^t, i_i^t))^t \right). \quad (11)$$

The function of variation of FAAN parameters $\tau_i \in T, i \in \mathbb{N}$ in the most general form is described as

$$\tau_i : E_r \rightarrow V \times \Delta \times \Lambda \times T. \quad (12)$$

The dynamics of the variation of parameters looks like

$$(v_i^{t+1}, \Delta^{t+1}, \lambda_i^{t+1}, \tau_i^{t+1}) = \tau_j(i_j)^t, \\ \tau_i^{t+1} \in T, \tau_j \in T. \quad (13)$$

The function of transition of FAAN into a new state is a result of two independent from each other functions (11) and (12), which can be presented in the following form:

$$(q^{t+1}, v_i^{t+1}, \Delta^{t+1}, \lambda_i^{t+1}, \tau_i^{t+1}) = \\ = \left(\sum_{i=1}^n (\delta(v_i^t, i_i^t))^t, \tau_j(i_j)^t, \tau_i^{t+1} \right), \quad (14)$$

which allows us to put in conformity the transition function to Moore automaton.

The initial state of FAAN is expressed through (11) and (13) with an empty set of signals at the input:

$\left(\sum_{i=1}^n (\delta(v_i^{t=0}, \emptyset)^t, \tau_j(\emptyset)^t) \right)$, where $t = 0$ — values in the initial moment of time.

Proceeding from all the enumerated above, FAAN can be put in conformity to Moore automaton (see the table).

Minimal realization of FAAN is presented in fig. 1.

4. Properties of synapses of FAAN

The following situations are possible, when the set of E_{in} and its subsets of the values of the input signals of i synapse are set $E_{in} = \bigcup_{i=1}^n E_i, E_i = \{e_i | e_i \in E_{in}\}$, n — number of bonds for which $w_i \neq 0$ is true at $\delta(w_i, e_j) = w_i \cdot e_j$.

For the information signals of i inputs (synapses) various situations are possible when they are set. Elements of various sets of E_{in} are equal among themselves

$$E_i \cup E_j = E_{in}, \forall e_i \in E_i, \forall e_j \in E_j ((e_i > e_j) \vee (e_i < e_j)) \Rightarrow \forall e_i \in E_i, \forall e_j \in E_j (e_i = e_j).$$

For example: to the 1st input the information about the color is sent (red, green etc.), to the 2nd input the information about the form (sphere, parallelepiped, etc.) is supplied.

Elements of not all the subsets are comparable between themselves $E_i \cup E_j = E_{in}, \exists e_i \in E_i, e_j \in E_j ((e_i > e_j) \vee (e_i \neq e_j) \vee (e_i < e_j))$. The signals sent to one of the inputs are not comparable with the other ones.

The order on the elements of the subsets can be non-rigorous $\forall e_i \in E_{in}, \exists e_j \in E_{in} ((e_i > e_j) \vee (e_i < e_j)) \Rightarrow (E_{in} \cap (U e_j) = \emptyset)$. All the signals supplied to the inputs are comparable with each other. The output of at least one neuron (input neuron) is connected with the input of at least one more neuron (for a minimal network, more than one neuron is necessary).

In a general case, a set of the output information signals of neuron E_{out} can have a power which is different from the power of the input signals of the subsequent neuron E_{in} , for example, the range of voltages on the output of 1...3 V of the first neuron is sent to the input of the second neuron, receiving signals in the range of 2...3 V. Then it is obvious, that the input signals in the range [1, 2] should not be considered as information ones.

The given case is formally described by application of the characteristic function to a set of the entrance information signals of every synapse: $\chi E_i(e_i), e_i \in E_i, \bigcup_{i=1}^n E_i = E_{in}$, where the values of function $\chi E_i(e_i) = 0$ are not considered in the given neuron.

On the sets of information signals of synapses $\chi E_i(e_i) \neq 0$, taken into account by the neuron receiving them $\bigcup_{i=1}^n E_i = E_{in}, |E_i| > 1$, a linear order can be set. In case if the order is not set, all the signals (or a part of them, in case, if a not strict order is set) will be indiscernible among themselves.

Summarizing all the described above, it is possible to conclude the following — the input information signals on different synapses, on the one hand, can be incomparable among themselves, on the other hand, the set of the signals of every synapse is described by a characteristic function. A result of this is definition of the optimal power of the sets of the information signals for a general case:

$$\bigcup_{i=1}^n E_i = E_{in}, |E_i| = |E_{out}| = 1.$$

Absence of a bond can be set through the value of the weight factors of synapses $w_i = 0$ at $\delta_i(w_i, i_j) = w_i \cdot i_j$. Since the set of the weight factors W is linearly ordered also for all the nonexistent bonds $w_i = 0$, then the elements of the set acquire the following values:

$\chi W_i(\emptyset) = 0$ for the absent bonds,

$\min(W_i) = k$, where $k \in \mathbb{N}$ is a certain constant.

$\min(W_i) \min(W_j) = k_w + h$, where $h \geq k$, $h \in \mathbb{N}$ is a certain constant and $W_i = \{w_j | w_{j+1} - w_j = h\}$.

The traditional approach to realization of the account of the weight of the input information signals of each i input is multiplication by its weight factor

$$\delta : w_i \cdot e_i \rightarrow \mathbb{N}, e_i \in E_i, \bigcup_{i=1}^n E_i = E_{in}, w_i \in W_i, \\ \bigcup_{i=1}^n W_i = W.$$

At a parity of elements of the set of $e_i \in E_{in}$, $w_j \in W$, set by a constant, identical for both sets, the maximal possible capacity of the sets in a limiting case is defined as follows:

$$\forall e_{i+1} \in E_{in}, \forall e_i \in E_{in}, \forall w_{j+1} \in W, \\ \forall w_j \in W((e_{i+1} - e_i = h) \wedge (w_{i+1} - w_i = h)) \Rightarrow \\ \Rightarrow |\text{Dom}\Delta| = |E_{in}| \cdot |W|.$$

Increase of the power of the set of the area of definition of $\text{Dom}\Delta$ FAAN at preservation of the power of the sets of E_{in} , W is possible at the expense of setting the task of a nonlinear dependence between the elements $w_j \in W$.

The first of the possible ways is to put in conformity to the elements of set W the elements of the linearly ordered set with a nonlinear dependence, for example: $W = \{\emptyset = 0, w_1 = 1, w_2 = n + 1, w_3 = n \cdot w_2 + 1, \dots, w_j = n \cdot w_{j-1} + 1\}$, where n — number of bonds for which $w_j \neq 0$ is fair. The above method increases maximally the power of the definition area of $\text{Dom}\Delta$ FAAN during a hardware realization.

Among the drawbacks of the approach are a full absence of a possibility to set a linear dependence by the input signals on the synapses, which limits its application in the hardware realizations intended for construction of networks on the basis of FAAN.

The second method can be defined as an introduction of nonlinearity between the subsets of set W , and it has the following scheme of the task:

$$W = \{\emptyset = 0, w_1 = 1, w_2 = 2 \cdot w_1, \dots, \\ w_i = i \cdot w_1, w_{i+1} = n \cdot w_i + 1, w_{i+2} = 2 \cdot w_{i+1}, \dots, \\ w_{i+j} = j \cdot w_{i+1}, \dots, w_{k+1} = n \cdot w_k + 1, \dots, \\ w_{k+l} = l \cdot w_{k+1}\}, \bigcup_{i=1}^l w_i = W_i, \bigcup_{i=1}^j w_j = W_j, \dots, \bigcup_{k+1}^l w_l = W_l, \bigcup_p W_p = W,$$

where n — number of bonds, for which $w_i \neq 0$ is fair.

The approach is a compromise between the possibility to set linear dependences at the account of the input signals on synapses and an increase of the power of the set of $\text{Dom}\Delta$ FAAN. The nonlinear dependence between the input signals E_{in} , also leads to an increase of the set of $\text{Dom}\Delta$ and can be described by one of two variants considered above for the set.

5. Formalism of the artificial neural networks on the basis of FAAN

Formalism of description of the network on the basis of FAAN is defined by a tuple $(E, W, Q, \Delta, \Lambda, T)$ without the use of N .

Let us demonstrate this. The used concepts of the input neurons can be extended not only to the information inputs (synapses), but also to the control inputs. Using the given condition, we will receive a single description of the control and information inputs, for the case of a hardware realization of the Harward architecture of construction of the computing systems.

Having defined that the output signal of each FAAN is sent to the input of all the other FAAN of the network, we get a fully connected network. Having set the values of the weighted factors of synapses $w_i \neq 0$, for the nonexistent bonds, we get a description of the artificial neuron network (ANN) without the use of N set.

The classes of the functions realized on FAAN depend directly on the power of set Q . Set Q is a linear ordered set of $\text{Dom}\Delta$ from one side and the area of values of Δ from the other side. For the case of $|E_{in}| = 1$ the following situations are possible. Since $\text{Dom}\Delta$ is linearly ordered and $\forall a \in \text{Dom}\Delta (\Delta(a_{i+1}) > \Delta(a_i))$, then at $|Q| = 1$ only realization of the threshold function is possible.

At $|Q| = 2$ realization of the threshold, linear and RBF functions is possible. The threshold function can be set by four methods, the first two of them are the following:

$$\forall \lambda \in \Lambda((\lambda(\max Q) = 1) \wedge (\lambda(q_i) = 0, q_i \neq \max Q)), \\ \forall \lambda \in \Lambda((\lambda(\min Q) = 0) \wedge (\lambda(q_i) = 1, q_i \neq \min Q)).$$

Inverting of the values of function λ from 0 to 1 gives two other ways.

The linear function is set as follows:

$$\forall q \in Q, \forall \lambda \in \Lambda(((\lambda(q_{i+1}) > \lambda(q_i)) \wedge (q_{i+1} > q_i)) \vee \\ \vee ((\lambda(q_{i+1}) < \lambda(q_i)) \wedge (q_{i+1} > q_i))).$$

RBF-function is set as follows:

$$\forall \lambda \in \Lambda(((\lambda(q_1) > \lambda(q_2)) \wedge (\lambda(q_1) > \lambda(\emptyset))) \vee \\ \vee ((\lambda(q_1) < \lambda(q_2)) \wedge (\lambda(q_1) < \lambda(\emptyset)))).$$

The minimal necessary power for setting of the periodic functions is $|Q| = 3$.

The hardware realization of setting of the periodic functions can be implemented by a modification of the function of account of the weighed input signals with application of operation of taking by the module $\Delta_i(\delta(w_i, i_j)) = \text{mod}_d(\sum_i \delta(w_i, i_j))$, where d — size of period on $\text{Dom}\Delta$.

A result of application of the modified function will be the value of the beginning of the current period. By setting of the corresponding values, E_{out} can be presented in the form of RBF, linear or threshold functions in the current period.

In the limiting case $|\text{Dom}\Delta| = |Q|$, which allows us to set not only linear, threshold, nonlinear, RBF, or periodic functions, but also almost periodic functions.

Inclusion into the FAAN model of the set of functions of variation of T parameters allows us to take into account in the algorithms of training of FAAN network the variations of the activation function in the process of training of the network.

This, for example, includes transition from the threshold functions to the linear functions, transition from the linear functions to the nonlinear functions (sigma function set on a discrete set) and RBF function; transition from RBF- and nonlinear functions to the periodic ones, and also transition from the periodic functions to almost periodic functions during training of ANN.

The available description of FAAN is applicable for description of the work of the synchronous artificial neural networks. The networks, in which all the neurons are activated (or not activated) during one moment of time, after which the data of the variables are nulled and a new cycle of operation of the network begins.

An attempt of an asynchronous description of the work of a single FAAN is impossible in view of absence of an object for comparison, which allows us to draw a conclusion about its practical uselessness. A hardware realization of the proposed model of FAAN (hybrid circuit solution) in the minimal composition demands for the realization a set of two units of registers (fig. 2).

Unlike individual FAAN, the networks on their basis can have an asynchronous mechanism of operation. Generalization of the FAAN networks for the case of an asynchronous operation of the network demands introduction of changes into expressions (9), (11).

The introduced changes should take into account a possibility of adjustment of the actuation time of the activation function of FAAN t_A and/or possibility of the account of the previous state by function Δ . The equation of variation of the states of FAAN described by equation (11) and the equation of the dynamics of output of symbols (9) for the asynchronous networks of FAAN acquires the following form:

$$o^{t_A} = \lambda_i(q^t)^{t_A}; \quad (15)$$

$$q^{t+1} = \sum_{i=1}^n (\delta(v_i^t, i_i^t)^t, q^t)^t. \quad (16)$$

The unit-by-unit composition of the hardware realization of FAAN for the asynchronous networks is presented in fig. 2.

Conclusion

The generalized model of FAAN makes it possible to model a wide set of the artificial neurons with various activation functions, which during a hardware realization of FAAN guarantees a relative universality of the technical solution and a wide set of opportunities for the network designing.

Construction of a model on the basis of the finite sets allows us to count the necessary number of the neurons (for the known algorithm of operation of the network), proceeding from the data about the sets, on the basis of which the FAAN networks are realized, or to implement an optimization of the available networks.

The networks constructed on the basis of the proposed models can recognize not only images, but also the mathematical operations of addition, subtraction, multiplication, etc. Realization of the mathematical operations allows us to apply a part of the network resources (having realized algorithm of training in a subset) for training of the other part of the network without application of the classical (Neiman and Harward) architectures.

The proposed FAAN model allows us to consider in the training algorithms of the variations of the activation function of FAAN and to construct an algorithm automatically generating a network with the optimal activation functions of the neurons. Besides, the generation of the FAAN networks (by means of application of the training algorithms), there is a possibility of setting an algorithm for operation of a network with a dynamically changing activation functions.

The author of the article expresses his deep gratitude to Igor Matiushkin, employee of the Department of Functional Electronics of NIIME and to Dmitry Telpukhov from IPPM of RAS for their valuable observations and discussion of the material of the article.

The work was done in cooperation with RFFI, project no. 17-07-00570 A.

References

1. Strukov D. B., Snider G. S., Stewart D. R., Williams R. S. The missing memristor found, *Nature LETTERS*, 2008, vol. 453, pp. 80–83, DOI: 10.1038/nature06932.
2. Merolla P. A., Arthur J. V., Alvarez-Icaza R. et al. A million spiking-neuron integrated circuit with a scalable communication network and interface, *Science*, 2014, vol. 354, issue 6197, pp. 668–673, DOI: 10.1126/science/1254642
3. Degans H. Imec demonstrates self-learning neuromorphic chip that composes music, available at: <http://www.imec-int.com>
4. McCulloch W. S., Pitts W. A. A logical Calculus of Ideas Immanent in Nervous Activity, *Bull. Mathematical Biophysics*, 1943, vol. 5, issue 4, pp. 115–133.
5. Koch C., Poggio T., McKenna T., Davis J. et al. (Eds.) *Multiplying with Synapses and Neurons in Single Neuron Computations*, Academic Press Inc., San Diego, CA, USA, 1992, Chapter 12, pp. 315–319.
6. Kandpal P. K., Mehta A. Comparison and Analysis of Multiplicative neuron and Multilayer Perceptrons using three different datasets, *International Journal of Engineering Research and General Science*, 2017, vol. 5, no. 3, pp. 58–70.
7. Shibzukhov Z. M., Cherednikov D. Y. O modeljah nejronov agregirujushhego tipa, *Mashinnoe obuchenie i analiz dannyh*, 2015, vol. 1, no. 12, pp. 1706–1716. DOI: 10.21469/22233792.1.12.06 (in Russian).
8. Hodashinsky I. A., Maljutin N. D. Modeli staticheskikh iskusstvennykh nejronov. Obzor, *Doklady TUSUR*, 2015, no. 2 (36), pp. 100–107 (in Russian).
9. Moore E. P. Gedanken-experiments on sequential machines, *Automata Studies, Annals of Mathematical Studies*, 1956, vol. 34, pp. 129–153.

Адрес редакции журнала: 107076, Москва, Стромынский пер., 4. Телефон редакции журнала (499) 269-5510. E-mail: nmst@novtex.ru
Журнал зарегистрирован в Федеральной службе по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия.

Свидетельство о регистрации ПИ № 77-18289 от 06.09.04.

Технический редактор Т. А. Шацкая. Корректор Е. В. Комиссарова.

Сдано в набор 21.05.2018. Подписано в печать 23.06.2018. Формат 60×88 1/8. Заказ МС0718. Цена договорная
Оригинал-макет ООО «Адвансед солюшнз». Отпечатано в ООО «Адвансед солюшнз». 119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: www.aov.ru

Рисунки к статье В. К. Илькова, А. О. Михалева, М. В. Майтамы
«КОММУТАТОРЫ НА АРСЕНИДЕ И НИТРИДЕ ГАЛЛИЯ»

V. K. Il'kov, A. O. Mikhalev, M. V. Maytama
«ARSENIDE AND NITRIDE GALLIUM SWITCHES»

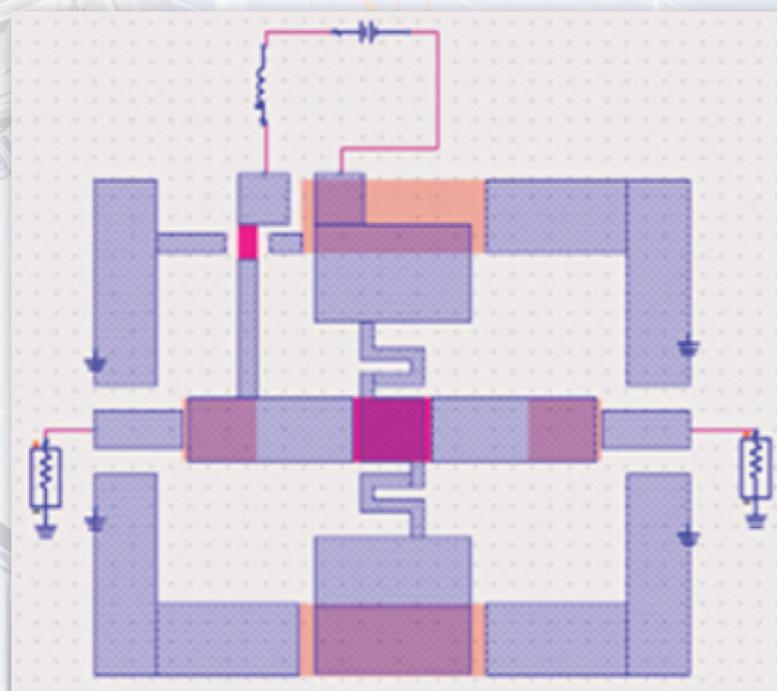


Рис. 4. Схема включения емкостного шунтирующего МЭМС-переключателя с креплением мембраны в виде меандра

Fig. 4. Switching circuit of a capacitive shunt MEMS switch with a meander membrane

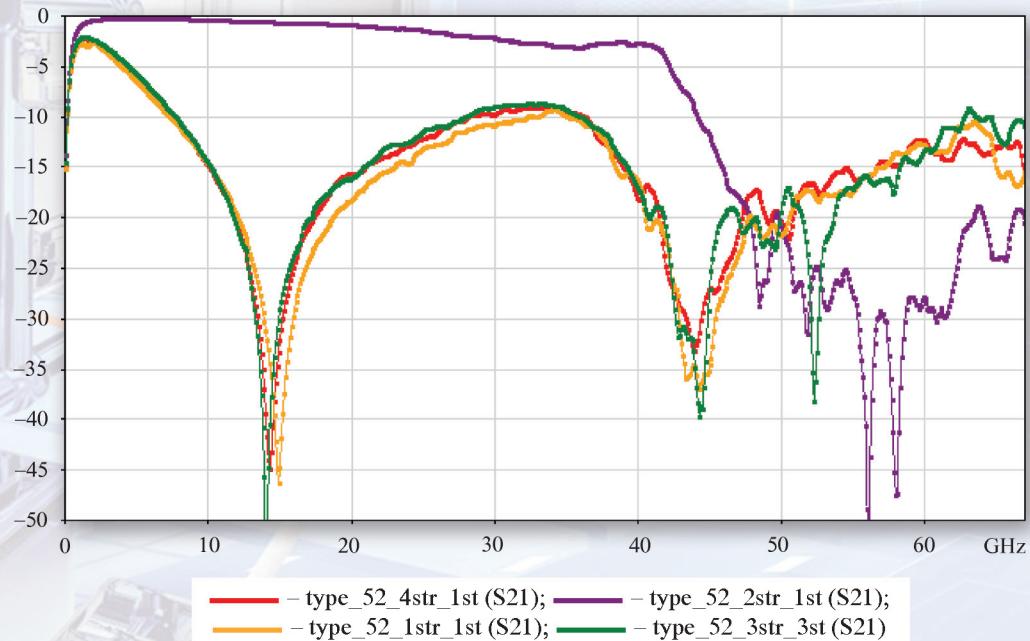


Рис. 6. Измеренная зависимость изоляции узкополосных СВЧ коммутаторов от частоты для двух СВЧ коммутаторов сигналов

Fig. 6. The measured dependence of isolation of a narrow-band microwave switch on frequency for two microwave signal switches



МЕЖДУНАРОДНЫЙ ФОРУМ МИКРОЭЛЕКТРОНИКА 2018

1-6 октября 2018 г.
г. Алушта
(Республика Крым)

Международный Форум «Микроэлектроника - 2018»

Приглашает Вас и Ваших сотрудников принять участие в Международной научной конференции:
«Микроэлектроника - ЭКБ и электронные модули»

Задачи Форума: комплексно рассмотреть актуальные вопросы разработки, производства и применения отечественной электронной компонентной базы и высокointегрированных модулей.

ПРИ ПОДДЕРЖКЕ



РАБОТАЙ В РОССИИ!

ОРГАНИЗАТОРЫ

Генеральный информационный партнёр



Оператор Форума: Компания «ПрофКонференции» • Тел.: +7 (495) 641-57-17 • Факс: +7 (495) 641-57-17 • E-mail: info@microelectronica.pro
Подробная информация и регистрация участников на официальном сайте Форума: microelectronica.pro