на правах рукописи

КАЛУГИН Александр Николаевич

РАЗРАБОТКА И ИССЛЕДОВАНИЕ МНОГОМЕРНЫХ ГЕНЕРАТОРОВ РАВНОМЕРНО РАСПРЕДЕЛЕННЫХ ПСЕВДОСЛУЧАЙНЫХ ВЕКТОРОВ, ОСНОВАННЫХ НА ПРЕДСТАВЛЕНИИ ДАННЫХ В АЛГЕБРАИЧЕСКИХ ПОЛЯХ

Специальность 05.13.17 – Теоретические основы информатики

Автореферат

диссертации на соискание ученой степени

кандидата физико-математических наук

Самара 2008

Работа выполнена на кафедре геоинформатики
государственного образовательного учреждения
высшего профессионального образования
"САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ им. ак. С.П. КОРОЛЕВА" (СГАУ),
и в ИНСТИТУТЕ СИСТЕМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК

Научный руководитель доктор физико-математических наук

В.М.Чернов

Официальные оппоненты: доктор физико-математических наук, профессор

С.Я.Шатских,

кандидат физико-математических наук, доцент

Э.И.Коломиец.

Ведущая организация: Омский государственный университет
им. Ф.М. Достоевского

Защита планируется "__20__" __июня__ в __10__ часов на заседании диссертационного совета Д 212.215.07 в Самарском государственном аэрокосмическом университете имени академика С.П.Королева по адресу: 443086, г. Самара, Московское шоссе, д. 34.

С диссертацией можно ознакомиться в библиотеке Самарского государственного аэрокосмического университета имени академика С.П.Королева.

Автореферат разослан "__19__" __мая__ 2008 г.

Ученый секретарь
диссертационного совета,
д.т.н., профессор И.В.Белоконов

Общая характеристика работы

Диссертация посвящена разработке и исследованию нового метода синтеза многомерных генераторов псевдослучайных векторов.

Актуальность исследования. Разработка методов построения и исследование свойств детерминированных последовательностей точек многомерного пространства, обладающих признаками "случайности", понимаемой относительно того или иного вероятностного критерия, является фундаментальной задачей. Для нее, с одной стороны, характерно использование развитых методов современной *абстрактной* математики (алгебра, теория чисел), а, с другой стороны, её *прикладная* значимость в информатике и вычислительной математике возрастает, в связи с бурным развитием методов численного моделирования с использованием высокоскоростных (в том числе и параллельных) вычислительных устройств и спецпроцессоров, развитием методов криптографии, использованием таких последовательностей при тестировании *VLSI* систем, в компьютерных играх, и.т.д.

Теория генераторов псевдослучайных последовательностей (генераторов случайных чисел, ГСЧ) — детерминированных алгоритмов, порождающих последовательности, свойства которых "имитируют" свойства последовательности реализаций независимых одинаково распределенных случайных величин — берет свое начало в 1946 году с предложенного Дж. Фон Нейманом метода "средин квадратов". Возникнув изначально как попытка синтезировать метод генерации псевдослучайной выборки, лишенный недостатков физических датчиков случайных чисел и таблиц случайных чисел, методология генерирования псевдослучайных последовательностей получила в дальнейшем бурное развитие как самостоятельное направление. Данные методы, исследовались многими авторами: П. Лемером (*P.Lehmer*), К.С. Таусвортом (*K.C.Tausworth*), Д. Кнутом (*D.Knuth*). Г. Нидеррайтером (*H.Niederreiter*), П.Хеллекалеком (*P.Hellekalek*), П. Леквие (*P.l'Ecuyer*), Дж. Марсальей (*G. Marsaglia*), С.Вегенкиттлом (*S. Wegenkittl*) и др. К настоящему времени разработаны различные методы генерации псевдослучайных последовательностей: линейный конгруэнтный генератор (*linear congruential generator*) , регистр сдвига с линейной обратной связью (*linear feedback shift register*), регистр сдвига с обобщенной обратной связью (*generalized feedback shift register*), генератор Фибоначчи с запаздываниями (*lagged-Fibonacci generator*), инверсный генератор (*inversive generator*), множественный рекурсивный генератор (*multiple recursive generator*) и др.

Большинство разработанных и детально исследованных на настоящий момент генераторов являются одномерными. Однако существуют вычислительные задачи (например, численное вычисление значения многомерного определенного интеграла по методу Монте Карло, задачи статистической физики и др.), требующие генерации последовательности псевдослучайных векторов многомерного пространства. Теория генерации псевдослучайных векторов является относительно новым направлением; к настоящему моменту разработано лишь несколько "естественно" многомерных генераторов: матричный конгруэнтный генератор (*matrix generator*), матричный инверсивный генератор (*matrix inverse generator*), матричный множественно-рекурсивный генератор (*multiple-recursive matrix generator*), генератор С. Вольфрама, основанный на использовании клеточного автомата. Достоинством данной группы методов является теоретически доказанное и положенное в основу метода генерирования соответствие порождаемого распределения теоретическому в многомерном пространстве размерности, равной размерности генерируемых векторов (относительно предварительно заданных критериев качества). К недостаткам этой группы можно отнести принципиальную невозможность эффективной параллельной реализации; так, например, для матрично-рекурсивного метода для генерации i-ой координаты очередного

псевдослучайного вектора, ($i = 0,1,...,\kappa - 1$, κ — размерность генератора) необходимы значения *всех* координат предыдущих h векторов ($h \in \mathbb{N}$ — параметр генератора, порядок матричного рекуррентного соотношения).

Альтернативным подходом к генерации многомерных (параллельных) псевдослучайных последовательностей являются методы "повышения размерности", основанные на распределении элементов одной или нескольких одномерных псевдослучайных последовательностей среди генерирующих процессоров (т. н. метод кругового обхода (*leapfrog*), метод блоков (*sequence splitting*), метод параметризации (*sequence parameterization and initialization*). Основной идеей, положенной в основу синтеза данных методов, является возможность их эффективной параллельной реализации на нескольких компьютерах / процессорах. Однако многомерное распределение синтезированных точек на выходе "параллелизованных" генераторов может иметь существенные отклонения от теоретического (уровень которых может варьироваться от едва заметных до абсурдно очевидных) в терминах широкого круга критериев. В качестве примера таких недостатков рассмотрим структуру множества элементов полного периода последовательности на выходе "параллелизованного" генератора *Randu*:

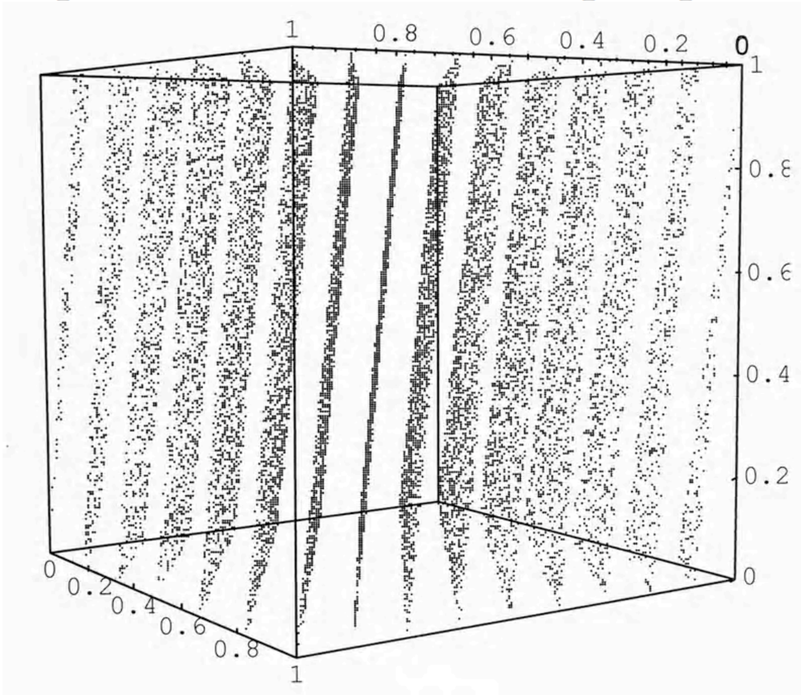

Рис. 1. Распределение трехмерных векторов, сформированных методом кругового обхода, из элементов последовательности Randu.

Актуальной и практически значимой представляется задача синтеза методов генерирования, комбинирующего положительные свойства как "естественно" многомерных, так и "параллельных" генераторов псевдослучайных последовательностей, лишенных описанных недостатков.

Целью диссертационной работы является разработка новых методов синтеза и исследование свойств многомерных генераторов, порождаемых рекуррентными процессами и базирующихся на концепции позиционных систем счисления в многомерных дискретных решетках.

Методы исследования. В диссертационной работе используются методы абстрактной алгебры, теории чисел, численных методов, теории вероятностной и математической статистики, теории случайных процессов.

Научная новизна работы. В диссертационной работе получены следующие новые научные результаты:

1) предложена новая схема синтеза генераторов равномерно распределенных псевдослучайных векторов, экстраполирующая методы синтеза одномерных ГСЧ на многомерный случай, базирующаяся на использовании канонических систем счисления;
2) получены аналитические оценки, характеризующие распределение последовательностей

точек многомерных пространств, сгенерированных обобщенным ГСЧ Таусворта, на полном и неполном периоде, статистическую независимость элементов последовательности на выходе генератора;

3) доказана возможность параллельной реализации разработанного обобщенного метода Таусворта и получены оценки ее вычислительной сложности.

Практическая значимость работы заключается в разработке эффективно реализуемых генераторов псевдослучайных векторов многомерного пространства, имеющих широкую область применения в задачах информатики и численного анализа.

Теоретическая значимость работы. В работе предложен новый общий подход к экстраполяции на многомерный случай методов синтеза одномерных генераторов псевдослучайных чисел, использующих представление данных в позиционных системах счисления; описана методология исследования обобщенных генераторов.

Апробация работы. Основные результаты диссертации докладывались на 10 международных, всероссийских и региональных научных конференциях, в частности: международной конференции "Automation, Control, and Information Technologies" (ACIT), Новосибирск, 2005; научно-технической конференции с международным участием "Перспективные информационные технологии в научных исследованиях проектировании и обучении" (ПИТ), Самара, 2006; 7-ой международной научной конференции "Monte Carlo and Quasi-Monte Carlo methods in Scientific Computing", Ульм, Германия, 2006; международной научной конференции "Graficon-2007", Москва, 2007.

Исследования по теме диссертационной работы выполнялись при поддержке РФФИ (*проекты №№, 03-01-00736, 06-01-00722*), Американского фонда гражданских исследований и развития в рамках российско-американской программы "Фундаментальные исследования и высшее образование", министерства образования и науки Самарской области (*грант 2005 года студентам, аспирантам и молодым ученым*).

Публикации. По тематике, непосредственно связанной с содержанием диссертации, опубликовано 10 работ, из них 4 опубликованы в ведущих рецензируемых научных журналах и изданиях, определенных Высшей аттестационной комиссией. Работы [1, 2, 3, 4, 5, 6, 7, 8] выполнены автором единолично, остальные работы написаны в соавторстве. На защиту выносятся только результаты, полученные лично соискателем.

Структура диссертации. Диссертационная работа, содержащая 159 стр., состоит из введения, пяти глав, заключения, списка использованной литературы, составляющего 107 наименований.

Краткое содержание работы

Во введении обоснована актуальность темы диссертационной работы, сформулированы ее цель и задачи, дан обзор научных работ по рассматриваемым вопросам, показана научная новизна, практическая и теоретическая значимость полученных результатов и сформулированы основные положения, выносимые на защиту.

В частности, отмечены основные методологические трудности синтеза и исследования генераторов псевдослучайных последовательностей, связанные с их предполагаемой универсальностью и отсутствием общепризнанного класса критериев качества. Приведены классификация и примеры наиболее часто используемых (в настоящее время) критериев качества для генераторов одномерных и многомерных псевдослучайных последовательностей.

Основная идея. Элементы последовательностей на выходе одномерных ГСЧ, используемых для численного моделирования (как результат конечности разрядной сетки

вычислительной платформы, выбранной для реализации алгоритма генерации), являются рациональными числами множества Λ:

$$x(n) \in \Lambda = [0;1) \cap \frac{1}{2^s}\mathbb{Z},\ n = 0,1,....$$

где s — длина машинного слова, точность разрядной сетки компьютера.

В двоичной системе счисления каждый элемент λ множества Λ может быть представлен в виде

$$\lambda = \sum_{r=0}^{s-1} a_r(\lambda) \cdot 2^{-j-1}, \qquad (1)$$

где элементы разложения a_r, $r = 0,1,...,s-1$ образуют бинарный s-мерный цифровой вектор $\vec{a}(\lambda)$: $\vec{a} = (a_0, a_1, ..., a_{s-1}) \in \{0,1\}^s$, $a_r \in \{0,1\}$, $r = 0,1,...,s-1$.

Таким образом, можно считать, что на выходе *любого одномерного* ГСЧ порождается последовательность бинарных *цифровых векторов* $\vec{a}(x(n)) \in \{0,1\}^s$ *размерности* s.

Концепция позиционных систем счисления может быть обобщена на многомерный случай. В 1975 г. Катаи (*I. Kàtai*) и Сабо (*J. Szabò*) в своей работе[1] доказали существование систем счисления вида $(-B \pm i, \{0 \ldots B^2\})$, $B \in Z^+$ и представимость любого целого гауссова числа (комплексного числа с целочисленными действительной и мнимой частью) в виде конечной суммы:

$$z = a + bi = \sum_{j=0}^{k} z_j(-B \pm i)^j,\ z_j \in \left\{0 \ldots B^2\right\},\ a,b \in Z,\ B \in Z^+,$$

с некоторым $k = k(z)$, зависящим от z.

Исследование таких позиционных систем счисления было продолжено Й. Тусвандлером (*J.Thuswaldner*), А. Пето (*A. Pethö*), С. Акиямой (*S. Akiyama*), А. Ковачем (*A. Kovács*). А. Ковач в своей работе[2] вводит понятие канонических систем счисления (*КСС*) в многомерных решетках Λ. Каждому элементу решетки $\vec{z} \in \Lambda$ в канонической системе счисления ставится в соответствие вектор цифр:

$$(\zeta_0, \zeta_1, \zeta_2, ...),\ \vec{a}_j = \zeta_j \vec{e} \in D.$$

Конструктивно доказано, что для любого $k \geq 2$, q-значные ($q \geq 2$, $q \in \mathbb{N}$) канонические системы счисления существуют.

Теория канонических систем счисления дает возможность использовать идеи, реализованные в одномерных генераторах псевдослучайных чисел, для построения их многомерных аналогов, а именно, при формальной замене в соотношении (1) основания обычной позиционной системы счисления на основание системы счисления в многомерной решетке заданной размерности, каждому цифровому вектору ставится в соответствие вектор многомерного пространства. Целесообразность такой замены с точки зрения качества равномерного распределения на выходе синтезированного генератора подлежит дальнейшему изучению, что и послужило мотивацией для постановки указанной выше цели работы и определило задачи диссертационного исследования.

Задачи диссертационной работы:

1) теоретическое обоснование возможности синтеза генераторов псевдослучайных последовательностей векторов (КСС-генераторов) с использованием теории канонических систем счисления в решетках, синтез обобщенного генератора Таусворта.

---

[1] Kátai, I. Canonical number systems for complex integers [Text] / I.Kátai, J.Szabo // Acta Sci. Math (Szeged) — 37— 1975.— pp. 255 — 260.

[2] Kovács A. On number expansions in lattices [Text] /A. Kovács // Proc. 5th Internation Conference on Applied Informatics/ Eger, Hungary. — 2001.

2) исследование свойств множества точек на выходе КСС-генераторов, возможности унифицирующего преобразования данного множества к виду, удобному для практического использования в задачах моделирования по методу Монте Карло;
3) аналитическое исследование многомерного распределения на выходе предложенного обобщенного генератора Таусворта (на полном и неполном периоде); исследование статистической независимости отсчетов генерируемой последовательности,
4) исследование свойств координатных последовательностей на выходе генератора, анализ влияния параметров генератора и начальных условий на его свойства;
5) численное исследование свойств последовательностей на выходе генератора на участке периода, сравнение предложенного генератора с существующими схемами.

Сформулированные выше задачи определяют структуру диссертации и содержание отдельных глав.

В **первой главе** представлены необходимые сведения из теории алгебраических полей, дискретных решеток. Рассматриваются некоторые свойства показательных и рекуррентных функций в конечных полях, тригонометрические суммы и суммы характеров с показательными функциями. Приводятся базовые определения и результаты из теории канонических систем счисления, теории множеств с малым отклонением.

В частности, в **разделе 1.3** рассматриваются определение дискретной решетки, положения теории приведенных базисов дискретных решеток, используемые для исследования обобщенных многомерных генераторов.

Определение 1.6[3]**.** Пусть $\{\vec{a}_0,\vec{a}_1,...,\vec{a}_{k-1}\}$ — множество линейно независимых векторов пространства $\mathbb{R}^k$. Линейная оболочка с целыми коэффициентами Λ векторов $\{\vec{a}_0,\vec{a}_1,...,\vec{a}_{k-1}\}$ называется *(дискретной) решеткой* в $\mathbb{R}^k$ с базисом $\mathbf{a}=\{\vec{a}_0,\vec{a}_1,...,\vec{a}_{k-1}\}$

$$\Lambda=\{\vec{\xi}\mid\vec{\xi}=\xi_0\vec{a}_0+\xi_1\vec{a}_1+...+\xi_{k-1}\vec{a}_{k-1}\},$$

где $\xi_i\in\mathbb{Z}$, $i=0,1,...,k-1$.

В **разделе 1.4** рассматриваются *канонические системы счисления* (*КСС*) в решетках, введенные в работах венгерского математика А. Ковача. Эти канонические системы счисления используются далее в диссертационной работе для синтеза многомерного обобщения одномерных ГСЧ.

Пусть Λ - решетка в $\mathbb{R}^\kappa$, $\kappa\in\mathbb{N}$, $M:\Lambda\to\Lambda$ — линейное отображение такое, что $\det(M)\neq 0$ и $\tilde{D}$ — конечное подмножество Λ, заданное соотношением

$$\tilde{D}=\{a\vec{e}\mid\vec{e}=(1,0,0,...,0)_{\mathbf{a}}\in\Lambda;\ a=0,1,...,|\det M|-1\}\ . \qquad (2)$$

Определение 1.10. Тройка $(\Lambda,M,\tilde{D})$ называется *канонической системой счисления,* если любой элемент $\vec{x}\in\Lambda$ может быть единственным образом представлен в виде:

$$\vec{x}=\sum_{i=0}^{l}M^i\vec{d}_i=\sum_{i=0}^{l}\xi_iM^i\vec{e}\ ,\ \vec{d}_i\in\tilde{D}\ ,\ l\in\mathbb{N}\ . \qquad (3)$$

Оператор M при этом называется *основанием системы счисления*, $\tilde{D}$ — *множеством цифр*, вектор $(\xi_0,\xi_1,...,\xi_l)$ - *КСС-кодом* элемента $\vec{x}$.

Лемма 1.1. (Классы Ковача) *Для множеств цифр $\tilde{D}$ вида (2) и различных значений размерности $\kappa\geq 2$ тройки $(\mathbb{Z}^\kappa,\mathbf{M}_f,\tilde{D})$ являются системами счисления для следующих классов многочленов $f\in\mathbb{Z}[x]$:*

$f_1=x^\kappa+c_1x+q$, *если и только если* $-1\leq c_1\leq q-2$, $q\geq 2$, $q=p$; (4)

$f_2=x^\kappa+px^{\kappa-1}+px^{\kappa-2}+...+px+p$, $2\leq p\in\mathbb{N}$, $q=p$; (5)

---

[3] Нумерация определений, утверждений и таблиц в автореферате соответствует нумерации в диссертационной работе

$$f_3 = x^\kappa + x^{\kappa-1} + x^{\kappa-2} + \ldots + x + p,\ 2 \le p \in \mathbb{N}\ ,\ q = p\,; \tag{6}$$

$$f_4 = x^\kappa + px^{\kappa-1} + p^2x^{\kappa-2} + \ldots + p^{\kappa-1}x + p^\kappa,\ 2 \le p \in \mathbb{N},\ q = p^\kappa\,. \tag{7}$$

В **Разделе 1.5** приводятся основные определения теории отклонений, множеств с малым отклонением.

Пусть $I^\kappa = [0,1)^\kappa$ — κ-мерный единичный гиперкуб. Пусть далее множество точек P — подмножество I^κ: $P = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N\}$, $P \subset I^\kappa = [0,1)^\kappa$. Пусть B — произвольное подмножество $B \subset I^\kappa$. Определим величину

$$A(B;P) = \sum_{n=1}^{N} c_B(\vec{x}_n),$$

где c_B — характеристическая функция множества B. Таким образом, $A(B;P)$ — функция-счетчик, значения которой равны количеству индексов n , $1 \le n \le N$ таких, что $\vec{x}_n \in B$.

Пусть B — непустое семейство измеримых по Лебегу подмножеств единичного гиперкуба I^κ.

Определение 1.1. Отклонением (*discrepancy*) множества точек P относительно семейства B называется величина, определенная соотношением.

$$D_N(\mathsf{B}, P) = \sup_{B \in \mathsf{B}} \left| \frac{A(B;P)}{Card(P)} - \lambda_k(B) \right|.$$

Определение 1.14. "Звездное" отклонение *(star-discrepancy)* $D_N^*(P) = D_N^*(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)$ множества точек P определяется соотношением

$$D_N^*(P) = D_N(\mathsf{J}^*; P), \tag{8}$$

где J^* — семейство всех подмножеств единичного гиперкуба I^k вида $\mathsf{J}^* = \prod_{i=1}^{k} [0; u_i)$.

Также в разделе 1.5 приводится определение и формулировка основных результатов из теории (t,m,k)-*сетей*.

Во **второй главе** рассматривается схема экстраполяции методов синтеза ГСЧ на многомерный случай (синтез КСС-генераторов).

Определение 2.1. Генераторы псевдослучайных последовательностей векторов многомерного пространства, использующие представление элементов многомерной решетки в канонических системах счисления, будем называть *КСС-генераторами* псевдослучайных векторов.

Схема состоит из трех этапов.

Этап 1. Генерация последовательности цифровых векторов с использованием некоторого рекуррентного процесса f (функции перехода).

$$\vec{Y}(n+m) = f(\vec{Y}(0), \vec{Y}(1), \ldots, \vec{Y}(n+m-1)),\ \vec{Y}(n) \in D^s = \left(\mathbb{Z}_q\right)^s,\ n = 0,1,\ldots.,\ s \in \mathbb{N}.$$

Этап 2. Интерпретация цифровых векторов

$$\vec{Y}(n) = (Y_0(n), Y_1(n), \ldots, Y_{s-1}(n)),$$

как КСС-кодов элементов в q-значной канонической системе счисления $(\mathbb{Z}^\kappa, \mathbf{M}, D)$.

$$\vec{x}(n) = \sum_{j=0}^{s-1} Y_j(n)(\mathbf{M}^{s-1-j}\vec{e})\ ,\ \vec{x}(n) \in \mathbb{Z}^\kappa,\ n = 0,1,\ldots. \tag{9}$$

Этап 3. Преобразование точки $\vec{x}(n) \in \mathbb{Z}^\kappa$, полученной на втором этапе, в точку регулярного множества I^κ (многомерного единичного гиперкуба):

$$\vec{u}(n) = g(\vec{x}(n)), \qquad \vec{u}(n) \in I^\kappa = [0,1)^\kappa.$$

Отметим, что на втором этапе цифровым векторам $\vec{Y}(n)$ ставится в соответствие элементы $\vec{x}(n)$ решетки $\mathbb{Z}^{\kappa}$. Рассмотрим множество точек $\vec{x}(n)$ решетки $\mathbb{Z}^{\kappa}$, соответствующих всем возможным s-мерным цифровым векторам.

Определение 2.3. Назовем *фундаментальной областью F КСС-генератора* множество всех возможных элементов решетки $\mathbb{Z}^{\kappa}$, имеющих в качестве КСС-кода цифровой вектор размерности s

$$F = \left\{ \vec{u} \,\middle|\, \vec{u} = \sum_{j=0}^{s-1} u_j (\mathbf{M}^i \vec{e}),\ u_i \in \{0,\ldots,q-1\} \right\}.$$

Фундаментальная область *КСС*-генератора представляет собой множество со сложной геометрической конфигурацией (см. рис. 2).

Рис. 2 — Примеры фундаментальных областей КСС-генераторов, использующих $\kappa = 3$-мерные канонические системы счисления $(\mathbb{Z}^{\kappa}, \mathbf{M}, \tilde{D})$ с различными основаниями.

Подобная форма фундаментальной области затрудняет практическое использование множества точек на выходе второго этапа КСС-генератора, т.к. в большинстве практических задач предполагается, что многомерные псевдослучайные векторы равномерно распределены в некоторой регулярной области (гиперкубе). В диссертационной работе исследуются свойства фундаментальной области, и доказывается, что существуют эффективные алгоритмы преобразования (унифицирующего) фундаментальной области к единичному гиперкубу.

Определение 2.4. Будем называть *покрытием решетки* $\mathbb{Z}^{\kappa}$ множество

$$Tes_{\Omega}(S) = \bigcup_{\vec{\omega} \in \Omega} (S + \vec{\omega}) \ , \ \mathrm{Tes}_{\Omega}(S) = \mathbb{Z}^{\kappa},$$

а множества $S_{\omega} = (S + \vec{\omega})$ *элементами покрытия* $Tes_{\Omega}(S)$, если выполняются условия:

$S_{\vec{\omega}_1} \cap S_{\vec{\omega}_2} \neq \varnothing$, если и только если $\vec{\omega}_1 = \vec{\omega}_2$, $\vec{\omega}_1, \vec{\omega}_2 \in \Omega$, где $\Omega \subset \mathbb{Z}^{\kappa}$ — некоторая решетка.

Теорема 2.1. *Пусть F — фундаментальная область КСС-генератора, тогда существует решетка $\Omega(F) \subset \mathbb{Z}^{\kappa}$ с базисом $W = \{\mathbf{M}^s \vec{e}, \mathbf{M}^{s+1} \vec{e}, \ldots, \mathbf{M}^{s+k-1} \vec{e}\}$, такая, что множество $\mathrm{Tes}_{\Omega}(F)$ является покрытием решетки $\mathbb{Z}^{\kappa}$.*

Определение 2.5. Назовем *унификацией фундаментальной области F* взаимнооднозначное отображение фундаментальной области КСС-генератора F в единичный гиперкуб I^{κ}: $\mathrm{P} : F \to I^{\kappa}$.

В диссертационной работе, рассматривается метод построения унифицирующего отображения, а именно, выделение гиперкуба из покрытия решетки $\mathbb{Z}^{\kappa}$ фундаментальными областями генератора. В этом случае отображение P удовлетворяет соотношению:

$$\mathrm{P}(u) = \frac{1}{q^t} \begin{cases} \vec{u}, \ при & \vec{u} \in q^t I^{\kappa}, \\ \vec{u}' = \vec{u} + \vec{\omega}, \vec{\omega} \in \Omega(F) \ : \left(\vec{u} \notin q^t I^{\kappa}\right) \wedge \left(\vec{u}' \in q^t I^{\kappa}\right) & при \ \vec{u} \notin q^t I^{\kappa}. \end{cases} \qquad (10)$$

В работе доказана следующая теорема:

Теорема 2.2. *Пусть задано количество цифр* q *и произвольная размерность канонической системы счисления* $\kappa \in \mathbb{N}$. *Для КСС-генераторов, использующих размерность цифрового вектора* $s = \kappa t$ *(*t *— параметр генератора, определяющий его период, и зависящий от используемой КСС) и канонические системы счисления, порожденные тремя из четырех классов Ковача (Лемма 1.1):*

— многочленами $f_2 = x^\kappa + px^{\kappa-1} + px^{\kappa-2} + ... + px + p$, $2 \le p \in \mathbb{N}$, $q = p$;

— многочленами $f_1 = x^\kappa + q$, *если и только если* $q \ge 2$, $q = p$;

— многочленами $f_4 = x^\kappa + px^{\kappa-1} + p^2x^{\kappa-2} + ... + p^{\kappa-1}x + p^\kappa$, $2 \le p \in \mathbb{N}$, $q = p^\kappa$., $2 \le p \in \mathbb{N}$,

$$(t = n \cdot \text{НОК}(\kappa, \kappa+1)/\kappa,\ n \in \mathbb{N});$$

возможно эффективное вычисление координат унифицированного образа точки $\vec{u} \in F$ *методом выделения куба из покрытия* $\text{Tes}(F)$ *решетки* $\mathbb{Z}^k$ *фундаментальными областями генератора с использованием соотношения:*

$$u_i' = q^{-t}\tilde{u}_i (\text{mod}\, q^t), \qquad (11)$$

где $\tilde{u}_i$ - наименьший неотрицательный вычет класса $u_i \ (\text{mod}\, q^t)$.

Также в работе показано, что для *КСС*-генераторов, использующих канонические системы счисления, порожденные многочленами $f_3 = x^\kappa + x^{\kappa-1} + x^{\kappa-2} + ... + x + p$, $2 \le p \in \mathbb{N}$, $q = p$, при $\kappa = 2,3,...,9$, $q \in \{2,3,5,7\}$, $t = 2,...,100$, отображение (10) не является инъективным, и, следственно, не существует унификации методом выделения гиперкуба из покрытия $\text{Tes}(F)$.

В последующих разделах работы производится детальное исследование реализации предложенной общей методики синтеза многомерных генераторов на примере генератора Таусворта.

Определение 2.2. Назовем обобщенным генератором Таусворта (генератором *LFSR-CNS*) КСС-генератор, использующий последовательность цифровых векторов $\{\vec{Y}(n)\}$, определенную соотношением,

$$\vec{Y}(n) = (y(nL), y(nL+1), ..., y(nL+s-1)),\ L \in \mathbb{N},\ s \in \mathbb{N},\ n = 0,1,...,$$

где $y(n)$ — линейная рекуррентная последовательность порядка m над полем $\mathbb{F}_q$ максимального периода $T = q^m - 1$, удовлетворяющая соотношению

$$y(n) + a_{m-1}y(n-1) + ... + a_0 y(n-m) = 0; \quad a_0, ..., a_{m-1} \in \mathbb{F}_q, \quad a_0 \ne 0, \quad (y(0), ..., y(m-1)) \ne \vec{0},$$

где L назовем *шагом генератора LFSR-CNS,* s *— размерностью цифрового вектора.*

Для такого обобщенного генератора возможно получение аналитических оценок для различных вероятностных критериев.

Третья глава работы посвящена аналитическому исследованию свойств многомерного распределения на выходе генератора *LFSR-CNS***.**

Основными теоретическими тестами псевдослучайных последовательностей являются: вычисление периода псевдослучайной последовательности; исследование равномерности псевдослучайной последовательности (соответствие теоретическому распределению); исследование "геометрической" структуры последовательности на выходе генератора; исследование автокорреляции псевдослучайной последовательности.

В третьей главе получены следующие основные результаты.

Лемма 3.1. *Если шаг* L *генератора LFSR-CNS удовлетворяет соотношению*

$$НОД(L, q^m - 1) = 1, \qquad (12)$$

то последовательность на выходе LFSR-CNS является последовательностью периода $T = q^m - 1$.

Для исследования равномерности и статистической независимости множества точек на полном периоде генератора *LFSR-CNS* рассмотрим множество векторов

$$P = \{\vec{X}_0, \vec{X}_1, ..., \vec{X}_{T-1}\},\ \vec{X}_n = \left(x_n, x_{n+1}, ..., x_{n+k-1}\right),\ n = 0,1,...,T-1,\ T = q^m - 1,$$

где $\{x_n\}$ -- выходная последовательность обобщенного генератора.

В случае если $k = 1$, исследование равномерности распределения элементов множества P соответствует исследованию равномерности распределения на выходе генератора *LFSR-CNS*. В случае если $k > 1$, исследование равномерности распределения точек множества P в многомерном кубе эквивалентно исследованию статистической независимости элементов на выходе генератора *LFSR-CNS*.

Для генератора *LFSR-CNS* два случая требуют различного подхода при исследовании: (1) случай низких размерностей ($k \le m/L$) и (2) случай высоких размерностей ($k > m/L$).

В работе доказаны следующие утверждения.

Лемма 3.1. *Пусть* $k \le m/L$, *пусть* $\vec{h} \in \frac{1}{q^t}\mathbb{Z}^{\kappa k} \cap I^{\kappa k}$. *Пусть далее* $Z(\vec{h})$ *- количество индексов* n, $0 \le n < T = q^m - 1$, *такое что* $\vec{X}_n = \vec{h}$, *тогда*

$$Z(\vec{h}) = \begin{cases} q^{m-ks}, \vec{h} \ne \vec{0}; \\ q^{m-ks} - 1, \vec{h} = \vec{0}. \end{cases} . \qquad (13)$$

Из соотношения (13) следует, что на выходе генератора *LFSR-CNS* генерируются все точки многомерного единичного куба, координаты которых являются рациональными числами со знаменателем q^t. Если $k = 1$, $m = s$, то на выходе генератора порождаются все точки множества $\frac{1}{q^t}\mathbb{Z}^{\kappa k} \cap I^{\kappa k}$, кроме нулевой, каждая по одному разу.

Теорема 3.1. *При* $k \le m/L$, $Card(P) = T$, *звездное отклонение* $D_T^{*(\kappa k)}(P)$ *множества точек* P *удовлетворяет неравенствам:*

$$1 - (1 - q^{-t})^{k\kappa} \le D_T^{*(\kappa k)} \le \frac{q^{tk\kappa}}{q^{tk\kappa} - 1}\left(1 - \left(1 - \frac{1}{q^t}\right)^{k\kappa}\right) + \frac{1}{q^{tk\kappa} - 1}. \qquad (14)$$

Неравенство в правой части соотношения (14) является следствием удаления нуля из множества точек куба $\frac{1}{q^t}\mathbb{Z}^{\kappa k} \cap I^{\kappa k}$. В случае если множество P дополнено нулевой точкой, соотношение (14) может быть переписано в виде:

$$D_T^{*(\kappa k)} = 1 - (1 - q^{-t})^{k\kappa}. \qquad (15)$$

Заметим, что отклонение множества точек на полном периоде, например, известного матричного генератора многомерных векторов, удовлетворяет соотношению (15). Если координаты точек на выходе генератора *LFSR-CNS* являются рациональными числами со знаменателем q^t, данное соотношение является лучшим возможным, и является следствием "дискретизации" координат точек на выходе генератора.

Если в качестве критерия равномерности используется *KCC*-отклонение, то справедливо утверждение.

Теорема 3.5. *При* $k = 1$, *для последовательности на выходе генератора LFSR-CNS, справедлива следующая верхняя оценка KCC-отклонения:*

$$D_N^{CNS}(P) \le C\left(\frac{m^2\sqrt{T}}{N}\right). \qquad (16)$$

Из соотношения (16) следует, что определенные аналитические гарантии качества распределения точек на выходе генератора *LFSR-CNS* могут быть получены для участка периода длины $N > \sqrt{T}$. При $N < \sqrt{T}$ оценка (16) является более слабой, чем тривиальная.

Для случая высоких размерностей (с использованием обобщения (t,m,k)-сетей) доказаны следующие утверждения:

Теорема 3.6. *k- мерное CNS-отклонение канонической (t,m,k)-сети P, $card(P) = N$ в F_Q^k по основанию 2 удовлетворяет соотношению:*

$$ND_N^{CNS}(P) \le 2^t \sum_{i=0}^{k-1}\binom{m-t}{i},$$

или

$$ND_N^{CNS}(P) \le \frac{2^t}{(k-1)!}\left(\frac{1}{\log 2}\right)^{k-1}(\log N)^{k-1} + O(2^t(\log N)^{k-2}).$$

Теорема 3.7. *Для $k > m/L$, рассмотрим множество из q^m точек*

$$P = \{\vec{0}, \vec{x}_0, \vec{x}_1, ..., \vec{x}_{T-1}\},$$

где согласно определению генератора LFSR-CNS

$$\vec{x}_n = (x_n, x_{n+1}, ..., x_{n+k-1}),$$

где x_n — элемент последовательности LFSR-CNS с индексом n. Множество P образует обобщенную (t,m,k)-сеть в q-значной канонической системе счисления $(\mathbb{Z}^\kappa, \mathbf{M}, D)$ с параметром t, удовлетворяющим соотношению:

$$t = m - r^{(k)}(A,s),$$

где $r^{(k)}(A,s)$ — величина, зависящая от характеристического многочлена A используемого рекуррентного соотношения.

Путем выбора характеристического многочлена A можно добиться увеличения или уменьшения величины $D_T^{CNS(k)}$ для конкретного генератора.

В **четвертой главе** проводится дополнительное аналитическое исследование распределения координатных последовательностей на выходе генератора *LFSR-CNS* и зависимости свойств распределения на выходе генератора от его параметров. Доказаны следующие основные утверждения.

Лемма 4.2. *Период координатных последовательностей на выходе генератора $LFSR-CNS$ равен периоду многомерной последовательности LFSR-CNS $T^{(j)} = T = q^m - 1$.*

Лемма 4.3. *Пусть $\eta \in \{aq^{-t} \mid a = 0,1,...,q^{t-1}\}$ и пусть $Z(\eta)$ — количество различных значений индекса n, $0 \le n < T$ таких, что $u^{(j)}(n) = \eta$. Тогда справедливы соотношения*

$$Z(\eta) = \begin{cases} q^{m-s+(\kappa-1)t}, & \text{если } \eta \ne 0, \\ q^{m-s+(\kappa-1)t} - 1, & \text{если } \eta = 0. \end{cases}$$

Лемма 4.4. *Для множества $\{u^{(j)}(n)\}$ справедливы следующие оценки отклонения:*

$$D_T^*(\{u^{(j)}(n)\}) = q^{-t},\ D_N^*(\{u^{(j)}(n)\}) \ge q^{-t},\ N < T.$$

Полученные выше значения отклонения также являются следствием дискретизации элементов последовательностей $u^{(j)}(n)$ и не могут быть улучшены для фиксированных q и t.

Для исследования статистической независимости элементов последовательности $\{u^{(j)}(n)\}$ в работе рассматривается последовательность векторов:

$$\vec{\upsilon}(n)=\left(u^{(j)}(n),...,u^{(j)}(n+k-1)\right),\ 0\le n<T,$$

для которой справедливо следующее утверждение.

Лемма 4.5. *Пусть* $k<m/s$, $\vec{h}\in[0,1)^k\cap q^{-t}\mathbb{Z}^k$ *и пусть* $Z(\vec{h})$ *— количество различных значений индекса* n, $0\le n<T$, *таких что* $\vec{\upsilon}(n)=\vec{h}$. *Тогда справедливы соотношения:*

$$Z(\vec{h})=\begin{cases}q^{m-ks+(\kappa-k)t}, & \text{если } \vec{h}\neq 0,\\ q^{m-ks+(\kappa-k)t}-1, & \text{если } \vec{h}=0.\end{cases} \qquad (17)$$

Из соотношений (17) следует, что звездное отклонение множества $\{\vec{\upsilon}(n)\}$, $0\le n<T$ удовлетворяет соотношению $D_T^*(\{\vec{\upsilon}(n)\})=1-\left(1-q^{-t}\right)^k$. То есть, в случае низких размерностей ($k<m/s$), значения отклонения являются следствием дискретизации значений элементов последовательности и не могут быть улучшены при фиксированных q и t.

Также в главе 4 проведено исследование зависимости свойств бинарного генератора *LFSR-CNS* от его параметров и получены следующие практические рекомендации:

- рекомендуется использовать значения $L=m$ и $s=\kappa\lfloor L/\kappa\rfloor$;
- вычислительная эффективность (в терминах количества арифметических операций) алгоритма генерирования зависит от количества ненулевых коэффициентов характеристического многочлена $A(z)$; путем выбора характеристического многочлена наперед заданные ограничения на вычислительную эффективность синтезированного алгоритма могут быть удовлетворены;
- при использовании участка генерируемой псевдослучайной последовательности длины $N\ll\sqrt{T}$, использование примитивных характеристических многочленов с числом ненулевых коэффициентов, превосходящим $m/2$ (где m - порядок многочлена), является предпочтительным использованию многочленов с малым числом ненулевых коэффициентов (например, трехчленов или пятичленов) с точки зрения критерия равномерности распределения элементов рассматриваемого участка псевдослучайной последовательности;
- при использовании участка псевдослучайной последовательности длины $N\ll\sqrt{T}$, предпочтительным, в терминах критерия равномерности распределения элементов последовательности, является использование генераторов $LFSR-CNS$, использующих каноническую систему счисления, порожденную вторым классом Ковача (5);
- при использовании участка псевдослучайной последовательности длины $N\ll\sqrt{T}$, начальное состояние генератора рекомендуется инициализировать с использованием дополнительной бинарной псевдослучайной последовательности; начальные состояния с отношением количества нулей и единиц близким к $1/2$.

Пятая глава работы посвящена статистическому исследованию свойств последовательности на выходе генератора *LFSR-CNS* на участке периода и исследованию эффективности параллельного алгоритма генерации *LFSR-CNS*.

В качестве обязательных этапов исследования любого генератора многими авторами рассматриваются теоретическое исследование последовательности на выходе генератора на полном и неполном периоде, статистическое тестирование участка псевдослучайной последовательности на неполном периоде, тестирование генератора в задачах, приближенных к реальным. Очевидным и принципиальным методологическим недостатком численного (статистического) исследования псевдослучайных последовательностей является отсутствие

ответа на вопрос о достаточности ("полноте") множества используемых статистических или практических тестов. В диссертационной работе, статистическое исследование свойств генератора *LFSR-CNS* проведено с использованием т.н. "физических тестов" (высотный корреляционный тест, тест множественного случайного блуждания). Данные тесты *пройдены генератором LFSR-CNS.*

Также в работе проведено численное исследование генератора с использованием вычисления многомерных интегралов рассматриваемых в известной монографии Н.М. Коробова[4]. В этих тестах результаты, показанные генератором, *превосходят результаты*, полученные с использованием методов Монте-Карло, использованных в работе Коробова. В результате численного исследования генератора *LFSR-CNS* с использованием библиотеки *TestU01* (наиболее полной библиотеки статистических тестов ГСЧ[5]), получены результаты, свидетельствующие, что характеристики синтезированного генератора *LFSR-CNS* превосходят характеристики стандартных генераторов, доступных пользователям различных платформ (см. таблицы 5.1, 5.2, 5.4).

В таблицах 5.1 и 5.2 приведены результаты "физических тестов". Суть данного класса тестов, является максимально приближенное к практическому использованию генератора вычисление по методу Монте-Карло значений физических величин, для которых известны точные аналитические значения. Точные значения величин ϕ и d, рассмотренных в работе равны 0.5. Диапазон значений, полученный, с использованием генератора *LFSR-CNS* содержит точное значение, чего нельзя сказать о стандартных генераторах, использованных при тестировании.

Таблица 5.1- Оценки величины $\tilde{\phi}$ для высотного корреляционного теста различных генераторов

Генератор	$\tilde{\phi}$	Генератор	$\tilde{\phi}$
RANLUX4	0,5001±0,0001	*R250*	**0,4989±0,0002**
RANMAR	0,5000±0,0001	*R89*	**0,4984±0,0001**
MZRAN	0,5000±0,0001	*LFSR-CNS*	0,5000±0.0001

Таблица 5.2– Оценки величины $\tilde{d}$ для теста множественного случайного блуждания

Генератор	$\tilde{d}$	Генератор	$\tilde{d}$
RANLUX4	0,5000±0.0001	*RANMAR*	0,5001±0.0001
RANLUX3	0,4999±0.0001	*MZRAN*	0,5000±0.0001
RANLUX2	0,5000±0.0001	*R250*	0,4984±0.0001
RANLUX1	0,4999±0.0001	*R89*	0,4981±0.0001
RANLUX0	0,4991±0.0001	*LFSR-CNS*	0,5000±0.0001

В Таблице 5.4 представлено количество непройденных тестов из наборов статистических тестов *SmallCrush, Crush, BigCrush* тестов библиотеки *TestU01* различными генераторами.

---

[4] Коробов, Н.М. Теоретико-числовые методы в приближенном анализе [Текст] / Н.М.Коробов / МЦНМО, Москва. — 2004. – 288 с.

[5] TestU01: A Software Library in ANSI C for Empirical Testing of Random Number Generators [Electronic resource] // http://www.iro. umontreal.ca/~lecuyer

Таблица 5.4– Результаты тестирования генератора LFSR-CNS с использованием библиотеки TestU01

Генератор	*SmallCrush*	*Crush*	*BigCrush*
LCG(2^{31},65539, 1)	11	106	нет данных
Java.util.Random	1	9	21
Unix-random-32	5	101	нет данных
Unix-random-64	4	57	нет данных
Unix-random-128	2	13	19
GFSR(250,103)	1	8	14
Matlab-rand	нет данных	5	8
WELL1024(a)	0	4	4
LFSR-CNS	0	4	4

Непройденные тесты – это тесты, связанные с вычислением линейной сложности генерируемых последовательностей. Данные тесты не пройдены в силу особенностей базового генератора Таусворта, что является общей характеристикой всех методов генерации, обобщающих метод Таусворта. В качестве примера может быть приведен комбинированный генератор WELL1024(a), не проходящий те же тесты.

Также в пятой главе приведены результаты сравнительного исследования вычислительной сложности генератора *LFSR-CNS* относительно методов параллелизации и естественно векторных генераторов.

Показано, что вычислительная сложность алгоритма генерирования *LFSR-CNS* эквивалентна вычислительной сложности базового генератора Таусворта, не зависит от количества компьютеров / процессоров, используемых для генерации, и является линейной относительно порядка рекуррентного соотношения. Для матричного генератора, даже в случае использования рекуррентного соотношения первого порядка, вычислительная сложность зависит квадратично от количества параллельных процессоров, используемых для вычислений.

В работе показано, что характеристики равномерного распределения на выходе генератора *LFSR-CNS* являются лучшими по сравнению с генераторами, основанными на параллелизации одномерных схем, и сравнимыми с характеристиками матричных генераторов.

Таким образом, на примере обобщенного генератора Таусворта, показано, что КСС-обобщение методов синтеза одномерных ГСЧ позволяет получить генераторы, обладающие преимуществами как "естественно"-многомерных генераторов псевдослучайных точек, так и методов, основанных на параллелизации одномерных ГСЧ.

На защиту выносятся следующие результаты:

1) общий метод синтеза псевдослучайных последовательностей векторов многомерного пространства, основанный на использовании канонических систем счисления в многомерных решетках;
2) исследование свойств фундаментальных областей предложенных многомерных генераторов, синтез эффективных алгоритмов унификации;
3) аналитическое исследование многомерного распределения на выходе обобщенного генератора Таусворта на полном и неполном периоде, исследование статистической независимости отсчетов генерируемой многомерной последовательности;
4) исследование свойств координатных последовательностей на выходе обобщенного генератора Таусворта, исследование зависимости свойств генератора от его параметров и начальных условий.

Основные результаты опубликованы в следующих работах:

1) **Калугин, А.Н.** Реализация вычисления свертки по модулю составных чисел Мерсенна [Текст] / А.Н.Калугин // Новые информационные технологии : тезисы докладов XI Международной студенческой школы-семинара Т. 1. / МГИЭМ — Москва, 2003. — С.354-355.
2) **Калугин, А.Н.** Алгоритм безошибочного вычисления свертки в расширениях конечных полей [Текст] / А.Н.Калугин //Компьютерная оптика. — 2003. — Выпуск 25. — С. 134-140.
3) **Kalouguine, A.N.** Fast parallel algorithms for convolution in canonical number systems [Текст] / A.N. Kalouguine// 7-th International conference on Pattern Recognition and Image Analysis: New information Technologies (PRIA'2004): Conference Proceedings (vol.1-3). Vol. 1 / St. Petersburg Electrotechnical University. - St. Petersburg, 2004. — pp. 252-255.
4) **Chernov, V.** Factorization Ambiguity in Algebraic Number Fields: Schönhage-Strassen Algorithm [Электронный ресурс] / V. Chernov, A. Kalouguine//Abstracts of the 4th European Congress of Mathematics — KTH, Stockholm — 2004.
5) **Kalouguine, A.N.** 3D generalization for LSFR random point generator [Текст] / A. Kalouguine, V. Chernov // Proceedings of The 2-nd IASTED International Multi-Conference on Automation, Control and Information Technology (ACIT 2005), conference «Signal and Image Processing» / ACTA Press. — Novosibirsk, 2005.
6) **Калугин, А.Н.** Трехмерное обобщение генератора LFSR случайных точек [Текст] / А.Н. Калугин //Компьютерная оптика. — 2005. — Выпуск 27. — С. 131-134.
7) **Калугин, А.Н.** Модификация многомерных псевдослучайных последовательностей с использованием пары двойственных LFSR-CNS генераторов [Текст] /А.Н. Калугин //Компьютерная оптика. — 2006. — Выпуск 28. — С. 112-118.
8) **Калугин, А.Н.** Генератор LFSR-CNS: Экспериментальные исследования многомерной псевдослучайной последовательности [Текст] /А.Н. Калугин // Перспективные информационные технологии в научных исследования, проектировании и обучении (ПИТ-2006). Труды научно-технической конференции с международным участием, Том 2 / Самара, 2006. — С.101-106.
9) **Калугин, А.Н.** Генератор LSFR-CNS: Аналитическое исследование равномерности распределения [Текст] / А. Н. Калугин //Компьютерная оптика. — 2007. — Выпуск 31. — С. 58-62.
10) **Kalouguine, A.N.** Fractal Fundamental Domains of Canonical Number Systems. Some Applications to Monte Carlo and Randomized Quasi-Monte Carlo Methods in Realistic Image Synthesis [Текст] / A.N. Kalouguine // Proceedings of the 17th International Conference, Graphicon-2007 Russia, Moscow, June 22-27, 2007/ МГУ, Москва. — 2007.

Подписано в печать 16.05. 2008

Формат 60 х 84 1/16

Бумага офсетная

Усл. печ. л. 1,0

Тираж 100 экз.