

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ОБРАЗОВАНИЯ
ОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ КОМПЬЮТЕРНЫХ НАУК

МАТЕМАТИЧЕСКИЕ
СТРУКТУРЫ
И
МОДЕЛИРОВАНИЕ

Выпуск 15

Омск 2005

Журнал «Математические структуры и моделирование» – Омск:
Омск. гос. техн. ун-т, 2005. – Вып. 15. – 126 с.

ISBN 5-8149-0249-3

Редакционная коллегия

Главный редактор

А.К. Гуц

д-р физ.-мат. наук, профессор

В.А. Топоногов

д-р физ.-мат. наук, профессор

Институт математики СО РАН (г. Новосибирск)

Н.Ф. Богаченко

к-т физ.-мат. наук

А.А. Магазёв

к-т физ.-мат. наук

Е.В. Палешева

к-т физ.-мат. наук

Художественное оформление

В.В. Коробицын

Адрес научной редакции

Россия, 644077, Омск - 77, пр.Мира, 55-А

Омский государственный университет

факультет компьютерных наук

E-mail: guts@univer.omsk.su

msm@cmm.univer.omsk.su

ISBN 5-8149-0249-3

© Омский госуниверситет, 2005

МАТЕМАТИЧЕСКИЕ СТРУКТУРЫ И МОДЕЛИРОВАНИЕ

В журнале публикуются статьи, в которых излагаются результаты исследований по фундаментальной и прикладной математике, теоретической физике и размышления, касающиеся окружающей нас природы и общества.

Публикуются также статьи по информационным технологиям, компьютерным наукам, защите информации, философии и истории математики.

Объекты исследования должны быть представлены в форме некоторых математических структур и моделей.

Журнал является реферируемым. Рефераты статей публикуются в журналах «Zentralblatt für Mathematik» (Германия) и «Mathematical Reviews» (США).

Электронная версия журнала представлена в сети «Интернет» по адресам:

<http://cmm.univer.omsk.su>
<http://univer.omskcity.ru>

Журнал издается на коммерческие средства факультета компьютерных наук Омского государственного университета.

Наш E-mail:

msm@cmm.univer.omsk.su

Подробную информацию можно найти на Web-сервере:

<http://cmm.univer.omsk.su>

СОДЕРЖАНИЕ

Фундаментальная математика

- G.A. Noskov. *Coarsely geodesic metrics on reductive groups (after H. Abels and G. A. Margulis)* 5
- Д.Н. Василенко, Е.В. Головачев, С.Н. Чуканов. *Линеаризация гладкой нелинейной системы управления по входной и выходной информации* 18
- И.В. Рымар, В.А. Симахин, В.А. Шапцев. *Непараметрическое прогнозирование случайных процессов* 25
- А.А. Чемёркин. *О топологиях, порождаемых сходимостью* 34

Прикладная математика и моделирование

- О.А. Голованова, В.А. Когут, Е.Ю. Ачкасова. *Идентификация кинетических параметров процесса кристаллизации оксалата кальция* 41
- В.В. Коробицын, Ю.В. Фролова. *Алгоритм численного решения дифференциальных уравнений с разрывной правой частью* 46

Теоретическая физика

- С.В. Белим. *Влияние эффектов дальнего действия на мультикритическое поведение неупорядоченных сжимаемых систем* 55
- А.И. Бреев, В.В. Клишевич. *О кватернионной форме уравнения Дирака* 66
- В.В. Варламов. *Точное решение для поля $(1, 0) \oplus (0, 1)$ в терминах функций на группе Пуанкаре* 74
- Е.В. Палешева. *Внешняя кривизна 3-мерного пространства и энергетические затраты, необходимые для образования 4-мерной кротовой норы* 92
- А.В. Шалупаев, В.В. Клишевич. *О полях Яно и Яно-Киллинга в моделях Фридмана* 102

продолжение на след.странице

Математическая экономика

О.С. Ёлкина, Е.В. Гуревич, А.К. Гуц.
*Математическое моделирование стратегий
экономического поведения людей на рынке
труда*.....107

А.К. Гуц, О.С. Ёлкина. *Описание равновесий
на рынке труда с помощью структур
Кулакова-Михайличенко*112

Защита информации

Е.Н. Дудоров. *Возможные варианты
построения интеллектуальной системы
обнаружения несанкционированной работы
программного обеспечения*116

COARSELY GEODESIC METRICS ON REDUCTIVE GROUPS (AFTER H. ABELS AND G. A. MARGULIS)

G.A. Noskov

В статье изучаются свойства функций длины на группах.

We are going to study the **length functions** on a group G , that is the functions $g \rightarrow |g| : G \rightarrow \mathbb{R}$, satisfying the following axioms:

- Positivity: $|1| = 0$ and $|g| > 0$ for all nonidentical $g \in G$;
- Triangle inequality: $|gh| \leq |g| + |h|$, $g, h \in G$;
- Symmetry: $|g| = |g^{-1}|$, $g \in G$.

Any length function gives rise to a left invariant metric d on G as usual: $d(g, h) \stackrel{\text{def}}{=} |g^{-1}h|$. And conversely, any left invariant metric defines a length function.

There are plenty of left invariant Riemannian metrics on any connected real Lie group G , although all of them give rise to a Lipschitz equivalent distance functions. On the other hand if Ω is a compact symmetric neighbourhood of identity $1 \in G$ then we can define a **word length function** on G by $|g| = \min\{i \in \mathbb{Z}_+ \mid g \in \Omega^i\}$. One can easily see that d is quasi-isometric to the metric induced by any left invariant Riemannian metric on G . Thus the problem of classification of metrics of above type up to a quasi-isometry is trivial - any two of them are quasi-isometric. Both two classes of metrics, introduced above, have a common feature – they are «coarsely geodesic» (see below the definition).

Recently H. Abels and G. A. Margulis gave much more refined classification of coarsely geodesic left invariant proper metrics on reductive Lie groups up to coarse equivalence [1]. They have defined a class of so called **normlike pseudometrics** on a reductive group, and have proved the following theorem.

Теорема 1. *Let $\mathbb{G}$ be a reductive $\mathbb{R}$ -group and $G = \mathbb{G}(\mathbb{R})^0$. Then any left invariant coarsely geodesic proper metric is bounded distance away from a unique normlike pseudometric on G .*

Copyright © 2005 **G.A. Noskov**.

Omsk Branch of Institute of Mathematics, and Mathematisches Institute der Heinrich-Heine-Universität Duesseldorf.

E-mail: noskov@private.omsk.su

This theorem gives the picture of a metric «in a large scale» and favorably compare to the following result of V. Berestovskiy which is «local» in nature [2] (for example it does not apply to word metrics):

Теорема 2. *Any interior left invariant metric on a Lie group is the Carnot-Caratheodory-Finsler one.*

We give an exposition of Abels & Margulis result, sacrificing some generality, but not the ideas.

1. Coarse world

Определение 1. A parameterized curve (not necessarily continuous!)

$$\alpha(t) : [0, t_0] \rightarrow X$$

in a metric space (X, d) is called a C -coarse geodesic, $C \geq 0$ if $d(\alpha(s), \alpha(t)) \stackrel{C}{=} |s - t|$ for all $s, t \in [0, t_0]$. (The symbol $\stackrel{C}{=}$ means equality up to an error not exceeding C .) The space (X, d) is called C -coarsely geodesic, if any two points $x, y \in X$ can be connected by a C -coarse geodesic.

A map $f : X \rightarrow Y$ of metric spaces is (A, B) -uniform, $A, B \geq 0$ if

$$\forall x, x' \in X : d(x, x') \leq A \Rightarrow d(f(x), f(x')) \leq B$$

and f is called **uniform** if $\forall A \geq 0$ there is $B \geq 0$, such that f is (A, B) -uniform. For example, any C -coarse geodesic is $(B, B + C)$ -uniform for any $B \geq 0$.

A map $f : X \rightarrow Y$ of metric spaces is **proper** if the preimage of a bounded set is bounded.

We say that the metrics d_1, d_2 on a space X are **Hausdorff** or **coarsely equivalent** iff $|d_1 - d_2|$ is a bounded function on $X \times X$.

2. Reflection groups

We recall notation and facts about finite groups generated by reflections and which we use later on. Sufficient references for this are [3, 4].

Let V be a Euclidean space, i.e. a finite dimensional vector space with an inner product. The **reflection** in a hyperplane H is the linear transformation $s_H : V \rightarrow V$ which is the identity on H and is multiplication by -1 on the (one-dimensional) orthogonal complement $H^\perp$. A **finite reflection group** is a finite group W of linear transformations generated by reflections. We call the fixed-point subspace $V_0 = V^W$ the **inessential** part of V , and its orthogonal complement V_1 the **essential** part of V . Let $\mathcal{H}$ denote the set of all reflecting hyperplanes H with $s_H \in W$. For a reflection $s \in S$, we denote its reflecting hyperplane by H^s . The connected components of the complement of the union $\bigcup_{H \in \mathcal{H}} H$ in $\mathbb{R}^n$ are called **Weyl chambers**, and the closure of

a Weyl chamber is called a **closed Weyl chamber**. Any closed Weyl chamber A^+ is a fundamental set for W in $\mathbb{R}^n$, i.e. each W -orbit has precisely one point in A^+ .

In particular we have a Coxeter projection $\mathfrak{c} : \mathbb{R}^n \rightarrow A^+$. It implies that W acts simply transitively on the set of Weyl chambers or, equivalently, any closed Weyl chamber can be uniquely represented as wA^+ , $w \in W$. The hyperplane L_s defines two closed halfspaces of $\mathbb{R}^n$ and those one, containing A^+ , we denote by L_s^+ . We fix a W -invariant inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$. For each reflection s there is a unique unit vector α_s , pointing to L_s^+ and orthogonal to L_s and we call this vector an s -root. In terms of roots $A^+ = \{x \in \mathbb{R}^n \mid \langle x, \alpha_s \rangle \geq 0 \ \forall s \in S_0\}$.

Any Weyl chamber is a unique irreducible intersection of halfspaces and we call the corresponding hyperplanes the **walls** of the chamber. The set S_0 of all reflections in the walls of A^+ generates W . The vectors $\{\alpha_s \mid s \in S_0\}$ form a basis of V_1 . Therefore $V_1^\perp \cap A^+$ is a simplicial cone of dimension $|S_0|$, and we have an orthogonal decomposition $A^+ = V_0 + (V_1 \cap A^+)$. We also recall that the angle between any two roots, corresponding to different walls of A^+ is obtuse.

We call convex cone A^{++} spanned by $\{\alpha_s \mid s \in S_0\}$ the **dual** of A^+ . The linear functional ℓ on V is called **positive** with respect to A^+ if it takes nonnegative values on A^{++} . Let $W_x = \{w \in W \mid wx = x\}$ denote the stabilizer of a point $x \in \mathbb{R}^n$ in W . It is known that W_x is generated by $W_x \cap S_0$ for any $x \in A^+$.

Лемма 1. (Supporting functionals) *Let $\|\cdot\|$ be a W -invariant norm on V . For any nonzero $x \in V$ one can find a linear functional ℓ_x on $\mathbb{R}^n$ such that*

$$\|\ell\| = 1, \quad \ell_x(x) = \|x\|,$$

and ℓ_x is positive with respect to any Weyl chamber containing x .

Доказательство. It follows from the Hahn-Banach theorem that there exists a linear functional ℓ_x on $\mathbb{R}^n$ such that $\|\ell_x\| = 1$ and $\ell_x(x) = \|x\|$. Averaging ℓ_x over the stabilizer W_x of x in W and using W -invariance of $\|\cdot\|$, we may assume ℓ_x to be invariant under W_x . Fix a Weyl chamber A^+ , containing x . It remains to prove that ℓ_x is positive with respect to A^+ , that is $\ell_x(\alpha_s) \geq 0$ for every $s \in S_0$, then by invariance $\ell_x(\alpha_s) = 0$. If $sx \neq x$, then

$$\ell_x(sx) \leq \|\ell_x\| \|sx\| = \|x\| = \ell_x(x),$$

i. e. $\ell_x(x) - \ell_x(sx)$ is positive. Moreover, $x - sx$ is a positive multiple of α_s and therefore $\ell_x(x) - \ell_x(sx)$ is a positive multiple of $\ell_x(\alpha_s)$, hence $\ell_x(\alpha_s)$ is positive and thus ℓ_x is positive with respect to A^+ . Since ℓ_x is W_x -invariant we conclude that ℓ_x is positive with respect to any wA^+ , $w \in W_x$. It is easy to see that any Weyl chamber, containing x is of the form wA^+ for some $w \in W_x$. ■

3. Length functions on $\mathbb{R}^n$ and stable norms

Let $|\cdot|$ be a length function on additive group $\mathbb{R}^n$. If $a \in \mathbb{R}^n$ then by the triangle inequality the sequence $|ma|$, $m \in \mathbb{N}$ is subadditive and therefore the following limit

$$\|a\| \stackrel{def}{=} \lim_{m \rightarrow \infty} \frac{|ma|}{m}$$

exists for every $a \in \mathbb{R}^n$ and clearly $|a| \geq \|a\|$. We call $\|a\|$ a **stable norm** of a . It follows from definition that the stable norm is nonnegative, symmetric, homogeneous, satisfies triangle inequality, but it might happen that the stable norm is not positive.

Определение 2. A length function $|\cdot|$ on $\mathbb{R}^n$ is **proper**, iff both $|\cdot|$ and any Euclidean metric on $\mathbb{R}^n$ have the same system of bounded subsets.

The definition is stronger than that of given in [1], but this does not affect the results we are going to prove.

We need a criterion for a stable norm to be positive. For example it is so if the length function is coarsely geodesic. We wish to strengthen this remark as follows:

Лемма 2. (When the stable norm is a norm) Suppose $\mathbb{R}^n$ is given with a proper length function $|\cdot|$ which is « C -uniform» for some $C \geq 0$ in a sense that for each $a \in \mathbb{R}^n$ there is a $(1, C)$ -uniform curve $\alpha : [0, |a|] \rightarrow \mathbb{R}^n$, starting at 0 and ending distance at most C from a . Then the associated stable norm $\|\cdot\|$ is a norm.

Доказательство. It follows easily from properness of the metric that condition of lemma holds for standard Euclidean length function $|\cdot|_e$ with some constant, say C' . In particular for each $a \in \mathbb{R}^n, m \in \mathbb{N}$ there is a $(1, C')$ -uniform (with respect to $|\cdot|_e$) curve $\alpha : [0, |a|] \rightarrow \mathbb{R}^n$, starting at 0 and ending distance at most C' from ma . The sequence of inequalities

$$m|a|_e = |ma|_e \stackrel{C'}{=} |\alpha(|ma|)|_e \leq \sum_{i=1}^{\lfloor |ma| \rfloor} |\alpha(i) - \alpha(i-1)|_e + C' \leq C'(\lfloor |ma| \rfloor + 1)$$

implies that if a is nonzero, then $|ma|$ grows at least linearly with m for nonzero a , hence $\|a\|$ is positive. ■

4. Groups with Cartan projection and stable norms

Определение 3. Suppose that we are given:

- 1) A group G with a coarsely geodesic length function $|\cdot|$;
- 2) Its subgroup A with a fixed isomorphism $A \simeq \mathbb{R}^n$;
- 3) Action of a reflection group W on A and a Weyl chamber A^+ .

We say that the map $a(g) : G \rightarrow A^+$ is a **Cartan projection** if the following conditions are satisfied:

- CP1) The restriction of the map $a(g)$ to A is a Coxeter projection;
- CP2) The map $a(g)$ coarsely preserves the norm $|\cdot|$;
- CP3) «Triangle inequality»: $a(g) + a(h) - a(gh) \in A^{++}$ for any $g, h \in G$.

We will call an assembly $\mathcal{G} = (G, |\cdot|, A, W, A^+, a(g))$ a **group with Cartan projection**. We fix such one throughout this section. Equally with the above notation, we write sometimes, $a(g) \stackrel{\text{def}}{=} a_g$ for the Cartan projection. Clearly, for any $w \in W$ the map $wa(g) : G \rightarrow wA^+$ is again a Cartan projection.

Лемма 3. If the length function $|\cdot|$ is proper on A then the Cartan projection is uniform.

Доказательство. Note that

$$a_{gh} - a_g \in a_h - A^{++}$$

and

$$a_{gh} - a_g \in -a_{h^{-1}} + A^{++}$$

(by Axiom CP3). By properness a_h and $a_{h^{-1}}$ are in a bounded subset of $\mathbb{R}^n$, hence so is the intersection of the right hand sides, since

$$(a - A^{++}) \cap (b + A^{++})$$

is compact for any $a, b \in \mathbb{R}^n$ and hence so is the intersection of the right hand sides. It follows that $|a_{gh} - a_g|$ is bounded whenever $|h|$ is bounded, that is the Cartan projection is uniform. ■

By construction of Section 3. associated to the length function on G is the stable norm on $A \simeq \mathbb{R}^n$ and we are interested when it is positive.

Лемма 4. *The stable norm is positive on A .*

Доказательство. Fix nonzero $a \in A$ and let $g(t)$, $0 \leq t \leq |ma|$ be a coarse geodesic in G from 0 to ma . Note that any coarse geodesic is $(1, C)$ -uniform for some $C > 0$. Since Cartan projection is uniform, it follows that the projected curve $a_{g(t)}$ is uniform again. It follows that $\|a\|$ is positive by Lemma 2. ■

Лемма 5. *The stable norm on A is W -invariant.*

Доказательство. Since the restriction of Cartan projection to A is a Coxeter projection, and it coarsely preserves the norm, we see that the restriction of length function to A is coarsely W -invariant. Hence the stable norm is W -invariant. ■

Определение 4. We extend a stable norm onto the whole G via Cartan projection: $\|g\| \stackrel{\text{def}}{=} \|a_g\|$, $g \in G$. Similarly, if ℓ is a functional on $\mathbb{R}^n$, we extend it onto G by precomposing with Cartan projection.

Лемма 6. *Stable norm on G satisfies the triangle inequality:*

$$\|gh\| \leq \|g\| + \|h\|$$

for any $g, h \in G$. Moreover for any functional ℓ on $\mathbb{R}^n$, positive with respect to A^+ , we have

$$\ell(gh) \leq \ell(g) + \ell(h)$$

for any $g, h \in G$.

Доказательство. Triangle inequality for ℓ immediately follows by application of ℓ to the inclusion in the axiom CP3 above. Let $g, h \in G$. If $\|a_{gh}\|$ is zero then the assertion is obvious, so we assume that it is nonzero. According to Lemma 1, there exists a linear functional ℓ on $\mathbb{R}^n$ such that $\|\ell\| = 1$, $\ell(a_{gh}) = \|a_{gh}\|$, and ℓ is positive with respect to A^+ . Then applying triangle inequality for ℓ we obtain

$$\|gh\| \stackrel{\text{def}}{=} \|a_{gh}\| = \ell(gh) \leq \ell(g) + \ell(h) \stackrel{\text{def}}{=} \ell(a_g) + \ell(a_h) \leq \|a_g\| + \|a_h\| \stackrel{\text{def}}{=} \|g\| + \|h\|.$$

■

5. Miraculous elements

Let $\mathcal{G} = (G, |\cdot|, A, W, A^+, a(g))$ be a group with Cartan projection and let C_0 be a constant, such that the length function is C_0 -coarsely geodesic and the Cartan projection C_0 -coarsely preserves the length function.

Определение 5. We say that $a \in \mathbb{R}^n$ is C -good, $C \geq 0$, if $\|a\| \geq |a| - C$. We say that a linear functional ℓ is **almost supporting** at a if $\|\ell\| = 1$ and $\ell_i(a) \geq \|a\| - 1$.

Our aim is to produce C -good elements g in «every direction».

Лемма 7. *There is a constant C depending only on $\mathcal{G}$ and on natural n_0 such that for any $a \in A^+$ with $\|a\| > 1$ and for any finite family $\mathcal{L}$ of linear functionals on $\mathbb{R}^n$ coarsely supporting at a , positive with respect to A^+ and of cardinality at most n_0 there is an element $g \in G$ such that,*

$$|a_g| \stackrel{C}{=} \|a\| \text{ and } \ell(a_g) \geq \|a\| - C, \ell \in \mathcal{L}. \quad (1)$$

Доказательство. Let $g(t) : [0, t_0] \rightarrow G$ be a C_0 -coarse geodesic in G from 0 to ma , in particular $t_0 \stackrel{C_0}{=} |ma|$. We define

$$t_j = \frac{t_0}{m} j$$

for $j = 0, \dots, m$ and we define

$$g_j \stackrel{\text{def}}{=} g(t_{j-1})^{-1} g(t_j),$$

for $1 \leq j \leq m$. We have $ma = g_1 \cdots g_m$ and

$$|g_j| \stackrel{\text{def}}{=} d(g(t_{j-1}), g(t_j)) \stackrel{C_0}{=} \frac{t_0}{m} \stackrel{C_0/m}{=} \frac{|ma|}{m}$$

for all j . In particular

$$|a_{g_j}| \stackrel{C_0}{=} |g_j| \stackrel{C_0}{=} \frac{|ma|}{m} \stackrel{1}{=} \|a\|$$

for $m \gg 0$ and we conclude that

$$|a_g| \stackrel{2C_0+1}{=} \|a\|$$

for all j and $m \gg 0$ (take into account that the composition of relations $\stackrel{C}{=}, \stackrel{D}{=}$ is $\stackrel{C+D}{=}$). Thus, any g_j for $m \gg 0$ fulfils the first condition in the Lemma. Further, for $\ell \in \mathcal{L}$

$$\ell(g_j) \leq \|a_{g_j}\| \leq |a_{g_j}| \leq |g_j| + C_0 \leq \|a\| + C_0 + 1 \quad (2)$$

for $m \gg 0$. Since any ℓ is coarsely supporting and by triangle inequality from Lemma 6 we have

$$m(\|a\| - 1) \leq \ell(ma) \leq \sum_{j=1}^m \ell(g_j). \quad (3)$$

We wish to derive from 2 inequalities above that

$$\ell(g_j) \geq \|a\| - C$$

for all $\ell \in \mathcal{L}$ and universal constant C . For any C and any $\ell \in \mathcal{L}$ we easily conclude from (2),(3) that

$$b_{\ell,m} \stackrel{\text{def}}{=} \#\{j \mid 1 \leq j \leq m, \ell(g_j) \leq \|a\| - C\} \leq \frac{m(C_0 + 2)}{C_0 + C + 1}.$$

Thus, taking $C \geq (n_0 + 1)(C_0 + 2)$ we obtain that

$$\frac{b_{\ell,m}}{m} < \frac{1}{n_0 + 1}$$

for any $\ell \in \mathcal{L}$ and for sufficiently large m . From this and the definition of $b_{\ell,m}$ we deduce the existence of j , $1 \leq j \leq m$, such that for all $\ell \in \mathcal{L}$, $\ell(g_j) > \|a\| - C$ and we take $g = g_j$. ■

6. Cartan projections of geodesic curves

Let $\mathcal{G} = (G, |\cdot|, A, W, A^+, a(g))$ be a group with Cartan projection and let C_0 be a constant, such that the length function is C_0 -coarsely geodesic and the Cartan projection C_0 -coarsely preserves the length function. Our aim is to construct a continuous curve, consisting entirely of C -good elements and along which the given finite family of functionals grow coarsely with time.

Лемма 8. *There exists a constant $C > 0$, depending only on $\mathcal{G}$ and on a natural number n_0 such that for any $a \in A^+$ and an arbitrary family $\mathcal{L}$ of coarsely supporting at a functionals, positive with respect to A^+ , and of cardinality at most n_0 there exists a continuous curve $\alpha : [0, t_0] \rightarrow A^+$, $t_0 \stackrel{C}{=} |a|$ starting at 0, such that*

$$|\alpha(t)| \stackrel{C}{=} t \quad \text{and} \quad \ell(\alpha(t)) \geq t - C, \quad t \in [0, t_0], \quad \ell \in \mathcal{L}. \quad (4)$$

Доказательство. Let D be a constant exceeding both C_0 and the constant given by Lemma 7. In particular, there is an element $g \in G$ such that

$$|a_g| \stackrel{D}{=} \|a\| \quad \text{and} \quad \ell(a_g) \geq \|a\| - D, \quad \ell \in \mathcal{L}. \quad (5)$$

Let $g(t) : [0, t_0] \rightarrow G$, be a D -coarse geodesic in G connecting 1 and g , in particular $t_0 \stackrel{D}{=} |g|$. Since the Cartan projection coarsely preserves the length and $g(t)$ is a D -coarsely geodesic we conclude that $|a_{g(t)}| \stackrel{D}{=} |g(t)| \stackrel{D}{=} t$ and thus $|a_{g(t)}| \stackrel{2D}{=} t$, hence the curve $a_{g(t)}$ satisfies the first assertion of the Lemma with $C = 2D$.

Let $g'(t) = g(t)^{-1}g$, so that $g = g(t)g'(t)$. For $\ell \in \mathcal{L}$ we have

$$\|a\| - D \leq \ell(g) \leq \ell(g(t)) + \ell(g'(t)) \leq \ell(g(t)) + |a_{g'(t)}| \stackrel{D}{=} \ell(g(t)) + |g'(t)| =$$

$$= \ell(g(t)) + d(g(t), g(t_0)) \stackrel{D}{=} \ell(g(t)) + t_0 - t.$$

The equalities $|g| \stackrel{D}{=} \|a\|$ and $t_0 \stackrel{D}{=} |g|$ imply $t_0 \stackrel{2D}{=} \|a\|$, and substituting this to the above we get

$$\ell(g(t)) \geq t - 5D, \quad 0 \leq t \leq t_0,$$

thus the curve $a_{g(t)}$ satisfies the second assertion with $C = 5D$.

To make $a_{g(t)}$ continuous we define the curve $\alpha(t)$ as the curve which coincides with $a_{g(t)}$ at the integral moments $j, 0 \leq j \leq [t_0]$ and at the moment t_0 and which is linear on the segments $[j, j+1], 0 \leq j \leq [t_0] - 1, [[t_0], t_0]$. First note $g(t)$ and $a(g)$ are uniform maps, so there is a constant E depending only on $\mathcal{G}$, such that that $|\alpha(t) - \alpha([t])| \leq E, t \in [0, t_0]$. Let $C = E + 5D + 1$. We have

$$||\alpha(t)| - t| \leq |\alpha(t) - \alpha([t])| + ||\alpha([t])| - t| \leq E + 2D,$$

hence $|\alpha(t)| \stackrel{C}{=} t$ for all $t \in [0, t_0]$. (We ask the reader to forgive the conflict of notations for length function on G and for absolute value for reals). Finally, for any $t \in [0, t_0]$ and any $\ell \in \mathcal{L}$ we have

$$\ell(\alpha(t)) = \ell(\alpha([t])) + \ell(\alpha(t) - \alpha([t])) \geq [t] - 5D - |\alpha(t) - \alpha([t])| \geq t - E - 5D - 1.$$

■

7. Main Theorem

Теорема 3. *Let $\mathcal{G}$ be a group with Cartan projection. Suppose that the length function is proper on G . Then*

$$\sup_{g \in G} ||g| - \|g|| < \infty.$$

In particular any left invariant, coarsely geodesic, proper metric on G is bounded distance away from a unique normlike pseudometric on G .

Доказательство. The main case is $g \in A \simeq \mathbb{R}^n$, from which the general case follows easily by definition of the stable norm on G . Since our length function is proper, $\|\cdot\|$ is a norm. Fix a real $r > 0$ and let $B_r \subset A$ be a unit ball of radius r about origin. In this proof we define a continuous map $\varphi : B_r \rightarrow \mathbb{R}^n$, such that $\varphi(B_r)$ consists entirely of C -good points and the image contains $B_{0,r-C}$. The last claim is proved using a topological argument, namely the degree of maps between spheres.

Applying standard arguments, one can show that there exists a triangulation $\mathcal{T}$ of the boundary ∂B_r such that each simplex of $\mathcal{T}$ has $\|\cdot\|$ -diameter $\leq 1/2$ and lies entirely in some Weyl chamber. Let $\mathcal{T}_b$ denote the barycentric subdivision of $\mathcal{T}$. For any simplex σ in either $\mathcal{T}$ or $\mathcal{T}_b$, we denote by V_σ the set of vertices in σ . Let σ_x denote the smallest simplex in $\mathcal{T}$ which contains a given $x \in \partial B_r$.

According to Lemma 1, for any nonzero $x \in \mathbb{R}^n$, one can find a linear functional ℓ_x on $\mathbb{R}^n$ such that $\|\ell_x\| = 1$, $\ell_x(x) = \|x\|$, and ℓ_x is positive with respect to any

Weyl chamber containing x . Note that ℓ_x is almost supporting at any point of σ_x . Indeed, for $y \in \sigma_x$ we have

$$\ell_x(y) = \|x\| + \ell_x(y - x) \geq \|y\| - 2\|y - x\| \geq \|y\| - 1.$$

For each vertex $v \in \mathcal{T}_b$ we define the set of functionals

$$\mathcal{L}_v = \{\ell_u : u \in V_{\sigma_v}\}.$$

Note that this set satisfies the conditions of the Lemma 8. Indeed, the functionals are of norm 1, they are almost supporting at v by previous remark, and if wA^+ is any Weyl chamber, containing v , then the simplex σ_v is contained in wA^+ too, hence all functionals are positive with respect to wA^+ . Now by Lemma 8 for some $C \geq 0$, depending only on $\mathcal{G}$ we find a continuous curve $\alpha : [0, t_v] \rightarrow \mathbb{R}^n$, such that

$$|\alpha_v(t)| \stackrel{C}{=} t, \quad 0 \leq t \leq t_v, \quad t_v \stackrel{C}{=} \|v\|, \quad \alpha(0) = 0,$$

and

$$\ell_u(\alpha_v(t)) \geq t - C \quad \text{for any } u \in V_{\sigma_v}.$$

We now define a continuous map $\varphi : B_r \rightarrow \mathbb{R}^n$ as follows. Take any $z \in \partial B_r$, and let σ be a simplex in $\mathcal{T}_b$ which contains z . Represent z as a convex linear combination

$$z = \sum_{v \in V(\sigma)} \lambda_{v,z} v,$$

and define

$$\varphi(sz) \stackrel{\text{def}}{=} \sum_{v \in V_\sigma} \alpha_v(s\lambda_{v,z}t_v), \quad 0 \leq s \leq 1.$$

It is clear that the nonzero coefficients $\lambda_{v,z}$ in the decomposition above do not depend on the choice of σ . Therefore $\varphi(tz)$ does not depend on the choice of σ either. Since the curves $\{\alpha_v(s)\}$ are continuous, the map $\varphi : B_r \rightarrow \mathbb{R}^n$ is continuous.

Claim. $\varphi(B_r)$ consists of C -good points.

Since $\mathcal{T}_b$ is the barycentric subdivision of $\mathcal{T}$, one can find a vertex u in σ such that $u \in V_{\sigma_v}$ for all $v \in V_\sigma$. Then it follows

$$\begin{aligned} \|\varphi(sz)\| &\geq \ell_u(\varphi(sz)) = \sum_{v \in V(\sigma)} \ell_u(\alpha_v(s\lambda_{v,z}a_v)) \geq \sum_{v \in V_\sigma} s\lambda_{v,z}a_v - C \geq \\ &\sum_{v \in V_\sigma} |\alpha_v(s\lambda_{v,z}a_v)| - C \geq \left| \sum_{v \in V_\sigma} \alpha_v(s\lambda_{v,z}a_v) \right| - C = |\varphi(sz)| - C. \end{aligned}$$

Let $\{\varphi_\lambda\}, \lambda \in [0, 1]$ be a linear homotopy between identity and ϕ . It follows from inequalities above that

$$\|\varphi_\lambda(z)\| \geq \ell_u(\varphi_\lambda(z)) = \ell_u(\lambda\varphi(z) + (1 - \lambda)z) > r - C$$

for any $z \in \partial B_r$ and $0 \leq \lambda \leq 1$. We get from above that

$$B_{r-C-1} \cap \varphi_\lambda(\partial B_r)$$

for any $0 \leq \lambda \leq 1$. The following topological lemma finish the proof.

Лемма 9. *Suppose that a continuous map $f : B_r \rightarrow \mathbb{R}^n$ satisfies the following property: $t B_{r-C-1} \cap f_\lambda(\partial B_r)$ for any $0 \leq \lambda \leq 1$, where f_λ is a linear homotopy between identity and f . Then the image $f(B_r)$ contains B_{r-C-1} for all r sufficiently large.*

Доказательство. Arguing by contradiction, suppose that $f(B_r)$ does not contain a point $y \in B_{r-C-1}$. By assumption the restrictions of Id and f to ∂B_r are homotopic as maps into $\mathbb{R}^n - \{y\}$. Composing with projection map of the last space onto the sphere S_r , we obtain that the identity map of the sphere is homotopic to the constant map - this contradicts to the well known fact (the continuous maps of the sphere into itself of different degree are not homotopic). This contradiction proves the Lemma. ■

Theorem is proved. ■

8. Examples of groups with Cartan projection

The notion of a group with a Cartan projection is motivated by the classical Cartan decomposition for semisimple or, more generally, reductive Lie groups. Let

$$G = \mathbb{G}(\mathbb{R})^0$$

be the connected component of the identity of the group $\mathbb{G}(\mathbb{R})$ of $\mathbb{R}$ -rational points of a reductive $\mathbb{R}$ -group $\mathbb{G}$. Fix a proper coarsely geodesic length function $|\cdot|$ on G . Let $\mathbb{A}$ be a maximal $\mathbb{R}$ -split torus in $\mathbb{G}$ and $A = \mathbb{A}(\mathbb{R})^0$. The group A is isomorphic to $\mathbb{R}^n$ where $n = \dim A$. Let $A^+ \subset A$ be a Weyl chamber for the Weyl group $W = \mathcal{N}_G(A)/\mathcal{Z}_G(A)$.

Теорема 4. *The assembly $\mathcal{G} = (G, |\cdot|, A, W, A^+, a(g))$ is a group with Cartan projection.*

Доказательство. It is well known that G admits a Cartan decomposition $G = KA^+K$, where K is a suitable maximal compact subgroup of G and that associated «Cartan projection»

$$a(g) : G \rightarrow A^+$$

is well defined. Moreover, if $w \in W$ and $a \in A^+$ then $w(a) \in KaK$, and this implies that the restriction of $a(g)$ onto A is a Coxeter projection, thus CP1) is fulfilled. Since K is compact and the length function is proper, it follows that the Cartan projection coarsely preserves the length function, hence CP2).

The axiom CP3) is rather nontrivial and relies on the presentation of positive linear functionals as linear combinations of highest weights of rational representations [5]. Let $\pi : G \rightarrow GL(V)$ be a rational representation of the group G defined and irreducible over $\mathbb{R}$. We decompose V into the direct sum

$$V = \bigoplus_{\chi \in X(A)} V_\chi, \quad V_\chi \neq \{0\}, \quad (6)$$

of the weight spaces

$$V_\chi \stackrel{\text{def}}{=} \{v \in V \mid \pi(a)v = \chi(a)v \text{ for any } a \in A\},$$

where $X(A)$ is the group of rational characters of A . Let $\mu_\pi \in X(A)$ be the highest weight of the representation π . Then

$$\chi(a) \leq \mu_\pi(a) \text{ if } a \in A^+ \text{ and } V_\chi \neq \{0\}. \quad (7)$$

It is well known that one can introduce a $\pi(K)$ -invariant inner product on V with respect to which the transformations $\pi(a), a \in A$, are self-adjoint. Then the subspaces V_χ in the decomposition (6) are mutually orthogonal, and it follows from (7) that $\|\pi(g)\| = \mu_\pi(a(g))$, for any $g \in G$, where the norm is taken with respect to the inner product just defined. As a consequence, we get that for any $g, h \in G$,

$$\log \mu_\pi(a(gh)) \leq \log \mu_\pi(a(g)) + \log \mu_\pi(a(h)). \quad (8)$$

It is well known that any linear functional ℓ on $A = \mathbb{R}^n$, which is positive with respect to A^+ , can be represented as a positive linear combination

$$\ell = \sum_{i=1}^m b_i \log \mu_{\pi_i}, \quad b_i \geq 0,$$

where $\pi_i, 1 \leq i \leq m$, are rational representations of G , defined and irreducible over $\mathbb{R}$. Then it follows from (8) that for any $g, h \in G$,

$$\ell(a(gh)) \leq \ell(a(g)) + \ell(a(h)). \quad (9)$$

Since ℓ is an arbitrary positive functional we get from this and the definition of A^{++} that the axiom CP3) is satisfied. ■

9. Generalizations, questions, problems

The notion of a group with a Cartan projection used above is not enough to treat the case of reductive groups over local fields. The generalization is given in [1] and here we give a sketch. Namely we must allow not only $\mathbb{R}^n$ but any closed cocompact group D of $\mathbb{R}^n$. An action of a reflection group W on $\mathbb{R}^n$ should leave D invariant and there must be a compact symmetric set $M \subseteq \mathbb{R}^n$, containing 0, such that for each $w \in W$ there is an inclusion

$$wA^+ \subseteq (D \cap wA^+) + M.$$

We say that the map

$$a(g) : G \rightarrow A^+$$

is a Cartan projection if CP1), CP2) are satisfied and CP3) is satisfied in the following relaxed form:

$$a(g) + a(h) - a(gh) \in A^{++} + M$$

for any $g, h \in G$. The generalization includes the case of invariant metrics on $\mathbb{Z}^n$ as well as invariant metrics on $\mathbb{R}^n$, but seems it does not include the case of $\mathbb{Z}^n$ -invariant metrics on $\mathbb{R}^n$, which were treated by D. Burago [6].

Question 1. Is the analog of Abels-Margulis theorem for lattices in a reductive group valid? (Presumably not). The same question for the metrics on a reductive group invariant under translations by the elements of the lattice. (The answer is presumably yes).

Question 2. What is the relation of the triangle inequality for stable norm to the Gelfand-Naimark theorem about singular values of the product of matrices?

Question 3. What is the analog of Abels-Margulis theorem for nilpotent, solvable, general Lie groups? Look at the S. Krat paper.

Question 4 What is the structure of the asymptotic cone of a reductive group with a normlike pseudometric? Presumably they are **Minkowski buildings**. The question is related to the results of Kleiner-Leeb, Thornton, Parreaud.

Question 5 What is the relation between Abels-Margulis and Berestovsky theorems? For example, could one calculate a normlike pseudometric associated to a Carnot-Caratheodory-Finslerian metric on a reductive group?

Question 6. Let M be a set provided with two interior metrics d_1 and d_2 . Assume that a group G acts cocompactly on M by isometries with respect to both metrics and

$$\lim_{d_2(x,y) \rightarrow \infty} \frac{d_1(x,y)}{d_2(x,y)} = 1. \quad (10)$$

Due to a result of D. Burago [6], if $G = \mathbb{Z}^n$, then d_1, d_2 are coarsely equivalent. This fact means that all metrics on M diverge linearly or stay within a finite distance from each other. Burago raised the question for which groups the same statement could be true. The Abels-Margulis result easily implies the positive answer for metrics on reductive groups. D. Burago suggested two different directions. The first is the case of semi-hyperbolic groups, i.e., groups of isometries of a space whose curvature is bounded from above by 0. The other one is the case of nilpotent groups and first of all, the Heisenberg group. In some cases the fact that two metrics cannot diverge more slowly than linearly could be described as the finiteness of the Gromov-Hausdorff distance between the group with induced metric and its asymptotic cone. In the case of the abelian group $\mathbb{Z}^n$ the asymptotic cone is $\mathbb{R}^n$ and it lies within a finite Gromov-Hausdorff distance from $\mathbb{Z}^n$. The Gromov-Hausdorff distance between Heisenberg groups and its asymptotic cone is finite [7].

ЛИТЕРАТУРА

1. Abels H., Margulis G.A. Coarsely geodesic metrics on reductive groups // Preprint, Bielefeld, 2003.
2. Berestovskii V.N. Homogeneous manifolds with an intrinsic metric.II // Siber. math. J. 1989. V. 30. P.180-191.
3. Brown K.S. Buildings. Springer Monographs in Mathematics. Berlin, Heidelberg: Springer-Verlag, 1996.

4. Humphreys J.E. Reflection groups and Coxeter groups. Cambridge Studies in Advanced Mathematics, 29. Cambridge: Cambridge University Press, 1990.
5. Vinberg E.B. Lie groups and Lie algebras, III. Springer-Verlag, 1994.
6. Burago D. Periodic Metrics // Advances in Soviet Mathematics. 1992. V. 9. P.205-210.
7. Krat S.A. Asymptotic properties of the Heisenberg group // Zapiski Nauchnykh Seminarov S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI). Geom. i Topol. 1999. V.261, №. 4. P.125-154.

ЛИНЕАРИЗАЦИЯ ГЛАДКОЙ НЕЛИНЕЙНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ПО ВХОДНОЙ И ВЫХОДНОЙ ИНФОРМАЦИИ

Д.Н. Василенко, Е.В. Головачев, С.Н. Чуканов

Рассматриваются вопросы линеаризации гладких нелинейных систем управления с использованием входной/выходной информации с целью последующего приведения к нормальной форме и классификации.

Для классификации гладких нелинейных систем управления их необходимо привести к нормальной форме. Из линейной алгебры известно, что если пара (A, C) линейной системы

$$\frac{dx}{dt} = Ax + bu; \quad y = x;$$

наблюдаема, то можно найти такие матрицу T и вектор k , что

$$T(A + kC)T^{-1} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$CT^{-1} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

что и определяет нормальную форму наблюдаемой линейной системы [1–3].

Если пара (A, b) управляемая, то существует такие матрица T и вектор k ,

$$T(A + bk)T^{-1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & 0 & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}; \quad Tk = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix},$$

что и определяет нормальную форму управляемой линейной системы.

Рассмотрим задачи линеаризации гладких нелинейных систем управления с использованием входной или выходной (ИО - input/output) информации с целью последующего приведения к нормальной форме и классификации.

1. Линеаризация систем по выходной информации (output linearization)

Пусть система описывается соотношениями:

$$\begin{aligned} \frac{dx}{dt} &= f(x); \quad y = h(x); \quad x \in R^n; \\ f(\cdot) &\in R^n; \quad y \in R, \quad h(\cdot) \in R. \end{aligned} \quad (1)$$

Поставим задачу нахождения для (1) такого линеаризующего преобразования $z = F(x)$ или $x = T(z) = F^{-1}(z)$, что

$$\frac{\partial F}{\partial x} f(F^{-1}(z)) - Az = a(Cz);$$

$$h(F^{-1}(z)) = Cz.$$

Потребуем выполнения следующих условий при преобразовании $x = T(z)$:

- (1) $\dim(\text{span}\{dh, dL_f h, \dots, dL_f^{n-1} h\}) = n$;
- (2) найдется векторное такое поле τ , что:

$$L_\tau h = \dots = L_\tau L_f^{n-2} h = 0; \quad L_\tau L_f^{n-1} h = 1.$$

Локально последнее условие в координатах векторов x и z можно записать в форме:

$$\begin{bmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial [L_f h]}{\partial x} \\ \frac{\partial [L_f^2 h]}{\partial x} \\ \dots \\ \frac{\partial [L_f^{n-1} h]}{\partial x} \end{bmatrix} \times \frac{\partial T}{\partial z} = \begin{bmatrix} \frac{\partial h}{\partial z} \\ \frac{\partial [L_f h]}{\partial z} \\ \frac{\partial [L_f^2 h]}{\partial z} \\ \dots \\ \frac{\partial [L_f^{n-1} h]}{\partial z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & * \\ 0 & 0 & 1 & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 1 & * & * & \dots & * \end{bmatrix}. \quad (2)$$

Первый столбец соотношения (2) представляет PDE для τ :

$$\left[\frac{\partial h}{\partial x}, \frac{\partial [L_f h]}{\partial x}, \frac{\partial [L_f^2 h]}{\partial x}, \dots, \frac{\partial [L_f^{n-1} h]}{\partial x} \right]^T \cdot \tau = [0, 0, \dots, 0, 1]^T$$

Отображение $x = T(z)$ связано с векторным полем τ соотношением

$$\frac{\partial T}{\partial z} = (\tau, -ad_f \tau, \dots, (-1)^{n-1} ad_f^{n-1} \tau) T(z) = T^* \left(\frac{\partial}{\partial z_1}, \frac{\partial}{\partial z_2}, \dots, \frac{\partial}{\partial z_n} \right),$$

$$f(x) = T^* \left(a_1(z_n) \frac{\partial}{\partial z_1} + (z_1 + a_2(z_n)) \frac{\partial}{\partial z_2} + \dots + (z_{n-1} + a_n(z_n)) \frac{\partial}{\partial z_n} \right) T^{-1}(x),$$

где

$$\begin{aligned}\tau(x) &= T^* \left(\frac{\partial}{\partial z_1} \right) T^{-1}(x); \\ ad_f^k \tau(x) &= (-1)^k T^* \left(\frac{\partial}{\partial z_{k+1}} \right) T^{-1}(x); \\ \forall k : 0 \leq k \leq n-1;\end{aligned}$$

для вектора $\tau = [\tau_1, \tau_2, \dots, \tau_n]$:

$$\frac{\partial x}{\partial z_i} = \tau_i(x(z)).$$

Продифференцируем выход:

$$\begin{aligned}y &= h(x) = z_n; \\ \frac{dy}{dt} &= L_f h(x) = z_{n-1} - a_n(z_n); \\ \frac{d^2 y}{dt^2} &= L_f^2 h(x) = z_{n-2} - a_{n-1}(z_n) - L_f a_n(z_n); \\ \frac{d^n y}{dt^n} &= L_f^n h(x) = -L_f^0 a_1 - L_f^1 a_2 - \dots - L_f^{n-1} a_n.\end{aligned}\tag{3}$$

Из соотношений (3) получим требуемое преобразование:

$$\begin{aligned}z_n &= h(x) = F_n(x); \\ z_{n-1} &= L_f h(x) + a_n(h(x)) = F_{n-1}(x); \\ z_{n-2} &= L_f^2 h(x) + a_{n-1}(h(x)) + L_f a_n(h(x)) = F_{n-2}(x); \\ &\dots \\ z_{n-i} &= L_f^i h(x) + a_{n-i+1}(h(x)) + L_f^{i-1} a_{n-i+2}(h(x)) + \dots + L_f^{i-1} a_n(h(x)) = F_{n-i}(x); \\ &\dots \\ z_1 &= L_f^{n-1} h(x) + a_2(h(x)) + L_f a_3(h(x)) + \dots + L_f^{n-2} a_n(h(x)) = F_1(x);\end{aligned}$$

и

$$\begin{aligned}\frac{dz_1}{dt} &= -a_1(z_n); \\ \frac{dz_2}{dt} &= z_3 - a_2(z_n); \\ \frac{dz_{n-1}}{dt} &= z_{n-2} - a_{n-1}(z_n); \\ \frac{dz_n}{dt} &= z_{n-1} - a_n(z_n); \\ L_f^n h(x) + a_1(h(x)) + L_f a_2(h(x)) + \dots + L_f^{n-1} a_n(h(x)) &= \\ = L_f^n z_n + a_1(z_n) + L_f a_2(z_n) + \dots + L_f^{n-1} a_n(z_n) &= 0.\end{aligned}\tag{4}$$

Соотношение (4) ограничивает выбор функций выхода $a_1(h(x)), \dots, a_n(h(x))$. Например, в случае $n=2$ и $a_2(h(x)) \equiv 0$ это соотношение примет форму:

$$f_1 \left(\frac{\partial \left[f_1 \left(\frac{\partial h}{\partial x_1} \right) + f_2 \left(\frac{\partial h}{\partial x_2} \right) \right]}{\partial x_1} \right) + f_2 \left(\frac{\partial \left[f_1 \left(\frac{\partial h}{\partial x_1} \right) + f_2 \left(\frac{\partial h}{\partial x_2} \right) \right]}{\partial x_2} \right) + a_1(h(x)) = 0.$$

Величина

$$\delta(h(x)) = (L_f^n h(x) + L_f^{n-1} a_n(h(x)) + \dots + L_f^1 a_2(h(x)) + L_f^0 a_1(h(x)))^2$$

может служить мерой «неинтегрируемости» системы (1); при $\delta(h(x)) = 0$ система является интегрируемой. Минимизация величины

$$\int_M \delta(h(x)) dx \rightarrow \min$$

может быть использована при выборе преобразований $z = F(x)$ или

$$x = T(z) = F^{-1}(z)$$

в области M .

2. Линеаризация систем формированием управления (input linearization)

Рассмотрим систему:

$$\frac{dx}{dt} = f(x) + g(x)u; \quad x \in R^n; \quad f(\cdot) \in R^n; \quad g(\cdot) \in R^n, u \in R. \quad (5)$$

Поставим задачу нахождения для системы (5) такого управления

$$u = \alpha(x) + \beta(x)v$$

и преобразования координат

$$z = F(x), \quad z \in R^n, \quad F(\cdot) \in R^n,$$

что

$$F_*(f + g\alpha) \cdot F^{-1}(z) = Az,$$

$$F_*(g\beta)_i \cdot F^{-1}(z) = b_i,$$

$$A \in R^{n \times n}, \quad b_i \in R.$$

Тогда исходная система (5) может быть переписана как

$$\frac{dz}{dt} = f + g\alpha + (g\beta)u$$

и преобразованная (линейная):

$$\frac{dz}{dt} = Az + bv.$$

Если система (5) линеаризуемая, то преобразование $z = F(x)$ удовлетворяет условиям:

$$\frac{\partial z_i}{\partial x} g(x) = L_g z_i(x) = 0; \quad 1 \leq i \leq n-1;$$

$$\frac{\partial z_n}{\partial x} g(x) = L_g z_n(x) = \beta^{-1}(x);$$

$$L_f z_i(x) = z_{i+1}(x); \quad 1 \leq i \leq n-1;$$

$$L_f z_n(x) = -\alpha(x)\beta^{-1}(x).$$

Если $\varphi(x) = z_1(x)$ то

$$z_{i+1}(x) = L_f^i \varphi(x)$$

и функция $\varphi(x)$ удовлетворяет соотношениям:

$$L_g \varphi(x) = L_g L_f \varphi(x) = \dots = L_g L_f^{n-2} \varphi(x) = 0;$$

$$L_g L_f^{n-1} \varphi(x) \neq 0.$$

В случае если $g = g_1$, проблема линеаризации в пространстве состояний разрешима, если

$$\dim(\text{span}\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-1} g\}) = n$$

и распределение

$$\{g, \text{ad}_f g, \dots, \text{ad}_f^{n-2} g\}$$

инволютивно. В этом случае требуемые α, β и $z = F(x)$ определяются как

$$\alpha(x) = -\frac{L_f^n \varphi(x)}{L_g L_f^{n-1} \varphi(x)};$$

$$\beta(x)v = (L_g L_f^{n-1} \varphi(x))^{-1};$$

$$z_i(x) = L_f^{i-1} \varphi(x);$$

$$z(x) = (\varphi(x), L_f \varphi(x), \dots, L_f^{n-1} \varphi(x)).$$

Например, в случае $n = 2$:

$$L_g \varphi(x) = g \left(\frac{\partial \varphi}{\partial x_1} \right) = 0;$$

$$L_g L_f \varphi(x) = g \left(\frac{\partial \left(f_1 \frac{\partial \varphi}{\partial x_1} + f_2 \frac{\partial \varphi}{\partial x_2} \right)}{\partial x_1} \right) \neq 0.$$

3. Линеаризация систем с множеством входов и выходов (MIMO - Multi Input/Multi Output)

Рассмотрим отображение

$$F : x \rightarrow \text{col}(\xi_1(x), \dots, \xi_m(x)),$$

где

$$\xi_i(x) = [\varphi_i(x), L_f \varphi_i(x), \dots, L_f^{k_i-1} \varphi_i(x)]^T;$$

вектор относительных степеней $k = (k_1, \dots, k_m)$ определяется из выполнения соотношений

$$\langle L_f^\alpha \varphi_i(x), g(i) \rangle \neq 0; \quad 0 \leq \alpha \leq k_i - 2; \quad 1 \leq i \leq m; \quad \langle L_f^{k_i-1} \varphi_i(x), g_i \rangle \neq 0;$$

$$A = \|a_{ij}(x)\| = \|\langle L_f^{k_i-1} \varphi_i(x), g_j \rangle\|$$

для MIMO системы:

$$\frac{dx}{dt} = f(x) + \sum_{i=1}^m g_i u_i;$$

$$y_i = \varphi_i(x).$$

Выберем α, β из уравнений:

$$A(x)\alpha(x) = -[L_f^{k_1} \varphi_1(x), \dots, L_f^{k_m} \varphi_m(x)]^T;$$

$$A(x)\beta(x) = I.$$

Каждая i -ая подсистема описывается системой, представленной в нормальной форме:

$$\frac{d\xi_i}{dt} \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \xi_i + \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix} v_i,$$

где $\xi_i \in R^{k_i}$.

4. Заключение

В дальнейших работах авторы предполагают построить численные методы решения PDE для линеаризации гладкой нелинейной системы управления с использованием входной или выходной информации.

ЛИТЕРАТУРА

1. Fujimoto K., Sugie T. Freedom in coordinate transformation for exact linearization and its application to transient behavior improvement // Automatica. 2001. № 37. P.137-144.
2. Isidori A., Krener A., Giorgi C., Monaco S. Nonlinear decoupling via feedback: a differential geometric approach // IEEE Transactions on Automatic control. 1981. № 26. P.331-345.
3. Sugie T., Fujimoto K. Controller design for an inverted pendulum based on approximate linearization // International Journal of Robust and Nonlinear Control. 1998. № 8(7). P.585-597.

НЕПАРАМЕТРИЧЕСКОЕ ПРОГНОЗИРОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ

И.В. Рымар, В.А. Симахин, В.А. Шапцев

Имеется выборка $(X_1, \dots, X_N) = \vec{X}$, принадлежащая стационарному эргодическому процессу $X(t)$ с функцией распределения (ф.р.) $F(X(t))$. Пусть $X(t_i) = X_i$ и $t_{i+1} - t_i = \Delta t$, т.е. $X(t_i) = X(t_0 + \Delta t(i - 1))$, $i = 1, 2, \dots$. Процесс $X(t)$ удовлетворяет условию сильного перемешивания. Нас интересует задача прогнозирования процесса $X(t)$ на k шагов вперед, т.е. требуется по наблюдениям $X_1, \dots, X_N$ оценить значение X_{N+K} , где $0 < k \leq m$.

Рассмотрим следующую задачу прогнозирования. Имеется выборка $(X_1, \dots, X_N) = \vec{X}$, принадлежащая стационарному эргодическому процессу $X(t)$ с функцией распределения (ф.р.) $F(X(t))$. Пусть $X(t_i) = X_i$ и $t_{i+1} - t_i = \Delta t$, т.е. $X(t_i) = X(t_0 + \Delta t(i - 1))$, $i = 1, 2, \dots$. Процесс $X(t)$ удовлетворяет условию сильного перемешивания (с.п.) [1]. Нас интересует задача прогнозирования процесса $X(t)$ на k шагов вперед, т.е. требуется по наблюдениям $X_1, \dots, X_N$ оценить значение X_{N+K} , где $0 < k \leq m$. Как известно [2], оптимальной в среднеквадратическом смысле оценкой прогноза $\hat{X}_{N+K}$ является регрессия

$$\hat{X}_{N+K} = \int X_{N+K} dF(X_{N+K} | \vec{X}), \quad (1)$$

где $F(X_{N+K} | \vec{X})$ — условная функция распределения.

В непараметрическом случае, когда $F(X(t))$ неизвестна, вычислить (1) невозможно. Поэтому для оценки прогноза в точке t_{N+K} используют различные методы аппроксимации функции (1). В непараметрическом случае в качестве оценки функционала (1) можно использовать следующую оценку [5, 6]:

$$\hat{X}_{N+K} = \int X_{N+K} dF_N(X_{N+K} | \vec{X}), \quad (2)$$

где $F_N(X_{N+K} | \vec{X})$ — непараметрическая оценка $F(X_{N+K} | \vec{X})$.

Рассмотрим, как можно ввести $F_N(X_{N+K} | \vec{X})$. Условную ф.р. $F(X_{N+K} | \vec{X})$ можно записать в виде

$$F(X_{N+K} | \vec{X}) = \frac{\frac{\partial^m}{\partial \vec{x}} F(X_{N+K} | \vec{X})}{\frac{\partial^m}{\partial \vec{x}} F(\vec{x})}, \quad (3)$$

где m — размерность вектора $\vec{X}$ при условии, что соответствующие производные существуют. Далее,

$$\frac{\partial^m}{\partial \vec{x}} F(\vec{X})$$

— m -ая плотность распределения, так как

$$F(X_{N+K} | \vec{X}) = \int \cdots \int_{m+1} c(X_{N+K} - t_1) c(x_m - t_2) \cdots \\ \cdots c(x_m - t_{m+1}) dF(t_1, \dots, t_{m+1}),$$

то

$$\frac{\partial^b}{\partial \vec{x}} F(X_{N+K} | \vec{X}) = \int \cdots \int_{m+1} c(X_{N+K} - t_1) \delta(x_m - t_2) \cdots \\ \cdots \delta(x_m - t_{m+1}) dF(t_1, \dots, t_{m+1}) dF(t_1, \dots, t_{m+1}),$$

где $c(x)$ — функция Хевисайда и $\delta(x)$ — дельта-функция. В качестве оценки $F_N(X_{N+K} | \vec{X})$ возьмем

$$F_N(X_{N+K} | \vec{X}) = \frac{\frac{\partial^m}{\partial \vec{x}} F_N(X_{N+K}, \vec{X})}{f_N(\vec{X})}, \quad (4)$$

где $F_N(\vec{X})$ и $f_N(\vec{X})$ — эмпирическая m -мерная ф.р. и непараметрическая оценка m -мерной плотности, т.е.

$$F_N(t) = \frac{1}{N-m} \sum_{i=1}^{N-m} \prod_{j=1}^m c(t_j - x_{i+j}), \quad (5)$$

а

$$F_N(t_{N+K}, t_1, \dots, t_m) = \frac{1}{N-m-k} \sum_{i=1}^{N-m-k} \prod_{j=1}^m c(t_j - x_{i+j}) c(t_{N+K} - x_{i+m+k}),$$

и $f_N(\vec{X})$ определена в [8],

$$f_N(\vec{t}) = \int_{-\infty}^{\infty} \frac{1}{h_N^m} \prod_{j=1}^m K\left(\frac{t_j - z_j}{h_N}\right) dF_N(\vec{Z}),$$

$F_N(\vec{Z})$ определена в [4]. Подставляя (4) в (2), получаем

$$\hat{X}_{N+K} = \frac{\frac{1}{N-m-k} \sum_{i=1}^{N-m-k} X_{i+m+k-1} \frac{1}{h_N^m} \prod_{j=1}^m K\left(\frac{X_{N-j+1}-X_{i+j}}{h_N}\right)}{\frac{1}{N-m} \sum_{j=1}^{N-m} \frac{1}{h_N^m} \prod_{j=1}^m K\left(\frac{X_{N-j+1}-X_{i+j-1}}{h_N}\right)} \quad (6)$$

или

$$\hat{X}_{N+K} = \frac{A_N}{B_N}, \quad (7)$$

где

$$B_N = \int \dots \int \tilde{\delta}_N(X_{N-j-1} - t_j) dF(t_1, \dots, t_m), \quad (8)$$

$\tilde{\delta}_N$ — дельта-последовательность [7]. Так как

$$\hat{X}_{N+K} = \int t_{N+K} dF(t_{N+K} | t_1, \dots, t_m),$$

$$F_N(t_{N+K} | t_1, \dots, t_m) = \frac{\frac{\partial^m}{\partial t_1, \dots, \partial t_m} F_N(t_{N+K}, t_1, \dots, t_m)}{f_N(t_1, \dots, t_m)}$$

и

$$\begin{aligned} \frac{\partial^m}{\partial t_1, \dots, \partial t_m} F_N(t_{N+K}, t_1, \dots, t_m) &= \\ &= \frac{1}{N-m-k} \sum_{i=1}^{N-m-k} c(t_{N+k} - t_{i+k+m}) \prod_{j=1}^m \delta(t_j - x_{i+j}), \end{aligned}$$

то

$$\begin{aligned} \frac{\partial^m}{\partial t_1, \dots, \partial t_m} \tilde{F}_N(t_{N+K}, t_1, \dots, t_m) &= \\ &= \frac{1}{N-m-k} \sum_{i=1}^{N-m-k} c(t_{N+k} - t_{i+k+m}) \prod_{j=1}^m \tilde{\delta}(t_j - x_{i+j}). \end{aligned}$$

Тогда A_N запишется в виде

$$A_N = \int_{m+2} \dots \int z \delta(t_{N+k} - z) \prod_{j=1}^m \delta_N(x_{N-j+1} - t_j) dF(t_{N+k}, t_1, \dots, t_k) dz. \quad (9)$$

В качестве δ_N -последовательности можно взять

$$\delta_N(x) = \frac{1}{h_N} K\left(\frac{x}{h_N}\right),$$

где $h_N \rightarrow 0$, $N \rightarrow \infty$ и $K(x)$ — ядро, удовлетворяющее стандартным условиям [7]. Интегрируя (9) по z , получим

$$A_N = \int \dots \int_{m+2} t_{N+k} \prod_{j=1}^m \delta_N(x_{N-j+1} - t_j) dF(t_{N+k}, t_1, \dots, t_m). \quad (10)$$

Теорема 1. Пусть $x_1, \dots, x_N$ — выборка стационарного процесса со слабой зависимостью и выполнены следующие условия:

- 1) $F(x(t))$ — непрерывна и имеет все частные производные второго порядка;
- 2) $\sum_{\tau=1}^{\infty} \alpha(\tau)^{\delta/\delta+2} < \infty$, $\delta > 0$ при с.п. ($\alpha(\tau)$ — коэффициент с.п.) [1];
- 3) $\sum_{\tau=1}^{\infty} \sqrt{\beta(\tau)} < \infty$ при р.с.п. ($\beta(\tau)$ — коэффициент р.с.п.);
- 4) Функция $K(u)$ удовлетворяет условиям:
 - а) $\int K^2(u) du < \infty$;
 - б) $\int K(u) du = 1$;
 - в) $\int uK(u) du = 0$;
- 5) $h_N^m N \rightarrow 0$, при $N \rightarrow \infty$;
- 6) $\int t_{N+K}^2 dF(t_{N+K}, x_N, \dots, x_{N-m}) < \infty$.

Тогда $Y_N = Nh_N^m \left[\hat{x}_{N+K} - \frac{EA_N}{EB_N} \right]$ имеет асимптотически нормальное распределение с нулевым средним и дисперсией G^2 , где G^2 определена ниже и E — знак математического ожидания.

Доказательство. Рассмотрим B_N .

$$\begin{aligned} E\left(B_N \middle| X_N, \dots, X_{N-m}\right) &= \int \dots \int_m \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) dF(t_1, \dots, t_m) = \\ &= \int \dots \int_m \prod_{j=1}^m \frac{1}{h_N^m} K\left(\frac{x_{N-j+1} - t_j}{h_N}\right) f(t_1, \dots, t_m) dt_1, \dots, dt_m. \end{aligned}$$

Делая замену переменной $(x_{N-j+1} - t_j)/h_N = u_j$, $t_j = (x_{N-j+1} - h_N u_j)$, $dt_j = -h_N du_j$, получаем:

$$E\left(B_N \middle| X_N, \dots, X_{N-m}\right) =$$

$$= \int \dots \int_m \prod_{j=1}^m K(u_j) f(x_N - h_N u_1, \dots, x_{N-m} - h_N u_m) du_1, \dots, du_m = \dots$$

Используя разложение функции $f(\dots)$ по формуле Тейлора и учитывая, что $h_N \rightarrow 0$:

$$\begin{aligned} f(x_N - h_N u_1, \dots, x_{N-m} - h_N u_m) &= f(x_N, \dots, x_{N-m}) + \\ &+ h_N \sum_{j=1}^m \frac{\partial}{\partial x_{N-j+1}} f(x_N, x_{N-j+1}, \dots, x_{N-m}) u_j + \\ &+ \frac{h_N^2}{2} \sum_{j=1}^m \sum_{i=1}^m \frac{\partial^2}{\partial x_{N-j+1} \times \partial x_{N-i+1}} f(x_N, \dots, x_{N-j+1}, x_{N-m}) u_i u_j + O[h_N^3 \vec{u}]. \end{aligned}$$

В результате будем иметь

$$\begin{aligned} f(X_N, \dots, X_{N-m}) &\left[\int K(u) du \right]^m + \\ &+ h_N \left[\int u K(u) du \right] I^{m-1} \sum_{j=1}^m \frac{\partial}{\partial X_{N-j+1}} f(x_N, \dots, x_{N-j+1}, x_{N-m}) + \\ &+ \frac{1}{2} h_N^2 \left[I^{m-1} \int u^2 K(u) du \sum_{j=1}^m \frac{\partial^2}{\partial X_{N-j+1}^2} f(\dots) + \right. \\ &\left. + \left(\int u K(u) du \right)^2 I^{m-2} \sum_{j=1}^m \sum_{\substack{i=1 \\ i \neq j}}^m \frac{\partial^2}{\partial X_{N-j+1} \partial X_{N-i+1}} f(\dots) \right] + o(h_N^3), \end{aligned}$$

где $I = \int K(u) du$.

$$\begin{aligned} E(B_N | X_N, \dots, X_{N-m}) &= f(X_N, \dots, X_{N-m}) + \\ &+ \frac{1}{2} h_N^2 \int u^2 K(u) du \sum_{j=1}^m \frac{\partial^2}{\partial X_{N-j+1}^2} f(X_N, \dots, X_{N-j+1}, \dots, X_{N-m}) + o(h_N^3). \end{aligned}$$

Найдем дисперсию B_N :

$$\begin{aligned} D(B_N | X_N, \dots, X_{N-m}) &= \int \dots \int_m \left[\prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) \right]^2 dF(t_1, \dots, t_m) - \\ &- \left(E(B_N | X_N, \dots, X_{N-m}) \right)^2 + 2 \sum_{\lambda=1}^{m(N)+t_k-t_1} \left[\int \dots \int_{m+\lambda} \left[\prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) \times \right. \right. \\ &\left. \left. \times \delta_N(X_{N-j+1} - t_{j+\lambda}) \right] dF_{m+\lambda}(t_1, \dots, t_m) - \left(E(B_N | X_N, \dots, X_{N-m}) \right)^2 \right], \end{aligned}$$

$$D\left(B_N \middle| X_N, \dots, X_{N-m}\right) \xrightarrow{N \rightarrow \infty} \frac{1}{N h_N^m} f\left(X_N, \dots, X_{N-m}\right) \left[\int K^2(u) du \right]^2.$$

Рассмотрим A_N . Из (8) имеем:

$$A_N = \int \dots \int_{m+1} t_{N+k} \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) dF_N(t_{N+k}, t_1, \dots, t_m).$$

Тогда

$$\begin{aligned} E\left(A_N \middle| X_N, \dots, X_{N-m}\right) &= \int \dots \int_{m+1} \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) dF(t_1, \dots, t_m) = \\ &= \int \dots \int_{m+1} t_{N+k} \prod_{j=1}^m \frac{1}{h_N^m} K\left(\frac{X_{N-j+1} - t_j}{h_N}\right) f(t_{N+k}, t_1, \dots, t_m) dt_{N+k} dt_1 \dots dt_m \end{aligned}$$

подстановкой

$$\frac{X_{N-j+1} - t_j}{h_N} = u_j, \quad j = \overline{1, m}$$

получаем

$$\int \dots \int_{m+1} t_{N+k} \prod_{j=1}^m K(u_j) f(X_N - h_N u_1, \dots, X_{N-m} - h_N u_m, t_{N+k}) dt_{N+k} du_1 \dots du_m.$$

Используя разложение функции $f(\dots)$ и так как $h_N \rightarrow 0$, будем иметь

$$\begin{aligned} &\int t_{N+k} f(t_{N+k}, X_N, \dots, X_{N-m}) dt_{N+k} \left[\int K(u) du \right]^2 + \\ &+ \left(\int K(u) du \right)^{m-1} h_N \int u K(u) du \sum_{j=1}^m \int t_{N+k} \times \\ &\quad \times \frac{\partial}{\partial X_{N-j+1}} f(t_{N+k}, X_N, \dots, X_{N-j+1}, \dots, X_{N-m}) dt_{N+k} + \\ &+ \frac{1}{2} h_N^2 \left(\int K(u) du \right)^{m-1} \left(\int u^2 K(u) du \right) \sum_{j=1}^m \int t_{N+k} \times \\ &\quad \times \frac{\partial^2}{\partial X_{N-j+1}^2} f(t_{N+k}, t_N, \dots, t_{N-m}) dt_{N+k} + \\ &+ \left(\int K(u) du \right)^{m-2} \left(\int u K(u) du \right)^2 \sum_{j=1}^m \sum_{\substack{i=1 \\ i \neq j}}^m \int t_{N+k} \times \\ &\quad \times \frac{\partial^2}{\partial X_{N-j+1} \partial X_{N-i+1}} f(t_{N+k}, t_N, \dots, t_{N-m}) dt_{N+k} + o(h_N^3). \end{aligned}$$

При условиях 4) теоремы получаем

$$E \left(A_N \middle| X_N, \dots, X_{N-m} \right) = \int t_{N+k} f(t_{N+k}, X_N, \dots, X_{N-m}) dt_{N+k} + \frac{1}{2} h_N^2 \times \\ \times \left(\int u^2 K(u) du \right) \sum_{j=1}^m \int t_{N+k} \frac{\partial^2}{\partial X_{N-j+1}^2} f(t_{N+k}, t_N, \dots, t_{N-m}) dt_{N+k} + o(h_N^3).$$

Найдем дисперсию A_N . Имеем

$$D \left(A_N \middle| X_N, \dots, X_{N-m} \right) = \int \dots \int_{m+1} t_{N+k}^2 \times \\ \times \prod_{j=1}^m \delta_N^2(X_{N-j+1} - t_j) dF(t_{N+k}, t_1, \dots, t_m) - \left(E \left(A_N \middle| X_N, \dots, X_{N-m} \right) \right)^2 + \\ + 2 \sum_{\lambda=1}^m \left[\int \dots \int_{m+2} t_{N+k} \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) t_{N+k+\lambda} \times \right. \\ \times \prod_{j=1}^m \delta_N(X_{N-j+1} - t_{j+\lambda}) dF_{m+2(\lambda+1)}(t_{N+k}, t_{N+k+\lambda}, t_1, \dots, t_m) - \\ \left. - \left(E \left(A_N \middle| X_N, \dots, X_{N-m} \right) \right)^2 \right] + \\ + 2 \sum_{\lambda=m+1}^{\infty} \left[\int \dots \int_{2(m+1)} t_{N+k} \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) t_{N+k+\lambda} \times \right. \\ \times \prod_{j=1}^m \delta_N(X_{N-j+1} - t_{j+\lambda}) dF_{2(m+1)}(t_{N+k}, t_{N+k+\lambda}, t_1, \dots, t_m, t_{1+\lambda}, \dots, t_{m+\lambda}) - \\ \left. - \left(E \left(A_N \middle| X_N, \dots, X_{N-m} \right) \right)^2 \right].$$

Можно показать, что при $N \rightarrow \infty$ и условиях теоремы 2), 3)

$$N h_N^2 D \left(A_N \middle| X_N, \dots, X_{N-m} \right) \xrightarrow{N \rightarrow \infty} \\ \xrightarrow{N \rightarrow \infty} \int t_{N+k}^2 f(t_{N+k}, X_N, \dots, X_{N-m}) dt_{N+k} \left[\int K^2(u) du \right]^m < \infty.$$

Найдем $\text{cov} \left(A_N, B_N \middle| X_N, \dots, X_{N-m} \right)$:

$$\text{cov} \left(A_N, B_N \middle| X_N, \dots, X_{N-m} \right) = \int \dots \int_{2m+1} t_{N+k} \prod_{j=1}^m \delta_N(X_{N-j+1} - t_j) t_{N+k+\lambda} \times \\ \times \delta_N(X_{N-j+1} - t_{j+m}) d\text{cov} [F_N(t_{N+k}, t_1, \dots, t_m), F_N(t_{m+1}, \dots, t_{2m})].$$

Можно показать (как и выше, см. [7]), что

$$\begin{aligned} \text{cov} \left(A_N, B_N \middle| X_N, \dots, X_{N-m} \right) &\xrightarrow{N \rightarrow \infty} \frac{1}{Nh_N^m} \left[\int K^2(u) du \right]^m \int t_{N+k} \times \\ &\times f(t_{N+k}, X_N, \dots, X_{N-m}) dt_{N+k} = \frac{1}{Nh_N^m} \left[\int K^2(u) du \right]^m \times \\ &\times f(X_N, \dots, X_{N-m}) \int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}) < \infty. \end{aligned}$$

При $N \rightarrow \infty$ имеем

$$\begin{aligned} E(A_N | X_N, \dots, X_{N-m}) &\cong f(X_N, \dots, X_{N-m}) \int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}), \\ E(B_N | X_N, \dots, X_{N-m}) &\cong f(X_N, \dots, X_{N-m}), \\ D(A_N | X_N, \dots, X_{N-m}) &\cong \frac{1}{Nh_N^m} \left[\int K^2(u) du \right]^m f(X_N, \dots, X_{N-m}) \times \\ &\times \int t_{N+k}^2 dF(t_{N+k} | X_N, \dots, X_{N-m}), \\ \text{cov}(A_N, B_N | X_N, \dots, X_{N-m}) &\cong \frac{1}{Nh_N^m} \left[\int K^2(u) du \right]^m f(X_N, \dots, X_{N-m}) \times \\ &\times \int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}). \end{aligned}$$

С учетом условий 2), 3) теоремы достаточно просто показать, что двумерная величина (A_N, B_N) удовлетворяет условиям многомерной центральной предельной теоремы, т.е. имеет двумерное нормальное распределение с вектором средних (EA_N, EB_N) и ковариационной матрицей

$$\begin{bmatrix} DA_N & \text{cov}(A_N, B_N) \\ \text{cov}(A_N, B_N) & DB_N \end{bmatrix},$$

$$D\hat{X}_{N+k} \cong (EB_N)^2 - 2EA_N(EB_N)^{-3} \text{cov}(B_N, A_N) + (EA_N)^2(EB_N)^{-4} DB_N.$$

Окончательно получаем:

$$\begin{aligned} D\hat{X}_{N+k} &\cong \frac{Q^m}{Nh_N^m} \left[f^{-2} \tilde{a}(t^2) - 2fa(\dots) \frac{1}{f^3} fa(\dots) + f^2 a^2(\dots) \frac{1}{f^4} f \right] = \\ &= \frac{Q^m}{Nh_N^m} \frac{1}{f} [\tilde{a}(t^2) - a(\dots)] = \\ &= \frac{Q^m}{Nh_N^m} \frac{1}{f} \left[\int t_{N+k}^2 dF(t_{N+k} | X_N, \dots, X_{N-m}) - \right. \\ &\quad \left. - \left(\int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}) \right)^2 \right], \end{aligned}$$

где

$$\begin{aligned} Q &= \int K^2(u) du, \\ f &= f(X_N, \dots, X_{N-m}), \\ \tilde{a}(t^2) &= \int t_{N+k}^2 dF(t_{N+k} | X_N, \dots, X_{N-m}), \\ a(\dots) &= \int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}). \end{aligned}$$

Таким образом, из теоремы 2.1.1 [7] непосредственно следует, что

$$Nh_N^m \left[\hat{X}_{N+k} - \frac{EA_N}{EB_N} \right]$$

будет иметь асимптотически нормальное распределение с нулевым средним и дисперсией

$$G^2 = \frac{Q^m}{f(X_N, \dots, X_{N-m})} \left[\int t_{N+k}^2 dF(t_{N+k} | X_N, \dots, X_{N-m}) - \left(\int t_{N+k} dF(t_{N+k} | X_N, \dots, X_{N-m}) \right)^2 \right],$$

что и требовалось доказать. ■

ЛИТЕРАТУРА

1. Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины. М.: Наука, 1965.
2. Левин Б.Р. Теоретические основы статистической радиотехники, т.2. М.: Сов. Радио, 1975.
3. Чуев Ю.В. и др. Прогнозирование количественных характеристик процессов. М.: Сов.Радио, 1975.
4. Дмитриев Ю.Г. и др. Непараметрическое оценивание функционалов по стационарным выборкам. Томск: Изд-во ТГУ, 1974.
5. Симахин В.А. Непараметрическое прогнозирование стационарных процессов // Экономико-математические методы в планировании и управлении. Челябинск, 1977.
6. Рымар Т.Н., Симахин В.А. Непараметрическое прогнозирование стационарных процессов со слабой зависимостью // Мат-лы IV-й Всесоюзной конф. «Перспективы и опыт внедрения статистических методов в АСУ ТП». Тула, 1990.
7. Добровидов А.В., Кошкин Г.М. Непараметрическое оценивание сигналов. М.: Наука, 1997. 334с.
8. Rosenblat M. Density estimates and Markov sequences // Nonparametric techniques in statistical inference. Cambridge, University Press, 1970. P.199-210.

О ТОПОЛОГИЯХ, ПОРОЖДАЕМЫХ СХОДИМОСТЬЮ

А.А. Чемёркин

Показано, что произвольный класс направленностей на множестве порождает на этом множестве топологию. Также доказано, что секвенциальная топология ассоциированная с локально выпуклой может не быть линейной.

Хорошо известно, что понятия сходимости и топологии тесно связаны, например, в [1] вводится понятие класса сходимости, который порождает топологию. В первом предложении данной статьи показано, что по произвольному классу направленностей можно построить сильнейшую топологию, в которой все эти направленности будут сходиться, причем класс этот может не являться классом сходимости. Далее в статье рассматривается топология, порожденная сходящимися в некоторой исходной топологии последовательностями, приводятся основные свойства такой топологии (полученные автором ранее и опубликованные в статье [2]), а также доказываются два новых результата.

Основные понятия функционального анализа и общей топологии используются в статье без предварительных пояснений (см., например, [1, 3–5]).

Для подмножества A векторного пространства X через $\text{l.h. } A$ и $\text{abs co } A$ обозначим соответственно линейную и абсолютно выпуклую оболочку множества A , через 0 будем обозначать нулевой элемент векторного пространства X . Если (X, τ) — топологическое пространство, то замыкание множества $A \subset X$ в топологии τ будем обозначать $\text{cl } A$ (или $\text{cl}_\tau A$), через $\text{Cl } X$ (или $\text{Cl}(X, \tau)$) обозначаем множество всех замкнутых подмножеств X , через $\mathcal{O}_x$ (или $\mathcal{O}_x^X$) — совокупность всех окрестностей точки $x \in X$.

1. Порождение топологии произвольным классом направленностей

Пусть X — непустое множество, $\mathcal{P}(X)$ — множество всех подмножеств множества X .

Определение 1. Пусть $J \neq \emptyset$, отношение σ на J называется направлением, если выполнены условия:

- 1) $\forall j \in J \ (j, j) \in \sigma$;
- 2) $\forall i, j, k \in J \ (i, j) \in \sigma, (j, k) \in \sigma \Rightarrow (i, k) \in \sigma$;
- 3) $\forall i, j \in J \ \exists k \in J : (k, i) \in \sigma, (k, j) \in \sigma$.

Далее вместо $(i, j) \in \sigma$ будем писать $i \succcurlyeq j$ (или $j \preccurlyeq i$). Множество J с заданным на нем направлением $\succcurlyeq$ называется направленным множеством.

Определение 2. Семейство $\{x_j\}_{j \in J}$ элементов из X называется направленностью, если его множество индексов J является направленным множеством.

Определение 3. Пусть X — топологическое пространство, $\{x_j\}_{j \in J}$ — направленность в X , будем говорить, что направленность $\{x_j\}_{j \in J}$ сходится к элементу $x \in X$ или, что x является пределом направленности $\{x_j\}_{j \in J}$ (пишем $x_j \rightarrow x$ и $x = \lim_{j \in J} x_j$), если

$$\forall V \in \mathcal{O}_x \ \exists j_V \in J : \forall j \succcurlyeq j_V \ x_j \in V.$$

Пусть Σ — некоторое множество направленностей элементов из X . Пусть также определен оператор $\lambda : \Sigma \rightarrow \mathcal{P}(X)$, для любой направленности $s \in \Sigma$ назовем $\lambda(s) \subset X$ множеством ее пределов.

Предложение 1. На X существует сильнейшая топология τ , в которой направленности $S \in \Sigma$ сходятся к элементам $\lambda(S)$, причем эта топология однозначно определяется парой (Σ, λ) .

Доказательство. Определим класс подмножеств τ множества X следующим образом

$$A \in \tau \Leftrightarrow \forall x \in A \ \forall \{x_j\}_{j \in J} \in \Sigma \ (x \in \lambda(\{x_j\}) \Rightarrow \exists j_0 \in J : \forall j \succcurlyeq j_0 \ x_j \in A).$$

Покажем, что τ является топологией. Очевидно, что $X, \emptyset \in \tau$. Пусть $A \in \tau$ и $B \in \tau$. Возьмем $x \in A \cap B$ и направленность $\{x_j\}_{j \in J} \in \Sigma$ такую, что $x \in \lambda(\{x_j\})$. Тогда найдется элемент $j_A \in J$ ($j_B \in J$) такой, что при $j \succcurlyeq j_A$ ($j \succcurlyeq j_B$) имеем $x_j \in A$ ($x_j \in B$). Далее зафиксируем $j_0 \succcurlyeq j_A, j_B$ и получим, что $x_j \in A \cap B$ при $j \succcurlyeq j_0$, то есть $A \cap B \in \tau$. Пусть теперь $\{A_i \mid i \in I\} \subset \tau$. Так как из принадлежности $x \in \bigcup_{i \in I} A_i$ следует существование такого $i \in I$, что $x \in A_i$, то имеем $\bigcup_{i \in I} A_i \in \tau$. Таким образом τ действительно является топологией. Остальные утверждения предложения вытекают непосредственно из определения топологии τ . ■

Замечание 1. В топологии τ могут, вообще говоря, сходиться направленности, не лежащие в Σ , возможно также, что для направленности $\{x_j\}_{j \in J} \in \Sigma$ найдется элемент $x \in X \setminus \lambda(\{x_j\})$ такой, что $\{x_j\}_{j \in J}$ сходится к x в топологии τ . То есть класс

$$\mathfrak{S} = \{ (S, s) \mid S \in \Sigma, s \in \lambda(S) \}$$

не является в общем случае классом сходимости (см. [1, с.107]). Следующий пример подтверждает сказанное.

Пример 1. Пусть $X = \mathbb{R}$, в Σ поместим все тождественные направленности. Для $\{x_j\}_{j \in J} \in \Sigma$ существует $x \in X$ такой, что $x_j = x$ ($j \in J$), положим

$$\lambda(\{x_j\}) = (x - 1, x) \cup (x, x + 1).$$

Тогда в топологии τ , построенной по (Σ, λ) открыты будут лишь X и $\emptyset$, то есть τ — антидискретная топология, однако в такой топологии произвольная направленность сходится к любой точке.

Пусть далее Y — топологическое пространство и $\mathfrak{F}$ — некоторое множество отображений X в Y . Определим

$$\mathfrak{S}_{\mathfrak{F}} = \{ (\{x_j\}_{j \in J}, x) \mid \forall f \in \mathfrak{F} \ f(x_j) \rightarrow f(x) \}.$$

Предложение 2. Класс $\mathfrak{S}_{\mathfrak{F}}$ является классом сходимости на X . То есть существует единственная топология τ на X , в которой направленность S сходится к элементу $s \in X$ тогда и только тогда, когда $(S, s) \in \mathfrak{S}_{\mathfrak{F}}$. Эта топология совпадает с инициальной топологией относительно отображений $f \in \mathfrak{F}$ (то есть со слабой топологией на X , при которой все отображения из $\mathfrak{F}$ непрерывны).

2. Секвенциальная топология

Определение 4. Подмножество A топологического пространства (X, τ) называется секвенциально замкнутым, если оно содержит предел каждой своей сходящейся последовательности. Наименьшее секвенциально замкнутое множество, содержащее множество A (оно, очевидно, совпадает с пересечением всех секвенциально замкнутых множеств, содержащих A), называется его секвенциальным замыканием и обозначается $\text{scl } A$.

Приведем основные свойства секвенциального замыкания в следующей лемме (см., например, [4, с.14], [6, с.4]).

Лемма 1. Для любых $A, B \subset X$ верно

1. $\text{scl } \emptyset = \emptyset$;
2. $A \subset \text{scl } A \subset \text{cl } A$;
3. A секвенциально замкнуто $\Leftrightarrow A = \text{scl } A$;
4. $\text{scl}(\text{scl } A) = \text{scl } A$;
5. $A \subset B \Rightarrow \text{scl } A \subset \text{scl } B$;
6. $\text{scl}(A \cup B) = \text{scl } A \cup \text{scl } B$;
7. $\text{scl}(A \cap B) \subset \text{scl } A \cap \text{scl } B$.

Из леммы 1 следует, что оператор, который каждому подмножеству X сопоставляет его секвенциальное замыкание, является оператором Куратовского, следовательно на X существует единственная топология,

операция замыкания в которой совпадает с операцией секвенциального замыкания в исходной топологии τ (см., например, [1, с.68]). Эту топологию будем называть *секвенциальной топологией, ассоциированной с τ* и обозначим $s\tau$. Замкнутые множества в топологии $s\tau$ — это множества, секвенциально замкнутые в τ , и для каждого $A \subset X$ $\text{scl}_\tau A = \text{cl}_{s\tau} A$.

Топология $s\tau$ может быть порождена в смысле предложения 1 парой (Σ, λ) , где класс Σ состоит из сходящихся в топологии τ последовательностей, а оператор λ каждой такой последовательности ставит в соответствие множество ее пределов в топологии τ .

Определение 5. Топологическое пространство (X, τ) называется секвенциальным, если топологии τ и $s\tau$ совпадают (см. также [7, с.94]).

Класс секвенциальных пространств достаточно широк. Например, любое топологическое пространство, удовлетворяющее первой аксиоме счетности (тем более полуметризуемое пространство), является секвенциальным (см. [1, с.105]). Однако наличие первой аксиомы счетности не является необходимым условием для секвенциальности пространства (см. [7, с.95]).

Многие свойства топологии $s\tau$ исследованы автором в статье [2], приведем здесь краткую сводку основных результатов:

1. Топология $s\tau$ мажорирует топологию τ , а сходящиеся последовательности в этих топологиях одни и те же.
2. Если на множестве X заданы две топологии τ и σ , сходящиеся последовательности в которых одни и те же, то $s\tau = s\sigma$.
3. Отображение $f : (X, \tau) \rightarrow (Y, \sigma)$ секвенциально непрерывно тогда и только тогда, когда оно непрерывно из $(X, s\tau)$ в $(Y, s\sigma)$.
4. Топологии τ и $s\tau$ на X совпадают тогда и только тогда, когда каждое секвенциально непрерывное отображение (X, τ) в любое пространство (Y, σ) непрерывно.

К моменту выхода вышеупомянутой статьи автору не удалось найти ответ на следующий вопрос: является ли секвенциальная топология, ассоциированная с линейной сама линейной. Отрицательный ответ дается в следующем предложении. Напомним прежде несколько определений из функционального анализа (см. [5]).

Определение 6. Полное метризуемое локально выпуклое пространство называется пространством Фреше.

Определение 7. Пусть семейство топологических векторных пространств $\{X_j\}_{j \in J}$ таково, что все X_j ($j \in J$) являются векторными подпространствами некоторого пространства X ,

$$X = \text{l.h.} \bigcup_{j \in J} X_j$$

и при $X_i \subset X_j$ вложение $X_i \rightarrow X_j$ непрерывно. Сильнейшая (локально выпуклая) линейная топология на X , при которой все вложения $X_j \rightarrow$

X непрерывна называется (локально выпуклой) индуктивной топологией на X относительно семейства $\{X_j\}_{j \in J}$, а само пространство X , наделенное этой топологией называется (локально выпуклым) индуктивным пределом семейства ТВП $\{X_j\}_{j \in J}$ (обозначается $(X = l.c. \operatorname{ind} \lim_{j \in J} X_j) \quad X = \operatorname{ind} \lim_{j \in J} X_j$).

Замечание 2. Если

$$X = l.c. \operatorname{ind} \lim_{j \in J} X_j,$$

то множества вида

$$\operatorname{abs co} \bigcup_{j \in J} U_j,$$

где $U_j \in \mathcal{O}_{\mathbb{O}}^{X_j}$ ($j \in J$), образуют базу окрестностей нуля в X (см. [5, с.36]).

Замечание 3. Индуктивный предел последовательности локально выпуклых пространств является локально выпуклым индуктивным пределом этих пространств (см. [5, с.36]).

Определение 8. Индуктивный предел

$$X = \operatorname{ind} \lim_{j \in J} X_j$$

называется регулярным, если из ограниченности множества B в X следует, что B содержится и ограничено в некотором X_j .

Определение 9. Индуктивный предел

$$X = \operatorname{ind} \lim_{j \in J} X_j$$

называется строгим, если $\{X_j\}_{j \in J}$ является направленностью с отношением $i \preccurlyeq j \Leftrightarrow X_i \subset X_j$ и при $i \preccurlyeq j$ топология, индуцированная из X_j на X_i , совпадает с исходной топологией на X_i .

Предложение 3. Справедливы следующие утверждения:

1. Пусть (X, τ) — регулярный индуктивный предел направленности секвенциальных локально выпуклых пространств $\{(X_j, \tau_j)\}_{j \in J}$, причем индуцированная из X на X_j топология совпадает с τ_j , тогда τ является финальной топологией относительно вложений $X_j \subset X$ (то есть сильнейшей топологией на X , при которой все вложения $X_j \subset X$ непрерывны).

2. Пусть (X, τ) — LF-пространство, то есть X — это строгий индуктивный предел возрастающей последовательности пространств Фреше $\{X_n\}_{n \in \mathbb{N}}$, метрику на X_n обозначим d_n и предположим, что выполнено следующее условие

$$\forall n \in \mathbb{N} \quad \forall x \in X_n \quad d_n(x, \mathbb{O}) \leq d_{n+1}(x, \mathbb{O}).$$

Тогда топология τ не является линейной и $\tau \neq \tau$.

Доказательство. 1. Достаточно показать следующую эквивалентность

$$A \in \text{Cl}(X, s\tau) \Leftrightarrow \forall j \in J \ A \cap X_j \in \text{Cl } X_j.$$

$\Rightarrow$) Так как $s\tau_j = \tau_j$, то достаточно показать, что $A \cap X_j$ секвенциально замкнуто в X_j . Пусть последовательность $\{x_n\}_{n \in \mathbb{N}} \subset A \cap X_j$ сходится в топологии τ_j к элементу $x \in X_j$. Тогда $x_n \rightarrow x$ в топологии τ (так как индуцированная из X на X_j топология совпадает с τ_j), откуда имеем, что $x_n \rightarrow x$ в $s\tau$. Таким образом $x \in A$, кроме того $x \in X_j$, следовательно $x \in A \cap X_j$ и $A \cap X_j \in \text{Cl } X_j$.

$\Leftarrow$) Пусть $\{x_n\}_{n \in \mathbb{N}} \subset A$ и $x_n \rightarrow x$ в топологии $s\tau$ (или в топологии τ , что равносильно). Покажем, что $x \in A$. Множество

$$B = \{x_n \mid n \in \mathbb{N}\} \cup \{x\}$$

ограниченно в X , следовательно, существует $j \in J$ такой, что $B \subset X_j$, откуда получаем сходимост $x_n \rightarrow x$ в топологии τ_j . Так как

$$A \cap X_j \in \text{Cl } X_j,$$

то $x \in A \cap X_j$, а значит $x \in A$.

2. Из утверждения 1 следует, что если топология $s\tau$ линейна, то $s\tau = \tau$, поэтому достаточно показать, что $s\tau \neq \tau$. Топология $s\tau$ — это финальная топология относительно вложений $X_n \subset X$, поэтому, как легко видеть, множество

$$B = \bigcup_{n \in \mathbb{N}} U_n,$$

где

$$U_n = \left\{ x \in X_n \mid d_n(x, \mathbb{O}) \leq \frac{1}{n} \right\},$$

является окрестностью нуля в топологии $s\tau$; покажем, что это множество не окрестность нуля в τ . Предположим противное, то есть существует множество

$$A = \text{abs co} \bigcup_{n \in \mathbb{N}} V_n,$$

где V_n — окрестность нуля в X_n ($n \in \mathbb{N}$), такое, что $A \subset B$. Покажем, что для любого $n \in \mathbb{N}$ верно включение $V_1 \subset U_n$. Пусть найдется $n \in \mathbb{N}$ такой, что V_1 не содержится в U_n , то есть существует $x \in V_1$, для которого

$$d_n(x, \mathbb{O}) > \frac{1}{n}.$$

Зафиксируем $y \in V_n \setminus X_{n-1}$. Далее рассмотрим множество элементов

$$x_t = tx + (1-t)y \quad (t \in \mathbb{R})$$

и непрерывную числовую функцию $f(t) = d_n(x_t, \mathbb{O})$, тогда

$$f(0) = d_n(y, \mathbb{O}) \leq \frac{1}{n}$$

и

$$f(1) = d_n(x, \mathbb{O}) > \frac{1}{n},$$

следовательно, по теореме о промежуточных значениях непрерывной функции найдется значение $s \in (0, 1)$ такое, что

$$f(s) > \frac{1}{n},$$

откуда имеем $x_s \notin U_n$. Так как $s \neq 0$, то $x_s \notin U_k$ ($k = 1, \dots, n-1$), кроме того $x_s \notin U_m$ при $m > n$, так как

$$d_m(x_s, \mathbb{O}) \geq d_n(x_s, \mathbb{O}) > \frac{1}{n} > \frac{1}{m}.$$

Таким образом $x_s \notin B$, однако $x_s \in A$. Следовательно наше предположение неверно, и для любого $n \in \mathbb{N}$ выполнено включение $V_1 \subset U_n$. Но тогда если $x \in V_1$, то для всякого $n \in \mathbb{N}$ имеем

$$d_1(x, \mathbb{O}) \leq d_n(x, \mathbb{O}) \leq \frac{1}{n},$$

то есть $x = \mathbb{O}$. Однако множество $V_1 = \{\mathbb{O}\}$ не является окрестностью нуля в X_1 , откуда следует, что наше предположение о существовании множества A также неверно, то есть $s\tau \neq \tau$, что и требовалось доказать. ■

Замечание 4. Условия 2-го утверждения предложения выполнены, например, для пространств $\mathbb{K}^\infty$ (где $\mathbb{K}$ — это $\mathbb{R}$ или $\mathbb{C}$) и $\mathcal{D}(\mathbb{R})$ (пространство гладких финитных функций Л. Шварца, см., например, [3, с.414]).

ЛИТЕРАТУРА

1. Келли Дж.Л. Общая топология. М.: Наука. Гл.ред.физ.-мат.лит., 1968. 384 с.
2. Чемёркин А.А. О некоторых свойствах секвенциального замыкания // Математические структуры и моделирование (Омск). 2003. Вып.11. С.21-27.
3. Эдвардс Р. Функциональный анализ. Теория и приложения. М.: Мир, 1969. 1071с.
4. Мельников Е.В. Топологические векторные пространства: Методические указания. Омск: ОмГУ, 1990. 43 с.
5. Мельников Е.В. Локально выпуклые пространства: Методические указания. Омск: ОмГУ, 1991. 48 с.
6. Мельников Е.В. Векторнозначные распределения и обобщенная корректность абстрактной задачи Коши. — Омский гос. ун-т. Омск, 1988. Деп. ВИНТИ 15.03.1988. №1994-B88, 79 с.
7. Энгелькинг Р. Общая топология. М.: Мир, 1986. 752 с.

ИДЕНТИФИКАЦИЯ КИНЕТИЧЕСКИХ ПАРАМЕТРОВ ПРОЦЕССА КРИСТАЛЛИЗАЦИИ ОКСАЛАТА КАЛЬЦИЯ

О.А. Голованова, В.А. Когут, Е.Ю. Ачкасова

Среди минераллообразований патогенного характера наиболее распространенными являются камни мочевой системы. Изучаются механизмы кристаллизации основной минералообразующей фазы – оксалата кальция.

Изучение кристаллизации малорастворимых соединений представляет собой значительный интерес, в первую очередь в связи с тем, что большинство физиогенных и патогенных минералов, образующихся в живых организмах, относится именно к этому типу соединений.

В настоящее время установлено, что патогенные биоминералы могут образовываться во многих тканях и органах человека [1]. Среди минераллообразований патогенного характера наиболее распространенными являются камни мочевой системы [2].

Механизмы формирования минеральной фазы устанавливаются в лабораторном эксперименте по кристаллизации основной минералообразующей фазы – оксалата кальция. При проведении эксперимента используется максимальное приближение условий эксперимента к физиологическим.

1. Проведение эксперимента

Основные этапы эксперимента сводятся к следующему:

1. Изучение кинетики кристаллизации минеральной фазы в отсутствии добавок и установление влияния степени пересыщения на кристаллизацию.
2. Изучение кинетики кристаллизации в присутствии добавок аминокислот (глутаминовой (*Gly*) и аспарагиновой (*Asp*)).
3. Определение параметров кристаллизации с помощью математической модели.

Скорость роста кристаллов определялась кондуктометрическим методом по изменению концентрации свободных ионов в растворе.

Процесс кристаллизации изучался как в растворах чистого оксалата кальция, так и в присутствии добавок аспарагиновой и глутаминовой аминокислот с целью определения их влияния на кинетику кристаллизации.

При помощи кондуктометра фиксировали изменение удельной электропроводности (УЭП) раствора во времени. Эксперименты проводили в диапазоне пресыщений с линейной зависимостью от УЭП.

Концентрации исходных растворов задавались исходя из расчета, что в конечном разбавлении пересыщение по малорастворимому оксалату кальция составит 5, 7 и 10 (это наиболее близкие значения к физиологическому раствору). Пересыщение рассчитывается как отношение концентрации ионов кальция (равной концентрации оксалат-ионов) в пересыщенном растворе к концентрации насыщенного раствора в условиях опыта (37°C).

2. Математическая модель

Для количественного анализа экспериментальных данных используется степень превращения α , определенная по формуле

$$\alpha(t) = \frac{(c_0 - c(t))}{(c_0 - c_s)} = \frac{V(t)}{V_{max}}, \quad (1)$$

где c_0 — это начальная концентрация соли, $c(t)$ — это концентрации в момент времени, а c_s — концентрация в состоянии равновесия, $V(t)$ — объем раствора в момент времени t , а V_{max} — это максимальный объем осадка, полученный в результате эксперимента.

Скорость осаждения зависит от площади поверхности кристаллов и пересыщения раствора [3]:

$$\frac{d\alpha}{dt} = k_1 A (c(t) - c_s)^n, \quad (2)$$

где A — это площадь поверхности, k_1 — коэффициент пропорциональности и n — порядок реакции.

Согласно [3], среднюю площадь поверхности можно представить в виде:

$$\bar{A} = \beta N(t)^{\frac{1}{3}} V(t)^{\frac{2}{3}}, \quad (3)$$

где $N(t)$ — это полное количество частиц в момент времени t , а β — коэффициент, учитывающий форму частицы.

Используя (1) и (3), получаем

$$\bar{A} = k_2 N(t)^{\frac{1}{3}} \alpha^{\frac{2}{3}}, \quad (4)$$

где $k_2 = \beta V_{max}^{\frac{2}{3}}$ — константа для любой начальной концентрации раствора.

Подставляя (4) в (2), и полагая, что $A = \bar{A}$, получаем

$$\frac{d\alpha}{dt} \alpha^{-\frac{2}{3}} = KN(t)^{\frac{1}{3}} (c(t) - c_s)^n. \quad (5)$$

На стадии роста кристаллов, число частиц N_t постоянно, тогда

$$\frac{d\alpha}{dt} \alpha^{-\frac{2}{3}} = K'(c(t) - c_s)^n, \quad (6)$$

где K' включает все константы.

Логарифмируя выражение (6), получаем:

$$\log \left(\frac{d\alpha}{dt} \right) - \frac{2}{3} \log \alpha = \log K' + n \log (c(t) - c_s). \quad (7)$$

В выражении (7) неизвестными величинами являются константы K' и n . Параметры $\log K'$ и n можно определить, используя метод наименьших квадратов.

3. Обработка экспериментальных данных

Значения УЭП, получаемые от прибора, имели недостаточную точность для вычисления производной с помощью разностной схемы или применения сглаживающего и дифференцирующего фильтров. Значения α монотонно возрастают, поэтому был использован следующий метод: по известным значениям α методом наименьших квадратов строился аппроксимирующий полином, в качестве производной от α выбиралась производная этого полинома.

Для проведения расчетов не разрабатывалось специальное ПО на языках программирования высокого уровня. Встроенных возможностей программы Microsoft Excel оказалось достаточно, аппроксимирующий полином строился с использованием функции линейной аппроксимации (ЛИНЕЙН). Быстродействие не является лимитирующим фактором, поскольку объем вычислений достаточно мал: число измерений УЭП не превышает 100.

4. Интерпретация результатов эксперимента

Из кинетических кривых, приведенных на рис. 1, видно, что степень превращения (α) монотонно увеличивается со временем, выходя на плато по завершении периода индукции. Очевидно, что период индукции тем меньше, чем больше пересыщение исходного раствора. Однако по числовому значению между α для всех трех пересыщений не обнаружено значимого отличия (3 параллельных опыта в каждой серии, $P = 0,95$). Значение α колеблется от 0,26 до 0,29.

Следующим этапом исследования было изучение влияния добавок аминокислот на процесс кристаллизации минеральной фазы. Добавление аминокислот вызывает уменьшение степени превращения, а вместе с тем и уменьшение константы скорости роста частиц твердой фазы, т.е. аминокислоты

Рис. 1. Кинетические кривые

Рис. 2. График зависимости скорости осаждения от пересыщения
(в логарифмическом масштабе)

Пересыщение/добавка	K' константа скорости реакции	n порядок реакции
5/-	10^{18}	6,2
7/-	10^{18}	6,6
10/-	10^{17}	6,7
7/ <i>Asp</i> (10ppm)	10^{19}	6,9
7/ <i>Gly</i> (10ppm)	$10^{19,5}$	7,2
5/ <i>Asp</i> (100ppm)	$10^{5,8}$	2,8
7/ <i>Asp</i> (100ppm)	$10^{4,7}$	2,7
10/ <i>Asp</i> (100ppm)	$10^{4,7}$	2,8
5/ <i>Gly</i> (100ppm)	$10^{5,8}$	2,8
7/ <i>Gly</i> (100ppm)	$10^{5,2}$	2,85
10/ <i>Gly</i> (100ppm)	$10^{5,0}$	2,9

Таблица 1. Кинетические параметры кристаллизации CaC_2O_4

выступают в роли ингибиторов роста кристаллов. График зависимости скорости осаждения от пересыщения приведен на рис. 2. Коэффициентами аппроксимирующей прямой являются n и $\log K'$.

Из данных, приведенных в таб. 1 видно, что увеличение концентрации аминокислот ведет к снижению константы и порядка процесса осаждения оксалата кальция. А это свидетельствует об ингибирующем влиянии аминокислот на рост частиц.

Ингибирующее действие данных аминокислот, можно объяснить абсорбцией её ионов на активных центрах кристаллов. Так аминокислоты в водных растворах в широком диапазоне pH находятся в виде цвиттер-ионов, способных принимать нужную конформацию, в которой две карбоксильные и одна аминогруппа могут взаимодействовать с минеральной поверхностью без дальнейшего изменения конформации [4].

5. Заключение

Обработка результатов кинетического эксперимента с помощью математической модели позволила определить основные характеристики процесса кристаллизации оксалата кальция (константы и порядка скорости роста кристаллов). Показано, что введение добавок аминокислот в модельный раствор ведет к снижению указанных характеристик. Данное явление объясняется ингибирующим эффектом аминокислот на процесс роста кристаллов и удовлетворительно согласуется с литературными данными.

ЛИТЕРАТУРА

1. Кораго А.А. Введение в биоминералогии. СПб.: Недра, 1992. 280 с.
2. Голованова О.А., Борбат В.Ф., Пятанова П.А. Комплексное изучение состава и строения почечных камней // Омский Научный Вестник. 2003. N.4. С.225-229.

3. Skrtic D., Markovic M., Komunjer Lj., Furder-Milhofer. Precipitation of calcium oxalates from high ionic strength solutions // Journal of Crystal Growth. 1984. N.66. P.431-440.
4. Fleming D.E., Bronswijk W., Ryall R. L. A comparative study of the adsorption of amino acids on to calcium minerals found in renal calculi // Clinical Science. 2001. V.101. P.159-168.

АЛГОРИТМ ЧИСЛЕННОГО РЕШЕНИЯ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С РАЗРЫВНОЙ ПРАВОЙ ЧАСТЬЮ

В.В. Коробицын, Ю.В. Фролова

Для систем обыкновенных дифференциальных уравнений $\dot{X}(t) = F(X)$ с разрывной правой частью предлагается алгоритм для нахождения численного решения. Вводится понятие динамической системы с клетчатой структурой. Доказывается теорема об аппроксимации решения системы схемой склеивания решений на границе между клетками. Описывается программа МЕР2, обеспечивающая исследования динамических систем такого класса. Приводятся примеры динамических систем и их фазовые портреты, построенные с помощью программы.

Проект МЕР2 (Modelling Evolution Processes 2) направлен на разработку программного обеспечения для анализа динамических систем. Модели процессов различной природы описываются динамическими системами:

- перенос биомассы в цепи питания организмов (растительность – травоядные – хищники);
- изменение экономических индексов в различных ситуациях;
- поведение индивидов в социальной среде.

Эти модели, как правило, описываются системами дифференциальных уравнений с разрывной правой частью. Такое описание обусловлено изменением поведения элементов системы в возникающих ситуациях. Например, если количество травоядных становится маленьким, то хищники становятся вегетарианцами и поедают растительность наравне с травоядными.

1. Динамическая система с клетчатой структурой

Будем рассматривать динамические системы вида

$$\dot{X}(t) = F(t, X) \tag{1}$$

с разрывной правой частью F . Решение таких систем не может быть представлено в классическом понимании как дифференцируемая кривая $X(t)$, скорость которой равна заданному векторному полю F . Иногда является естественным принять некоторое $X(t)$ как решение системы (1), пренебрегая тем, что $\dot{X}(t) \neq F(X(t))$, например скользящий режим в [1].

Пусть Ω' — некоторая открытая область в R^n и $F(X)$, $X \in \Omega'$ разрывное векторное поле в Ω' . Обозначим через C_F и D_F подмножества точек Ω' , где F является непрерывной и разрывной соответственно. Поскольку нас не интересует поведение траектории динамической системы вблизи границы области, то мы будем рассматривать систему (1) в некоторой подобласти $\Omega \subset \Omega'$, замыкание $\bar{\Omega}$ которой является компактом и находится внутри Ω' .

Если динамическая система составлена из конечного числа «хороших» систем, то есть множество D_F состоит из конечного числа кусочно-гладких гиперповерхностей, тогда будем говорить о динамической системе с клетчатой структурой.

Определение 1. Множество C будем называть клеткой динамической системы $\dot{X}(t) = F(X)$, если C является линейно-связным подмножеством множества C_F .

Для всех кусочно-гладких систем множество C_F может быть представлено как объединение конечного числа клеток C_i , $i = 1, \dots, n$

$$C_F = \bigcup_{i=1}^n C_i.$$

Определение 2. Динамическая система $\dot{X} = F(t, X)$ имеет клетчатую структуру C_s , если объединение замыканий клеток $\bar{C}_i$, $i = 1, \dots, n$ совпадает с замыканием области определения $\bar{\Omega}$ функции $F(t, X)$:

$$C_s : \bigcup_{i=1}^n \bar{C}_i = \bar{\Omega}.$$

Отметим, что для динамической системы с клетчатой структурой подмножество D_F точек, где функция $F(t, X)$ разрывна, составлена из точек замыкания всех клеток:

$$D_F = \bigcup_{i=1}^n \bar{C}_i \setminus \bigcup_{i=1}^n C_i.$$

2. Алгоритм численного решения динамических систем с клетчатой структурой

Поиск численного решения динамической системы с клетчатой структурой состоит из двух этапов. Во-первых, для любой клетки C_i , $i = 1, \dots, n$ численное решение может быть найден классическими методами, например методом Рунге-Кутты 4(5). Во-вторых, решение должно быть склеено на множестве D_F , где динамическая система не имеет гладкого решения.

Численная схема метода Рунге-Кутты 4(5) имеет вид [2]

$$\begin{aligned} K_1 &= F(t, X(t)), \\ K_2 &= F(t + h/2, X(t) + K_1 h/2), \\ K_3 &= F(t + h/2, X(t) + K_2 h/2), \\ K_4 &= F(t + h, X(t) + K_3 h), \\ X(t + h) &= X(t) + (K_1 + 2K_2 + 2K_3 + K_4)h/6. \end{aligned}$$

Здесь $X(t)$ — известное значение функции в момент времени t , $X(t + h)$ — неизвестное значение в момент $t + h$, h — шаг интегрирования. Этот метод имеет четвертый порядок точности, является устойчивым и сходящимся.

Используя оценку Ричардсона, относительная погрешность численного решения определяется формулой

$$d = \frac{\|X_h - X_{2h}\|}{\|X_h\|} \cdot \frac{1}{2^{k-1} - 1},$$

здесь X_h и X_{2h} — значения решения, найденные двумя шагами длины h и одним шагом длины $2h$. Параметр k — порядок метода ($k = 5$). Для нахождения оптимального значения шага интегрирования воспользуемся оценкой

$$s = 0.9 \cdot \left(\frac{\varepsilon}{d}\right)^{\frac{1}{k}}.$$

Эта оценка гарантирует приближение относительной погрешности к заданному значению точности ε . Но шаг интегрирования не может меняться слишком быстро, поэтому его изменение ограничивается минимальным и максимальным значениями изменений

$$s_1 = \max(\min(s, s_{\min}), s_{\max}),$$

здесь $s_{\min} = 0.1$ и $s_{\max} = 5.0$ — экспериментальные значения параметров. Тогда новое значение шага интегрирования вычисляется по формуле

$$h_{\text{new}} = h \cdot s_1.$$

Этот алгоритм позволяет управлять погрешностью численного решения в пределах области непрерывности C_i функции $F(t, X)$, но, когда кривая решения пересекает границу между клетками, тогда этот алгоритм не может обеспечить приемлемое решение.

Пусть найденное решение $X(t)$ лежит в клетке C_i , а значение в некоторый последующий момент времени $X^*(t + h) = X(t) + h \cdot F(t, X(t))$ принадлежит другой клетке C_j , $j \neq i$. Полученное решение $X^*(t + h)$ показывает, что решение приблизилось к границе клетки. Мы предлагаем алгоритм склеивания решения на границе между клетками.

Алгоритм (поиск решения на границе между клетками).

Шаг 1. (Идти к границе.) Найдем шаг интегрирования h_1 такой, что точка $X_1 = X(t_1)$, $t_1 = t + h_1$ находится в клетке C_i и расстояние до точки решения на границе $X_b \in D_F$ меньше, чем $\varepsilon \cdot \|X_1\|$:

$$\|X_1 - X_b\| \leq \varepsilon \cdot \|X_1\|,$$

где ε — заданная точность. Тогда мы говорим, что точка X_1 является приближением точки X_b .

Шаг 2. (Точка X_b является стационарной?)

Если $(F(t_1, X_1), n(X_b)) \cdot (n(X_b), F(t_2, X_2)) < 0$, то найдена стационарная точка и программа завершается, иначе перейти к шагу 3. (Здесь $n(X_b)$ — вектор нормали к поверхности разрыва функции $F(t, X)$ в точке X_b , $(\cdot, \cdot)$ — скалярное произведение. Предложенное условие означает, что вектора $F(t_1, X_1)$ и $F(t_2, X_2)$ направлены в сторону пересечения с поверхностью разрыва, а это означает, что решение не покинет поверхности разрыва.)

Шаг 3. (Идти в клетку C_j .) Пусть следующий шаг интегрирования h_2 будет таким, что точка $X(t_1 + h_2/2) = X_b$, тогда решение в клетке C_j находим с помощью схемы

$$\begin{aligned} K_1 &= F(t_1, X_1), \\ K_2 &= F(t_1 + h_2, X_1 + h_2 K_1), \\ X(t_1 + h_2) \approx X_2 &= X_1 + h_2 \cdot \frac{K_1 + K_2}{2}. \end{aligned} \quad (2)$$

(Здесь $X_1 + h_2 K_1$ представляет первое приближение точки X_2 . Среднее значение скорости кривой в точке X_b задается вектором $V = (K_1 + K_2)/2$.)

Далее докажем теорему об аппроксимации решения схемой (2). Введем некоторые обозначения. Пусть функция $\bar{X}(t)$ представляет точное решение динамической системы $\dot{X} = F(t, X)$. Рассмотрим решение около границы между клетками C_i и C_j . Будем считать, что решение в граничной точке X_b является непрерывным, но негладким. Кроме того, будем считать, что точка X_b не является стационарной. Обозначим $X_1 = \bar{X}(t_1) \in C_i$, $X_2 = \bar{X}(t_2) \in C_j$, $h_t = t_2 - t_1$, $a \leq t_1 < t_2 \leq b$. Тогда граничная точка определяется формулой $X_b = \bar{X}(t_1 + \theta h_t)$, $\theta \in (0, 1)$. Пусть точка X_2^* — приближенное решение, полученное с помощью численной схемы (2). Обозначим отклонение от точного решения $\Delta X_2^* = X_2 - X_2^*$.

Теорема 1. Пусть векторное поле $F(t, X)$ ограничено: $\|F(t, X)\| \leq M$ для всех $X \in \Omega$, $a \leq t \leq b$, дифференцируемо по X в некоторой окрестности точки X_2 , и производная в точке X_2 ограничена $\|\frac{dF}{dX}(t_2, X_2)\| = L < \infty$. Тогда отклонение приближенного решения X_2^* от точного значения X_2 определяется формулой $\Delta X_2^* = R \cdot h_t + O(h_t^2)$, где $R \leq 2M(\theta - 1/2)$.

Доказательство. Пусть $X_b^- = \lim_{\tau \rightarrow \theta-0} \bar{X}(t_1 + \tau h_t)$ и $X_b^+ = \lim_{\tau \rightarrow \theta+0} \bar{X}(t_1 + \tau h_t)$.

Поскольку решение $X(t)$ является гладким в клетках C_i , C_j , то значения X_b^- и X_b^+ можно представить в виде

$$X_b^- = X_1 + \theta h_t F(t_1, X_1) + O(h_t^2), \quad (3)$$

$$X_b^+ = X_2 - (1 - \theta) h_t F(t_2, X_2) + O(h_t^2). \quad (4)$$

Так как решение $\bar{X}(t)$ является непрерывным в точке X_b , то эти значения совпадают $X_b^- = X_b^+ = X_b$. Приравняем правые части выражений (3) и (4),

получим уравнение для X_2 :

$$X_2 = X_1 + (\theta F(t_1, X_1) + (1 - \theta)F(t_2, X_2))h_t + O(h_t^2).$$

Вычислим приближенное значение $\tilde{X}_2 = X_1 + h_t F(t_1, X_1)$. Найдем отклонение приближенного значения $\tilde{X}_2$ от точного решения X_2 :

$$\Delta \tilde{X}_2 = X_2 - \tilde{X}_2 = (1 - \theta)h_t(-F(t_1, X_1) + F(t_2, X_2)) + O(h_t^2).$$

Пусть X_2^* — это приближенное значение точки X_2 , полученное по схеме (2):

$$X_2^* = X_1 + h_t \frac{F(t_1, X_1) + F(t_2, \tilde{X}_2)}{2}.$$

Вычислим отклонение точки X_2^* от точного значения X_2

$$\Delta X_2^* = X_2 - X_2^* = \left(\theta - \frac{1}{2}\right)h_t F(t_1, X_1) + h_t \left((1 - \theta)F(t_2, X_2) - \frac{1}{2}F(t_2, \tilde{X}_2)\right) + O(h_t^2).$$

Используя формулу Тейлора, представим функцию $F(t, X)$ в окрестности точки X_2 в виде

$$F(t_2, \tilde{X}_2) - F(t_2, X_2) = L\Delta \tilde{X}_2 + O(\|\Delta \tilde{X}_2\|^2).$$

Это сделать возможно поскольку в этой окрестности функция дифференцируема и ее производная ограничена. Кроме того, очевидно, что $O(\|\Delta \tilde{X}_2\|^2) = O(h_t^2)$. Получаем

$$\Delta X_2^* = h_t \left(\left(\theta - \frac{1}{2}\right) + (\theta - 1)\frac{h_t L}{2} \right) (F(t_1, X_1) - F(t_2, X_2)) + O(h_t^2),$$

и мы можем записать

$$\Delta X_2^* = R \cdot h_t + O(h_t^2),$$

где $R = (\theta - 1/2) \cdot (F(t_1, X_1) - F(t_2, X_2))$. Поскольку функция $F(t, X)$ ограничена числом M , то $\|F(t_1, X_1) - F(t_2, X_2)\| \leq 2M$, поэтому $R \leq 2M(\theta - 1/2)$. Теорема 1 доказана. ■

Замечание 1. Если в условиях теоремы 1 положить, что шаг h_t выбран таким образом, что $\theta - 1/2 \leq h_t$. Тогда $\Delta X_2^* = O(h_t^2)$.

3. Программа MER2 для исследования динамических систем с клетчатой структурой

Программа MER2 реализует предложенный выше алгоритм численного решения систем дифференциальных уравнений с разрывной правой частью. Она выполняет следующие функции:

- ввод системы с клетчатой структурой в графическом режиме;
- сохранение и загрузка файла данных системы. Формат файла текстовый, и файл может быть отредактирован при необходимости текстовым редактором;
- вывод графиков решений системы;
- построение проекций фазового пространства на плоскость.

Модель трехуровневой цепи питания. Для демонстрации возможностей программы МЕР2, представим результаты исследования модели трехуровневой цепи питания [3, 4]. Эта модель описывает изменение биомассы растительности (R), травоядных (C) и хищников (P). Хищники могут питаться как травоядными, так и растительностью. Динамика популяций описывается системой дифференциальных уравнений

$$\begin{aligned}\frac{dR}{dt} &= R \left(r \left(1 - \frac{R}{K} \right) - \frac{\lambda_{RC}C}{1 + h_{RC}\lambda_{RC}R} - \frac{u_{RP}\lambda_{RP}P}{1 + u_{RP}\lambda_{RP}h_{RP}R + u_{CP}\lambda_{CP}h_{CP}C} \right), \\ \frac{dC}{dt} &= C \left(\frac{e_{RC}\lambda_{RC}R}{1 + h_{RC}\lambda_{RC}R} - \frac{u_{CP}\lambda_{CP}P}{1 + u_{RP}\lambda_{RP}h_{RP}R + u_{CP}\lambda_{CP}h_{CP}C} - m_C \right), \\ \frac{dP}{dt} &= P \left(\frac{u_{RP}\lambda_{RP}e_{RP}R + u_{CP}\lambda_{CP}e_{CP}C}{1 + u_{RP}\lambda_{RP}h_{RP}R + u_{CP}\lambda_{CP}h_{CP}C} - m_P \right).\end{aligned}\quad (5)$$

Определения и значения параметров представлены в таблице 1.

Рис. 1. Топология цепи питания адаптивных хищников.

Результаты моделирования приведены на рисунках 2 и 3. На рисунках представлены траектории решений в трехмерном фазовом пространстве (R,C,P), спроектированные на плоскость (R,C). Начальные значения переменных выбраны произвольно.

Согласно модели V. Křivan и S. Diehl хищник выбирает свою диету по следующему правилу. Возможно два варианта питания: 1) растительность более продуктивна для питания хищников, чем травоядные; 2) травоядные более продуктивны для питания хищников, чем растительность. В первом случае имеет место неравенство:

$$\frac{e_{RP}}{h_{RP}} > \frac{e_{CP}}{h_{CP}}. \quad (6)$$

Оптимальная стратегия питания хищников состоит в поглощении растительности ($u_{RP} = 1$), травоядные же включены в диету хищников только тогда, когда плотность растительности меньше переключательной плотности

$$R^* = \frac{e_{CP}}{\lambda_{RP}(e_{RP}h_{CP} - e_{CP}h_{RP})}.$$

Топология цепи питания не является фиксированной, она меняется от исключительно конкурентной цепи ($u_{RP} = 1$; $u_{CP} = 0$ в модели (5)) до цепи

Таблица 1. Определения параметров и их значения, используемые при моделировании.

Параметр	Определение	Значение (единицы)
r	Максимальная скорость роста растительности	0.3 час^{-1}
K	Вместимость растительности в окружающей среде	10 мг С/л
λ_{RC}	Скорость поиска растительности травоядным	0.037 л/(мг час)
λ_{RP}	Скорость поиска растительности хищником	0.025 л/(мг час)
λ_{CP}	Скорость поиска травоядного хищником	0.025 л/(мг час)
u_{RP}	Вероятность нападения хищников на растительность	1 (безразмерна)
u_{CP}	Вероятность нападения хищников на травоядных	1 (безразмерна)
h_{RC}	Время необходимое травоядным для усвоения растительности	3 час
h_{RP}	Время необходимое хищникам для усвоения растительности	4 час
h_{CP}	Время необходимое хищникам для усвоения травоядных	4 час
e_{RC}	Эффективность усвоения биомассы растительности для травоядных	0.6 (безразмерна)
e_{RP}	Эффективность усвоения биомассы растительности для хищников	0.36 (безразмерна)
e_{CP}	Эффективность усвоения биомассы травоядного для хищников	0.6 (безразмерна)
m_C	Скорость отмирания травоядных	0.03 h^{-1}
m_P	Скорость отмирания хищников	0.0275 h^{-1}

всеядного хищника ($u_{RP} = u_{CP} = 1$). Изменение стратегии происходит при переходе через переключательную плотность растительности R^* (рис. 1(A)).

В случае когда травоядные являются более питательными для хищников, чем растительность, неравенство (6) инвертируется. Когда плотность травоядных меньше переключательной плотности

$$C^* = \frac{e_{RP}}{\lambda_{CP}(e_{CP}h_{RP} - e_{RP}h_{CP})},$$

хищники выбирают стратегию всеядного ($u_{RP} = u_{CP} = 1$), а когда больше переключательной плотности, то они потребляют только травоядных ($u_{RP} = 0$; $u_{CP} = 1$). Таким образом, топология цепи питания не является фиксированной и переключается между цепью питания «настоящего» хищника и цепью питания всеядного хищника, зависящей от плотности травоядных (рис. 1(B)).

4. Выводы и заключение

Результаты моделирования системы трехуровневой цепи питания в двух режимах приведены на рис. 2, 3. В фазовом пространстве системы имеется две клетки. В первом режиме клетки разделены плоскостью $R = R^*$, а во втором — плоскостью $C = C^*$.

Полученные фазовые портреты подтверждают целесообразность введения клетчатой структуры в динамической системе (5). Решения системы имеют выраженные особенности на границе между клетками. Заранее определенное геометрическое расположение клеток значительно уменьшает затраты на

исследование системы. В частности, предложенный численный алгоритм поиска решения на границе клеток дает более точное решение, чем классический метод Рунге-Кутты, что продемонстрировано на графике решений системы двумя методами (рис. 4).

Рис. 2. Фазовое пространство системы в первом режиме. Параметры системы: $e_{RP} = 0.2$, $e_{CP} = 0.1$, $R^* = 10$, остальные указаны в таблице 1.

ЛИТЕРАТУРА

1. Филиппов А.Ф. Дифференциальные уравнения с разрывной правой частью. – М.: Наука. Главная редакция физико-математической литературы, 1985.
2. Хайрер Э., Нёрсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М.: Мир, 1990.
3. Křivan, V. Optimal Intraguild Foraging and Population Stability // Theoretical Population Biology. 2000. V.58. P.79-94.
4. Křivan, V., Diehl, S. Adaptive omnivory and species coexistence in tri-trophic food webs // Theoretical Population Biology. 2005. V.67. P.85-99.

Рис. 3. Фазовое пространство системы во втором режиме. Параметры системы: $e_{RP} = 0.1$, $e_{CP} = 0.3$, $C^* = 5$, остальные указаны в таблице 1.

Рис. 4. Кривые решения системы двумя методами: 1) классический метод Рунге-Кутты; 2) предложенный алгоритм. Параметры системы: $e_{RP} = 0.1$, $e_{CP} = 0.3$, $C^* = 5$, остальные указаны в таблице 1. Начальные значения: $R(0) = 15$, $C(0) = 1$, $P(0) = 7$.

ВЛИЯНИЕ ЭФФЕКТОВ ДАЛЬНОДЕЙСТВИЯ НА МУЛЬТИКРИТИЧЕСКОЕ ПОВЕДЕНИЕ НЕУПОРЯДОЧЕННЫХ СЖИМАЕМЫХ СИСТЕМ

С.В. Белим

В рамках теоретико-полевого подхода рассмотрено мультикритическое поведение систем, описываемых двумя флуктуирующими параметрами порядка, с замороженными дефектами структуры при различных значениях параметра дального действия непосредственно в трехмерном пространстве. Выявлены устойчивые фиксированные точки ренормгруппового преобразования. Определены типы мультикритического поведения при различных значениях параметра дального действия.

Как хорошо известно, критические свойства систем задаются малым количеством параметров. Таких как размерность, симметрия параметра порядка и скорость убывания взаимодействия с расстоянием. Особый интерес представляют системы, в которых кроме обычного ближнего действия присутствуют эффекты дального действия. В классической изингоподобной системе взаимодействие между флуктуациями убывает экспоненциально с расстоянием по закону $\exp(-r/r_0)$, в связи с чем рассматривается взаимодействие только между ближайшими соседями, в связи с чем данные системы можно охарактеризовать как ближнего действия. При убывании взаимодействия с расстоянием r по закону r^{-D-a} , где D – размерность пространства, уже нельзя ограничиваться взаимодействием между ближайшими соседями, и возникают эффекты дального действия. Как показано в работе [1] в рамках ε -разложения ($\varepsilon = 2a - D$) для значений параметра дального действия в интервале $1,5 < a < 2$ наблюдается негауссово критическое поведение, отличное от критического поведения ближнего действия систем и существенно зависящее от параметра дального действия a . В интервале значений $a \leq 1,5$ в системе наблюдается гауссово критическое поведение. Как показали расчеты непосредственно в трехмерном пространстве для критического поведения однородных систем с эффектами дального действия подтверждается предсказание ε -разложения [2]. Введение в систему замороженных примесей существенно изменяет картину критического поведения [3]. Так для

неупорядоченных систем также как и для однородных систем при значениях параметра дальногодействия $a \leq 1,5$ наблюдается гауссово критическое поведение. Однако интервал $1,5 < a < 2$ разбивается на два. При значениях $1,8 \leq a < 2$ критическое поведение носит негауссовый характер, существенно зависящий от параметра дальногодействия a . Однако при $1,5 < a < 1,8$ в системе происходит срыв на фазовый переход первого рода. Упругие деформации приводят как к смене режима критического поведения так и к появлению на фазовой диаграмме вещества трикритических линий [4]. Однако интервалы значений параметра дальногодействия, в рамках которых критическое поведение качественно носит одинаковый характер остаются неизменными.

Как было показано в работе [5] в однородных системах, описываемых двумя флуктуирующими параметрами порядка эффекты дальногодействия могут приводить к изменению режима мультикритического поведения. Для значений параметра дальногодействия $2 > a > 1,6$ наблюдается бикритическое поведение, тогда как для $1,5 < a \leq 1,6$ поведение становится тетракритическим. В интервале значений параметра порядка $a \leq 1,5$ устойчивой становится гауссова критическая точка.

Присутствие замороженных точечных дефектов структуры приводит к изменению режима поведения близкодействующих систем как в бикритической, так и в тетракритической области [6]. В указанной работе показано, что влияние δ -коррелированных примесей приводит к развязыванию параметров порядка в мультикритических точках. В работе [7] выявлено, что упругие деформации для однородных систем приводят к смене бикритического поведения тетракритическим. Для неупорядоченных систем [2] деформационные степени свободы не меняя типа мультикритического поведения изменяют режим тетракритического поведения.

Предметом данной статьи является исследование влияния эффектов дальногодействия эффектов на неупорядоченные сжимаемые системы, описываемые двумя параметрами порядка.

При структурных фазовых переходах с отсутствием пьезоэффекта в парафазе упругие деформации играют роль вторичного параметра порядка, флуктуации которого в большинстве случаев не являются критическими [8]. В связи с тем, что в критической области основной вклад в стрикционные эффекты дает зависимость обменного интеграла от расстояния, рассматриваются лишь упруго-изотропные системы.

Репличный гамильтониан неупорядоченной сжимаемой системы с эффектами дальногодействия имеет вид:

$$\begin{aligned}
 H_0 = & \frac{1}{2} \int d^D q (\tau_1 + q^a) \sum_{a=1}^m \Phi_q^a \Phi_{-q}^a + \frac{1}{2} \int d^D q (\tau_2 + q^a) \sum_{a=1}^m \Psi_q^a \Psi_{-q}^a + \\
 & + u_{01} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Phi_{q_3}^b \Phi_{-q_1-q_2-q_3}^b) + \\
 & + u_{02} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Psi_{q_1}^a \Psi_{q_2}^a) (\Psi_{q_3}^b \Psi_{-q_1-q_2-q_3}^b) +
 \end{aligned}$$

$$\begin{aligned}
& +2u_{03} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Psi_{q_3}^b \Psi_{-q_1-q_2-q_3}^b) - \\
& - \frac{\delta_{01}}{2} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Phi_{q_3}^a \Phi_{-q_1-q_2-q_3}^a) - \\
& - \frac{\delta_{02}}{2} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Psi_{q_1}^a \Psi_{q_2}^a) (\Psi_{q_3}^a \Psi_{-q_1-q_2-q_3}^a) - \\
& - \delta_{03} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Psi_{q_3}^a \Psi_{-q_1-q_2-q_3}^a) + \\
& + g_1 \int d^D q_1 d^D q_2 y_{q_1} \sum_{a=1}^m \Phi_{q_2}^a \Phi_{-q_1-q_2}^a + g_2 \int d^D q_1 d^D q_2 y_{q_1} \sum_{a=1}^m \Psi_{q_2}^a \Psi_{-q_1-q_2}^a + \\
& + \frac{g_1^0}{\Omega} y_0 \int d^D q \sum_{a=1}^m \Phi_q^a \Phi_{-q}^a + \frac{g_2^0}{\Omega} y_0 \int d^D q \sum_{a=1}^m \Psi_q^a \Psi_{-q}^a + 2\beta \int d^D q y_q y_{-q} + 2\frac{\beta_0}{\Omega} y_0^2,
\end{aligned} \tag{1}$$

где Φ и Ψ — m -мерные флуктуирующие параметры порядка, u_{01} и u_{02} — положительные константы,

$$\tau_1 \sim |T - T_{c1}|/T_{c1},$$

$$\tau_2 \sim |T - T_{c2}|/T_{c2},$$

T_{c1} и T_{c2} — температуры фазового перехода для первого и второго параметра порядка соответственно,

$$y(x) = \sum_{\alpha=1}^3 u_{\alpha\alpha}(x),$$

где $u_{\alpha\beta}$ — тензор деформаций, g_1 и g_2 — параметры квадратичной стрикции, β — постоянная, характеризующая упругие свойства кристалла, D — размерность пространства. Свойства исходной системы могут быть получены в пределе $m \rightarrow 0$. Неотрицательные константы δ_{01} , δ_{02} , δ_{03} описывают взаимодействие критических флуктуаций через поле примесей. Взаимодействие примесей с упругими деформациями носит линейный характер и при усреднении по примесям приводит к переопределению констант δ_{01} , δ_{02} , δ_{03} . В данном гамильтониане уже проведено интегрирование по слагаемым, зависящим от нефлуктуирующих переменных, не взаимодействующих с параметром порядка. В (1) выделены слагаемые y_0 , описывающие однородные деформации. Как показано в работе [9], такое разделение необходимо, так как неоднородные деформации y_q отвечают за обмен акустическими фононами и приводят к эффектам дальнего действия, которые отсутствуют при однородных деформациях.

Определим эффективный гамильтониан системы, зависящий только от сильно флуктуирующих параметров порядка Φ и Ψ , следующим образом:

$$\exp\{-H[\Phi, \Psi]\} = B \int \exp\{-H_0[\Phi, \Psi, y]\} \prod dy_q. \tag{2}$$

Если эксперимент осуществляется при постоянном объеме, то y_0 является константой, интегрирование в (2) проводится только по неоднородным деформациям и однородные деформации вклада в эффективный гамильтониан не вносят. При постоянном давлении в гамильтониан добавляется слагаемое $P\Omega$, объем представляется в терминах компонент тензора деформации в виде

$$\Omega = \Omega_0[1 + \sum_{\alpha=1} u_{\alpha\alpha} + \sum_{\alpha \neq \beta} u_{\alpha\alpha} u_{\beta\beta} + O(u^3)] \quad (3)$$

и интегрирование в (2) осуществляется также и по однородным деформациям. Как отмечено в [10], учет в (3) квадратичных слагаемых может оказаться важным в случае высоких давлений и кристаллов с большими стрикционными эффектами. В результате:

$$\begin{aligned} H = & \frac{1}{2} \int d^D q (\tau_1 + q^2) \sum_{a=1}^m \Phi_q^a \Phi_{-q}^a + \frac{1}{2} \int d^D q (\tau_2 + q^2) \sum_{a=1}^m \Psi_q^a \Psi_{-q}^a + \\ & + v_{01} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Phi_{q_3}^b \Phi_{-q_1-q_2-q_3}^b) + \\ & + v_{02} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Psi_{q_1}^a \Psi_{q_2}^a) (\Psi_{q_3}^b \Psi_{-q_1-q_2-q_3}^b) + \\ & + 2v_{03} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Psi_{q_3}^b \Psi_{-q_1-q_2-q_3}^b) - \\ & - \frac{\delta_{01}}{2} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Phi_{q_3}^a \Phi_{-q_1-q_2-q_3}^a) - \\ & - \frac{\delta_{02}}{2} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Psi_{q_1}^a \Psi_{q_2}^a) (\Psi_{q_3}^a \Psi_{-q_1-q_2-q_3}^a) - \\ & - \delta_{03} \int d^D q_1 d^D q_2 d^D q_3 \sum_{a=1}^m (\Phi_{q_1}^a \Phi_{q_2}^a) (\Psi_{q_3}^a \Psi_{-q_1-q_2-q_3}^a) + \\ & + \frac{z_1^2 - w_1^2}{2} \int d^D q_1 d^D q_2 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{-q_1}^a) (\Phi_{q_2}^b \Phi_{-q_2}^b) + \\ & + \frac{z_2^2 - w_2^2}{2} \int d^D q_1 d^D q_2 \sum_{a=1}^m (\Psi_{q_1}^a \Psi_{-q_1}^a) (\Psi_{q_2}^a \Psi_{-q_2}^a) + \\ & + (z_1 z_2 - w_1 w_2) \int d^D q_1 d^D q_2 \sum_{a,b=1}^m (\Phi_{q_1}^a \Phi_{-q_1}^a) (\Psi_{q_2}^b \Psi_{-q_2}^b), \\ & v_{01} = u_{01} - z_1^2/2, \quad v_{02} = u_{02} - z_2^2/2, \quad v_{03} = u_{03} - z_1 z_2/2, \\ & z_1 = \frac{g_1}{\sqrt{\beta}}, \quad z_2 = \frac{g_2}{\sqrt{\beta}}, \quad w_1 = \frac{g_1^0}{\sqrt{\beta_0}}, \quad w_2 = \frac{g_2^0}{\sqrt{\beta_0}}. \end{aligned} \quad (4)$$

Данный гамильтониан приводит к широкому разнообразию

мультикритических точек. Возможно как бикритическое

$$(v_3 + (z_1 z_2 - w_1 w_2 - \delta_3)/2)^2 > (v_1 + (z_1^2 - w_1^2 - \delta_1)/2)(v_2 + (z_2^2 - w_2^2 - \delta_2)/2),$$

так и тетракритическое

$$(v_3 + (z_1 z_2 - w_1 w_2 - \delta_3)/2)^2 < (v_1 + (z_1^2 - w_1^2 - \delta_1)/2)(v_2 + (z_2^2 - w_2^2 - \delta_2)/2)$$

поведение. В первом случае в мультикритической точке пересекаются две линии фазовых переходов второго рода и одна линия фазовых переходов первого рода, во втором — четыре линии фазовых переходов второго рода. В непосредственной окрестности мультикритической точки система демонстрирует специфическое критическое поведение, характеризующееся конкуренцией типов упорядочения. При этом в бикритической точке происходит вытеснение одного критического параметра другим, тетракритическая же точка допускает существование смешанной фазы с сосуществующими типами упорядочения. Кроме того стрикционные эффекты могут приводить к мультикритическим точкам более высокого порядка.

В рамках теоретико-полевого подхода [11] асимптотическое критическое поведение и структура фазовых диаграмм во флуктуационной области определяется ренормгрупповым уравнением Каллана-Симанчика для вершинных частей неприводимых функций Грина. Для вычисления β - и γ -функций как функций входящих в уравнение Каллана-Симанчика перенормированных вершин взаимодействия $u_1, u_2, u_3, \delta_1, \delta_2, \delta_3, g_1, g_2, g_1^{(0)}, g_2^{(0)}$ или более удобных для определения мультикритического поведения модели комплексных вершин $z_1, z_2, w_1, w_2, v_1, v_2, v_3, \delta_1, \delta_2, \delta_3$ был применен стандартный метод, основанный на диаграммной технике Фейнмана и процедуре перенормировки [12]. В результате, в рамках двухпетлевого приближения были получены следующие выражения для β -функций:

$$\begin{aligned} \beta_{v_1} = & -v_1 + 36v_1^2 + 4v_3^2 - 24v_1\delta_1 - 1728 \left(2\tilde{J} - 1 - \frac{2}{9}\tilde{G} \right) v_1^3 - \\ & - 192 \left(2\tilde{J} - 1 - \frac{2}{9}\tilde{G} \right) v_1 v_3^2 - 64(2\tilde{J} - 1)v_3^3 + \\ & + 96 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_1 v_3 \delta_3 + 32(2\tilde{J} - 1)v_3^2 \delta_3 + \\ & + 2304(2\tilde{J} - 1 - \frac{1}{6}\tilde{G})v_1^2 \delta_1 - 672 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_1 \delta_1^2 + 16(2\tilde{J} - 1)v_3^2 \delta_1, \\ \beta_{v_2} = & -v_2 + 36v_2^2 + 4v_3^2 - 24v_2\delta_2 - 1728 \left(2\tilde{J} - 1 - \frac{2}{9}\tilde{G} \right) v_2^3 - \\ & - 192 \left(2\tilde{J} - 1 - \frac{2}{9}\tilde{G} \right) v_2 v_3^2 - 64(2\tilde{J} - 1)v_3^3 + 96 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_2 v_3 \delta_3 + \\ & + 32(2\tilde{J} - 1)v_3^2 \delta_3 + 2304 \left(2\tilde{J} - 1 - \frac{1}{6}\tilde{G} \right) v_2^2 \delta_2 - \\ & - 672 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_2 \delta_2^2 + 16(2\tilde{J} - 1)v_3^2 \delta_2, \end{aligned} \quad (5)$$

$$\begin{aligned}
\beta_{v3} &= -v_3 + 16v_3^2 + 12v_1v_3 + 12v_2v_3 - 4v_3\delta_1 - 4v_3\delta_2 - 16v_3\delta_3 - \\
&- 320\left(2\tilde{J} - 1 - \frac{2}{5}\tilde{G}\right)v_3^3 - 288\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_1^2v_3 - \\
&- 288\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_2^2v_3 - 576(2\tilde{J} - 1)v_1v_3^2 - 576(2\tilde{J} - 1)v_2v_3^2 + \\
&+ 448\left(2\tilde{J} - 1 - \frac{2}{7}\tilde{G}\right)v_3^2\delta_3 + 192(2\tilde{J} - 1)v_3^2\delta_1 + 192(2\tilde{J} - 1)v_3^2\delta_2 - \\
&- 48\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_3\delta_1^2 - 48\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_3\delta_2^2 + \\
&+ 432\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_1v_3\delta_1 + 432\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_2v_3\delta_2 + \\
&+ 576(2\tilde{J} - 1)v_1v_3\delta_3 + 576(2\tilde{J} - 1)v_2v_3\delta_3 - 192(2\tilde{J} - 1)v_3\delta_3^2 - \\
&- 192(2\tilde{J} - 1)v_3\delta_1\delta_3 - 192(2\tilde{J} - 1)v_3\delta_2\delta_3, \\
\beta_{\delta 1} &= -\delta_1 + 16\delta_1^2 - 24v_1\delta_1 - 8v_3\delta_3 - 352\left(2\tilde{J} - 1 - \frac{1}{22}\tilde{G}\right)\delta_1^3 + \\
&+ 128(2\tilde{J} - 1)v_3^2\delta_3 - 192(2\tilde{J} - 1)v_3\delta_3^2 - 1152\left(2\tilde{J} - 1 - \frac{1}{3}\tilde{G}\right)v_1\delta_1^2 + \\
&+ 576\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_1^2\delta_1 + 64(2\tilde{J} - 1 - 2\tilde{G})v_3^2\delta_1 - \\
&- 192\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_3\delta_1\delta_3, \\
\beta_{\delta 2} &= -\delta_2 + 16\delta_2^2 - 24v_2\delta_2 - 8v_3\delta_3 - 352\left(2\tilde{J} - 1 - \frac{1}{22}\tilde{G}\right)\delta_2^3 + \\
&+ 128(2\tilde{J} - 1)v_3^2\delta_3 - 192(2\tilde{J} - 1)v_3\delta_3^2 - 1152\left(2\tilde{J} - 1 - \frac{1}{3}\tilde{G}\right)v_2\delta_2^2 + \\
&+ 576\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_2^2\delta_2 + 64(2\tilde{J} - 1 - 2\tilde{G})v_3^2\delta_2 - \\
&- 192\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_3\delta_2\delta_3, \\
\beta_{\delta 3} &= -\delta_3 + 8\delta_3^2 + 12v_1\delta_3 + 12v_2\delta_3 + 4v_3\delta_1 + 4v_3\delta_2 + 4\delta_1\delta_3 + 4\delta_2\delta_3 - \\
&- 64(2\tilde{J} - 1)\delta_3^3 + 288(2\tilde{J} - 1)v_1\delta_3^2 + 288(2\tilde{J} - 1)v_2\delta_3^2 + 288(2\tilde{J} - 1)v_1^2\delta_3 + \\
&+ 288(2\tilde{J} - 1)v_2^2\delta_3 + 48\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)\delta_1^2\delta_3 + 48\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)\delta_2^2\delta_3 + \\
&+ 96(2\tilde{J} - 1)\delta_1\delta_3^2 + 96(2\tilde{J} - 1)\delta_2\delta_3^2 + 64(2\tilde{J} - 1)v_3^2\delta_1 + 64(2\tilde{J} - 1)v_3^2\delta_2 - \\
&- 32(2\tilde{J} - 1)v_3\delta_1^2 - 32(2\tilde{J} - 1)v_3\delta_2^2 + 64(2\tilde{J} - 1 - 2\tilde{G})v_3^2\delta_3 - \\
&- 192(2\tilde{J} - 1 - 2\tilde{G})v_3\delta_3^2 - 288\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_1\delta_1\delta_3 - \\
&- 288\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_2\delta_2\delta_3 - 128(2\tilde{J} - 1)v_3\delta_1\delta_3 - 128(2\tilde{J} - 1)v_3\delta_2\delta_3, \\
\beta_{z1} &= -z_1 + 24v_1z_1 + 2z_1^3 - 16\delta_1z_1 - 4\delta_3z_2 + 2z_1z_2^2 + 4v_3z_2 - \\
&- 576\left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}\right)v_1^2z_1 - 32(2\tilde{J} - 1)v_3^2z_1 - 16(2\tilde{J} - 1 - \tilde{G})v_3^2z_2 +
\end{aligned}$$

$$\begin{aligned}
& + 120 \left(2\tilde{J} - 1 - \frac{8}{5}\tilde{G} \right) v_1 z_1 \delta_1 - 32 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) z_1 \delta_1^2 - 16(2\tilde{J} - 1) z_1 \delta_3^2 - \\
& - 32(2\tilde{J} - 1) z_2 \delta_3^2 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 z_1 \delta_3 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 z_2 \delta_3, \\
\beta_{z_2} & = -z_2 + 24v_2 z_2 + 2z_2^3 - 16\delta_2 z_2 - 4\delta_3 z_1 + 2z_2 z_1^2 + 4v_3 z_1 - \\
& - 576 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_2^2 z_2 - 32(2\tilde{J} - 1) v_3^2 z_2 - 16(2\tilde{J} - 1 - \tilde{G}) v_3^2 z_1 + \\
& + 120 \left(2\tilde{J} - 1 - \frac{8}{5}\tilde{G} \right) v_2 z_2 \delta_2 - 32(2\tilde{J} - 1 - \frac{2}{3}\tilde{G}) z_2 \delta_2^2 - 16(2\tilde{J} - 1) z_2 \delta_3^2 - \\
& - 32(2\tilde{J} - 1) z_1 \delta_3^2 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 z_2 \delta_3 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 z_1 \delta_3, \\
\beta_{w_1} & = -w_1 + 24v_1 w_1 + 2z_1^2 w_1 - 2w_1^3 - 16\delta_1 w_1 - 4\delta_3 w_2 + 2w_1 z_2^2 + \\
& + 4v_3 w_2 - 576 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_1^2 w_1 - \\
& - 32(2\tilde{J} - 1) v_3^2 w_1 - 16(2\tilde{J} - 1 - \tilde{G}) v_3^2 w_2 + 120 \left(2\tilde{J} - 1 - \frac{8}{5}\tilde{G} \right) v_1 w_1 \delta_1 - \\
& - 32 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) w_1 \delta_1^2 - 16(2\tilde{J} - 1) w_1 \delta_3^2 - 32(2\tilde{J} - 1) w_2 \delta_3^2 + \\
& + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 w_1 \delta_3 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 w_2 \delta_3, \\
\beta_{w_2} & = -w_2 + 24v_2 w_2 + 2z_2^2 w_2 - 2w_2^3 - 16\delta_2 w_2 - 4\delta_3 w_1 + 2w_2 z_1^2 + \\
& + 4v_3 w_1 - 576 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) v_2^2 w_2 - \\
& - 32(2\tilde{J} - 1) v_3^2 w_2 - 16(2\tilde{J} - 1 - \tilde{G}) v_3^2 w_1 + 120 \left(2\tilde{J} - 1 - \frac{8}{5}\tilde{G} \right) v_1 w_2 \delta_1 - \\
& - 32 \left(2\tilde{J} - 1 - \frac{2}{3}\tilde{G} \right) w_2 \delta_1^2 - 16(2\tilde{J} - 1) w_2 \delta_3^2 - 32(2\tilde{J} - 1) w_1 \delta_3^2 + \\
& + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 w_2 \delta_3 + 128(2\tilde{J} - 1 - 5\tilde{G}) v_3 w_1 \delta_3.
\end{aligned}$$

Известно, что ряды теории возмущений являются асимптотическими, а вершины взаимодействия флуктуаций параметров порядка во флуктуационной области достаточно велики, чтобы можно было непосредственно применять выражения (5). Поэтому с целью извлечения из полученных выражений нужной физической информации был применен обобщенный на многопараметрический случай метод Паде-Бореля. При этом прямое и обратное преобразования Бореля имеют вид

$$\begin{aligned}
& f(v_1, v_2, v_3, \delta_1, \delta_2, \delta_3, z_1, z_2, w_1, w_2) = \\
& = \sum_{i_1, \dots, i_{10}} c_{i_1 \dots i_{10}} v_1^{i_1} v_2^{i_2} v_3^{i_3} \delta_1^{i_4} \delta_2^{i_5} \delta_3^{i_6} z_1^{i_7} z_2^{i_8} w_1^{i_9} w_2^{i_{10}} = \\
& = \int_0^\infty e^{-t} F(v_1 t, v_2 t, v_3 t, \delta_1 t, \delta_2 t, \delta_3 t, z_1 t, z_2 t, w_1 t, w_2 t) dt, \\
& F(v_1, v_2, v_3, \delta_1, \delta_2, \delta_3, z_1, z_2, w_1, w_2) =
\end{aligned} \tag{6}$$

$$= \sum_{i_1, \dots, i_{10}} \frac{c_{i_1, \dots, i_7}}{(i_1 + \dots + i_{10})!} v_1^{i_1} v_2^{i_2} v_3^{i_3} \delta_1^{i_4} \delta_2^{i_5} \delta_3^{i_6} z_1^{i_7} z_2^{i_8} w_1^{i_9} w_2^{i_{10}}.$$

Для аналитического продолжения борелевского образа функции вводится ряд по вспомогательной переменной θ :

$$\begin{aligned} \tilde{F}(v_1, v_2, v_3, \delta_1, \delta_2, \delta_3, z_1, z_2, w_1, w_2, \theta) = \\ = \sum_{k=0}^{\infty} \theta^k \sum_{i_1, \dots, i_{10}} \frac{c_{i_1 \dots i_{10}}}{k!} v_1^{i_1} v_2^{i_2} v_3^{i_3} \delta_1^{i_4} \delta_2^{i_5} \delta_3^{i_6} z_1^{i_7} z_2^{i_8} w_1^{i_9} w_2^{i_{10}} \delta_{(i_1 + \dots + i_{10}), k}, \end{aligned} \quad (7)$$

к которому применяется аппроксимация Паде $[L/M]$ в точке $\theta = 1$.

В двухпетлевом приближении для вычисления β -функций был использован аппроксимант $[2/1]$. Природа критического поведения определяется существованием устойчивой фиксированной точки, удовлетворяющей системе уравнений:

$$\beta_i(v_1^*, v_2^*, v_3^*, \delta_1^*, \delta_2^*, \delta_3^*, z_1^*, z_2^*, w_1^*, w_2^*) = 0 \quad (i = 1, \dots, 10). \quad (8)$$

Требование устойчивости фиксированной точки сводится к условию, чтобы собственные значения b_i матрицы

$$\begin{aligned} B_{i,j} = \frac{\partial \beta_i(v_1^*, v_2^*, v_3^*, \delta_1^*, \delta_2^*, \delta_3^*, z_1^*, z_2^*, w_1^*, w_2^*)}{\partial v_j} \\ (v_i, v_j \equiv v_1^*, v_2^*, v_3^*, \delta_1^*, \delta_2^*, \delta_3^*, z_1^*, z_2^*, w_1^*, w_2^*) \end{aligned} \quad (9)$$

лежали в правой комплексной полуплоскости.

Полученная система просуммированных β -функций содержит широкое разнообразие фиксированных точек, лежащих в физической области значений вершин ($v_i \geq 0$, $\delta_i \geq 0$, $i = 1, 2, 3$). Полный анализ фиксированных точек, соответствующих критическому поведению только одного параметра порядка, приведен в работе [3]. Фиксированные точки однородной системы, описываемой двумя параметрами порядка, с эффектами дальнего действия были получены в работе [5]. Фиксированные точки, отвечающие за мультикритическое поведение сжимаемых однородных приведены в Таблице 1, неупорядоченных систем — в Таблице 2. В обеих таблицах введен дополнительный параметр

$$\begin{aligned} p = (v_3 + (z_1 z_2 - w_1 w_2 - \delta_3)/2)^2 - \\ - (v_1 + (z_1^2 - w_1^2 - \delta_1)/2)(v_2 + (z_2^2 - w_2^2 - \delta_2)/2), \end{aligned}$$

определяющий тип мультикритического поведения. При значениях $p > 0$ в системе наблюдается бикритическое поведение, в обратном случае $p \leq 0$ — тетракритическое.

Анализ значений фиксированных точек однородных сжимаемых систем с эффектами дальнего действия и их устойчивости позволяет сделать ряд выводов. Для всех значений параметра дальнего действия $1,5 < a < 2$ на фазовой

Таблица 1. Значения фиксированных точек и собственных значений матрицы устойчивости однородных систем

N	v_1^*, v_2^*, v_3^*	z_1^*, z_2^*	w_1^*, w_2^*	b_1, b_2, b_3	b_4, b_5	b_6, b_7	p
$a = 1, 6$							
1,1	0,027427	0,224319	0,111301	0,157	1,118	1,798	-0,002092
	0,027427	-0,224319	-0,111301	0,738	0,138	0,256	
	0,026699			0,919			
1,2	0,027427	0,224319	0,224319	0,157	1,118	-1,050	-0,000039
	0,027427	-0,224319	-0,224319	0,738	0,138	-1,188	
	0,026699			0,919			
$a = 1, 7$							
2,1	0,031287	0,248013	0	0,113	1,675	2,546	-0,003849
	0,031287	-0,248013	0	0,629	0,095	0,041	
	0,031334			0,809			
2,2	0,031287	0,248013	0,248013	0,113	1,675	-1,580	0,000039
	0,031287	-0,248013	-0,248013	0,629	0,095	-1,675	
	0,031334			0,809			
$a = 1, 8$							
3,1	0,033682	0,266919	0	0,090	1,831	2,980	-0,004802
	0,033682	-0,266919	0	0,571	0,104	0,115	
	0,034575			0,753			
3,2	0,033682	0,266919	0,266919	0,090	1,831	-1,954	0,000061
	0,033682	-0,266919	-0,266919	0,571	0,104	-2,079	
	0,034575			0,753			
$a = 1, 9$							
4,1	0,035842	0,297071	0	0,069	2,079	3,765	-0,079943
	0,035842	-0,297071	0	0,505	0,125	0,049	
	0,039202			0,702			
4,2	0,035842	0,297071	0,297071	0,069	2,079	-1,954	0,000252
	0,035842	-0,297071	-0,297071	0,505	0,125	-2,079	
	0,039202			0,702			

диаграмме вещества может наблюдаться тетракритическое поведение ($p < 0$), для этого константы, характеризующие стрикционное взаимодействие флуктуирующих параметров порядка с деформационными степенями свободы, должны быть разного знака. Для систем, характеризующихся стрикционными константами одного знака, устойчивых фиксированных точек не существует, что свидетельствует о срыве на фазовый переход первого рода. Устойчивые фиксированные точки (ФТ), описывающие тетракритическое поведение системы в интервале значений параметра порядка $1,6 < a < 2$ (ФТ 2.1, 3.1, 4.1), характеризуются нулевыми значениями эффективных зарядов w_1^* и w_2^* , свидетельствующих о том, что данная тетракритическая точка должна наблюдаться при постоянном объеме системы. При значении параметра

Таблица 2. Значения фиксированных точек и собственных значений матрицы устойчивости неупорядоченных систем

N	v_1^*, v_2^*, v_3^*	$\delta_1^*, \delta_2^*, \delta_3^*$	b_1, b_2, b_3	b_4, b_5, b_6
$a = 1,6$				
1,3	0	-0,630515	11,225	22,854
	0	-0,630515	15,069	22,854
	0	-3,713534	15,069	17,439
$a = 1,7$				
2,3	0	-0,320512	2,239	11,106
	0	-0,320512	2,239	11,116
	0	-1,932190	5,197	9,250
$a = 1,8$				
3,3	0	-0,256437	2,068	8,642
	0	-0,256437	2,068	8,642
	0	-1,545749	3,890	7,441
$a = 1,9$				
4,3	0	-0,223270	1,224	3,149
	0	-0,223270	1,338	7,349
	0	-0,790114	1,338	7,349

дальнодействия $a = 1,6$ эффективные заряды w_1^* и w_2^* отличны от нуля (ФТ 1.1). По-видимому данное отличие обусловлено тем, что в отсутствие упругих деформаций система демонстрирует тетракритическое поведение при $a = 1,6$ и бикритическое при остальных значениях параметра дальнодействия $1,6 < a < 2$. Кроме тетракритических точек упругие деформации приводят к возможности появления на фазовой диаграмме вещества критических точек более высокого порядка. Так фиксированные точки 1,2, 2,2, 3,2, 4,2 характеризуются равенством эффективных зарядов $z_1 = w_1$ и $z_2 = w_2$, то есть в них пересекаются по две линии трикритического поведения, в следствие чего они соответствуют критическим точкам четвертого порядка. Данные фиксированные точки не являются устойчивыми в рамках выбранного приближения, однако для полного разрешения вопроса об их устойчивости необходим учет в исходном гамильтониане кубических слагаемых по деформационным степеням свободы.

Как хорошо видно из Таблицы 2 фиксированные точки неупорядоченной «жесткой» системы лежат в нефизической области значений ($\delta_1, \delta_2 < 0$). Более того все устойчивые фиксированные точки имеют достаточно большие значения по модулю, для которых не применимы положения используемой теории. Отсюда можно сделать вывод о неустойчивости всех фиксированных точек, лежащих в физической области значений эффективных зарядов, и, как следствие, об отсутствии на фазовой диаграмме вещества мультিকритических точек. Нахождение устойчивых фиксированных

точек сжимаемых неупорядоченных систем лишено смысла, так как они характеризуются теми же значениями эффективных зарядов v_i , δ_i , что и несжимаемые системы, а потому также будут лежать в нефизической области. Таким образом введение в систему замороженных дефектов структуры приводит к размытию мультикритического поведения.

ЛИТЕРАТУРА

1. Fisher M. E., Ma S.-k., Nickel B. G. Critical exponents for long-range interaction // Phys. Rev. Lett. 1972. V.29. P.917.
2. Белим С.В. Влияние эффектов дальнего действия на критическое поведение трехмерных систем // Письма в ЖЭТФ - 2003 - № 77 - В. 2 - С. 118-120.
3. Белим С.В. Влияние упругих деформаций на критическое поведение неупорядоченных систем с эффектами дальнего действия // ЖЭТФ. 2003. Т.124, Вып.6. С.507-513.
4. Белим С.В. Прудников В.В. Трикритическое поведение сжимаемых систем с замороженными дефектами структуры // ФТТ. 2001. Т.43, Вып.7. С.1299.
5. Белим С.В. Влияние эффектов дальнего действия на мультикритическое поведение однородных систем // ЖЭТФ. 2004. Т.125, Вып.1. С.152.
6. Прудников В.В., Прудников П.В., Федоренко А.А. // ФТТ. 2000. Т.42, N.1. С.158.
7. Белим С.В. Влияние стрикционных эффектов на мультикритическое поведение однородных систем // Письма в ЖЭТФ. 2002. N.75, Вып.9. С.547-550.
8. Imry Y. Tricritical Points in Compressible Magnetic Systems // Phys. Rev. Lett. B. 1974. V.33. P.1304.
9. Ларкин А.И., Пикин С.А. // ЖЭТФ. 1969. Т.56. С.1664.
10. Bergman D.J., Halperin B.I. Critical behavior of an Ising model on a cubic compressible lattice // Phys.Rev.B. 1976. V.13, N.4. P.2145.
11. Amit D. Field theory the renormalization group and critical phenomena. New York: McGraw-Hill, 1976.
12. Zinn-Justin J. Quantum field theory and critical phenomena. Oxford: Clarendon Press, 1989.

О КВАТЕРНИОННОЙ ФОРМЕ УРАВНЕНИЯ ДИРАКА

А.И. Бреев, В.В. Клишевич

Мы рассматриваем кватернионные формы уравнения Дирака и их связи с уравнением Гельмгольца. Дается доказательство формулы Коши.

Одно из направлений, связанное с различными обобщениями уравнения Дирака, основывается на различном представлении оператора Дирака как в плоском, так и в искривленном пространстве. В работе [1] предложена формулировка теории уравнения Дирака в случае, когда волновая функция рассматривается как полный набор антисимметричных тензорных полей по отношению к группе общекоординатных преобразований. В работе [2] дана матричная формулировка этого формализма, приспособленная для исследований соответствующей квантовой теории. Интересным аспектом построенной конструкции является факт «точного» извлечения квадратного корня из оператора Клейна-Гордона. Иначе говоря, квадрат оператора Дирака тождественно совпадает с оператором Клейна-Гордона как в плоском, так и в искривленном пространстве. В классической теории уравнения Дирака на римановом многообразии это свойство не наблюдается [3].

В данной статье мы изучаем одно специальное представление, которое было введено в работе Кравченко [4]. Мы работаем с уравнением Дирака в кватернионной форме. Фактически, рассмотренное представление есть аналог уравнения Дирака в трехмерном евклидовом пространстве. Основным интерес к данной конструкции связан с различными приложениями в математической физике, в частности, с уравнением Гельмгольца и в кватернионном анализе, который рассматривается как некоторое обобщение комплексного анализа.

Содержание нашей статьи следующее. В разделе 1 приведены краткие сведения о кватернионах. В разделе 2 приведено уравнение Гельмгольца и рассмотрены условия излучения Зоммерфельда, которые обычно используют для выделения физически значимых решений. Аналог этих условий используется в разделе 3 для введенной кватернионной формы уравнения Дирака. Раздел 3 основной. Здесь мы определяем правую и левую кватернионные формы оператора Дирака, представляем эти формы в виде скалярной и векторной части, а также находим фундаментальные решения. В заключительной части 4 дано доказательство кватернионной интегральной формулы Коши, которая есть аналог известной формулы комплексного анализа.

1. Кватернионы

Кватернион a может быть представлен в виде линейной комбинации своих компонент ($\{a_k\} \subset \mathbb{R}$ для действительных кватернионов и $\{a_k\} \subset \mathbb{C}$ для комплексных кватернионов) и элементов ортонормированного базиса i_k следующим образом

$$a = \sum_{k=0}^3 a_k i_k,$$

где $i_0 = 1$ и $\{i_k \mid k = 1, 2, 3\}$ – кватернионные мнимые единицы, обладающие свойствами: $i_k^2 = -1$, $k = 1, 2, 3$ и

$$i_1 i_2 = -i_2 i_1 = i_3, \quad i_2 i_3 = -i_3 i_2 = i_1, \quad i_3 i_1 = -i_1 i_3 = i_2.$$

В некоторых случаях кватернион a рассматривают в виде суммы скалярной $\text{Sc}(a) = a_0$ и векторной части

$$\text{Vec}(a) = \vec{a} = a_1 i_1 + a_2 i_2 + a_3 i_3.$$

Используя это представление, мы можем определить результат умножения двух произвольных кватернионов a и b в форме

$$a \cdot b = a_0 b_0 - \langle \vec{a}, \vec{b} \rangle + a_0 \vec{b} + \vec{a} b_0 + [\vec{a} \times \vec{b}], \quad (1)$$

где $\langle \cdot, \cdot \rangle$ – скалярное умножение, $[\cdot \times \cdot]$ – векторное.

Кватернион сопряженный кватерниону $a = a_0 + \vec{a}$ определяется как

$$\bar{a} = a_0 - \vec{a}.$$

Из (1) легко получить, что

$$a \cdot \bar{a} = a_0^2 + a_1^2 + a_2^2 + a_3^2 = |a|^2. \quad (2)$$

Сопряженные кватернионы обладают свойством

$$\overline{a \cdot b} = \bar{b} \cdot \bar{a}.$$

Отметим, что

$$|a \cdot b| = |a| \cdot |b|.$$

Из (2) следует, что каждый ненулевой действительный кватернион a обратим и обратный кватернион дается формулой $a^{-1} = \bar{a}/|a|^2$.

Модуль в (2) может быть введен как

$$|a| = \sqrt{|a_0|^2 + |a_1|^2 + |a_2|^2 + |a_3|^2}$$

или в виде

$$|a|^2 = |\text{Re } a|^2 + |\text{Im } a|^2.$$

Здесь $|a_k|^2 = a_k a_k^*$ и «*» означает комплексное сопряжение.

2. Уравнение Гельмгольца

Рассмотрим уравнение Гельмгольца

$$\Delta u + \nu^2 u = -f(x). \quad (3)$$

В [5] показано, что

$$\theta_{\pm\nu}(x) = -\frac{e^{\pm i\nu|x|}}{4\pi|x|}$$

есть фундаментальное решение оператора Гельмгольца:

$$(\Delta + \nu^2)\theta_{\pm\nu} = \delta. \quad (4)$$

Чтобы выделить класс единственности решения для уравнения Гельмгольца в неограниченных областях, являющихся внешностью ограниченных областей, нужно накладывать дополнительные ограничения на поведение решений на бесконечности. Одними из таких ограничений являются условия излучения Зоммерфельда

$$\frac{\partial u(x)}{\partial|x|} \mp i\nu u(x) = o\left(\frac{1}{|x|}\right), |x| \rightarrow \infty \quad (5)$$

$$u(x) = O\left(\frac{1}{|x|}\right). \quad (6)$$

Знак минус в формуле (5) соответствует расходящимся волнам, а знак плюс — сходящимся.

Решение уравнения Гельмгольца (3) удовлетворяющее условиям излучения (5) и (6) можно рассматривать как амплитуду установившегося колебания, полученную с помощью предельного перехода из неустойчившихся колебаний, вызванных периодическим внешним возмущением с частотой ν и амплитудой $f(x)$

3. Кватернионная форма оператора Дирака

Пусть задана функция f принимает значения в $\mathbb{H}(\mathbb{C})$ и дифференцируема в области $\Omega \subset \mathbb{R}^3$. Следуя статье [4], определим оператор Moisil-Theodoresco

$$D_l f = \sum_{k=1}^3 i_k \partial_k f, \quad (7)$$

где

$$\partial_k = \frac{\partial}{\partial x_k}.$$

Определение 1. Оператор (7) называется (левой) кватернионной формой трехмерного оператора Дирака.

Получим представление оператора (7) в виде векторной и скалярной части

$$\begin{aligned}
D_l = \sum_{k=1}^3 i_k \partial_k f &= i_1 \left(\frac{\partial f_0}{\partial x_1} + \frac{\partial f_1}{\partial x_1} i_1 + \frac{\partial f_2}{\partial x_1} i_2 + \frac{\partial f_3}{\partial x_1} i_3 \right) + \\
&+ i_2 \left(\frac{\partial f_0}{\partial x_2} + \frac{\partial f_1}{\partial x_2} i_1 + \frac{\partial f_2}{\partial x_2} i_2 + \frac{\partial f_3}{\partial x_2} i_3 \right) + \\
&+ i_3 \left(\frac{\partial f_0}{\partial x_3} + \frac{\partial f_1}{\partial x_3} i_1 + \frac{\partial f_2}{\partial x_3} i_2 + \frac{\partial f_3}{\partial x_3} i_3 \right) = \\
&= - \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \frac{\partial f_3}{\partial x_3} \right) + \left(\frac{\partial f_0}{\partial x_1} i_1 + \frac{\partial f_0}{\partial x_2} i_2 + \frac{\partial f_0}{\partial x_3} i_3 \right) + \\
&+ \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) i_1 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) i_2 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) i_3 = \\
&= \left(-\operatorname{div} \vec{f}, \operatorname{grad} f_0 + \operatorname{rot} \vec{f} \right), \tag{8}
\end{aligned}$$

где $\vec{f}$ и f_0 — векторная и скалярная часть кватернионной функции f соответственно. Аналогично определяется правый оператор Moisil-Theodoresco:

$$D_r f = \sum_{k=1}^3 (\partial_k f) i_k. \tag{9}$$

Определение 2. Оператор (9) называется (правой) кватернионной формой трехмерного оператора Дирака.

Поступая также как в (8) получим

$$D_r f = \left(-\operatorname{div} \vec{f}, \operatorname{grad} f_0 - \operatorname{rot} \vec{f} \right).$$

Далее мы будем работать только с D_l и для этой формы оператора Дирака будем использовать символ D . Заметим, что

$$\begin{aligned}
D_l^2 f = D_r^2 f &= i_1 \frac{\partial}{\partial x_1} \left(i_1 \frac{\partial f}{\partial x_1} \right) + i_1 \frac{\partial}{\partial x_1} \left(i_2 \frac{\partial f}{\partial x_2} \right) + i_1 \frac{\partial}{\partial x_1} \left(i_3 \frac{\partial f}{\partial x_3} \right) + \\
&+ i_2 \frac{\partial}{\partial x_2} \left(i_1 \frac{\partial f}{\partial x_1} \right) + i_2 \frac{\partial}{\partial x_2} \left(i_2 \frac{\partial f}{\partial x_2} \right) + i_2 \frac{\partial}{\partial x_2} \left(i_3 \frac{\partial f}{\partial x_3} \right) + \\
&+ i_3 \frac{\partial}{\partial x_3} \left(i_1 \frac{\partial f}{\partial x_1} \right) + i_3 \frac{\partial}{\partial x_3} \left(i_2 \frac{\partial f}{\partial x_2} \right) + i_3 \frac{\partial}{\partial x_3} \left(i_3 \frac{\partial f}{\partial x_3} \right) = \\
&= - \left(\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2} \right) = -\Delta f.
\end{aligned}$$

Введем оператор

$$D_{\pm\nu} = D \pm \nu I,$$

где ν комплексное число. С помощью этого оператора можно представить оператор Гельмгольца в форме

$$\Delta + \nu^2 = -(D + \nu)(D - \nu) = -D_\nu D_{-\nu}. \tag{10}$$

Лемма 1. Фундаментальное решение оператора $D_{\pm\nu}$ имеет вид

$$\mathcal{K}_{\pm\nu} = -(D \mp \nu)\theta_\nu. \quad (11)$$

Доказательство. Из формулы (10) следует, что

$$(\Delta + \nu^2)\theta_\nu = -D_\nu D_{-\nu}\theta_\nu = \delta.$$

Функция $-(D - \nu)\theta_\nu$ является фундаментальным решением оператора D_ν :

$$\mathcal{K}_\nu = -(D - \nu)\theta_\nu. \quad (12)$$

Заменим ν на $-\nu$ и учитывая формулу (4) получим

$$-(D - \nu)(D + \nu)\theta_{-\nu} = -(D - \nu)(D + \nu)\theta_\nu = D_{-\nu}(-(D + \nu)\theta_\nu) = \delta.$$

То есть функция $-(D + \nu)\theta_\nu$ является фундаментальным решением оператора $D_{-\nu}$. Учитывая (12), получим (11). ■

Найдем представление $\mathcal{K}_\nu$ в виде скалярной и векторной части. В силу формул (11), (8), а также того, что θ_ν является скаляром, получим:

$$\begin{aligned} \mathcal{K}_{\pm\nu} &= -(D \mp \nu)\theta_\nu = -D\theta_\nu \pm \nu\theta_\nu = \\ &= \operatorname{div}(\theta_\nu) - \operatorname{grad}(\theta_\nu) - \operatorname{rot}(\theta_\nu) \pm \nu\theta_\nu = (\pm\nu\theta_\nu, -\operatorname{grad}(\theta_\nu)). \end{aligned} \quad (13)$$

Отсюда, в частности, следует что

$$D_r\mathcal{K}_{\pm\nu} = D\mathcal{K}_{\pm\nu} \quad (14)$$

(так как D_r отличается от D только знаком перед ротором, а из предыдущей формулы видно что ротор $\mathcal{K}_{\pm\nu}$ равен нулю).

Отдельно вычислим градиент

$$\begin{aligned} \operatorname{grad}(\theta_\nu) &= \sum_{k=1}^3 (\partial_k \theta_\nu) i_k = - \sum_{k=1}^3 \frac{i\nu x_k - \frac{x_k}{|x|}}{4\pi|x|^2} i_k e^{i\nu|x|} = \\ &= - \frac{e^{i\nu|x|}}{4\pi|x|} \left(i\nu \frac{\vec{x}}{|x|} - \frac{\vec{x}}{|x|^2} \right) = -\theta_\nu \left(\frac{\vec{x}}{|x|^2} - i\nu \frac{\vec{x}}{|x|} \right). \end{aligned} \quad (15)$$

Таким образом, учитывая (13), получаем выражение для $\mathcal{K}_{\pm\nu}$ в явном виде:

$$\mathcal{K}_{\pm\nu} = \left(\pm\nu + \frac{\vec{x}}{|x|^2} - i\nu \frac{\vec{x}}{|x|} \right) \theta_\nu.$$

Приведем без доказательства важную теорему, которая будет использована в дальнейшем

Теорема 1. (Кватернионная формула Стокса) Рассмотрим область Ω лежащую в $\mathbb{R}^3$ с кусочно-гладкой границей $\Gamma = \partial\Omega$. Пусть f и g есть кватернионные функции (из $\mathbb{H}(\mathbb{C})$) дифференцируемые в этой области и непрерывные на ее границе. Тогда

$$\int_{\Omega} ((D_r f(y))g(y) + f(y)(Dg(y))) dy = \int_{\Gamma} f(y)\vec{n}(y)g(y)d\Gamma_y, \quad (16)$$

где $\vec{n}(y)$ — внешняя нормаль к поверхности Γ : $\vec{n}(y) = \sum_{k=1}^3 i_k n_k(y)$.

4. Кватернионная интегральная формула Коши

Рассмотрим интегральный оператор

$$K_\nu[f](x) = - \int_\Gamma \mathcal{K}_\nu(x-y) \vec{n}(y) f(y) d\Gamma_y, \quad x \in \mathbb{R}^3 \setminus \Gamma, \quad (17)$$

который может быть рассмотрен как аналог интегрального оператора Коши вследствие следующей теоремы приведенной ниже.

Лемма 2. *Справедлива формула*

$$\mathcal{K}_\nu(\vec{x} - \vec{y}) = -\mathcal{K}_\nu(\vec{y} - \vec{x}) + 2\nu\theta_\nu(\vec{y} - \vec{x}). \quad (18)$$

Доказательство. Используя (13) и (15) получим:

$$\begin{aligned} \mathcal{K}_\nu(\vec{x} - \vec{y}) &= -D\theta_\nu(\vec{x} - \vec{y}) + \nu\theta_\nu(\vec{x} - \vec{y}) = \\ &= -\text{grad } \theta_\nu(\vec{x} - \vec{y}) + \nu\theta_\nu(\vec{x} - \vec{y}) = \\ &= \left(\frac{\vec{x} - \vec{y}}{|\vec{x} - \vec{y}|^2} - i\nu \frac{\vec{x} - \vec{y}}{|\vec{x} - \vec{y}|} \right) \theta_\nu(\vec{x} - \vec{y}) + \nu\theta_\nu(\vec{x} - \vec{y}) = \\ &= - \left(\frac{\vec{y} - \vec{x}}{|\vec{y} - \vec{x}|^2} - i\nu \frac{\vec{y} - \vec{x}}{|\vec{y} - \vec{x}|} \right) \theta_\nu(\vec{y} - \vec{x}) + \nu\theta_\nu(\vec{y} - \vec{x}) = \\ &= - \left(\left(\frac{\vec{y} - \vec{x}}{|\vec{y} - \vec{x}|^2} - i\nu \frac{\vec{y} - \vec{x}}{|\vec{y} - \vec{x}|} \right) \theta_\nu(\vec{y} - \vec{x}) + \nu\theta_\nu(\vec{y} - \vec{x}) \right) + \\ &+ 2\nu\theta_\nu(\vec{y} - \vec{x}) = -\mathcal{K}_\nu(\vec{y} - \vec{x}) + 2\nu\theta_\nu(\vec{y} - \vec{x}). \end{aligned}$$

■

Теорема 2. (Кватернионная интегральная формула Коши) Пусть функция $f \in \mathbb{H}((C))$ дифференцируема в области Ω из $\mathbb{R}^3$ и непрерывна на ее границе. Тогда если $f \in \ker D_\nu(\Omega)$ то

$$f(x) = K_\nu[f](x), \quad \forall x \in \Omega.$$

Доказательство. Применим к (17) кватернионную формулу Стокса (16):

$$\begin{aligned} K_\nu[f](x) &= - \int_\Gamma \mathcal{K}_\nu(\vec{x} - \vec{y}) \vec{n}(\vec{y}) f(\vec{y}) d\Gamma_y = \\ &= - \int_\Omega (\{D_r \mathcal{K}_\nu(\vec{x} - \vec{y})\} f(\vec{y}) + \mathcal{K}_\nu(\vec{x} - \vec{y}) \{Df(\vec{y})\}). \end{aligned} \quad (19)$$

Воспользовавшись формулами (19), (14), леммой 2 и учитывая, что $D_\nu f = 0$,

преобразуем (19):

$$\begin{aligned}
 K_\nu[f](x) &= - \int_{\Omega} (\{D\mathcal{K}_\nu(\vec{x} - \vec{y})\}f(\vec{y}) + \mathcal{K}_\nu(\vec{x} - \vec{y})\{Df(\vec{y})\}) dy = \\
 &= - \int_{\Omega} (\{D_\nu\mathcal{K}_\nu(\vec{x} - \vec{y}) - \nu\mathcal{K}_\nu(\vec{x} - \vec{y})\}f(\vec{y}) + \\
 &\quad + \mathcal{K}_\nu(\vec{x} - \vec{y})\{D_\nu f(\vec{y}) - \nu f(\vec{y})\}) dy = \\
 &= - \int_{\Omega} (\{D_\nu(-\mathcal{K}_\nu(\vec{y} - \vec{x}) + 2\nu\theta_\nu(\vec{y} - \vec{x})) - \\
 &\quad - \nu\mathcal{K}_\nu(\vec{x} - \vec{y})\}f(\vec{y}) - \nu\mathcal{K}_\nu(\vec{x} - \vec{y})f(\vec{y})) dy = \\
 &= - \int_{\Omega} (\{-\delta(\vec{y} - \vec{x}) + 2\nu D_\nu\theta_\nu(\vec{y} - \vec{x})\}f(\vec{y}) - \\
 &\quad - 2\nu\mathcal{K}_\nu(\vec{x} - \vec{y})f(\vec{y})) dy = \\
 &= f(\vec{x}) - 2\nu \int_{\Omega} \{D_\nu\theta_\nu(\vec{y} - \vec{x})\}f(\vec{y})dy + 2\nu \int_{\Omega} \mathcal{K}_\nu(\vec{x} - \vec{y})f(\vec{y})dy.
 \end{aligned} \tag{20}$$

Вычислим первый интеграл последнего выражения:

$$\int_{\Omega} \{D_\nu\theta_\nu(\vec{y} - \vec{x})\}f(\vec{y})dy = \int_{\Omega} \{D\theta_\nu(\vec{y} - \vec{x}) + \nu\theta_\nu(\vec{y} - \vec{x})\}f(\vec{y})dy. \tag{21}$$

Воспользовавшись формулами (11) и (19) вычислим $D\theta_\nu(\vec{y} - \vec{x})$:

$$\begin{aligned}
 D\theta_\nu(\vec{y} - \vec{x}) &= -\mathcal{K}_\nu(\vec{y} - \vec{x}) + \nu\theta_\nu(\vec{y} - \vec{x}) = \\
 &= \mathcal{K}_\nu(\vec{x} - \vec{y}) - 2\nu\theta_\nu(\vec{y} - \vec{x}) + \nu\theta_\nu(\vec{y} - \vec{x}) = \\
 &= \mathcal{K}_\nu(\vec{x} - \vec{y}) - \nu\theta_\nu(\vec{y} - \vec{x}).
 \end{aligned}$$

Подставляя последнее выражение в интеграл (21) получим:

$$\int_{\Omega} \{D_\nu\theta_\nu(\vec{y} - \vec{x})\}f(\vec{y})dy = \int_{\Omega} \mathcal{K}_\nu(\vec{x} - \vec{y})f(\vec{y})dy.$$

Теперь подставим этот интеграл в формулу (20) и получим

$$K_\nu[f](x) = f(x).$$

Что и требовалось доказать. ■

5. Заключение

В статье рассмотрено кватернионное представление уравнения Дирака, которое связано с уравнением Гельмгольца. Как известно, уравнение Гельмгольца есть трехмерный аналог стационарного уравнения Шредингера для свободной нерелятивистской частицы. В этом контексте можно придать физический смысл доказанной кватернионной формуле Коши, которая является естественным обобщением интегральной формулы Коши из комплексного анализа. Физически значимые решения (убывающие на бесконечности) выделяются с помощью условий излучения Зоммерфельда и удовлетворяют задаче на собственные значения в интегральной форме.

ЛИТЕРАТУРА

1. Benn I.M., Tucker R.W. Weyl scaling and conformal properties of the Kahler equation
// Phys.Lett. B. 1983. V.132. N.4. P.325
2. Ционенко Д.А. Уравнение Дирака-Кэлера в неевклидовом пространстве-времени.
// Вести нац.акад.наук Белоруссии. 2003, N.1. С.81.
3. Klishevich V.V. On the existense of the second Dirac operator // Class. Quantum Grav. V.17. P.305
Klishevich V.V. Group conjugations of Dirac operators as an invariant of the Riemannian manifold // Class. Quantum Grav. V.19. P.4287.
4. Kravchenko V.V. and Castillo R.P. An analogue of the Sommerfeld radiation condition for the Dirac operator. // Math. Meth. Appl. Sci. 2002. V.25. P.1383.
5. Владимиров В.С. Уравнения математической физики. М.: Наука, 1984.

ТОЧНОЕ РЕШЕНИЕ ДЛЯ ПОЛЯ $(1, 0) \oplus (0, 1)$ В ТЕРМИНАХ ФУНКЦИЙ НА ГРУППЕ ПУАНКАРЕ

В.В. Варламов

В работе определяется безмассовое скалярное поле спина 1 на 8-мерном конфигурационном пространстве; данное пространство является прямым произведением пространства Минковского и двумерной комплексной сферы. Полевые уравнения для поля со спином 1 получаются из диракоподобного лагранжиана, разделенного на две составляющие, одна из которых связана с группой трансляций, а другая – с группой Лоренца. Показано, что диракоподобная форма уравнений Максвелла (электродинамика в так называемой формулировке Майорана-Оппенгеймера) следует непосредственно из полевых уравнений для составляющей группы трансляций. Поле фотона представляется функциями Биденхарна на группе Пуанкаре. Полученные результаты позволяют нам рассматривать поля Дирака и Максвелла на равных основаниях, как функции на группе Пуанкаре.

Как известно, любая величина, которая преобразуется линейно под действием преобразований Лоренца, является спинором. По этой причине спинорные величины рассматриваются как фундаментальные в квантовой теории поля и основные уравнения для таких величин должны быть записаны в спинорной форме. Спинорная форма уравнений Дирака была впервые дана Ван дер Варденом [1], где было показано, что теория Дирака может быть полностью описана в этой форме (см. также [2]). В свою очередь, спинорная формулировка уравнений Максвелла изучалась Лапортом и Уленбеком [3]. В 1936 г. Румер [7] показал, что спинорные формы уравнений Дирака и Максвелла выглядят очень схоже¹. Далее, Майорана [8] и Оппенгеймер [9] предложили рассматривать максвелловскую теорию электромагнетизма как волновую механику фотона. Они ввели волновую функцию вида $\psi = \mathbf{E} - i\mathbf{B}$, удовлетворяющую безмассовым

Copyright © 2005 **В.В. Варламов**.

Сибирский государственный индустриальный университет.

E-mail: root@varlamov.kemerovo.su

¹Конечно, не существует эквивалентности между уравнениями Дирака и Максвелла, как это утверждает Камполаттаро [4, 5] (см. также дискуссию, касающуюся этого вопроса [6]). Главное различие между этими уравнениями содержится в том, что поля Дирака и Максвелла имеют различные спинтензорные размерности. Данные поля преобразуются соответственно в рамках конечномерных представлений $(1/2, 0) \oplus (0, 1/2)$ и $(1, 0) \oplus (0, 1)$ группы Лоренца.

диракоподобным уравнениям². Уравнения Максвелла в диракоподобной форме рассматривались многими авторами на протяжении длительного времени [12–24]. Интерес к электродинамике в формулировке Майораны-Оппенгеймера возрос в последние годы [25–33].

В современной теоретической физике общепризнанно, что большинство «элементарных» частиц имеют очень сложную внутреннюю структуру, т.е. элементарные частицы представляются пространственно протяженными. По этой причине реальные частицы нельзя рассматривать как точечные объекты (конечно, в общем эти более реалистичные поля являются нелокальными полями). В стороне от струнных моделей, имеется другой путь для описания пространственно протяженных частиц, предложенный Финкельштейном [34], который показал, что модели элементарных частиц с внутренними степенями свободы могут быть описаны на многообразиях более широких чем пространство-время Минковского (однородные пространства группы Пуанкаре $\mathcal{P}$). Все однородные пространства группы $\mathcal{P}$, которые содержат пространство-время, были исследованы Финкельштейном [34], Бакри и Кильбергом [35]. В 1964г. под влиянием теории полюсов Редже, Люрса предложил построить квантовую теорию поля на групповом многообразии группы $\mathcal{P}$ [36]. Построения квантово-полевых теорий на различных однородных пространствах группы $\mathcal{P}$ рассматривались в работах [37–46].

Главной целью настоящей статьи является синтез двух упомянутых выше направлений (диракоподобная формулировка уравнений Максвелла и квантовая теория поля на группе Пуанкаре). Поле Максвелла представляется функциями Биденхарна [47] на групповом многообразии $\mathcal{M}_{10}$ (данное многообразие является прямым произведением пространства-времени и группового многообразия группы Лоренца). Показывается, что общий вид волновой функции наследует свою структуру от полупрямого произведения $SL(2, \mathbb{C}) \odot T_4$ и по этой причине поле Максвелла на $\mathcal{M}_{10}$ определяется факторизацией $\psi(x)\psi(\mathbf{g})$, где $x \in T_4$, $\mathbf{g} \in SL(2, \mathbb{C})$. Очевидно, что диракоподобная форма уравнений Максвелла должна быть получена из диракоподобного лагранжиана. С использованием лагранжева формализма на касательном расслоении $T\mathcal{M}_{10}$ многообразия $\mathcal{M}_{10}$ находятся полевые уравнения отдельно для частей $\psi(x)$ и $\psi(\mathbf{g})$. Решения полевых уравнений для $\psi(x)$ получены в приближении плоской волны³. В свою очередь, решения полевых уравнений с $\psi(\mathbf{g})$ найдены в форме разложений по присоединенным гиперсферическим функциям⁴.

²В противоположность методу Гупта-Блейлера [10, 11], где квантуется ненаблюдаемый вектор-потенциал A_μ , одно из главных преимуществ электродинамики Майораны-Оппенгеймера заключается в том, что она оперирует с наблюдаемыми величинами, такими как электрическое и магнитное поля.

³Как известно, экспоненты определяют унитарные представления подгруппы трансляций T_4 . В некотором смысле, функции e^{ikx} могут пониматься как «матричные элементы» группы T_4 .

⁴Матричные элементы как спинорной, так и главной серий представлений группы Лоренца выражаются через гиперсферические функции [48] (см. Приложение).

1. Функции на группе Пуанкаре

Рассмотрим некоторые основные факты, касающиеся группы Пуанкаре $\mathcal{P}$. Прежде всего, группа $\mathcal{P}$ имеет то же число связных компонент, что и группа Лоренца. Будем рассматривать только компоненту $\mathcal{P}_+^\uparrow$, соответствующую связной компоненте $L_+^\uparrow$ (так называемая специальная группа Лоренца [49]). Как известно, универсальная накрывающая $\overline{\mathcal{P}_+^\uparrow}$ группы $\mathcal{P}_+^\uparrow$ определяется полупрямым произведением

$$\overline{\mathcal{P}_+^\uparrow} = SL(2, \mathbb{C}) \odot T_4 \simeq \mathbf{Spin}_+(1, 3) \odot T_4,$$

где T_4 — подгруппа четырехмерных трансляций.

Каждое преобразование $T_\alpha \in \mathcal{P}_+^\uparrow$ определяется множеством параметров $\alpha(\alpha_1, \dots, \alpha_{10})$, которое может быть представлено точкой пространства $\mathcal{M}_{10}$. Пространство $\mathcal{M}_{10}$ обладает локально евклидовыми свойствами, следовательно, $\mathcal{M}_{10}$ является *групповым многообразием группы Пуанкаре*. Легко видеть, что множество α может быть разбито на два подмножества, $\alpha(x_1, x_2, x_3, x_4, \mathfrak{g}_1, \mathfrak{g}_2, \mathfrak{g}_3, \mathfrak{g}_4, \mathfrak{g}_5, \mathfrak{g}_6)$, где $x_i \in T_4$ — параметры подгруппы трансляций, $\mathfrak{g}_j$ — параметры группы $SL(2, \mathbb{C})$. В свою очередь, преобразование $T_{\mathfrak{g}}$ определяется множеством $\mathfrak{g}(\mathfrak{g}_1, \dots, \mathfrak{g}_6)$, которое может быть представлено точкой шестимерного подмногообразия $\mathcal{L}_6 \subset \mathcal{M}_{10}$, называемого *групповым многообразием группы Лоренца*.

В настоящей статье мы ограничимся рассмотрением конечномерных представлений группы Пуанкаре. Группа четырехмерных трансляций T_4 является абелевой группой, образованной прямым произведением четырех одномерных групп трансляций, каждая из которых изоморфна аддитивной группе вещественных чисел. Отсюда следует, что все неприводимые представления группы T_4 являются одномерными и выражаются через экспоненту. В свою очередь, как было показано Наймарком [50], спинорные представления исчерпывают по существу все конечномерные представления группы $SL(2, \mathbb{C})$. Любое спинорное представление группы $SL(2, \mathbb{C})$ может быть определено в пространстве симметрических полиномов следующего вида:

$$p(z_0, z_1, \bar{z}_0, \bar{z}_1) = \sum_{\substack{(\alpha_1, \dots, \alpha_k) \\ (\dot{\alpha}_1, \dots, \dot{\alpha}_r)}} \frac{1}{k! r!} a^{\alpha_1 \dots \alpha_k \dot{\alpha}_1 \dots \dot{\alpha}_r} z_{\alpha_1} \dots z_{\alpha_k} \bar{z}_{\dot{\alpha}_1} \dots \bar{z}_{\dot{\alpha}_r} \quad (1)$$

$$(\alpha_i, \dot{\alpha}_i = 0, 1),$$

где числа $a^{\alpha_1 \dots \alpha_k \dot{\alpha}_1 \dots \dot{\alpha}_r}$ не изменяются при перестановках индексов. Выражения (1) следует понимать как *функции на группе Лоренца*. Когда коэффициенты $a^{\alpha_1 \dots \alpha_k \dot{\alpha}_1 \dots \dot{\alpha}_r}$ в (1) зависят от переменных $x_i \in T_4$ ($i = 1, 2, 3, 4$), мы приходим функциям Биденхарна [47]:

$$p(x, z, \bar{z}) = \sum_{\substack{(\alpha_1, \dots, \alpha_k) \\ (\dot{\alpha}_1, \dots, \dot{\alpha}_r)}} \frac{1}{k! r!} a^{\alpha_1 \dots \alpha_k \dot{\alpha}_1 \dots \dot{\alpha}_r}(x) z_{\alpha_1} \dots z_{\alpha_k} \bar{z}_{\dot{\alpha}_1} \dots \bar{z}_{\dot{\alpha}_r} \quad (2)$$

$$(\alpha_i, \dot{\alpha}_i = 0, 1)$$

Функции (2) следует рассматривать как *функции на группе Пуанкаре*. Некоторые приложения этих функций содержатся в [45, 51]. Представления группы Пуанкаре $SL(2, \mathbb{C}) \odot T(4)$ реализуются посредством функций (2).

Пусть $\mathcal{L}(\alpha)$ – лагранжиан на групповом многообразии $\mathcal{M}_{10}$ группы Пуанкаре (другими словами, $\mathcal{L}(\alpha)$ является 10-мерной функцией точки), где α – множество параметров данной группы. Тогда интеграл от $\mathcal{L}(\alpha)$ по некоторому 10-мерному объему Ω группового многообразия будем называть *действием на группе Пуанкаре*:

$$A = \int_{\Omega} d\alpha \mathcal{L}(\alpha),$$

где $d\alpha$ – мера Хаара⁵ на группе $\mathcal{P}$.

Пусть $\psi(\alpha)$ – функция на многообразии $\mathcal{M}_{10}$ (сейчас достаточно предположить, что $\psi(\alpha)$ является квадратично интегрируемой функцией на группе Пуанкаре) и пусть

$$\frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial \alpha} \frac{\partial \mathcal{L}}{\partial \frac{\partial \psi}{\partial \alpha}} = 0 \quad (3)$$

– уравнения Эйлера-Лагранжа на $\mathcal{M}_{10}$ (более точно говоря, уравнения (3) действуют в касательном расслоении $T\mathcal{M}_{10} = \bigcup_{\alpha \in \mathcal{M}_{10}} T_{\alpha}\mathcal{M}_{10}$ многообразия $\mathcal{M}_{10}$, см. [52]). Введем теперь лагранжиан $\mathcal{L}(\alpha)$, зависящий от полевой функции $\psi(\alpha)$ следующим образом:

$$\mathcal{L}(\alpha) = -\frac{1}{2} \left(\psi^*(\alpha) B_{\mu} \frac{\partial \psi(\alpha)}{\partial \alpha_{\mu}} - \frac{\partial \psi^*(\alpha)}{\partial \alpha_{\mu}} B_{\mu} \psi(\alpha) \right) - \kappa \psi^*(\alpha) B_{11} \psi(\alpha),$$

где B_{ν} ($\nu = 1, 2, \dots, 10$) – квадратные матрицы. Число строк и столбцов этих матриц равно числу компонент функции $\psi(\alpha)$, κ – отличная от нуля вещественная постоянная.

Далее, если B_{11} невырождена, то можно ввести матрицы

$$\Pi_{\mu} = B_{11}^{-1} B_{\mu}, \quad \mu = 1, 2, \dots, 10,$$

и представить лагранжиан $\mathcal{L}(\alpha)$ в виде

$$\mathcal{L}(\alpha) = -\frac{1}{2} \left(\bar{\psi}(\alpha) \Pi_{\mu} \frac{\partial \psi(\alpha)}{\partial \alpha_{\mu}} - \frac{\partial \bar{\psi}(\alpha)}{\partial \alpha_{\mu}} \Pi_{\mu} \psi(\alpha) \right) - \kappa \bar{\psi}(\alpha) \psi(\alpha),$$

где

$$\bar{\psi}(\alpha) = \psi^*(\alpha) B_{11}.$$

В случае безмассового поля $(j, 0) \oplus (0, j)$ рассмотрим на групповом многообразии $\mathcal{M}_{10}$ лагранжиан вида

$$\mathcal{L}(\alpha) = -\frac{1}{2} \left(\bar{\psi}(\alpha) \Pi_{\mu} \frac{\partial \psi(\alpha)}{\partial \alpha_{\mu}} - \frac{\partial \bar{\psi}(\alpha)}{\partial \alpha_{\mu}} \Pi_{\mu} \psi(\alpha) \right). \quad (4)$$

⁵Инвариантная мера $d\alpha$ на группе Пуанкаре может быть факторизована в виде $d\alpha = dx d\mathbf{g}$, где $d\mathbf{g}$ – мера Хаара на группе Лоренца.

Как прямое следствие определения (2) релятивистская волновая функция $\psi(\alpha)$ на групповом многообразии $\mathcal{M}_{10}$ представляется следующей факторизацией:

$$\psi(\alpha) = \psi(x)\psi(\mathbf{g}) = \psi(x_1, x_2, x_3, x_4)\psi(\varphi, \epsilon, \theta, \tau, \phi, \varepsilon), \quad (5)$$

где $\psi(x_i)$ — функция, зависящая от параметров подгруппы T_4 ,

$$x_i \in T_4 \quad (i = 1, \dots, 4),$$

а $\psi(\mathbf{g})$ — функция на группе Лоренца, где шесть параметров этой группы определяются углами Эйлера $\varphi, \epsilon, \theta, \tau, \phi, \varepsilon$, которые образуют комплексные углы вида

$$\varphi^c = \varphi - i\epsilon, \quad \theta^c = \theta - i\tau, \quad \phi^c = \phi - i\varepsilon$$

(см. Приложение).

2. Поле $(1, 0) \oplus (0, 1)$

Как уже отмечалось выше, диракоподобная форма уравнений Максвелла должна быть получена из диракоподобного лагранжиана. Это одно из главных утверждений, которое мы докажем в данной статье. С этой целью перепишем лагранжиан (4) в виде

$$\begin{aligned} \mathcal{L}(\alpha) = & -\frac{1}{2} \left(\bar{\psi}(\alpha) \Gamma_\mu \frac{\partial \psi(\alpha)}{\partial x_\mu} - \frac{\partial \bar{\psi}(\alpha)}{\partial x_\mu} \Gamma_\mu \psi(\alpha) \right) - \\ & - \frac{1}{2} \left(\bar{\psi}(\alpha) \Upsilon_\nu \frac{\partial \psi(\alpha)}{\partial \mathbf{g}_\nu} - \frac{\partial \bar{\psi}(\alpha)}{\partial \mathbf{g}_\nu} \Upsilon_\nu \psi(\alpha) \right), \end{aligned} \quad (6)$$

где

$$\psi(\alpha) = \psi(x)\psi(\mathbf{g}) \quad (\mu = 0, 1, 2, 3, \nu = 1, \dots, 6),$$

и

$$\Gamma_0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \Gamma_1 = \begin{pmatrix} 0 & -\alpha_1 \\ \alpha_1 & 0 \end{pmatrix}, \quad \Gamma_2 = \begin{pmatrix} 0 & -\alpha_2 \\ \alpha_2 & 0 \end{pmatrix}, \quad \Gamma_3 = \begin{pmatrix} 0 & -\alpha_3 \\ \alpha_3 & 0 \end{pmatrix},$$

$$\Upsilon_1 = \begin{pmatrix} 0 & \Lambda_1^* \\ \Lambda_1 & 0 \end{pmatrix}, \quad \Upsilon_2 = \begin{pmatrix} 0 & \Lambda_2^* \\ \Lambda_2 & 0 \end{pmatrix}, \quad \Upsilon_3 = \begin{pmatrix} 0 & \Lambda_3^* \\ \Lambda_3 & 0 \end{pmatrix}, \quad (7)$$

$$\Upsilon_4 = \begin{pmatrix} 0 & i\Lambda_1^* \\ i\Lambda_1 & 0 \end{pmatrix}, \quad \Upsilon_5 = \begin{pmatrix} 0 & i\Lambda_2^* \\ i\Lambda_2 & 0 \end{pmatrix}, \quad \Upsilon_6 = \begin{pmatrix} 0 & i\Lambda_3^* \\ i\Lambda_3 & 0 \end{pmatrix}, \quad (8)$$

где

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (9)$$

$$\Lambda_1 = \frac{c_{11}}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Lambda_2 = \frac{c_{11}}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \Lambda_3 = c_{11} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (10)$$

Варьируя независимо $\psi(x)$ и $\bar{\psi}(x)$ в лагранжиане (6), приходим к уравнениям

$$\Gamma_\mu \frac{\partial \psi(x)}{\partial x_\mu} = 0, \quad (11)$$

$$\Gamma_\mu^T \frac{\partial \bar{\psi}(x)}{\partial x_\mu} = 0. \quad (12)$$

Уравнение (11) может быть переписано следующим образом:

$$\left[\frac{i\hbar}{c} \frac{\partial}{\partial t} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} - i\hbar \frac{\partial}{\partial \mathbf{x}} \begin{pmatrix} 0 & -\alpha_i \\ \alpha_i & 0 \end{pmatrix} \right] \begin{pmatrix} \psi(x) \\ \psi^*(x) \end{pmatrix} = 0, \quad (13)$$

где

$$\begin{pmatrix} \psi(x) \\ \psi^*(x) \end{pmatrix} = \begin{pmatrix} \mathbf{E} - i\mathbf{B} \\ \mathbf{E} + i\mathbf{B} \end{pmatrix} = \begin{pmatrix} E_1 - iB_1 \\ E_2 - iB_2 \\ E_3 - iB_3 \\ E_1 + iB_1 \\ E_2 + iB_2 \\ E_3 + iB_3 \end{pmatrix}.$$

Из уравнения (13) следует, что

$$\left(\frac{i\hbar}{c} \frac{\partial}{\partial t} - i\hbar \alpha_i \frac{\partial}{\partial \mathbf{x}} \right) \psi(x) = 0, \quad (14)$$

$$\left(\frac{i\hbar}{c} \frac{\partial}{\partial t} + i\hbar \alpha_i \frac{\partial}{\partial \mathbf{x}} \right) \psi^*(x) = 0. \quad (15)$$

Последние уравнения в совокупности с условиями поперечности

$$\mathbf{p} \cdot \psi = 0, \quad \mathbf{p} \cdot \psi^* = 0$$

совпадают с уравнениями Максвелла. Действительно, принимая во внимание уравнение $(\mathbf{p} \cdot \alpha)\psi = \hbar \nabla \times \psi$, получим

$$\frac{i\hbar}{c} \frac{\partial \psi}{\partial t} = -\hbar \nabla \times \psi, \quad (16)$$

$$-i\hbar \nabla \cdot \psi = 0. \quad (17)$$

Откуда

$$\begin{aligned} \nabla \times (\mathbf{E} - i\mathbf{B}) &= -\frac{i}{c} \frac{\partial (\mathbf{E} - i\mathbf{B})}{\partial t}, \\ \nabla \cdot (\mathbf{E} - i\mathbf{B}) &= 0 \end{aligned}$$

(константа $\hbar$ сокращается). Разделяя вещественную и мнимую части, получим уравнения Максвелла

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \times \mathbf{B} &= \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}, \\ \nabla \cdot \mathbf{E} &= 0, \\ \nabla \cdot \mathbf{B} &= 0. \end{aligned}$$

Легко проверить, что мы снова придем к уравнениям Максвелла исходя из уравнений

$$\left(\frac{i\hbar}{c}\frac{\partial}{\partial t} + i\hbar\alpha_i\frac{\partial}{\partial \mathbf{x}}\right)\psi^*(x) = 0, \quad (18)$$

$$-i\hbar\nabla \cdot \psi^*(x) = 0. \quad (19)$$

Несмотря на то, что уравнения (14) и (15) приводят к одним и тем же уравнениям Максвелла, физическая интерпретация данных уравнений различна (см. [28, 32]). А именно, уравнения (14) и (15) есть уравнения с отрицательной и положительной спиральностью, соответственно⁶.

Как обычно, сопряженная волновая функция

$$\bar{\psi}(x) = \psi^\dagger(x)\Gamma_0 = (\psi(x), \psi^*(x))$$

соответствует античастице (данное обстоятельство является прямым следствием диракоподобности лагранжиана (6)). Следовательно, мы пришли к очень спорному заключению, что уравнение (12) описывает античастицу (антифотон) и, более того, отсюда следует существование тока и заряда для фотонного поля. На первый взгляд, мы пришли к резкому противоречию с общепризнанным положением, гласящим, что фотон является истинно нейтральной частицей. Однако, легко проверить, что уравнения (12) также приводят к уравнениям Максвелла. Это означает, что фотон совпадает со своей «античастицей». Следуя стандартной процедуре, приведенной во многих учебниках, мы можем определить «заряд» фотона следующим выражением:

$$Q \sim \int d\mathbf{x} \bar{\psi}\Gamma_0\psi, \quad (20)$$

где $\bar{\psi}\Gamma_0\psi = 2(\mathbf{E}^2 + \mathbf{B}^2)$. Однако, Ньютон и Вигнер [57] показали, что для фотона не существует локализованных состояний. Следовательно, интеграл в

⁶Интересно отметить, что соответствие между спиновыми состояниями фотона и комплексификацией группы $SU(2)$ представляется локальным изоморфизмом вида $SU(2) \otimes SU(2) \simeq SL(2, \mathbb{C})$. Как известно, корневая подгруппа поупростой группы Ли O_4 (группа вращений четырехмерного пространства) является нормальным делителем группы O_4 . По этой причине шестипараметрическая группа O_4 является полупростой и представляется прямым произведением двух трехпараметрических унимодулярных групп. По аналогии с группой O_4 , двукратное накрытие $SL(2, \mathbb{C})$ собственной группы Лоренца (группа вращений четырехмерного пространственно-временного континуума) является полупростой и представляется прямым произведением двух трехпараметрических специальных унимодулярных групп, $SL(2, \mathbb{C}) \simeq SU(2) \otimes SU(2)$. Явный вид данного изоморфизма может быть получен посредством комплексификации группы $SU(2)$, т.е. $SL(2, \mathbb{C}) \simeq \text{complex}(SU(2)) \simeq SU(2) \otimes SU(2)$ [53]. Более того, в работах [54–56] группа Лоренца представляется произведением $SU_R(2) \otimes SU_L(2)$, а спиноры $\psi(p^\mu) = (\phi_R(p^\mu), \phi_L(p^\mu))^T$ преобразуются в рамках пространства представления $(j_1, j_2) \oplus (j_2, j_1)$. Компоненты $\phi_R(p^\mu)$ and $\phi_L(p^\mu)$ соответствуют различным спиральным состояниям (право- и лево-поляризованные спиноры). Отсюда следует аналогия с фотонными спиновыми состояниями. А именно, операторы $X = J + iK$ и $Y = J - iK$ соответствуют право- и лево-поляризованным состояниям фотона, где J и K – генераторы вращений и лоренцевых сдвигов, соответственно.

правой части выражения (20) представляет собой неопределяемое выражение. Поскольку интеграл (20) не существует в общем, то «заряд» фотона не может рассматриваться как постоянная величина (в противоположность электронному полю, которое имеет локализованные состояния [57] и хорошо определенный постоянный заряд). По существу, можно сказать, что «заряд» фотона равен энергии $\mathbf{E}^2 + \mathbf{B}^2$ для γ -кванта.

Таким образом, уравнение (13) приводит к двум диракоподобным уравнениям (14) и (15), которые в комбинации с условиями поперечности (17) и (19), эквивалентны уравнениям Максвелла. Представим решения уравнения (14) в виде плоской волны

$$\psi(x) = \varepsilon(\mathbf{k}) \exp[i\hbar^{-1}(\mathbf{k} \cdot \mathbf{x} - \omega t)]. \quad (21)$$

После подстановки формы решения (21) в (14) приходим к следующей задаче на собственные значения:

$$-c \begin{pmatrix} 0 & ik_3 & -ik_2 \\ -ik_3 & 0 & ik_1 \\ ik_2 & -ik_1 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \omega \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}.$$

Легко проверить, что мы приходим к той же задаче на собственные значения, начиная с уравнения (15). Вековое уравнение имеет решения $\omega = \pm c|\mathbf{k}|, 0$.

Следовательно, решения уравнения (13) в приближении плоской волны задаются функциями

$$\begin{aligned} \psi_{\pm}(\mathbf{k}; \mathbf{x}, t) &= \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_{\pm}(\mathbf{k}) \\ \varepsilon_{\pm}(\mathbf{k}) \end{pmatrix} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)], \\ \psi_0(\mathbf{k}; \mathbf{x}) &= \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_0(\mathbf{k}) \\ \varepsilon_0(\mathbf{k}) \end{pmatrix} \exp[i\mathbf{k} \cdot \mathbf{x}] \end{aligned}$$

и комплексно-сопряженными функциями $\psi_{\pm}^*(\mathbf{k}; \mathbf{x}, t)$, $\psi_0^*(\mathbf{k}; \mathbf{x})$ ($\mathbf{E} + i\mathbf{B}$), соответствующими положительной спиральности; здесь $\omega = c|\mathbf{k}|$ и $\varepsilon_{\lambda}(\mathbf{k})$ ($\lambda = \pm, 0$) — векторы поляризации фотона:

$$\begin{aligned} \varepsilon_{\pm}(\mathbf{k}) &= \{2|\mathbf{k}|^2(k_1^2 + k_2^2)\}^{-\frac{1}{2}} \begin{bmatrix} -k_1 k_3 \pm ik_2 |\mathbf{k}| \\ -k_2 k_3 \mp ik_1 |\mathbf{k}| \\ k_1^2 + k_2^2 \end{bmatrix}, \\ \varepsilon_0(\mathbf{k}) &= |\mathbf{k}|^{-1} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}. \end{aligned}$$

Варьируя теперь $\psi(\mathbf{g})$ и $\bar{\psi}(\mathbf{g})$ в лагранжиане (6), приходим к уравнениям

$$\begin{aligned} \Upsilon_{\nu} \frac{\partial \psi(\mathbf{g})}{\partial \mathbf{g}_{\nu}} &= 0, \\ \Upsilon_{\nu}^T \frac{\partial \bar{\psi}(\mathbf{g})}{\partial \mathbf{g}_{\nu}} &= 0. \end{aligned} \quad (22)$$

Последние уравнения записаны в параметрах группы $SL(2, \mathbb{C})$. Поскольку универсальная накрывающая $SL(2, \mathbb{C})$ собственной группы Лоренца является комплексификацией группы $SU(2)$ (см., например, [58]), то будет более удобным выразить шесть параметров группы $SL(2, \mathbb{C})$ через три параметра a_1, a_2, a_3 группы $SU(2)$. Очевидно, что $\mathfrak{g}_1 = a_1, \mathfrak{g}_2 = a_2, \mathfrak{g}_3 = a_3, \mathfrak{g}_4 = ia_1, \mathfrak{g}_5 = ia_2, \mathfrak{g}_6 = ia_3$. Принимая во внимание структуру матриц Υ_ν , задаваемой выражениями (7)–(8), перепишем первое уравнение из (22) следующим образом:

$$\begin{aligned} \sum_{k=1}^3 \Lambda_k \frac{\partial \psi}{\partial a_k} - i \sum_{k=1}^3 \Lambda_k \frac{\partial \psi}{\partial a_k^*} &= 0, \\ \sum_{k=1}^3 \Lambda_k^* \frac{\partial \dot{\psi}}{\partial \tilde{a}_k} + i \sum_{k=1}^3 \Lambda_k^* \frac{\partial \dot{\psi}}{\partial \tilde{a}_k^*} &= 0, \end{aligned} \quad (23)$$

где $a_1^* = -i\mathfrak{g}_4, a_2^* = -i\mathfrak{g}_5, a_3^* = -i\mathfrak{g}_6$, а $\tilde{a}_k, \tilde{a}_k^*$ – параметры, соответствующие дуальному базису. По существу, уравнения (23) определены в трехмерном комплексном пространстве $\mathbb{C}^3$. В свою очередь, пространство $\mathbb{C}^3$ изометрично шестимерному бивекторному пространству $\mathbb{R}^6$ (пространство параметров группы Лоренца [59]). Бивекторное пространство $\mathbb{R}^6$ является касательным пространством группового многообразия $\mathfrak{L}_6$ группы Лоренца, т.е. групповое многообразие $\mathfrak{L}_6$ в каждой своей точке локально эквивалентно пространству $\mathbb{R}^6$. Таким образом, для всех $\mathfrak{g} \in \mathfrak{L}_6$ имеем $T_{\mathfrak{g}}\mathfrak{L}_6 \simeq \mathbb{R}^6$. Учитывая явный вид⁷ матриц Λ_i , задаваемый выражениями (10), перепишем систему (23) в следующем виде:

$$\begin{aligned} \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_1} - i \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_2} + \frac{\partial \psi_1}{\partial x_3} - i \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_1^*} - \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_2^*} - i \frac{\partial \psi_1}{\partial x_3^*} &= 0, \\ \frac{\partial \psi_1}{\partial x_1} + \frac{\partial \psi_3}{\partial x_1} + i \frac{\partial \psi_1}{\partial x_2} - i \frac{\partial \psi_3}{\partial x_2} - i \frac{\partial \psi_1}{\partial x_1^*} - i \frac{\partial \psi_3}{\partial x_1^*} - \frac{\partial \psi_1}{\partial x_2^*} + \frac{\partial \psi_3}{\partial x_2^*} &= 0, \\ \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_1} + i \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_2} - \frac{\partial \psi_3}{\partial x_3} - i \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_1^*} + \frac{\sqrt{2}}{2} \frac{\partial \psi_2}{\partial x_2^*} + i \frac{\partial \psi_3}{\partial x_3^*} &= 0, \\ \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_1} - i \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_2} + \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_3} + i \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_1^*} + \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_2^*} + i \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_3^*} &= 0, \\ \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_1} + \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_1} + i \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_2} - i \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_2} + i \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_1^*} + i \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_1^*} - \frac{\partial \dot{\psi}_1}{\partial \tilde{x}_2^*} + \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_2^*} &= 0, \\ \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_1} + i \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_2} - \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_3} + i \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_1^*} - \frac{\sqrt{2}}{2} \frac{\partial \dot{\psi}_2}{\partial \tilde{x}_2^*} - i \frac{\partial \dot{\psi}_3}{\partial \tilde{x}_3^*} &= 0. \end{aligned} \quad (24)$$

⁷Матрицы (10) могут быть получены при $l = 1$ из более общих выражений (62)–(67), данных в [62].

Разделение переменных в системе (24) реализуется посредством факторизации:

$$\begin{aligned}\psi_1 &= \mathbf{f}_{1,1}^l(r) \mathfrak{M}_l^1(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \psi_2 &= \mathbf{f}_{1,0}^l(r) \mathfrak{M}_l^0(0, 0, \theta, \tau, 0, 0), \\ \psi_3 &= \mathbf{f}_{1,-1}^l(r) \mathfrak{M}_l^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \dot{\psi}_1 &= \mathbf{f}_{1,1}^i(r^*) \mathfrak{M}_i^1(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \dot{\psi}_2 &= \mathbf{f}_{1,0}^i(r^*) \mathfrak{M}_i^0(0, 0, \theta, \tau, 0, 0), \\ \dot{\psi}_3 &= \mathbf{f}_{1,-1}^i(r^*) \mathfrak{M}_i^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0),\end{aligned}$$

где $\mathfrak{M}_l^m(\varphi, \epsilon, \theta, \tau, 0, 0)$ ($\mathfrak{M}_i^m(\varphi, \epsilon, \theta, \tau, 0, 0)$) — присоединенные гиперсферические функции, определенные на поверхности двумерной комплексной сферы радиуса r , $\mathbf{f}_{lmk}^l(r)$ и $\mathbf{f}_{lmk}^i(r^*)$ — радиальные функции (более подробно о двумерной комплексной сфере см. [60–62])⁸.

Повторяя для случая $l = 1$ все преобразования, проведенные для общей релятивистски-инвариантной системы [62], получим

$$\begin{aligned}2 \frac{d\mathbf{f}_{1,1}^l(r)}{dr} - \frac{1}{r} \mathbf{f}_{1,1}^l(r) - \frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,0}^l(r) &= 0, \\ -\frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,-1}^l(r) + \frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,1}^l(r) &= 0, \\ -2 \frac{d\mathbf{f}_{1,-1}^l(r)}{dr} + \frac{1}{r} \mathbf{f}_{1,-1}^l(r) + \frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,0}^l(r) &= 0, \\ 2 \frac{d\mathbf{f}_{1,1}^i(r^*)}{dr^*} - \frac{1}{r^*} \mathbf{f}_{1,1}^i(r^*) - \frac{\sqrt{2i(i+1)}}{r^*} \mathbf{f}_{1,0}^i(r^*) &= 0,\end{aligned}$$

⁸Выбирая двумерную комплексную сферу в качестве внутреннего спинового пространства, мы видим, что конфигурационное пространство $\mathcal{M}_{10} = \mathbb{R}^{1,3} \times \mathfrak{L}_6$ сводится к $\mathcal{M}_8 = \mathbb{R}^{1,3} \times \mathbb{C}^2$. Бакри и Кильберг [35] утверждали, что восьмимерное конфигурационное пространство $\mathcal{M}_8$ является наиболее подходящим для описания как полуцелого, так и целого спина. $\mathcal{M}_8$ является однородным пространством группы Пуанкаре. Действительно, пространство $\mathbb{C}^2$ гомеоморфно расширенной комплексной плоскости $\mathbb{C} \cup \infty$, которая представляет абсолют (множество бесконечно удаленных точек) пространства Лобачевского $S^{1,2}$. При этом группа дробно-линейных преобразований плоскости $\mathbb{C} \cup \infty$ изоморфна группе движений пространства $S^{1,2}$. В свою очередь, пространство Лобачевского $S^{1,2}$ является абсолютм мира Минковского $\mathbb{R}^{1,3}$ и, следовательно, группа дробно-линейных преобразований плоскости $\mathbb{C} \cup \infty$ дважды покрывает группу вращений пространства $\mathbb{R}^{1,3}$, т.е. группу Лоренца. Нетрудно видеть, что двумерная комплексная сфера является комплексификацией небесной сферы Пенроуза [63]. Далее, используя каноническую проекцию $\pi : \mathbb{C}_*^2 \rightarrow S^2$, где $\mathbb{C}_*^2 = \mathbb{C}^2 / \{0, 0\}$ и S^2 — двумерная вещественная сфера, видим, что конфигурационное пространство $\mathcal{M}_8 = \mathbb{R}^{1,3} \times \mathbb{C}^2$ сводится к $\mathcal{M}_6 = \mathbb{R}^{1,3} \times S^2$. Вещественная 2-сфера S^2 имеет наименьшую возможную размерность среди всех однородных пространств группы Лоренца. По этой причине $\mathcal{M}_6$ является минимальным однородным пространством группы Пуанкаре (пространственно-временные трансляции действуют тривиально на сфере S^2). Полевые модели на конфигурационном пространстве $\mathcal{M}_6$ рассматривались в недавних работах [41–43]. Дрекслер [43] рассматривал 2-сферу как «спиновую оболочку» $S_{r=2s}^2$ радиуса $r = 2s$, где $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$

$$\begin{aligned}
& -\frac{i\sqrt{2l(l+1)}}{r^*} \mathbf{f}_{1,-1}^i(r^*) + \frac{\sqrt{2l(l+1)}}{r^*} \mathbf{f}_{1,1}^i(r^*) = 0, \\
& -2\frac{d\mathbf{f}_{1,-1}^i(r^*)}{dr^*} + \frac{1}{r^*} \mathbf{f}_{1,-1}^i(r^*) + \frac{\sqrt{2l(l+1)}}{r^*} \mathbf{f}_{1,0}^i(r^*) = 0.
\end{aligned} \tag{25}$$

Из второго и пятого уравнений следует $\mathbf{f}_{1,-1}^l(r) = \mathbf{f}_{1,1}^l(r)$ и $\mathbf{f}_{1,-1}^i(r^*) = \mathbf{f}_{1,1}^i(r^*)$. Учитывая эти соотношения запишем систему (25) в следующем виде:

$$\begin{aligned}
& 2\frac{d\mathbf{f}_{1,1}^l(r)}{dr} - \frac{1}{r} \mathbf{f}_{1,1}^l(r) - \frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,0}^l(r) = 0, \\
& -2\frac{d\mathbf{f}_{1,-1}^l(r)}{dr} + \frac{1}{r} \mathbf{f}_{1,-1}^l(r) + \frac{\sqrt{2l(l+1)}}{r} \mathbf{f}_{1,0}^l(r) = 0, \\
& 2\frac{d\mathbf{f}_{1,1}^i(r^*)}{dr^*} - \frac{1}{r^*} \mathbf{f}_{1,1}^i(r^*) - \frac{\sqrt{2l(l+1)}}{r^*} \mathbf{f}_{1,0}^i(r^*) = 0, \\
& -2\frac{d\mathbf{f}_{1,-1}^i(r^*)}{dr^*} + \frac{1}{r^*} \mathbf{f}_{1,-1}^i(r^*) + \frac{\sqrt{2l(l+1)}}{r^*} \mathbf{f}_{1,0}^i(r^*) = 0.
\end{aligned} \tag{26}$$

Легко видеть, что первое уравнение эквивалентно второму, а третье уравнение эквивалентно четвертому. Таким образом, мы приходим к следующей системе неоднородных дифференциальных уравнений первого порядка:

$$\begin{aligned}
& 2r\frac{d\mathbf{f}_{1,1}^l(r)}{dr} - \mathbf{f}_{1,1}^l(r) - \sqrt{2l(l+1)}\mathbf{f}_{1,0}^l(r) = 0, \\
& 2r^*\frac{d\mathbf{f}_{1,1}^i(r^*)}{dr^*} - \mathbf{f}_{1,1}^i(r^*) - \sqrt{2l(l+1)}\mathbf{f}_{1,0}^i(r^*) = 0,
\end{aligned}$$

где функции $\mathbf{f}_{1,0}^l(r)$ и $\mathbf{f}_{1,0}^i(r^*)$ понимаются как неоднородные части. Решения данных уравнений выражаются через элементарные функции:

$$\begin{aligned}
\mathbf{f}_{1,1}^l(r) &= C\sqrt{r} + \sqrt{2l(l+1)}r, \\
\mathbf{f}_{1,1}^i(r^*) &= \dot{C}\sqrt{r^*} + \sqrt{2l(l+1)}r^*.
\end{aligned}$$

Следовательно, решения для радиальной части имеют вид:

$$\begin{aligned}
\mathbf{f}_{1,1}^l(r) &= \mathbf{f}_{1,-1}^l(r) = C\sqrt{r} + \sqrt{2l(l+1)}r, \\
\mathbf{f}_{1,0}^l(r) &= \sqrt{2l(l+1)}r, \\
\mathbf{f}_{1,1}^i(r^*) &= \mathbf{f}_{1,-1}^i(r^*) = \dot{C}\sqrt{r^*} + \sqrt{2l(l+1)}r^*, \\
\mathbf{f}_{1,0}^i(r^*) &= \sqrt{2l(l+1)}r^*.
\end{aligned}$$

Следовательно, решениями $SL(2, \mathbb{C})$ -полевых уравнений (24) являются

$$\begin{aligned}\psi_1(r, \varphi^c, \theta^c) &= \mathbf{f}_{1,1}^l(r) \mathfrak{M}_l^1(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \psi_2(r, \varphi^c, \theta^c) &= \mathbf{f}_{1,0}^l(r) \mathfrak{M}_l^0(0, 0, \theta, \tau, 0, 0), \\ \psi_3(r, \varphi^c, \theta^c) &= \mathbf{f}_{1,-1}^l(r) \mathfrak{M}_l^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \dot{\psi}_1(r^*, \dot{\varphi}^c, \dot{\theta}^c) &= \mathbf{f}_{1,1}^{\dot{l}}(r^*) \mathfrak{M}_{\dot{l}}^1(\varphi, \epsilon, \theta, \tau, 0, 0), \\ \dot{\psi}_2(r^*, \dot{\varphi}^c, \dot{\theta}^c) &= \mathbf{f}_{1,0}^{\dot{l}}(r^*) \mathfrak{M}_{\dot{l}}^0(0, 0, \theta, \tau, 0, 0), \\ \dot{\psi}_3(r^*, \dot{\varphi}^c, \dot{\theta}^c) &= \mathbf{f}_{1,-1}^{\dot{l}}(r^*) \mathfrak{M}_{\dot{l}}^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0),\end{aligned}$$

где

$$\begin{aligned}l &= 1, 2, 3, \dots, \\ \dot{l} &= 1, 2, 3, \dots,\end{aligned}$$

$$\mathfrak{M}_l^{\pm 1}(\varphi, \epsilon, \theta, \tau, 0, 0) = e^{\mp(\epsilon + i\varphi)} Z_l^{\pm 1}(\theta, \tau),$$

$$\begin{aligned}Z_l^{\pm 1}(\theta, \tau) &= \cos^{2l} \frac{\theta}{2} \cosh^{2l} \frac{\tau}{2} \sum_{k=-l}^l i^{\pm 1-k} \tan^{\pm 1-k} \frac{\theta}{2} \tanh^{-k} \frac{\tau}{2} \times \\ &\times {}_2F_1\left(\begin{matrix} \pm 1 - l + 1, 1 - l - k \\ \pm 1 - k + 1 \end{matrix} \middle| i^2 \tan^2 \frac{\theta}{2} \right) {}_2F_1\left(\begin{matrix} -l + 1, 1 - l - k \\ -k + 1 \end{matrix} \middle| \tanh^2 \frac{\tau}{2} \right),\end{aligned}$$

$$\mathfrak{M}_l^0(0, 0, \theta, \tau, 0, 0) = Z_l^0(\theta, \tau),$$

$$\begin{aligned}Z_l^0(\theta, \tau) &= \cos^{2l} \frac{\theta}{2} \cosh^{2l} \frac{\tau}{2} \sum_{k=-l}^l i^{-k} \tan^{-k} \frac{\theta}{2} \tanh^{-k} \frac{\tau}{2} \times \\ &\times {}_2F_1\left(\begin{matrix} -l + 1, 1 - l - k \\ -k + 1 \end{matrix} \middle| i^2 \tan^2 \frac{\theta}{2} \right) {}_2F_1\left(\begin{matrix} -l + 1, 1 - l - k \\ -k + 1 \end{matrix} \middle| \tanh^2 \frac{\tau}{2} \right),\end{aligned}$$

$$\mathfrak{M}_{\dot{l}}^{\pm 1}(\varphi, \epsilon, \theta, \tau, 0, 0) = e^{\mp(\epsilon - i\varphi)} Z_{\dot{l}}^{\pm 1}(\theta, \tau),$$

$$\begin{aligned}Z_{\dot{l}}^{\pm 1}(\theta, \tau) &= \cos^{2\dot{l}} \frac{\theta}{2} \cosh^{2\dot{l}} \frac{\tau}{2} \sum_{\dot{k}=-\dot{l}}^{\dot{l}} i^{\pm 1-\dot{k}} \tan^{\pm 1-\dot{k}} \frac{\theta}{2} \tanh^{-\dot{k}} \frac{\tau}{2} \times \\ &\times {}_2F_1\left(\begin{matrix} \pm 1 - \dot{l} + 1, 1 - \dot{l} - \dot{k} \\ \pm 1 - \dot{k} + 1 \end{matrix} \middle| i^2 \tan^2 \frac{\theta}{2} \right) {}_2F_1\left(\begin{matrix} -\dot{l} + 1, 1 - \dot{l} - \dot{k} \\ -\dot{k} + 1 \end{matrix} \middle| \tanh^2 \frac{\tau}{2} \right),\end{aligned}$$

$$\mathfrak{M}_{\dot{l}}^0(0, 0, \theta, \tau, 0, 0) = Z_{\dot{l}}^0(\theta, \tau),$$

$$Z_l^0(\theta, \tau) = \cos^{2i} \frac{\theta}{2} \cosh^{2i} \frac{\tau}{2} \sum_{k=-i}^i i^{-k} \tan^{-k} \frac{\theta}{2} \tanh^{-k} \frac{\tau}{2} \times \\ \times {}_2F_1 \left(\begin{matrix} -i+1, 1-i-k \\ -k+1 \end{matrix} \middle| i^2 \tan^2 \frac{\theta}{2} \right) {}_2F_1 \left(\begin{matrix} -i+1, 1-i-k \\ -k+1 \end{matrix} \middle| \tanh^2 \frac{\tau}{2} \right),$$

где $\mathfrak{M}_l^0(0, 0, \theta, \tau, 0, 0)$ ($Z_l^0(\theta, \tau)$) — зональные гипертсферические функции (см. [48]).

Таким образом, в согласии с факторизацией (5) явный вид релятивистской волновой функции $\psi(\alpha) = \psi(x)\psi(\mathbf{g})$ на группе Пуанкаре в случае представления $(1, 0) \oplus (0, 1)$ (поле Максвелла) определяется следующими выражениями (полное множество):

$$\psi_1(\alpha) = \psi_+(\mathbf{k}; \mathbf{x}, t)\psi_1(\mathbf{g}) = \\ = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_+(\mathbf{k}) \\ \varepsilon_+(\mathbf{k}) \end{pmatrix} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \mathbf{f}_{1,1}^l(r) \mathfrak{M}_l^1(\varphi, \epsilon, \theta, \tau, 0, 0),$$

$$\psi_0(\alpha) = \psi_0(\mathbf{k}; \mathbf{x})\psi_0(\mathbf{g}) = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_0(\mathbf{k}) \\ \varepsilon_0(\mathbf{k}) \end{pmatrix} \exp[i\mathbf{k} \cdot \mathbf{x}] \mathbf{f}_{1,0}^l(r) \mathfrak{M}_l(0, 0, \theta, \tau, 0, 0),$$

$$\psi_{-1}(\alpha) = \psi_-(\mathbf{k}; \mathbf{x}, t)\psi_{-1}(\mathbf{g}) = \\ = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_-(\mathbf{k}) \\ \varepsilon_-(\mathbf{k}) \end{pmatrix} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \mathbf{f}_{1,-1}^l(r) \mathfrak{M}_l^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0),$$

$$\dot{\psi}_1(\alpha) = \psi_+^*(\mathbf{k}; \mathbf{x}, t)\dot{\psi}_1(\mathbf{g}) = \\ = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_+^*(\mathbf{k}) \\ \varepsilon_+^*(\mathbf{k}) \end{pmatrix} \exp[-i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \mathbf{f}_{1,1}^l(r^*) \mathfrak{M}_l^1(\varphi, \epsilon, \theta, \tau, 0, 0),$$

$$\dot{\psi}_0(\alpha) = \psi_0^*(\mathbf{k}; \mathbf{x})\dot{\psi}_0(\mathbf{g}) = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_0^*(\mathbf{k}) \\ \varepsilon_0^*(\mathbf{k}) \end{pmatrix} \exp[-i\mathbf{k} \cdot \mathbf{x}] \mathbf{f}_{1,0}^l(r^*) \mathfrak{M}_l(0, 0, \theta, \tau, 0, 0),$$

$$\dot{\psi}_{-1}(\alpha) = \psi_-^*(\mathbf{k}; \mathbf{x}, t)\dot{\psi}_{-1}(\mathbf{g}) = \\ = \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_-^*(\mathbf{k}) \\ \varepsilon_-^*(\mathbf{k}) \end{pmatrix} \exp[-i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \mathbf{f}_{1,-1}^l(r^*) \mathfrak{M}_l^{-1}(\varphi, \epsilon, \theta, \tau, 0, 0). \quad (27)$$

Множество (27) состоит из поперечных решений $\psi_{\pm 1}(\alpha)$ (положительная энергия), $\dot{\psi}_{\pm 1}(\alpha)$ (отрицательная энергия) и продольных решений $\psi_0(\alpha)$, $\dot{\psi}_0(\alpha)$. Решения с отрицательной энергией $\dot{\psi}_{\pm 1}(\alpha)$ должны быть опущены, поскольку фотоны не имеют античастиц. Продольные решения $\psi_0(\alpha)$ и $\dot{\psi}_0(\alpha)$ не дают вклада в реальное фотонное поле благодаря условиям поперечности (17) и (19).

Следовательно, любой реальный фотон должен описываться только решением $\psi_{\pm 1}(\alpha)$:

$$\begin{aligned}\psi_{\pm 1}(\alpha) &= \psi_{\pm}(\mathbf{k}; \mathbf{x}, t) \psi_{\pm 1}(\mathbf{g}) = \\ &= \{2(2\pi)^3\}^{-\frac{1}{2}} \begin{pmatrix} \varepsilon_{\pm}(\mathbf{k}) \\ \varepsilon_{\pm}(\mathbf{k}) \end{pmatrix} \exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)] \mathbf{f}_{1, \pm 1}^l(r) \mathfrak{M}_l^{\pm 1}(\varphi, \epsilon, \theta, \tau, 0, 0).\end{aligned}$$

Таким образом, мы нашли множество решений, определяющее поле Максвелла $(1, 0) \oplus (0, 1)$ на группе Пуанкаре (или, эквивалентно, на групповом многообразии $\mathcal{M}_8$). Следует отметить, что ранее полученные решения для поля $(1/2, 0) \oplus (0, 1/2)$ (поле Дирака) [64] имеют аналогичную математическую структуру, т.е. они являются функциями на группе Пуанкаре. Данное обстоятельство позволяет рассматривать поля $(1/2, 0) \oplus (0, 1/2)$ и $(1, 0) \oplus (0, 1)$ на равных основаниях, с единой теоретико-групповой точки зрения.

Приложение. Комплексная 2-сфера и гиперсферические функции

Построим в пространстве $\mathbb{C}^3$ двумерную комплексную сферу из величин

$$z_k = x_k + iy_k, \quad z_k^* = x_k - iy_k$$

следующим образом:

$$\mathbf{z}^2 = z_1^2 + z_2^2 + z_3^2 = \mathbf{x}^2 - \mathbf{y}^2 + 2i\mathbf{x}\mathbf{y} = r^2, \quad (\text{A.1})$$

и ее комплексно-сопряженную (дуальную) сферу

$$\mathbf{z}^{*2} = z_1^{*2} + z_2^{*2} + z_3^{*2} = \mathbf{x}^2 - \mathbf{y}^2 - 2i\mathbf{x}\mathbf{y} = r^{*2}. \quad (\text{A.2})$$

Как известно, величины $\mathbf{x}^2 - \mathbf{y}^2$, $\mathbf{x}\mathbf{y}$ инвариантны относительно преобразований Лоренца, поскольку поверхность комплексной сферы инвариантна (операторы Казимира группы Лоренца конструируются из таких величин, см. также (A.4)). Более того, поскольку вещественная и мнимая части комплексной 2-сферы преобразуются подобно электрическому и магнитному полям, соответственно, то инвариантность величины $\mathbf{z}^2 \sim (\mathbf{E} + i\mathbf{B})^2$ под действием преобразований Лоренца очевидна.

Группа $SL(2, \mathbb{C})$ всех комплексных матриц

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

второго порядка с определителем $\alpha\delta - \gamma\beta = 1$ является *комплексификацией* группы $SU(2)$. Группа $SU(2)$ является одной из вещественных форм группы $SL(2, \mathbb{C})$. Переход от $SU(2)$ к $SL(2, \mathbb{C})$ реализуется посредством

комплексификации трех вещественных параметров φ, θ, ψ (углы Эйлера). Пусть $\theta^c = \theta - i\tau$, $\varphi^c = \varphi - i\epsilon$, $\psi^c = \psi - i\varepsilon$ – комплексные углы Эйлера, где

$$\begin{aligned} 0 &\leq \operatorname{Re} \theta^c = \theta \leq \pi, & -\infty < \operatorname{Im} \theta^c = \tau < +\infty, \\ 0 &\leq \operatorname{Re} \varphi^c = \varphi < 2\pi, & -\infty < \operatorname{Im} \varphi^c = \epsilon < +\infty, \\ -2\pi &\leq \operatorname{Re} \psi^c = \psi < 2\pi, & -\infty < \operatorname{Im} \psi^c = \varepsilon < +\infty. \end{aligned} \quad (\text{A.3})$$

Как известно, группа Лоренца имеет два независимых оператора Казимира

$$\begin{aligned} X^2 &= X_1^2 + X_2^2 + X_3^2 = \frac{1}{4}(A^2 - B^2 + 2iAB), \\ Y^2 &= Y_1^2 + Y_2^2 + Y_3^2 = \frac{1}{4}(\tilde{A}^2 - \tilde{B}^2 - 2i\tilde{A}\tilde{B}). \end{aligned} \quad (\text{A.4})$$

Используя параметры (A.3), получим для операторов Казимира следующие выражения:

$$\begin{aligned} X^2 &= \frac{\partial^2}{\partial \theta^{c2}} + \cot \theta^c \frac{\partial}{\partial \theta^c} + \frac{1}{\sin^2 \theta^c} \left[\frac{\partial^2}{\partial \varphi^{c2}} - 2 \cos \theta^c \frac{\partial}{\partial \varphi^c} \frac{\partial}{\partial \psi^c} + \frac{\partial^2}{\partial \psi^{c2}} \right], \\ Y^2 &= \frac{\partial^2}{\partial \dot{\theta}^{c2}} + \cot \dot{\theta}^c \frac{\partial}{\partial \dot{\theta}^c} + \frac{1}{\sin^2 \dot{\theta}^c} \left[\frac{\partial^2}{\partial \dot{\varphi}^{c2}} - 2 \cos \dot{\theta}^c \frac{\partial}{\partial \dot{\varphi}^c} \frac{\partial}{\partial \dot{\psi}^c} + \frac{\partial^2}{\partial \dot{\psi}^{c2}} \right]. \end{aligned} \quad (\text{A.5})$$

Матричные элементы унитарных неприводимых представлений группы Лоренца являются собственными функциями операторов (A.5):

$$\begin{aligned} [X^2 + l(l+1)] \mathfrak{M}_{mn}^l(\varphi^c, \theta^c, \psi^c) &= 0, \\ [Y^2 + i(i+1)] \mathfrak{M}_{\dot{m}\dot{n}}^i(\dot{\varphi}^c, \dot{\theta}^c, \dot{\psi}^c) &= 0, \end{aligned} \quad (\text{A.6})$$

где

$$\begin{aligned} \mathfrak{M}_{mn}^l(\varphi^c, \theta^c, \psi^c) &= e^{-i(m\varphi^c + n\psi^c)} Z_{mn}^l(\theta^c), \\ \mathfrak{M}_{\dot{m}\dot{n}}^i(\dot{\varphi}^c, \dot{\theta}^c, \dot{\psi}^c) &= e^{-i(\dot{m}\dot{\varphi}^c + \dot{n}\dot{\psi}^c)} Z_{\dot{m}\dot{n}}^i(\dot{\theta}^c). \end{aligned} \quad (\text{A.7})$$

Подставляя функции (A.7) в (A.6) и учитывая операторы (A.5), получим комплексный аналог уравнений Лежандра:

$$\left[(1 - z^2) \frac{d^2}{dz^2} - 2z \frac{d}{dz} - \frac{m^2 + n^2 - 2mnz}{1 - z^2} + l(l+1) \right] Z_{mn}^l = 0, \quad (\text{A.8})$$

$$\left[(1 - \dot{z}^{*2}) \frac{d^2}{d\dot{z}^{*2}} - 2\dot{z}^* \frac{d}{d\dot{z}^*} - \frac{\dot{m}^2 + \dot{n}^2 - 2\dot{m}\dot{n}\dot{z}^*}{1 - \dot{z}^{*2}} + i(i+1) \right] Z_{\dot{m}\dot{n}}^i = 0, \quad (\text{A.9})$$

где $z = \cos \theta^c$ и $\dot{z}^* = \cos \dot{\theta}^c$. Последние уравнения имеют три особые точки -1 ,

$+1, \infty$. Решения уравнения (A.8) имеют вид

$$\begin{aligned}
 Z_{mn}^l = & \sum_{k=-l}^l i^{m-k} \sqrt{\Gamma(l-m+1)\Gamma(l+m+1)\Gamma(l-k+1)\Gamma(l+k+1)} \times \\
 & \times \cos^{2l} \frac{\theta}{2} \tan^{m-k} \frac{\theta}{2} \times \\
 & \times \sum_{j=\max(0, k-m)}^{\min(l-m, l+k)} \frac{i^{2j} \tan^{2j} \frac{\theta}{2}}{\Gamma(j+1)\Gamma(l-m-j+1)\Gamma(l+k-j+1)\Gamma(m-k+j+1)} \times \\
 & \times \sqrt{\Gamma(l-n+1)\Gamma(l+n+1)\Gamma(l-k+1)\Gamma(l+k+1)} \cosh^{2l} \frac{\tau}{2} \tanh^{n-k} \frac{\tau}{2} \times \\
 & \times \sum_{s=\max(0, k-n)}^{\min(l-n, l+k)} \frac{\tanh^{2s} \frac{\tau}{2}}{\Gamma(s+1)\Gamma(l-n-s+1)\Gamma(l+k-s+1)\Gamma(n-k+s+1)}. \quad (\text{A.10})
 \end{aligned}$$

Будем называть функции Z_{mn}^l в (A.10) *гиперсферическими функциями*⁹. Функции Z_{mn}^l могут быть записаны через гипергеометрические ряды следующим образом:

$$\begin{aligned}
 Z_{mn}^l = & \cos^{2l} \frac{\theta}{2} \cosh^{2l} \frac{\tau}{2} \sum_{k=-l}^l i^{m-k} \tan^{m-k} \frac{\theta}{2} \tanh^{n-k} \frac{\tau}{2} \times \\
 & {}_2F_1 \left(\begin{matrix} m-l+1, 1-l-k \\ m-k+1 \end{matrix} \middle| i^2 \tan^2 \frac{\theta}{2} \right) {}_2F_1 \left(\begin{matrix} n-l+1, 1-l-k \\ n-k+1 \end{matrix} \middle| \tanh^2 \frac{\tau}{2} \right). \quad (\text{A.11})
 \end{aligned}$$

Следовательно, матричные элементы выражаются посредством функции (*обобщенная гиперсферическая функция*)

$$\mathfrak{M}_{mn}^l(\mathfrak{g}) = e^{-m(\epsilon+i\varphi)} Z_{mn}^l(\cos \theta^c) e^{-n(\epsilon+i\psi)}, \quad (\text{A.12})$$

где

$$Z_{mn}^l(\cos \theta^c) = \sum_{k=-l}^l P_{mk}^l(\cos \theta) \mathfrak{P}_{kn}^l(\cosh \tau), \quad (\text{A.13})$$

здесь $P_{mn}^l(\cos \theta)$ – обобщенная сферическая функция на группе $SU(2)$ (см. [66]), а функция $\mathfrak{P}_{mn}^l$ является аналогом обобщенной сферической функции для группы $QU(2)$ (так называемая функция Якоби [58]). $QU(2)$ является группой квазиунитарных унитарных матриц второго порядка. Как и группа $SU(2)$, группа $QU(2)$ есть одна из вещественных форм группы $SL(2, \mathbb{C})$

⁹Гиперсферические функции (или гиперсферические гармоники) известны в математике давно (см., например, [65]). Эти функции являются обобщением обычных сферических функций на случай n -мерных евклидовых пространств. По этой причине мы оставим это имя (гиперсферические функции) и для случая псевдоевклидовых пространств.

($QU(2)$ некомпактна). Группа $QU(2)$ изоморфна группе $SU(1, 1) \sim SL(2, \mathbb{R})$ (трехмерная группа Лоренца). Присоединенные гипersферические функции получаются из (A.12) при $n = 0$:

$$\mathfrak{M}_l^m(\mathfrak{g}) = e^{m(\epsilon + i\varphi)} Z_l^m(\cos \theta^c).$$

ЛИТЕРАТУРА

1. B. L. van der Waerden, *Nachr. d. Ges. d. Wiss. Göttingen*, 100 (1929).
2. W. L. Bade, H. Jehle, *Rev. Mod. Phys.* **25**, 714 (1953).
3. O. Laporte, G. E. Uhlenbeck, *Phys. Rev.* **37**, 1380 (1931).
4. A. A. Campolattaro, *Int. J. Theor. Phys.* **29**, 141, 477 (1990).
5. J. Vaz, Jr., W. A. Rodrigues, Jr., *Int. J. Theor. Phys.* **32**, 945 (1993).
6. A. Gsponer, *Int. J. Theor. Phys.* **41**, 689 (2002).
7. Ю. Б. Румер, *Спинорный анализ* (ОНТИ, М.-Л., 1936).
8. E. Majorana, *Scientific Papers*, unpublished, deposited at the “Domus Galileana”, Pisa, quaderno **2**, p.101/1; **3**, p.11, 160; **15**, p.16; **17**, p.83, 159.
9. J. R. Oppenheimer, *Phys. Rev.* **38**, 725 (1931).
10. S. N. Gupta, *Proc. Phys. Soc.* **A63**, 681 (1950).
11. K. Bleuler, *Helv. Phys. Acta* **23**, 567 (1950).
12. W. J. Archibald, *Can. J. Phys.* **33**, 565 (1955).
13. S. A. Bludman, *Phys. Rev.* **107**, 1163 (1957).
14. T. Ohmura, *Prog. Theor. Phys.* **16**, 684 (1956).
15. R. H. Good, *Phys. Rev.* **105**, 1914 (1957).
16. J. S. Lomont, *Phys. Rev.* **111**, 1710 (1958).
17. A. A. Borhgardt, *Sov. Phys. JETP* **34**, 334 (1958).
18. H. E. Moses, *Phys. Rev.* **113**, 1670 (1959).
19. M. Sachs, S. L. Schwebel, *J. Math. Phys.* **3**, 843 (1962).
20. R. Mignani, E. Recami, M. Baldo, *Lettere al Nuovo Cimento* **11**, 568 (1974).
21. H. Bacry, *Nuov. Cim.* **A32**, 448 (1976).
22. A. Da Silveira, *Z. Naturforsch* **A34**, 646 (1979).
23. E. Giannetto, *Lettere al Nuovo Cimento* **44**, 140 (1985).
24. K. Ljolje, *Fortschr. Phys.* **36**, 9 (1988).
25. H. Sallhofer, *Z. Naturforsch* **A45**, 1361 (1990).
26. V. M. Simulik, *Theor. Math. Phys.* **87**, 386 (1991).
27. T. Inagaki, *Phys. Rev* **A49**, 2839 (1994).
28. I. Białynicki-Birula, *Acta Phys. Pol.* **A86**, 97 (1994).
29. J. F. Sipe, *Phys. Rev.* **A52**, 1875 (1995).
30. S. Bruce, *Nuov. Cim.* **B110**, 115 (1995).
31. V. V. Dvoeglazov, *Nuov. Cim.* **112**, 847 (1997).
32. A. Gersten, *Found. Phys. Lett.* **12**, 291 (1998).
33. S. Esposito, *Found. Phys.* **28**, 231 (1998).
34. D. Finkelstein, *Phys. Rev.* **100**, 924 (1955).

35. H. Bacry, A. Kihlberg, *J. Math. Phys.* **10**, 2132 (1969).
36. F. Lurçat, *Physics* **1**, 95 (1964).
37. A. Kihlberg, *Ann. Inst. Henri Poincaré* **13**, 57 (1970).
38. C. P. Boyer, G. N. Fleming, *J. Math. Phys.* **15**, 1007 (1974).
39. H. Arodź, *Acta Phys. Pol.* **B7**, 177 (1976).
40. M. Toller, *J. Math. Phys.* **37**, 2694 (1996).
41. S. M. Kuzenko, S. L. Lyakhovich, A. Yu. Segal, *Int. J. Mod. Phys.* **A10**, 1529 (1995).
42. S. L. Lyakhovich, A. Yu. Segal, A. A. Sharapov, *Phys. Rev.* **D54**, 5223 (1996).
43. W. Drechsler, *J. Math. Phys.* **38**, 5531 (1997).
44. A. A. Deriglazov, D. M. Gitman, *Mod. Phys. Lett.* **A14**, 709 (1999).
45. D. M. Gitman, A. L. Shelepin, *Int. J. Theor. Phys.* **40**, 603 (2001).
46. J. -Y. Grandpeix, F. Lurçat, *Found. Phys.* **32**, 109 (2002); *ibid.* **32**, 133 (2002).
47. L. C. Biedenharn, H. W. Braden, P. Truini, H. van Dam, *J. Phys. A: Math. Gen.* **21**, 3593 (1988).
48. V. V. Varlamov, *Hyperspherical Functions and Harmonic Analysis on the Lorentz Group*, Mathematical Physics Research at the Cutting Edge (Ed. C. V. Benton) pp.193-250 (Nova Science Publishers, New York, 2004).
49. Ю. Б. Румер, А. И. Фет, *Теория групп и квантованные поля* (М.:Наука, 1977).
50. М. А. Наймарк, *Линейные представления группы Лоренца* (М.:Физматлит, 1958).
51. M. A. Vasiliev, *Int. J. Mod. Phys.* **D5**, 763 (1996).
52. В. И. Арнольд, *Математические методы классической механики* (М.:Наука, 1989).
53. V. V. Varlamov, *Hadronic J.* **25**, 481 (2002).
54. D.V. Ahluwalia, D.J. Ernst, *Int. J. Mod. Phys.* **E2**, 397 (1993).
55. V.V. Dvoeglazov, *Nuov. Cim.* **B111**, 483 (1996).
56. L. Ryder, *Quantum Field Theory* (Cambridge University Press, Cambridge, 1985).
57. T. D. Newton, E. P. Wigner, *Rev. Mod. Phys.* **21**, 400 (1949).
58. Н. Я. Виленкин, *Специальные функции и теория представлений групп* (М.:Наука, 1965).
59. А. З. Петров, *Пространства Эйнштейна* (М.:Физматгиз, 1961).
60. M. Huszar, J. Smorodinsky, *Preprint JINR No. E2-5020*, Dubna (1970).
61. Я. А. Смородинский, М. Хусар, *ТМФ* **4**, 328 (1970).
62. V. V. Varlamov, *Int. J. Theor. Phys.* **42**, 583 (2003).
63. Р. Пенроуз, В. Риндлер, *Спиноры и пространство-время* (М.:Мир, 1987).
64. V. V. Varlamov, *J. Phys. A: Math. Gen.*, **37**, 5467 (2004).
65. Н. Bateman, A. Erdélyi, *Higher Transcendental Functions, vol. II* (Mc Grow-Hill Book Company, New York, 1953).
66. И. М. Гельфанд, Р. А. Минлос, З. Я. Шапиро, *Представления группы вращений и группы Лоренца, их применения* (М.:Физматлит, 1958).

ВНЕШНЯЯ КРИВИЗНА 3-МЕРНОГО ПРОСТРАНСТВА И ЭНЕРГЕТИЧЕСКИЕ ЗАТРАТЫ, НЕОБХОДИМЫЕ ДЛЯ ОБРАЗОВАНИЯ 4-МЕРНОЙ КРОТОВОЙ НОРЫ

Е.В. Палешева

В работе исследуется возможность уменьшения затрат энергии, необходимых для образования 4-мерной кротовой норы. Соответствующее изменение скачка плотности энергии может быть достигнуто только если внешняя кривизна 3-мерного пространства не удовлетворяет условию непрерывности.

В последние годы наблюдается увеличение числа работ, связанных с существованием времениподобных замкнутых гладких кривых. К настоящему моменту можно уже говорить о нескольких типах 4-мерных пространств, допускающих машину времени. К таким пространствам относятся, например, пространства гёделевского типа — решения описывающие вращающуюся вселенную. Если в таком пространстве-времени существуют временные петли, то это изначальное свойство 4-мерной геометрии. К другому типу 4-мерных пространств, которые, как считают некоторые авторы, допускают еще один вариант машины времени, относятся 4-мерные пространства с 3-мерной кротовой норой. В такой модели, по предположению Кипа Торна, временная петля появляется за счет механического движения одного из концов кротовой норы, причем времениподобные замкнутые гладкие кривые изначально могут и не существовать в данном пространстве-времени. Впрочем, если следовать результатам, представленным в [3–6], в работе Торна имеются некоторые неточности и поэтому предложенная им модель в действительности не может описывать машину времени. Но в некоторых 4-мерных пространствах с 3-мерной кротовой норой временные петли являются свойством 4-мерной геометрии и вот в таких пространствах машина времени уже будет реальностью, а не ошибочным домыслом, как и в метриках гёделевского типа. Следует также отметить что исследования, касающиеся решений с машиной времени, проводятся не только в рамках классической теории относительности, но и в теории суперструн, а также в квантовой

теории поля. Кроме всего вышесказанного следует упомянуть как о машине времени в пространствах с 4-мерной кротовой норой, так и о машине времени в 5-мерном пространстве-времени, допускающем пружинные слои. Основные результаты, касающиеся этих моделей, можно найти в [1–3]. В данных работах, например, рассматривался процесс рождения 4-мерной кротовой норы. Было замечено, что в этом случае 3-мерное пространство постепенно «разрывается» на два куска. Другими словами, рождение 4-мерной кротовой норы можно рассматривать как процесс, в ходе которого связное 3-мерное пространство становится несвязным. Последнее, в свою очередь, осуществляется как стягивания границы отрываемой области в точку, после чего 3-мерное пространство «разрывается» на две связные области. Такой подход позволил автору работ [1–3] рассчитать средний скачок плотности энергии в отрываемой области. Данная величина оказалась обратно пропорциональна площади характерного сечения отрываемой области. Впрочем при этом предполагалась непрерывность внешней кривизны 3-мерного пространства. В данной работе мы попробуем отказаться от последнего условия, возможно это позволит нам уменьшить скачок плотности энергии. Учитывая, что нам придется использовать такие понятия как внешняя кривизна, мы, следуя работе [7], приведем вывод уравнений Гаусса-Кодацци и их следствий¹.

1. Уравнения Гаусса-Кодацци

Рассмотрим пространственно-подобную гиперповерхность, метрический тензор которой будем определять выражением

$$\gamma_{\alpha\beta} = -g_{\alpha\beta} + \frac{g_{0\alpha}g_{0\beta}}{g_{00}}. \quad (1)$$

В качестве сигнатуры 4-мерного пространства-времени выберем $\langle +, -, -, - \rangle$. Заметим также, что если справедливо соотношение (1), то

$$\gamma^{\alpha\beta} = -g^{\alpha\beta}. \quad (2)$$

На выбранной гиперповерхности зададим систему координат $\langle e^1, e^2, e^3 \rangle$, такую что e^α также являются базисными векторами 4-мерного пространства-времени. Для удобства вычислений будем предполагать, что вектор нормали к гиперповерхности соответствует вектору e^0 . Такая система координат является синхронной, т.е.

$$g_{00} = 1, \quad g_{0\alpha} = 0,$$

и соответствует гауссовой нормальной системе координат на выбранной пространственно-подобной гиперповерхности.

¹ Дело в том, что в работе Мизнера, Торна и Уиллера используется сигнатура $\langle -, +, +, + \rangle$, отличная от используемой в этой статье $\langle +, -, -, - \rangle$, и кроме того расщепление пространства-времени также отлично от применяемого нами. Поэтому нам приходится заново повторить вывод уравнений Гаусса-Кодацци.

В произвольной системе отсчета ковариантная производная вектора нормали n связана с базисными векторами на гиперповерхности следующим образом:

$$\nabla_\alpha n = -K_{\alpha\beta} e^\beta, \quad (3)$$

$$\nabla_\alpha n = -K_\alpha^\beta e_\beta, \quad (4)$$

где величина $K_{\alpha\beta}$ называется тензором внешней кривизны данной гиперповерхности. Тогда, в силу соотношений (1) и (2),

$$K_{\alpha\beta} = -e_\beta \cdot \nabla_\alpha n, \quad (5)$$

$$K_\alpha^\beta = g^{\beta\mu} K_{\alpha\mu}, \quad (6)$$

где $g^{\beta\mu}$ — компоненты метрического тензора пространства-времени.

Продолжая следовать результатам работы [7], а также учитывая синхронность системы отсчета, мы можем записать уравнения Гаусса-Вайнгартена:

$${}^{(4)}\nabla_\alpha e_\beta = K_{\alpha\beta} \cdot n + {}^{(3)}\Gamma_{\alpha\beta}^\mu e_\mu. \quad (7)$$

Здесь ${}^{(3)}\Gamma_{\alpha\beta}^\mu$ — символы Кристоффеля на гиперповерхности. В дальнейшем вместо ${}^{(4)}\nabla$ будем писать ∇ .

Для вычисления компонент 4-мерного тензора Римана воспользуемся оператором кривизны (см. [8])

$$R(\mathbf{a}, \mathbf{b}) = [\nabla_{\mathbf{a}}, \nabla_{\mathbf{b}}] - \nabla_{[\mathbf{a}, \mathbf{b}]}$$

Как известно,

$$R_{\alpha\beta\gamma\delta} = e_\alpha \cdot R(e_\gamma, e_\delta) e_\beta$$

и

$$R_{\beta\gamma\delta}^\alpha = e^\alpha \cdot R(e_\gamma, e_\delta) e_\beta.$$

Итак, вычислим $R_{\gamma\alpha\beta}^\mu$. Так как для любых α и β величина²

$$[e_\alpha, e_\beta] = {}^{(4)}\nabla_\alpha e_\beta - {}^{(4)}\nabla_\beta e_\alpha = {}^{(4)}\Gamma_{\alpha\beta}^k e_k - {}^{(4)}\Gamma_{\beta\alpha}^k e_k$$

равна нулю, то

$$R(e_\alpha, e_\beta) e_\gamma = \nabla_\alpha \nabla_\beta e_\gamma - \nabla_\beta \nabla_\alpha e_\gamma$$

и в силу уравнений Гаусса-Вайнгартена мы имеем:

$$R(e_\alpha, e_\beta) e_\gamma = \nabla_\alpha (K_{\beta\gamma} n + {}^{(3)}\Gamma_{\beta\gamma}^\mu e_\mu) - \nabla_\beta (K_{\alpha\gamma} n + {}^{(3)}\Gamma_{\alpha\gamma}^\mu e_\mu). \quad (8)$$

Применяя ковариантное дифференцирование в уравнении (8) и учитывая соотношения (3) мы получим

$$R(e_\alpha, e_\beta) e_\gamma = (K_{\beta\gamma|\alpha} - K_{\alpha\gamma|\beta}) n + (K_{\alpha\gamma} K_\beta^\mu - K_{\beta\gamma} K_\alpha^\mu) e_\mu + {}^{(3)}R_{\gamma\alpha\beta}^\mu e_\mu, \quad (9)$$

²Здесь e_k вектора 4-мерного базиса.

где

$$K_{\beta\gamma|\alpha} = \frac{\partial K_{\beta\gamma}}{\partial x^\alpha} - {}^{(3)}\Gamma_{\alpha\gamma}^\mu K_{\beta\mu} - {}^{(3)}\Gamma_{\alpha\beta}^\mu K_{\gamma\mu}.$$

Далее, применяя выражение (9) и соотношения

$${}^{(4)}R_{\gamma\alpha\beta}^0 = e^0 R(e_\alpha e_\beta) e_\gamma,$$

$${}^{(4)}R_{\gamma\alpha\beta}^\mu = e^\mu R(e_\alpha e_\beta) e_\gamma,$$

находим уравнения Гаусса-Кодацци:

$$\begin{aligned} {}^{(4)}R_{\gamma\alpha\beta}^0 &= K_{\beta\gamma|\alpha} - K_{\alpha\gamma|\beta}, \\ {}^{(4)}R_{\gamma\alpha\beta}^\mu &= K_{\alpha\gamma} K_\beta^\mu - K_{\beta\gamma} K_\alpha^\mu + {}^{(3)}R_{\gamma\alpha\beta}^\mu. \end{aligned} \tag{10}$$

2. Скалярная кривизна

Для того, чтобы получить соотношение, определяющее скалярную кривизну пространства-времени через внешнюю и скалярную кривизны 3-мерного пространства, одних только уравнений Гаусса-Кодацци не достаточно. Необходимо еще найти выражение, определяющее величину ${}^{(4)}R_{\gamma 0\beta}^0$. Для этого воспользуемся равенством

$${}^{(4)}R_{\gamma 0\beta}^0 = {}^{(4)}R_{0\gamma 0\beta} = {}^{(4)}R_{\beta 0\gamma 0},$$

выполненным в нашем случае. Так как

$${}^{(4)}R_{\beta 0\gamma 0} = e_\beta \cdot R(e_\gamma, n)n,$$

то

$${}^{(4)}R_{\gamma 0\beta}^0 = e_\beta \cdot R(e_\gamma, n)n.$$

Найдем действие оператора $R(e_\gamma, n)$ на вектор нормали. По определению оператора кривизны получаем:

$$R(e_\gamma, n)n = \nabla_\gamma \nabla_n n - \nabla_n \nabla_\gamma n - \nabla_{[e_\gamma, n]}. \tag{11}$$

В силу того, что $[e_\gamma, n] = 0$, а также $\nabla_n n = 0$, уравнение (11) равносильно следующему:

$$R(e_\gamma, n)n = -\nabla_n \nabla_\gamma n. \tag{12}$$

Учитывая связь (4) между производной нормали и тензором внешней кривизны, преобразуем выражение (12) к следующему виду:

$$R(e_\gamma, n)n = \nabla_n (K_\gamma^\mu e_\mu). \tag{13}$$

После ковариантного дифференцирования уравнение (13) преобразуется в

$$R(e_\gamma, n)n = \frac{\partial K_\gamma^\mu}{\partial x^0} e_\mu + K_\gamma^\mu \nabla_n e_\mu. \tag{14}$$

Далее, так как $[n, e_\gamma] = 0$, то справедливо соотношение

$$\nabla_n e_\gamma - \nabla_\gamma n = 0,$$

учитывая которое перепишем (14):

$$R(e_\gamma, n)n = \frac{\partial K_\gamma^\mu}{\partial x^0} e_\mu + K_\gamma^\mu \nabla_\mu n. \quad (15)$$

Применяя вновь в уравнении (15) равенство (4), окончательно получим, что

$$R(e_\gamma, n)n = \frac{\partial K_\gamma^\mu}{\partial x^0} e_\mu - K_\gamma^\mu K_\mu^\nu e_\nu. \quad (16)$$

В результате, домножив скалярно соотношение (16) на вектор e_β , будем иметь:

$${}^{(4)}R_{\gamma 0\beta}^0 = g_{\beta\mu} \frac{\partial K_\gamma^\mu}{\partial x^0} - K_\gamma^\mu K_{\beta\mu}. \quad (17)$$

Так как

$$K_\gamma^\mu = g^{\mu\alpha} K_{\gamma\alpha},$$

то соотношение (17) равносильно следующему:

$${}^{(4)}R_{\gamma 0\beta}^0 = g_{\beta\mu} \frac{\partial g^{\mu\alpha}}{\partial x^0} K_{\gamma\alpha} + \frac{\partial K_{\gamma\beta}}{\partial x^0} - K_\gamma^\mu K_{\beta\mu}. \quad (18)$$

Учитывая равенство (см. [8])

$$\frac{\partial g^{\mu\alpha}}{\partial x^0} = 2K^{\mu\alpha}$$

и меняя порядок суммирования приведем выражение (18) к виду

$${}^{(4)}R_{\gamma 0\beta}^0 = K_{\beta\mu} K_\gamma^\mu + \frac{\partial K_{\gamma\beta}}{\partial x^0}. \quad (19)$$

Далее, используя уравнения Гаусса-Кодацци и соотношение (19) мы можем вычислить величину ${}^{(4)}R_{\gamma\beta}$. В силу того, что

$${}^{(4)}R_{\gamma\beta} = {}^{(4)}R_{\gamma k\beta}^k = {}^{(4)}R_{\gamma 0\beta}^0 + {}^{(4)}R_{\gamma\mu\beta}^\mu,$$

получаем:

$${}^{(4)}R_{\gamma\beta} = K_{\beta\mu} K_\gamma^\mu + \frac{\partial K_{\gamma\beta}}{\partial x^0} + K_{\mu\gamma} K_\beta^\mu - K_{\beta\gamma} K_\mu^\mu + {}^{(3)}R_{\gamma\mu\beta}^\mu. \quad (20)$$

Так как

$$K_{\beta\mu} K_\gamma^\mu = K_{\gamma\mu} K_\beta^\mu$$

и $K_{\mu\gamma} = K_{\gamma\mu}$, то равенство (20) эквивалентно выражению

$${}^{(4)}R_{\gamma\beta} = 2K_{\gamma\mu} K_\beta^\mu + \frac{\partial K_{\gamma\beta}}{\partial x^0} - K_{\beta\gamma} K_\mu^\mu + {}^{(3)}R_{\gamma\beta}. \quad (21)$$

Для получения соотношения, определяющего скалярную кривизну пространства-времени через скалярную кривизну и внешнюю кривизну 3-мерного пространства, осталось вычислить компоненту ${}^{(4)}R_{00}$. Поскольку

$${}^{(4)}R_{00} = {}^{(4)}R_{000}^0 + {}^{(4)}R_{0\mu 0}^\mu,$$

найдем сначала ${}^{(4)}R_{000}^0$. Но так как

$${}^{(4)}R_{000}^0 = e^0 \cdot R(n, n)n,$$

а

$$R(n, n)n = \nabla_n \nabla_n n - \nabla_n \nabla_n n - \nabla_{[n, n]} = 0,$$

то ${}^{(4)}R_{000}^0 = 0$.

Для получения уравнения, определяющего компоненту ${}^{(4)}R_{0\mu 0}^\mu$, воспользуемся равенством

$${}^{(4)}R_{0\alpha 0}^\mu = e^\mu \cdot R(e_\alpha, n)n.$$

Применяя формулу (16) будем иметь:

$${}^{(4)}R_{0\alpha 0}^\mu = \frac{\partial K_\alpha^\mu}{\partial x^0} - K_\alpha^\nu K_\nu^\mu = \frac{\partial K_\alpha^\mu}{\partial x^0} - K_{\alpha\nu} K^{\nu\mu}.$$

Таким образом, получаем:

$${}^{(4)}R_{0\mu 0}^\mu = \frac{\partial K_\mu^\mu}{\partial x^0} - K_{\mu\nu} K^{\mu\nu}.$$

В результате, находим следующее соотношение:

$${}^{(4)}R_{00} = \frac{\partial K_\mu^\mu}{\partial x^0} - K_{\mu\nu} K^{\mu\nu}. \quad (22)$$

Учитывая (22) и (21) получим выражение для скалярной кривизны пространства-времени, определяемой в нашем случае соотношением

$${}^{(4)}R = {}^{(4)}R_{00} + g^{\gamma\beta} {}^{(4)}R_{\gamma\beta}. \quad (23)$$

Вводя обозначение

$$K = K_\mu^\mu$$

и используя равенства

$$g^{\gamma\beta} K_{\gamma\mu} K_\beta^\mu = K_{\gamma\mu} K^{\gamma\mu}, \quad g^{\gamma\beta} K_{\beta\gamma} K = K^2,$$

после подстановки в уравнение (23) компонент ${}^{(4)}R_{00}$ и ${}^{(4)}R_{\gamma\beta}$ получим выражение

$${}^{(4)}R = \frac{\partial K}{\partial x^0} + K_{\mu\nu} K^{\mu\nu} - K^2 + g^{\gamma\beta} \frac{\partial K_{\gamma\beta}}{\partial x^0} - {}^{(3)}R. \quad (24)$$

$$t = \frac{1}{2}$$

$$\gamma_{\alpha\beta}(t), \quad 0 \leq t \leq 1$$

Рис. 1. Процесс рождения 4-мерной кротовой норы как разрыв связности 3-мерного пространства.

С помощью соотношений $K = g^{\gamma\beta} K_{\beta\gamma}$ и

$$\frac{\partial g^{\gamma\beta}}{\partial x^0} = 2K^{\gamma\beta}$$

нетрудно проверить справедливость равенства

$$g^{\gamma\beta} K_{\beta\gamma} = \frac{\partial K}{\partial x^0} - 2K^{\gamma\beta} K_{\gamma\beta},$$

подставляя которое в уравнение (24) получим искомое выражение для скалярной кривизны:

$${}^{(4)}R = 2\frac{\partial K}{\partial x^0} - K_{\mu\nu}K^{\mu\nu} - K^2 - {}^{(3)}R. \quad (25)$$

3. Скачок плотности энергии

Как известно, плотность энергии $\varepsilon = T_0^0$ определяется уравнением

$$\kappa T_0^0 = {}^{(4)}R_0^0 - \frac{1}{2}R. \quad (26)$$

Так как в выбранной нами системе отсчета справедливо равенство $R_0^0 = R_{00}$, то подставляя в уравнение (26) соотношения (22) и (25), мы получим:

$$\kappa\varepsilon = {}^{(3)}R + K^2 - K_{\mu\nu}K^{\mu\nu}. \quad (27)$$

Нетрудно видеть, что хотя в выражение для скалярной кривизны входит производная

$$\frac{\partial K}{\partial x^0},$$

а следовательно и вторые производные по времени от $g_{\alpha\beta}$, соответствующие компоненты отсутствуют в уравнении, определяющем плотность энергии.

Следуя работе [3] определим скачок плотности энергии в момент образования 4-мерной кротовой норы. Данный процесс можно рассматривать как отрыв от 3-мерного пространства M_0 области D_0 (см. рис. 1). Последнее определяется как стягивание границы области D_0 в точку. Для соответствующего изменения связности 3-мерного пространства в [3] рассматривается семейство римановых метрик $\gamma_{\alpha\beta}(t)$, $t \in [0, 1]$, заданных на пространстве M_0 . При этом параметр t не обязательно совпадает с координатой времени. Предполагалось, что данное семейство метрик удовлетворяет следующим условиям:

1. $\gamma_{\alpha\beta}(t)$ при $0 \leq t < 1/2$ поле класса C^2 , при $t \geq 1/2$ на границе ∂D_0 имеет разрывы производных первого рода;
2. $\partial\gamma_{\alpha\beta}/\partial n$ непрерывны;
3. $\gamma_{\alpha\beta}(t) = \gamma_{\alpha\beta}(0)$ вне ε -окрестности области D_0 ;
4. пространство $\langle M_0, \gamma_{\alpha\beta}(t) \rangle$ при $t > 1/2$ имеет неотрицательную кривизну;
5. пространство $\langle M_0, \gamma_{\alpha\beta}(t) \rangle$ при $t \in [0, 1]$ допускает регулярное единичное киллингово поле ξ_t .

Нетрудно заметить, что непрерывность производных $\partial\gamma_{\alpha\beta}/\partial n$ в выбранной системе отсчета соответствует непрерывность первых производных по времени от $g_{\alpha\beta}$, а значит и непрерывности тензора внешней кривизны $K_{\alpha\beta}$. Поэтому в процессе вычисления скачка плотности энергии при переходе через точку $t = 1/2$ с данным условием все слагаемые, содержащие компоненты тензора внешней кривизны аннулируются. В результате, в [3] показано, что при данных условиях

$$\langle \delta\varepsilon \rangle \sim \frac{1}{\sigma},$$

в силу полученного соотношения

$$\langle \delta^{(3)}R \rangle \sim \frac{1}{\sigma}. \quad (28)$$

Справедливость результата (28) вынуждает нас отказаться от непрерывности тензора внешней кривизны, если мы хотим уменьшить скачок плотности энергии. Поэтому мы рассматриваем семейство метрик, удовлетворяющих условиям

1. $\gamma_{\alpha\beta}(t)$ при $0 \leq t < 1/2$ поле класса C^2 , при $t \geq 1/2$ на границе ∂D_0 имеет разрывы производных первого рода;
2. $\gamma_{\alpha\beta}(t) = \gamma_{\alpha\beta}(0)$ вне ε -окрестности области D_0 ;
3. пространство $\langle M_0, \gamma_{\alpha\beta}(t) \rangle$ при $t > 1/2$ имеет неотрицательную кривизну;

4. пространство $\langle M_0, \gamma_{\alpha\beta}(t) \rangle$ при $t \in [0, 1]$ допускает регулярное единичное киллингово поле ξ_t .

В результате, так как

$$K^2 - K_{\mu\delta}K^{\mu\delta} = \frac{1}{4} (\gamma^{\alpha\beta}\gamma^{\mu\nu} - \gamma^{\alpha\mu}\gamma^{\beta\nu}) \frac{\partial\gamma_{\alpha\beta}}{\partial x^0} \frac{\partial\gamma_{\mu\nu}}{\partial x^0}, \quad (29)$$

получаем:

$$\langle \delta (K^2 - K_{\mu\delta}K^{\mu\delta}) \rangle = \frac{1}{4} \left\langle (\gamma^{\alpha\beta}\gamma^{\mu\nu} - \gamma^{\alpha\mu}\gamma^{\beta\nu}) \delta \left(\frac{\partial\gamma_{\alpha\beta}}{\partial x^0} \frac{\partial\gamma_{\mu\nu}}{\partial x^0} \right) \right\rangle. \quad (30)$$

Поэтому, уменьшение скачка плотности энергии возможно например в том случае, когда

$$\left\langle \delta \frac{\partial\gamma_{\alpha\beta}}{\partial x^0} \right\rangle \sim \frac{1}{\sqrt{\sigma}}.$$

Рассмотрим случай диагональной метрики. Тогда

$$\begin{aligned} \langle \delta (K^2 - K_{\mu\delta}K^{\mu\delta}) \rangle &= \frac{1}{2} \left\langle \delta \left(\frac{\partial \ln(\gamma_{11})}{\partial x^0} \frac{\partial \ln(\gamma_{22})}{\partial x^0} \right) \right\rangle + \\ &+ \frac{1}{2} \left\langle \left(\delta \frac{\partial \ln(\gamma_{11})}{\partial x^0} \frac{\partial \ln(\gamma_{33})}{\partial x^0} \right) \right\rangle + \\ &+ \frac{1}{2} \left\langle \left(\delta \frac{\partial \ln(\gamma_{22})}{\partial x^0} \frac{\partial \ln(\gamma_{33})}{\partial x^0} \right) \right\rangle. \end{aligned} \quad (31)$$

Если, например, производные

$$\frac{\partial\gamma_{11}}{\partial x^0}, \frac{\partial\gamma_{22}}{\partial x^0}$$

непрерывны и

$$\frac{\partial\gamma_{33}}{\partial x^0} \sim \frac{1}{\sigma},$$

то в силу соотношения (31) получим:

$$\langle \delta (K^2 - K_{\mu\delta}K^{\mu\delta}) \rangle = \frac{1}{2} \left\langle \delta \frac{\partial \ln(\gamma_{33})}{\partial x^0} \right\rangle \left(\frac{\partial \ln(\gamma_{11})}{\partial x^0} + \frac{\partial \ln(\gamma_{22})}{\partial x^0} \right), \quad (32)$$

при дополнительном условии

$$\left(\frac{\partial \ln(\gamma_{11})}{\partial x^0} + \frac{\partial \ln(\gamma_{22})}{\partial x^0} \right) = const.$$

При этом знак величины может оказаться как отрицательным так и положительным. Другими словами, надлежащим изменением производных по времени от компонент метрического тензора пространства-времени мы можем либо уменьшить, либо увеличить скачок плотности энергии. При этом, можно даже попытаться свести затраты энергии к выделению энергии.

Заключение

В данной работе мы показали, что усредненный скачок плотности энергии, возникающий в момент образования 4-мерной кротовой норы, можно уменьшить или увеличить. В результате трудности, связанные с проблемой больших затрат энергии необходимых для образования 4-мерной ручки, переходят в трудности иного рода: как изменить внешнюю кривизну 3-мерного пространства (другими словами, как изменить «способ» вложения 3-мерного пространства в пространство-время).

Мы до сих пор рассматривали только процедуру отрыва 3-мерной области от пространства M_0 . Но в действительности с рождением кротовой норы связан также и процесс приклеивания данной области к многообразию M_0 . И при этом также возникнет скачок энергии. Хотя, как нетрудно догадаться, эти процессы «взаимобратны». Это означает, что если в одном случае происходит выделение энергии, то в другом требуются ее затраты. Поэтому возникает следующее предположение. Если при отрыве области D_0 происходит выделение энергии, то можно попробовать каким-то образом сохранить хоть какую-то ее часть для того, чтобы использовать для приклеивания данной области обратно к пространству M_0 . Конечно, процесс диссипации энергии неизбежен, но мы по крайней мере сможем свести к минимуму проблему связанную с необходимыми на образование 4-мерной кротовой норы затратами энергии.

ЛИТЕРАТУРА

1. Гуц А.К. Изменение топологии физического пространства в замкнутой вселенной // Известия вузов. Физика. 1982. №5. С. 23-26.
2. Гуц А.К. Нарушение связности физического пространства // Известия вузов. Физика. 1983. №8. С. 3-6.
3. Гуц А.К. Элементы теории времени. Омск: Изд-во Наследие. Диалог-Сибирь, 2004.
4. Guts A.K. Time machine as four-dimensional wormhole. Los Alamos e-print gr-qc/9612064 (1996).
5. Konstantinov M. Ju. The Principle of Self-Consistency as a consequence of the Principle of Minimal Action. Los Alamos e-print gr-qc/9510039 (1995).
6. Константинов М.Ю. О кинематических свойствах топологически нетривиальных моделей пространства-времени // Известия вузов. Физика. 1992. №12. С.84-88.
7. Мизнер Ч., Торн К., Уиллер Дж. Гравитация. Т.2. М.: Мир, 1997.
8. Лайтман А., Пресс В., Прайс Р., Тюкольски С. Сборник задач по теории относительности и гравитации. М.: Мир, 1979.

О ПОЛЯХ ЯНО И ЯНО-КИЛЛИНГА В МОДЕЛЯХ ФРИДМАНА

А.В. Шалупаев, В.В. Клишевич

В работе получены все решения уравнений Яно и Яно-Киллинга для открытой и закрытой моделей Фридмана. Представленные выводы сформулированы в виде двух лемм. Обсуждаются некоторые применения полученных решений.

Модели Фридмана [1] занимают важное место в теоретических исследованиях по современной космологии. В основе моделей лежат предположения об однородности и изотропии распределения вещества по пространству. Современные астрономические данные не противоречат такому предположению. Существенным свойством моделей являются их нестационарность. По-видимому, именно это свойство дает правильное объяснение явления красного смещения [2].

Изучение в литературе моделей Фридмана ведется с разных точек зрения [3–5]. В зависимости от выбора в метрике масштабного фактора, получают модели, описывающие те или иные распределения вещества в пространстве. Тема нашей статьи связана с уравнением Дирака и его операторов симметрии в пространствах Фридмана. Общий вид операторов симметрии для уравнения Дирака и уравнения полей, по которым можно построить операторы симметрии, впервые были получены в работе В.Н. Шаповалова [6] и независимо в работах Картера и МакЛенагана [7, 8]. Фактически построение оператора симметрии для уравнения Дирака в произвольном римановом пространстве сводится к вопросу о существовании специальных векторных и тензорных полей, которые называются полями Киллинга, Яно и Яно-Киллинга (общие определения можно найти также в [9]). Зная решение уравнений на поля Киллинга, Яно и Яно-Киллинга в конкретном искривленном пространстве, можно построить операторы симметрии для уравнения Дирака в явном виде. Изучив алгебру, которую образуют операторы симметрии, мы имеем возможность исследовать вопросы интегрируемости уравнения Дирака и построения его решений.

Краткое содержание нашей работы следующее. В разделах 1, 2 мы рассматриваем уравнения на поля Яно и Яно-Киллинга и указываем необходимые алгебраические условия существования этих полей в

произвольном римановом пространстве. В разделе 3, опираясь на необходимые условия, мы строим решения указанных уравнений в открытой и закрытой моделях Фридмана. Результаты сформулированы в виде двух лемм.

Элемент длины в закрытой изотропной модели записывается в виде

$$ds^2 = a^2(\eta)(d\eta^2 - d\chi^2 - \sin^2 \chi(d\theta^2 + \sin^2 \theta d\varphi^2)). \quad (1)$$

Элемент длины в открытой изотропной модели записывается в виде

$$ds^2 = a^2(\eta)(d\eta^2 - d\chi^2 - \sinh^2 \chi(d\theta^2 + \sin^2 \theta d\varphi^2)). \quad (2)$$

Функция $a(\eta)$ называется масштабным фактором.

1. Векторное поле Яно

Векторное поле Яно определяется из уравнений

$$f_{i;j} = \frac{1}{4}g_{ij}f^k{}_{;k}, \quad f_i = \eta_{,i}. \quad (3)$$

Рассмотрим некоторые очевидные свойства этого векторного поля, подробно доказанные в работе [10].

Свойство 1. $f_{i;j} = f_{j;i}$. Следует из формулы (3).

Свойство 2. $f_{i,j} = f_{j,i}$. Следует из свойства 1 в силу симметричности символов Кристоффеля по нижним индексам.

Свойство 3. Если все компоненты вектора Яно не зависят от координаты x_j , то компонента f_j суть константа. Следует из свойства 2.

Свойство 4. Если компонента вектора Яно f_j суть константа, то все остальные компоненты не зависят от координаты x_j . Следует из свойства 2.

Свойство 5. Поднимая индекс i в формуле (3), получим $f^i{}_{;j} = 0$, $i \neq j$.

Свойство 6. Положив в предыдущей формуле $i = j = 1, 2, 3, 4$, получим $f^1{}_{;1} = f^2{}_{;2} = f^3{}_{;3} = f^4{}_{;4} = \Theta$. С учетом последнего свойства уравнение (3) может быть переписано в виде $f_{i;j} = \Theta g_{ij}$, а с использованием свойства 2 уравнению (3) можно придать следующий вид:

$$f_{i;j} + f_{j;i} = 2\Theta g_{ij}, \quad \Theta = \frac{1}{4}f^n{}_{;n}. \quad (4)$$

По терминологии монографии [9], векторное поле, удовлетворяющее уравнению (4) называется конформно-киллинговым, а по терминологии монографии [11] – обобщенным векторным полем Киллинга. В той же монографии [11, стр.72] указаны условия совместности этих уравнений в общем случае. Эти условия представляют собой соотношения на компоненты поля и их производные до определенного порядка. Покажем, что в нашей ситуации эти условия представляют собой соотношения только на компоненты поля.

Свойство 7. (Необходимое условие существования векторного поля Яно).

Сворачивая по индексам i и j формулу Риччи $f_{i;j;k} - f_{i;k;j} = f_n R_{ijk}^n$ и подставляя в нее формулу (3), получим $f_{i;j;k} = -\frac{1}{3}g_{ij}f_n R_k^n$. Используя формулу Риччи получим соотношение

$$f_n(g_{ij}R_k^n - g_{ik}R_j^n + 3R_{ijk}^n) = 0. \quad (5)$$

Тензор Римана определяется формулой:

$$R_{jkl}^i = \Gamma_{jl,k}^i - \Gamma_{jk,l}^i + \Gamma_{nk}^i \Gamma_{jl}^n - \Gamma_{nl}^i \Gamma_{jk}^n, \quad (6)$$

а тензор Риччи – сверткой по первому и третьему индексам:

$$R_{jk} = R_{jnk}^n. \quad (7)$$

Обобщим формулу (5) на произвольную размерность многообразия. Для этого следует использовать уравнение на вектор Яно в n -мерном пространстве, оно имеет вид:

$$f_{i;j} = \frac{1}{n}g_{ij}f_{;k}^k. \quad (8)$$

Формула (5) в n -мерном случае выглядит следующим образом:

$$f_n(g_{ij}R_k^n - g_{ik}R_j^n + (n-1)R_{ijk}^n) = 0. \quad (9)$$

2. Тензорное поле Яно-Киллинга

Антисимметричное тензорное поле Яно-Киллинга определяется из уравнений

$$f_{ij} + f_{ji} = 0, \quad f_{ij;k} = e_{ijkl}g^l, \quad (f_{ij;k} + f_{ik;j} = 0). \quad (10)$$

Рассмотрим некоторые очевидные свойства этого тензорного поля, подробно доказанные в работе [10].

Свойство 1. $f_{ij;k} = f_{jk;i} = f_{ki;j}$, то есть индексы можно циклировать. Следует из формулы (10).

Свойство 2. $f_{ij;k} = 0$ при совпадении любых двух индексов. Следует из формулы (10).

Свойство 3. $3f_{ij;k} = f_{ij;k} + f_{jk;i} + f_{ki;j} = f_{ij,k} + f_{jk,i} + f_{ki,j}$. Следует из антисимметрии этого тензора и симметрии символов Кристоффеля по нижним индексам.

Свойство 4. (Необходимое условие существования тензорного поля Яно – Киллинга). Рассматривая два уравнения на поле Яно – Киллинга, можно получить соотношение

$$R_{jji}^n f_{in} + R_{iij}^n f_{jn} = 0, \quad (11)$$

которое должно выполняться для всех индексов $i \neq j$.

Свойство 5. (Необходимое условие существования тензорного поля Яно – Киллинга). Рассматривая тройку уравнений на поле Яно – Киллинга, можно получить соотношение

$$(R_{kij}^n + R_{jik}^n)f_{in} - R_{iik}^n f_{jn} - R_{iij}^n f_{kn} = 0, \quad (12)$$

которое должно выполняться для всех различных индексов i, j, k . Отметим, что формула (11) есть частный случай формулы (12), если положить $k = j$.

3. Поля Яно и Яно-Киллинга в моделях Фридмана

При нахождении решений уравнений на поля Яно и Яно-Киллинга мы существенно использовали необходимые условия существования этих полей (5), (11) и (12), а также отмеченные их свойства. Приводим окончательные результаты. Всюду указываем только ненулевые компоненты.

Лемма 1. *Метрики Фридмана (1) допускают единственное нетривиальное векторное поле Яно и Яно-Киллинга только при значении масштабного фактора*

$$a = a_0 \exp(c\eta), \quad (13)$$

ненулевая компонента поля Яно имеет вид

$$f_1 = f_0 \exp(2c\eta), \quad (14)$$

ненулевые компоненты поля Яно-Киллинга для метрики (1) имеют вид

$$f_{23} = -(c_1 \cos(\varphi) - c_2 \sin(\varphi)) \sin(\chi) \exp(3c\eta), \quad (15)$$

$$f_{24} = (c_1 \sin(\varphi) + c_2 \cos(\varphi)) \sin(\chi) \sin(\theta) \cos(\theta) \exp(3c\eta), \quad (16)$$

$$f_{34} = (c_1 \sin(\varphi) - c_2 \cos(\varphi)) \sin(\chi)^2 \cos(\chi) \sin(\theta)^2 \exp(3c\eta) + c_3 \sin(\theta) \sin(\chi)^3 \exp(3c\eta), \quad (17)$$

здесь $a_0, c, c_1, c_2, c_3, f_0$ – произвольные постоянные.

Доказательство. Из необходимых алгебраических условий (5) следует, что если $a = a_0 \exp(c\eta)$, a_0, c – постоянные, то $f_2 = f_3 = f_4 = 0$. Далее из системы уравнений (3) находим общий вид поля Яно в виде (14). При $a \neq a_0 \exp(c\eta)$ системы (3) и (5) допускают только тривиальное решение. Используя необходимых алгебраических условий на поле Яно-Киллинга (12) следует, что если $a = a_0 \exp(c\eta)$, a_0, c – постоянные, то $f_{12} = f_{13} = f_{14} = 0$. Используя систему уравнений (10) находим общий вид поля Яно-Киллинга в виде (15-17). При $a \neq a_0 \exp(c\eta)$ системы (10) и (12) допускают только тривиальное решение. ■

Лемма 2. *Метрики Фридмана (2) допускают нетривиальное тензорное поле Яно и Яно-Киллинга только при значении масштабного фактора*

$$a = a_0 \exp(c\eta), \quad (18)$$

ненулевая компонента поля Яно имеет вид

$$f_1 = f_0 \exp(2c\eta), \quad (19)$$

ненулевые компоненты поля Яно-Киллинга для метрики (2) имеют вид

$$f_{23} = -(c_1 \cos(\varphi) - c_2 \sin(\varphi)) \sinh(\chi) \exp(3c\eta), \quad (20)$$

$$f_{24} = (c_1 \sin(\varphi) + c_2 \cos(\varphi)) \sinh(\chi) \sin(\theta) \cos(\theta) \exp(3c\eta), \quad (21)$$

$$f_{34} = -(c_1 \sin(\varphi) - c_2 \cos(\varphi)) \sinh(\chi)^2 \cosh(\chi) \sin(\theta)^2 \exp(3c\eta) + c_3 \sin(\theta) \sinh(\chi)^3 \exp(3c\eta); \quad (22)$$

здесь a_0, c, c_1, c_2, c_3 – произвольные постоянные.

Доказательство. Из необходимых алгебраических условий (5) следует, что если $a = a_0 \exp(c\eta)$, a_0, c – постоянные, то $f_2 = f_3 = f_4 = 0$. Далее из системы уравнений (3) находим общий вид поля Яно в виде (19). При $a \neq a_0 \exp(c\eta)$ системы (3) и (5) допускают только тривиальное решение. Из необходимых алгебраических условий (12) следует, что если $a = a_0 \exp(c\eta)$, a_0, c – постоянные, то $f_{12} = f_{13} = f_{14} = 0$. Далее из системы уравнений (10) находим общий вид поля Яно-Киллинга в виде (20)-(22). При $a \neq a_0 \exp(c\eta)$ системы (10) и (12) допускают только тривиальное решение. ■

ЛИТЕРАТУРА

1. Фридман А.А. О кривизне пространства. Петроград. 1922.
2. Ландау Л.Д., Лифшиц Е.М. Теория поля. Том II. М.: Наука. 1988.
3. Розенталь И.Л. Элементарные частицы и космология // УФН. 1997. Т.167, N.8. С.802-810.
4. Гинзбург В.Л. О некоторых успехах физики и астрономии за последние 3 года // УФН. 2002. Т.172, N.2. С.213-219.
5. Глинер Э.Б. Раздувающаяся Вселенная и вакуумоподобное состояние физической среды // УФН. 2002. Т.172, N.2. С.221-228.
6. Шаповалов В.Н. Симметрия уравнения Дирака-Фока // Известия вузов. Физика. 1975. N.6. С.57-63.
7. Carter B. Killing tensor quantum numbers and conserved currents in curved space // Phys.Rev. 1977. V.16D. P.3395-3414.
8. Carter B. and McLenaghan R.G. Generalized total angular momentum operator for the Dirac equation in curved space-time // Phys.Rev. 1979. V.19 D. P.1093-1097.
9. Яно К., Бохнер С. Кривизна и числа Бетти. М.: ИЛ. 1957.
10. Клишевич В.В. К вопросу о существовании спинорных операторов симметрии для уравнения Дирака // Известия вузов. Физика. 2000. Т.43, N.10. С.87-91.
11. Ибрагимов Н.Х. Группы преобразований в математической физике. М.: Наука. 1983.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СТРАТЕГИЙ ЭКОНОМИЧЕСКОГО ПОВЕДЕНИЯ ЛЮДЕЙ НА РЫНКЕ ТРУДА

О.С. Ёлкина, Е.В. Гуревич, А.К. Гуц

Дается математическая модель стратегий экономического поведения людей на рынке труда в рамках теории уравнений в частных производных. Показано, что стратегии экономического поведения индивидов описываются уравнением параболического типа.

Формирование рыночно организованной системы общественного производства предполагает наличие особых институтов: организаций, норм, правил и законов поведения, ранее не существовавших в жизни России. С вхождением рыночных отношений в жизнь каждого человека, большое значение приобретает знание механизмов и стереотипов экономического поведения людей. В сфере труда именно это поведение формирует занятость, что отражает степень готовности людей к «рыночным» законам распределения труда и дохода.

Одной из стадий формирования рыночных отношений является появление такого понятия, как «рынок труда». В настоящее время существует множество определений данной категории, мы придерживаемся мнения И.С.Масловой «рынок труда» – органическая сфера рыночной экономики, выполняющая функцию опосредования через куплю-продажу рабочей силы соединения вещественных и человеческого факторов производства, поддержания их сбалансированности в условиях многообразия форм собственности на средства производства и преимущественного права граждан на распоряжение своей рабочей силой и способностями» [1].

1. Стратегии экономического поведения индивидов на рынке труда

Каждый индивид имеет свою стратегию S экономического поведения на рынке труда, то есть устойчивую направленность, связанную с выбором желаемого

положения в сфере занятости, складывающегося исходя из получения определенного дохода при затрате определенных трудовых затрат.

Разнообразие стратегий экономического поведения людей складывается из взаимодействия трех факторов:

- величина дохода D ,
- величина труда T
- возможная область приложения труда как условия необходимого соединения труда и дохода S .

Причем последний фактор выступает не только местом, где возможно необходимое соединение, но и формирует привлекательные или непривлекательные соединения, а также формирует и изменяет условия соединения, создает возможности для формирования новых. Этот фактор способствует так же удовлетворению и таких потребностей, как достижение престижного статуса, социального положения.

Экономическое поведение людей формируется под влиянием широкой сферы экономической деятельности и, соответственно, должно охватывать все возможные варианты соединения труда и дохода – *типы стратегий*, которые оптимизируют поведение человека на рынке труда. Как показали исследования, проведенные О.С.Ёлкиной и В.С.Половинко [2], эти варианты формируют *матрицу стратегий* – некоторую область Ω в пространстве с осями D, T (доход, труд), которую более наглядно можно представить в следующем виде.

Рис. 1. Область Ω , представляющая стратегии экономического поведения индивидов на рынке труда

Каждый участок в Ω – это определенная область приложения труда (это не всегда конкретное предприятие или организация, это может быть и иная объективно существующая определенная сфера деятельности), **внутри**¹ которой действует строго определенный обмен труда на доход или выгоды. Этот обмен может быть эквивалентным, когда величина труда и величина дохода уравниваются:

- минимум дохода при минимуме труда ($\min D / \min T$);
- средний уровень дохода при среднем уровне труда ($\text{ср.} D / \text{ср.} T$);
- максимальный доход при максимальном уровне труда ($\max D / \max T$),

или же не эквивалентным:

- максимальный доход при минимуме труда ($\max D / \min T$);
- средний уровень дохода при максимальном уровне труда ($\text{ср.} D / \max T$);
- минимальный уровень дохода при среднем уровне труда ($\min D / \text{ср.} T$)
- и т.д.

когда величина дохода не соответствует величине трудовых усилий. Это объясняется тем, что существуют препятствия, которые тем или иным образом постоянно ограничивают хотя бы одну из переменных. Эти ограничения могут существовать как со стороны индивида, так и со стороны приоритетной (для индивида) сферы приложения труда. В конечном итоге двусторонние ограничения приводят к такому положению индивида на рынке труда, которое является для него оптимальным с трех сторон: со стороны приложения трудовых усилий, со стороны дохода и со стороны сферы приложения труда.

2. Уравнение для стратегий экономического поведения

Для того чтобы построить математическую модель, описывающую стратегии поведения индивида на рынке труда, мы обратим особое внимание на фактор S , называемый областью приложения труда (см. § 1). В силу сказанного о нем в § 1 фактор S является функцией, зависящей от двух переменных – от величины дохода D и от величины вкладываемого труда T . Тем самым индивид получает «точку расположения» (D, T) в области стратегий Ω , которая делится на сферы (прямоугольники на рис.1) приложения труда. Более того, число $S(D, T)$ как бы уточняет значимость «места» (D, T) с точки зрения возможного места приложения труда. Поскольку фактор S формирует границы в стратегиях, отделяя их одну от другой, и создает ограничения (ресурсные и т.д.) для переливов рабочей силы [2, с.64], то значение $S(D, T) = 0$ или условие $S|_L = 0$ на «кривой» $L \subset \Omega$ можно трактовать как отсутствие рабочей силы в точке (D, T) или соответственно как отсутствие «перелива» рабочей силы через L .

¹Мы особо обращаем внимание на слово «внутри», полагая, что на рынке труда могут существовать ситуации, когда выход за пределы данных участков невозможен (или почти невозможен). В какой-то степени это моменты стабилизации рынка труда. Излагаемая ниже математическая модель «отлавливает» это обстоятельство как «обертон» в «музыке рынка труда».

Позволим себе далее называть стратегией экономического поведения функцию $S = S(D, T)$.

Предположим, что индивид, находящийся в точке $(D, T) \in \Omega$, меняет свою стратегию и тем самым начинает перемещаться в другую точку $(D + dD, T + dT)$ области Ω . Как измерить разницу между стратегиями $S(D + dD, T + dT)$ и $S(D, T)$? Трудность заключается в том, что для перемещения индивида нет явно выделенного направления – точка $(D + dD, T + dT)$ может находиться относительно точки (D, T) на плоскости (D, T) в любом месте. А хотелось бы измерить интегральную разницу, интегральную разницу при смене стратегии.

Примем, что точка $(D + dD, T + dT)$ находится на окружности O очень малого радиуса ε с центром в исходной точке (D, T) . Тогда интегральная разница при смене стратегии равна

$$\begin{aligned} \delta &= \frac{1}{\varepsilon^2} \int_0^{2\pi} [S(D + r \cos \phi, T + r \sin \phi) - S(D, T)] d\phi = \\ &= \int_0^{2\pi} \left[S(D, T) + \frac{\partial S}{\partial D}(D, T) \varepsilon \cos \phi + \frac{\partial S}{\partial T}(D, T) \varepsilon \sin \phi + \frac{1}{2} \left[\frac{\partial^2 S}{\partial D^2}(D, T) \varepsilon^2 \cos^2 \phi + \right. \right. \\ &\quad \left. \left. + 2 \frac{\partial^2 S}{\partial D \partial T}(D, T) \varepsilon \cos \phi \sin \phi + \frac{\partial^2 S}{\partial T^2}(D, T) \varepsilon^2 \sin^2 \phi \right] + o(\varepsilon^3) \right] - S(D, T) d\phi. \end{aligned}$$

Интегрируя по ϕ , получаем

$$\delta = \pi \left[\frac{\partial^2 S}{\partial D^2} + \frac{\partial^2 S}{\partial T^2} \right] (D, T) + o(\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} \pi \Delta S.$$

Таким образом, разница при смене описывается лапласианом от S [3, с.96]. Вспоминая, что стратегия S должна еще зависеть от времени t , а процесс смены индивидами стратегий экономического поведения скорее всего походит на диффузию, поскольку похож на медленное расползание, вкрапление в ближайшие области приложения труда, то мы приходим к важному заключению: уравнение, которое описывает стратегию $S = S(t, D, T)$, в простейшем случае, – это уравнение диффузии:

$$\frac{\partial S}{\partial t} = a^2 \Delta S. \quad (1)$$

В более общем случае, мы должны рассматривать уже уравнение вида

$$\frac{\partial S}{\partial t} = a^2 \Delta S + f(t, D, T). \quad (2)$$

Каковы краевые условия для рассматриваемой задачи описания стратегий экономического поведения на рынках труда? Помимо очевидного начального условия

$$S(0, D, T) = S_0(D, T), \quad (3)$$

мы можем предположить, что индивиды, как правило, не выходят на границы прямоугольника Ω . Иначе говоря, имеем граничное условие

$$S|_{\partial\Omega \times [0, \infty)} = 0. \quad (4)$$

3. Стратегии поведения для Ω в случае однородного уравнения

Пусть $\Omega = [0, l] \times [0, m]$. Тогда, как известно, решение краевой задачи (1),(3),(4) имеет вид

$$S(t, D, T) = \frac{4}{lm} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} e^{-\pi^2 a^2 \left(\frac{k^2}{l^2} + \frac{j^2}{m^2} \right) t} \sin \frac{k\pi D}{l} \sin \frac{j\pi T}{m}.$$

Мы видим, что на основную моду $k = j = 1$ стратегий поведения на рынке труда

$$\sin \frac{\pi D}{l} \sin \frac{\pi T}{m}$$

накладываются дополнительные моды, одна из которых $k = j = 3$, разбивает прямоугольник $\Omega = [0, l] \times [0, m]$ на девять прямоугольников с нулевым значением функции стратегий $S(D, T)$ на границах этих девяти прямоугольников, удивительным образом соответствуя матрице стратегий Ω , изображенной на рис.1, и слову «внутри», на которое мы обратили внимание читателя в § 1.

Иначе говоря, наша модель с одной стороны отражает тонкое наблюдение специалистов-экономистов [2] в сфере рынка труда, а с другой – предсказывает массу других тонких сторон (другие моды) в стратегиях поведения людей на рынке труда, которые, как хочется надеяться, будут обнаружены теоретиками и практиками экономической науки.

ЛИТЕРАТУРА

1. Российский рынок труда. М.: Изд-во Ин-та экономики РАН, 1993. 188с.
2. Ёлкина О.С., Половинко В.С. Экономическое поведение работников на рынке труда. Омск: Изд-во ОмГПУ, 2001. 278 с.
3. Пу Т. Нелинейная экономическая динамика. Ижевск: Издательский дом «Удмуртский университет», 2000. 200 с.

ОПИСАНИЕ РАВНОВЕСИЙ НА РЫНКЕ ТРУДА С ПОМОЩЬЮ СТРУКТУР КУЛАКОВА-МИХАЙЛИЧЕНКО

А.К. Гуц, О.С. Ёлкина

Предлагается математическое описание равновесий на рынке труда, опирающееся на теорию фундаментальных отношений Ю.И.Кулакова. Показан вклад структур Кулакова-Михайличенко в формирование точек равновесия на рынке труда.

Одним из самых удивительных открытий в современной науке является выявление универсальной роли при описании различных явлений в природе и обществе так называемых фундаментальных отношений (структур) Кулакова, существование которых впервые было обнаружено Ю.И.Кулаковым в физике и геометрии [1, 2, 5]. Со временем они были найдены в психологии и социологии [2] и, наконец, в микро- и макроэкономике [3–5].

Естественно было ожидать, что структуры Кулакова проявят себя и в теории рынка труда. Именно это и показано в данной заметке. При этом мы постарались выявить не просто универсально-фундаментальный характер с точки зрения теории Кулакова рыночно-трудовых отношений, но и выявить вклад структур Кулакова в формирование точек равновесия на рынке труда.

1. Структуры Кулакова

Рынок труда складывается под воздействием двух систем: спроса на труд W/P и предложение труда L . Иначе говоря, допуская упрощение, мы можем сказать, что имеется множество $\mathcal{S}$ работодателей, обеспечивающих спрос на труд, и множество $\mathcal{P}$ работников, предлагающих себя в качестве рабочей силы, т.е. обеспечивающие предложение труда. Мы констатируем наличие отношения $f(i, \alpha), i \in \mathcal{S}, \alpha \in \mathcal{P}$ на рынке труда $\mathcal{S} \times \mathcal{P}$, где $f : \mathcal{S} \times \mathcal{P} \rightarrow \mathbb{R}$ – вещественная функция.

Теория Ю.И.Кулакова указывает на наличие в любой предметной сфере (геометрия, механика, макроэкономика и пр.) следующих симметрии и закона Φ : допускается, что существует пара натуральных чисел (s, r) – ранг структуры

отношений, такая, что на рынке труда любые s работодателей взаимозаменяемы на любых других s работодателей и любые r работодателей взаимозаменяемы на любых других r работников. Кроме того, постулируется существование функции Φ (закона для структуры Кулакова ранга (s, r)) такой, что

$$\Phi(f(i_1, \alpha_1), f(i_2, \alpha_1), \dots, f(i_s, \alpha_1), f(i_1, \alpha_2), \dots, f(i_s, \alpha_r)) = 0$$

для любого набора $(i_1, \dots, i_s)$ по s работодателей и набора $(\alpha_1, \dots, \alpha_r)$ по r работников.

В результате этого постулата удастся найти вид отношения f и вид закона Φ (см. подробности в [1]).

2. Графические зависимости для равновесия на рынке труда

Ситуация на рынке труда иллюстрируется следующими графиками:

Рис. 1. Графическое описание равновесия на рынке труда

Здесь:

- кривая D – это спрос на труд. Зависимость простая; чем выше цена на трудовые ресурсы, тем меньше спрос на них;
- кривая S – это предложение труда. Чем выше цена на труд, тем больше людей предлагают свой труд на рынке труда.
- точка равновесия M . Когда спрос и предложение на труд равны, устанавливается относительное равновесие. Это равновесие не может находиться длительное время в устойчивом состоянии.

Возможны перемещения кривых S и D вверх и вниз. При этом происходит смещение точки равновесия M , и устанавливается новое относительное равновесие M_1 .

Каждая из двух приведенных кривых не задается посредством точной математической формулы. Например, очевидно, что кривая D может

задаваться как

$$W/P = \frac{c}{L}, \quad (1)$$

а кривая S –

$$W/P = c_1 \cdot L^2 + c_2. \quad (2)$$

Соотношение (1), будучи переписанным в виде

$$c = (W/P)L,$$

сразу обнаруживает структуру Кулакова рода (2,2). Роль отношения по Кулакову играет здесь величина c . Это отношение между работниками (L) и работодателями, определяющими оплату труда (W/P). Взаимозаменяемость работников и взаимозаменяемость работодателей определяет тип симметрии, которая и порождает системы отношений по Кулакову.

При выбранных переменных соотношение (2) не подпадает ни под одну из формул для фундаментальных отношений Кулакова. Тем не менее, универсальный характер системы отношений, выявленной Кулаковым обнаруживается. Но для этого нужно перейти к обобщению теории Кулакова и описывать ситуацию равновесия M не с помощью одного отношения $f : \mathcal{S} \times \mathcal{P} \rightarrow \mathbb{R}$, а естественно с помощью пары отношений $f : \mathcal{S} \times \mathcal{P} \rightarrow \mathbb{R}^2$.

3. Появление пары отношений Кулакова-Михайличенко при описании равновесий на рынке труда

Выше было дано описание рынка труда в виде двух кривых. Их пересечение задает точку равновесия.

Ось абсцисс X – величина требующегося труда (количество работников) (L)
Ось ординат Y – величина реально заработной платы (цена на труд) (W/P). Кривая D показывает, каким будет спрос на труд при определенной величине заработной платы. Показывает, что более низкой заработной плате соответствует больший спрос на труд и наоборот. Кривая S – показывает величину предложения труда.

Кривые D и S можно получить также на основе принципа симметрий трудовых отношений. Для этого нужно использовать теорию двуметрических структур Михайличенко, обобщающую теорию фундаментальных отношений Кулакова. Г.Г.Михайличенко рассматривает не одно, как у Ю.И.Кулакова, отношение, а сразу два различных отношений.

В статье [6] дается описание всех двуметрических структур ранга $(n+1, 2)$.

Например, при $n = 2$ имеем следующую пару трудовых отношений:

$$f_1(X, Y; A, B) = XA, \quad f_2(X, Y; A, B) = YA + B.$$

Предположим, что трудовые отношения фиксированы, т.е.

$$f_1(X, Y; A, B) = XA = \text{const} = a,$$

$$f_2(X, Y; A, B) = YA + B = \text{const} = b.$$

Тогда, если принять, что $A = Y, B = -X$ (в 4-мерном пространстве выделяется 2-мерная плоскость), то

$$X = (f_1 - a)/Y,$$

$$X = Y^2 + b.$$

Переобозначив $Y = L, X = W/P$, находим конкретное выражение для кривые D и S . При этом возникает естественный вопрос, какой смысл имеют координаты A и B ? Оставляя этот вопрос пока без ответа, мы получаем, что фундаментальные отношения f_1, f_2 с точки зрения теории Кулакова-Михайличенко тесно связаны с такими основными величинами, характеризующими рынок труда как количеством работников L и величиной реально заработной платы (цена на труд) W/P :

$$W/P = f_1/A,$$

$$L = (f_2 - B)/A.$$

ЛИТЕРАТУРА

1. Кулаков Ю.И. Теория физических структур. М., 2004. 847 с.
2. Гуц А.К., Коробицын В.В., Лаптев А.А. и др. Математические модели социальных систем: Учебное пособие. Омск: ОмГУ, 2000. 256 с.
3. Добренко М.А., Гуц А.К. Первичные структуры отношений Кулакова в микроэкономике // Математические структуры и моделирование. 2003. Вып.11. С.88-96.
4. Добренко М.А., Гуц А.К. Макроэкономические первичные структуры отношений Кулакова // Математические структуры и моделирование. 2003. Вып.12. С.130-133.
5. Гуц А.К., Паутова Л.А., Фролова Ю.В. Математическая социология: Учебное пособие. Омск: Издательство Наследие. Диалог-Сибирь, 2003. 192 с.
6. Михайличенко Г.Г. Двуметрические физические структуры ранга $(n_1, 2)$ // СМЖ. 1993. Т.34. N.2. С.132-143.

ВОЗМОЖНЫЕ ВАРИАНТЫ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ НЕСАНКЦИОНИРОВАННОЙ РАБОТЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Е.Н. Дудоров

Представлен анализ основных методов обнаружения несанкционированной работы программного обеспечения. Рассмотрен возможный подход к построению интеллектуальной системы выявления подозрительной активности исполняемого кода.

В настоящее время большое внимание уделяется вопросам обеспечения информационной безопасности автоматизированных систем обработки информации, предназначенных для управления жизненно важными объектами страны. К таким объектам можно отнести системы телекоммуникаций, атомные станции, системы управления воздушным и наземным транспортом, банковские системы, системы обработки и хранения секретной и конфиденциальной информации и т.д.

Указанные автоматизированные системы также называются критическими, так как любое искажение, несанкционированное использование, потеря циркулирующей в них информации может привести к неисправимым последствиям. В целях предупреждения негативных ситуаций необходимо постоянно совершенствовать методы и средства обеспечения безопасности подобных компьютерных систем. Любая явная или скрытая угроза направлена на нарушение целостности вычислительных процессов, происходящих в автоматизированной системе. В качестве возможного подхода к обеспечению стабильной работы предлагается внедрить в компьютерную систему интеллектуальную среду, позволяющую анализировать и контролировать процессы, сопутствующие запуску и работе программного обеспечения.

Стремительное развитие информационных технологий привело к появлению бесчисленного множества программных средств скрытого информационного воздействия. При создании нового программного продукта в большинстве случаев используется уже готовые программные решения, порой содержащие в себе заимствованные ошибки программирования, а также дополнительные программные модули, способные нанести существенный вред при обработке

информации. Данное обстоятельство приводит к необходимости создания гибкого аппарата принятия решений о несанкционированном поведении внедряемого программного обеспечения.

В качестве решения данной задачи предлагается использовать наработки таких направлений искусственного интеллекта как теория распознавания образов, нечеткая логика и искусственные нейронные сети.

Рассмотрим наиболее распространенные подходы к выявлению подозрительной активности программного обеспечения.

1. Сигнатурный метод анализа

Метод основан на том, что большинство вредоносных воздействий на систему известны и развиваются по схожим сценариям. В данном подходе сигнатуры подозрительной активности определяют характерные особенности, условия, устройства и взаимосвязь событий, которые ведут к попыткам нарушения безопасности системы или способствуют им. Простейшим методом реализации сигнатурного анализа является поддержание системой безопасности базы данных сигнатур вторжений. Последовательность действий, выполняемая пользователем или программой во время выполнения, сравнивается с известными сигнатурами.

Необходимо отметить, что непосредственное сравнение сигнатуры вторжения с регистрируемой активностью малоэффективно в связи с тем, что регистрируемые данные, относящиеся к атаке, часто бывают зашумлены вследствие вариаций действий нарушителя во время атаки или мутаций сценария.

2. Контроль работы программ по профилям

Наиболее характерные особенности автоматизированных систем обработки информации (АС) является наличие большого количества разнородных компонентов, сложных взаимно переплетающихся связей, развитой системы математического обеспечения, предназначенной для обработки огромных информационных потоков. Любая АС характеризуется множеством состояний. В компьютерных системах для описания поведения субъектов по отношению к объектам используют понятие «профили».

Профиль представляет собой образ, который создается на базе множества состояний протекающих процессов в АС. Иначе, профили это набор статистических характеристик (частоты встречаемости событий, временные интервалы работы CPU и т.д.), вычисленных по наблюдениям за действиями субъекта по отношению к объекту, а также некоторой статистической моделью такого поведения.

Структура профиля может быть представлена в следующем виде: имя переменной; отражаемые действия; имеющиеся исключения; данные использования ресурсов; период измерений; порог допустимых значений; субъект; объект; значение последнего наблюдения модели и т.д. Статистические

сведения о программах можно получить лишь путем мониторинга за его работой в конкретной среде и на конкретной платформе.

Любое отклонение используемого профиля от эталонного, рассматривается как проявление подозрительной активности программы. Недостатком данного метода является: во-первых, факт того, что «статистические» системы могут быть с течением времени «обучены» нарушителями так, чтобы подозрительные действия рассматривались как нормальные. Во-вторых, «статистические» системы не чувствительны к порядку следования событий. В-третьих, очень трудно задать граничные (пороговые) значения характеристик, отслеживаемых системой обнаружения аномалий, чтобы адекватно идентифицировать подозрительную деятельность.

3. Использование прогнозируемых шаблонов

Этот способ позволяет «предсказывать» будущие события на основе уже происшедших. Данный метод имеет несколько преимуществ. Во-первых, правила, базирующиеся на последовательности шаблонов, могут определять подозрительную активность, которую трудно идентифицировать традиционными методами. Во-вторых, системы, построенные на этой модели, очень высоко приспособлены к изменениям. Это связано с тем, что редкие шаблоны непрерывно удаляются, оставляя наиболее часто используемые шаблоны. В третьих, подозрительные действия могут быть обнаружены в течение нескольких секунд после генерации события.

Недостаток данного метода состоит в том, что если в базе знаний не описаны некоторые сценарии осуществления несанкционированной деятельности, то такие действия не будут определены как подозрительные.

4. Метод обнаружения опасных комбинаций безопасных событий

Подразумевается распознавание последовательности, на первый взгляд, «безобидных» действий, суммарным результатом которых является нарушение целостности вычислительного процесса и доведение системы до неработоспособного состояния. Сложность распознавания таких событий заключается в большом количестве комбинаций опасных событий.

5. Анализ перехода системы из состояния в состояние

В данном подходе подозрительная активность представляется как последовательность переходов контролируемой системы из состояния в состояние. Состояние шаблона атаки соответствует состоянию системы и связано с утверждениями, которые должны быть удовлетворены для последующего перехода из состояния в состояние. Возможные состояния связаны дугами, представляющими события, необходимые для перехода из состояния в состояние. Типы допустимых событий встроены в модель. Данная

модель может определить только атаку, состоящую из последовательных событий, не позволяя выразить атаки с более сложной их структурой. Более того, не существует общего целевого механизма в случае частичного распознавания атак.

6. Нечеткие сети Петри

Метод, использующий формализм сетей Петри для обнаружения нарушителей, является развитием метода анализа изменений состояний. Сети Петри представляют собой математическую модель для представления структуры и анализа динамики функционирования систем в терминах «условие-событие». К настоящему времени известно большое количество разновидностей сетей Петри, к которым, в первую очередь, следует отнести временные СП, СП с разноцветными маркерами и дугами, алгебраические СП, Е-сети. Данные классы моделей позволяют представить структуру и динамику функционирования исследуемых систем в условиях отсутствия влияния различных факторов неопределенности [7]. Указанный метод позволяет получить более точные результаты в области обнаружения нарушителя, однако требует значительных вычислений.

7. Контроль превышения пороговой величины частоты событий

Метод наиболее часто используется для распознавания атак типа «отказ в обслуживании». Для анализа частоты событий и установления для них пороговых величин системе распознавания требуются сведения о нормальном, рабочем состоянии АС, которые для каждой системы уникальны. Использование данного подхода требует глубокого анализа всех параметров АС, а так же составление алгоритма действий на случай превышения пороговых величин.

8. Статистический анализ последовательности системных вызовов

Для обнаружения несанкционированной деятельности исследуется применимость различных моделей анализа данных к последовательностям системных вызовов, сгенерированных исследуемой программой во время исполнения. Эти данные могут рассматриваться как стационарные (отсутствие корреляции между различными запусками программ), но не независимыми (часто имеются различные распределения последовательностей вызовов в начале выполнения программы, в ее конце и на определенных участках) [5]. Наиболее часто используются следующие методы анализа данных:

- простая нумерация последовательностей. Метод заключается в нумерации последовательностей системных вызовов, обнаруженных в трассе программы при нормальном исполнении и вторжении.

– относительная частота последовательностей. Метод основан на частотном распределении последовательность системных вызовов. Каждая последовательность характеризуется значением, определяемым на основе того, как часто она встречается в нормальном поведении. Таким образом, каждое выполнение программы может быть представлено гистограммой последовательностей вызовов. Каждая гистограмма определяет точку в многомерном пространстве. Множество нормальных выполнений программы образует центроиды, относительно расстояния до которых можно делать заключения о правильности текущего выполнения.

– информационный анализ. Метод состоит в выделении особенностей больших наборов данных, т.е. необходимо определить наиболее компактное описание нормального поведения на основе перебора максимально возможных шаблонов, отражающих нормальное поведение пользователя. На основе определенных особенностей необходимо сделать обобщение для определения нормальных шаблонов, пропущенных в тренировочных данных.

– машина конечных состояний (скрытая модель Маркова). Подход состоит в разработке машины конечных состояний для распознавания "языка-трассы" программы. Для решения данной задачи существует много техник, основанных на использовании как детерминистских, так и вероятностных автоматов. Обычно определяется частота, с которой появляются отдельные символы (системные вызовы) в зависимости от некоторого числа предыдущих символов. Состояния соответствуют истории системы, а переходы показывают вероятность появления следующего символа. При этом алгоритмы базируются на стационарности данных.

Одной из основных моделей построения таких систем является Скрытая Модель Маркова (СММ). Состояния СММ представляют некоторые ненаблюдаемые условия моделируемой системы. В каждом состоянии существует вероятность некоторого наблюдаемого вывода системы и отдельная вероятность, определяющая следующее состояние. Наличие в модели отдельного распределения вероятности вывода для каждого состояния и возможности изменений состояний во времени позволяет представить в модели нестационарные последовательности. Модель является очень мощной, но требует большого объема вычислений.

9. Продукционные /экспертные системы

Продукционные /экспертные системы обнаружения вторжений представляют собой специализированные автоматы обработки знаний для интерактивного и кооперативного решения проблем распознавания на естественном профессиональном языке со способностями приобретения, хранения и представления знаний в форме алгоритмических программ с одной стороны и неалгоритмических фактов и правил, с другой стороны [1].

Основное преимущество данных систем состоит в том, что они не отвергают и не заменяют традиционного подхода к программированию. Они отличаются от традиционных программ тем, что ориентированы на решение

неформализованных задач и обладают следующими особенностями:

- алгоритм решений не известен заранее, а строится самой экспертной системой с помощью символических рассуждений, базирующихся на эвристических приемах;
- ясность полученных решений, то есть система «осознает» в терминах пользователя, как она получила решение;
- способность анализа и объяснения своих действий и знаний;
- способность приобретения новых знаний от пользователя-эксперта, не знающего программирования, и изменения в соответствии с ними своего поведения.

Таким образом, экспертная система состоит из набора правил, которые охватывают знание человека-эксперта. На основе анализа данных, получаемых от модулей слежения, экспертная система может предоставлять заключение об аномальной работе программного обеспечения. Экспертные системы допускают объединение огромного опыта, накопленного человеком, в компьютерном приложении, которое затем использует эти знания для идентификации деятельности, которая соответствует несанкционированному поведению программ. Наряду с достоинствами есть и недостатки. Наличие базы знаний требует высокой квалификации для ее постоянного обновления для того, чтобы оставаться актуальной. [2]

Системы на основе фиксированных правил страдают от неспособности обнаруживать сценарии несанкционированной деятельности, которые имеют место в течение продолжительного времени. Одним из путей устранения названной проблемы является использование нейронных сетей.

10. Нечеткая логика

Нечеткая логика (fuzzy logic) позволяет описывать правила в незавершенном, «размытом» режиме, в котором правила основываются на знаниях и весах событий, позволяющих предположить вероятность атаки. Использование аппарата нечеткой логики при составлении правил экспертной системы позволяет ввести нечеткость в определение вредоносных воздействий и сетевых вторжений. Ввиду новизны и сложности технологии на сегодняшний день известно ограниченное число специализированных приложений.

11. Нейронные сети

В отличие от экспертных систем, которые могут дать пользователю определенный ответ, соответствуют или нет рассматриваемые характеристики характеристикам, заложенным в базе данных правил, нейросеть проводит анализ информации и предоставляет возможность оценить, согласуются ли данные с характеристиками, которые она научена распознавать. Наиболее важное преимущество нейросетей при обнаружении подозрительной активности программы заключается в способности «изучать» характеристики

несанкционированной деятельности и идентифицировать процессы, которые не похожи на те, что наблюдались в системе прежде.

Общая идея использования нейронных сетей заключается в следующем [10,11]: искусственная нейросеть состоит из набора элементарных элементов, которые взаимосвязаны друг с другом и трансформируют набор входных данных к набору желаемых выходных данных. Результат преобразования определяется характеристиками элементов и весами, соответствующими взаимосвязям между ними. Путем видоизменения соединений между узлами сети можно адаптироваться к желательным выходным результатам. Обучение сети включает многократную подачу примеров на ее входы. Сеть отгадывает выход, сравнивает его с предложенным правильным и выполняет внутреннюю коррекцию сети, если выход был неправильным. Этот процесс повторяется для каждого примера. Для оценки выхода сети, необходим критерий, определяющий достаточность соответствия. Обычно считается достаточным, если сеть дает правильный ответ в 90 случаев.

Искусственные нейросети представляют значительный потенциал для решения задач, связанных с обнаружением злоупотреблений, внешних атак, направленных против сетевых ресурсов. Но, до сих пор остаются неразрешенными следующие проблемы:

- если для обучения используется небольшое множество примеров, результаты обнаружения являются приемлемыми; большое множество примеров значительно расширяет пространство решений, что приводит к неприемлемому уровню ошибок в распознавании;
- нейросети не позволяют увидеть логику принятия решения;
- нейросети требуют четкой настройки топологии и параметров сети, что во многом требует от специалиста высокой квалификации в данной области.

В настоящее время большинство из перечисленных методов применяются в системах выявления атак (Intrusion Detection System IDS). Эти системы решают задачу мониторинга информационной системы на сетевом, системном и прикладном уровнях с целью выявления нарушений безопасности и оперативного реагирования на них. Наиболее известные продукты: RealSecure, Snort NetRanger, NIDES. Однако стоит отметить, что подобные системы, как правило, ориентированы на отслеживание состояния сетевого взаимодействия отдельных хостов или сегментов сети. Разработка и создание интеллектуальных систем распознавания подозрительной активности программного обеспечения является отдельным, перспективным направлением в организации комплексной защиты АС.

В этой связи предлагается в качестве возможного подхода при построении систем выявления подозрительной активности ПО использовать наработки таких направлений искусственного интеллекта как нейронные сети и аппарата нечеткой логики на базе экспертной системы. Данный подход позволит за счет нейросети фильтровать большой объем вычислительных процессов, выделять подозрительные признаки деятельности и передавать данные экспертной системе для последующего принятия решения о нештатной работе программы. Структура возможной системы обнаружения подозрительной активности ПО

представлена на рисунке 1.

Рисунок 1. Основные элементы системы обнаружения подозрительной активности ПО

В составе интеллектуальной системы присутствуют следующие компоненты:

Сенсоры слежения – обеспечивают сбор данных из контролируемого пространства.

Подсистема обнаружения подозрительной активности – непосредственно занимается обнаружением подозрительной активности исполняемого кода, основываясь на критериях обнаружения.

База знаний – в зависимости от методов, используемых в системе, может содержать профили вычислительной системы, сигнатуры атак, сценарии осуществления несанкционированной деятельности и т.д.

Подсистема принятия решений – контролирует все компоненты системы обнаружения, хранит конфигурацию и инициирует сенсоры и подсистему обнаружения подозрительной активности, обеспечивает их критериями обнаружения, принимает решение о блокировании опасных процессов.

Хранилище данных о подозрительной активности – обеспечивает хранение данных, собранных в процессе функционирования подсистемы обнаружения подозрительной активности.

Подсистема реагирования – система, осуществляющая реагирование на обнаруженные несанкционированной деятельности программы.

Теоретическая правомерность использования нейросетей для идентификации подозрительной активности программ подтверждается результатами, полученными в ходе проведенных исследований при построении интеллектуальных систем обнаружения компьютерных вирусов [9, 10]. Предлагаемая сеть была предназначена для изучения характеристик деятельности обычной системы и идентификации статистических отклонений от нормы, что позволяло говорить о наличии вируса.

Таким образом, предложенный подход может быть реализован при построении систем выявления подозрительной активности программ.

ЛИТЕРАТУРА

1. Белов В.И. *Теория фазовых измерительных систем* / Под. ред. проф. Г.Н.Глазова. Томск: ТГАСУР, 1994.
2. Белозерский Л.А. Конспект лекций по курсу «Основы построения систем распознавания образов». Донецкий государственный институт искусственного интеллекта, 1997.
3. Дж. Кеннеди. Нейросетевые технологии в диагностике аномальной сетевой активности (перевод А.В. Лукацкого, Ю.Ю. Цаплевым, В.П. Сахоровым)// Сборник статей НИП «Информзащита», 2000.
4. . Корнеев В.В. Использование нейросетей для аутентификации классов пользователей, операторов, идентификации управляющих воздействий и технологических процессов в компьютерах // Сб. докл. Республиканской научно-технической конференции «Методы и технические средства обеспечения безопасности информации» 17-19 окт. 1995. с. 95-97.

-
5. Корнеев В.В., Масалович А.И., Савельева Е.В., Шашаев А.Е.. Распознавание программных модулей и обнаружение несанкционированных действий с применением аппарата нейросетей. <http://banana.stack.net:16000/db/msg/23820.html>
 6. Корт С.С. Теоретические основы защиты информации. Учебное пособие. М.: Гелиос АРВ, 2004. 240 с., ил.
 7. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия - Телеком, 2001. 382 с.: ил.
 8. Чернолучский И.Г. Методы принятия решений. СПб.: БХВ-Петербург, 2005. 416 с.: ил.
 9. Минаев Ю.Н., Филимонова О.Ю., Бенамеур Лиес. Методы и алгоритмы решения задач идентификации и прогнозирования в условиях неопределенности в нейросетевом логическом базисе. М.: Горячая линия - Телеком, 2003. 205 с.: ил.
 10. Denault, M., Gritzalis, D., Karagiannis, D., and Spirakis, P. (1994). Intrusion Detection: Approach and Performance Issues of the SECURENET System. In *Computers and Security* Vol. 13, No. 6, pp. 495-507
 11. Fox, Kevin L., Henning, Rhonda R., and Reed, Jonathan H. (1990). A Neural Network Approach Towards Intrusion Detection. In *Proceedings of the 13th National Computer Security Conference*.
 12. Hampshire J., Perlmutter B. A. . Equivalence Proofs for Multy-Layer Perceptron Classifiers and the Bayesian Discriminant Function. Carnegie Mellon University, Pittsburg, 1997.
 13. Сайт с ПО SNORT. <http://www.snort.org>.
 14. Сайт с ПО NIDS. <http://www.bgnet.no/giva> .

Математические структуры и моделирование

Вып. 15

Журнал

Редактор Е.В. Брусницына

Подписано в печать 10.11.2005.
ОП. Формат 60 × 84 1/8. Печ.л. 23,5. Уч.-изд.л. 21,0.
Тираж 120 экз.

Отпечатано в Полиграфическом центре КАН
644050, г. Омск, пр. Мира, 32, ком.11, тел. (381-2) 65-47-31
Лицензия ПЛД № 58-47 от 21.04.97 г.