

ISSN 2222-8772

МАТЕМАТИЧЕСКИЕ
СТРУКТУРЫ
И
МОДЕЛИРОВАНИЕ

№ 1(45)
2018



**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. Ф.М. ДОСТОЕВСКОГО»**

**МАТЕМАТИЧЕСКИЕ
СТРУКТУРЫ
И
МОДЕЛИРОВАНИЕ**

№ 1(45)

Омск
2018

Математические структуры и моделирование. — Омск : Омский государственный университет, 2018. — № 1(45). — 171 с.
ISSN 2222-8772 (print)
ISSN 2222-8799 (online)

Редакционная коллегия

А. К. Гуц	главный редактор, председатель редакционной коллегии, доктор физ.-мат. наук, профессор, зав. кафедрой кибернетики, Омский государственный университет им. Ф.М. Достоевского.
Д. Н. Лавров	ответственный за выпуск редактор, зам. глав. редактора, канд. техн. наук, доцент, зав. каф. компьютерных технологий и сетей, Омский государственный университет им. Ф.М. Достоевского.
Н. Ф. Богаченко	технический редактор, зам. глав. редактора, канд. физ.-мат. наук, доцент, Омский государственный университет им. Ф.М. Достоевского.
С. В. Белим	д.ф.-м.н., профессор, зав. кафедрой информационной безопасности, проректор по научной работе, Омский государственный университет им. Ф.М. Достоевского.
В. П. Голубятников	доктор физ.-мат. наук, профессор Новосибирского государственного университета, главный научный сотрудник Института математики СО РАН, г. Новосибирск.
С. И. Горлов	доктор физ.-мат. наук, профессор, ректор Нижневартковского государственного университета.
А. Г. Гринь	доктор физ.-мат. наук, профессор, кафедра кибернетики, Омский государственный университет им. Ф.М. Достоевского.
В. А. Еровенко	доктор физ.-мат. наук, профессор, зав. кафедрой общей математики и информатики Белорусского государственного университета, г. Минск, Республика Беларусь.
B. Zilber	Dr.Sc. (Phys.-Math.), Professor of Mathematical Logic, Mathematical Institute, University of Oxford, UK.
А. Н. Кабанов	канд. физ.-мат. наук, кафедра кибернетики, Омский государственный университет им. Ф.М. Достоевского.
А. В. Копыльцов	д.т.н., профессор, кафедра информационных систем, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина).
А. Г. Коробейников	д.т.н., профессор, зам. директора по науке Санкт-Петербургского филиала Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В. Пушкова РАН.
П. А. Корчагин	доктор техн. наук, профессор, проректор по научной работе, Сибирская государственная автомобильно-дорожная академия (СибАДИ).
V. Kreinovich	Ph.D. (Phys.-Math.), Professor, Computer Science Department, University of Texas at El Paso, Texas, USA.
А. В. Плетюхов	д.ф.-м.н., профессор кафедры общей и теоретической физики Брестского государственного университета им. А.С. Пушкина, Республика Беларусь.
Л. Б. Соколинский	д.ф.-м.н., профессор, проректор по информатизации, зав. кафедрой системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет), г. Челябинск.
A. A. Fedorenko	Ph.D. (Phys.-Math.), Researcher (CR1) at the French National Centre of Scientific Research (CNRS) Laboratoire de Physique de l'ENS-Lyon, France.
A. Jadczyk	Ph.D., Professor, Researcher, Laboratoire de Physique, Universite de Toulouse III et CNRS, France.

Учредитель

Федеральное государственное бюджетное образовательное учреждение высшего образования «Омский государственный университет им. Ф. М. Достоевского».
Свидетельство о регистрации средства массовой информации ПИ № ФС77-72200 от 15 января 2018 г. выдано Роскомнадзором.

Адрес редакции, издателя и типографии

644077, Омская обл., г. Омск,
пр-т Мира, д. 55а.

Дата выхода в свет: 30.03.2018.

Тираж 100 экз.

Свободная цена.

МАТЕМАТИЧЕСКИЕ СТРУКТУРЫ И МОДЕЛИРОВАНИЕ

Журнал основан в 1998 году. В журнале публикуются статьи, в которых излагаются результаты исследований по фундаментальной и прикладной математике, теоретической физике, компьютерным наукам, философии и истории математики и информатики, а также размышления, касающиеся окружающей нас природы и общества. Объекты исследования должны быть представлены в форме некоторых математических структур и моделей.

Все статьи журнала проходят обязательное рецензирование. Рефераты статей журнала опубликованы в «Реферативном журнале» и «Mathematical Reviews» (США). Журнал индексируется в РИНЦ (elibrary.ru) и «Zentralblatt für Mathematik» (Германия). Журнал входит в Перечень рецензируемых научных изданий ВАК РФ, в которых должны быть опубликованы основные результаты диссертаций на соискание учёных степеней (Приказ Минобрнауки России от 25 июля 2014 г. № 793).

Все статьи в журнале публикуются под лицензией Attribution 4.0 International (CC-BY).

Электронная версия журнала представлена в сети:

<http://msm.univer.omsk.su>
<http://msm.omsu.ru>

Подписной индекс по каталогу «Пресса России»: 94082

Электронная почта главного редактора:

guts@omsu.ru

Электронная почта выпускающего редактора:

lavrov@omsu.ru

СОДЕРЖАНИЕ

Фундаментальная математика и физика

- А.Г. Гринь. *Критерий нормальной сходимости для симметрических функций от зависимых величин* 5
- А.В. Левичев, А.Ю. Пальянов. *Анализ в космических расслоениях на основе групп $U(1,1)$ и $U(2)$: случаи $SU(2,2)$ -действий в их 2- и 4-накрытиях* 12
- В.А. Еровенко. *«Синдром Саймона» в проблеме надёжности компьютерных доказательств* 23
- Н.В. Михайлова. *Методологический прагматизм проблемно-ориентированного обучения математике* 30
- А.К. Гуц. *Корреляция рефлексии и межвременных переходов* 37
- V. Kreinovich, S. Soloviev. *Can Mass Be Negative?* 43
- J. McClure, O. Kosheleva, V. Kreinovich. *The Sums of $\mathbf{m}_i \cdot \mathbf{v}_i$ and $\mathbf{m}_i \cdot \mathbf{v}_i^2$ Are Preserved, Why not Sum of $\mathbf{m}_i \cdot \mathbf{v}_i^3$: A Pedagogical Remark* .. 49

Прикладная математика и моделирование

- Л.А. Володченкова, А.К. Гуц. *Моделирование равновесной эволюции формирования лесного биоценоза на сплошных вырубках* 52
- Г.Д. Анисимова. *Об асимптотическом поведении решений задачи Коши для системы функционально-дифференциальных уравнений гиперболического типа* 57
- Ю.М. Краковский, А.Н. Лузгин. *Алгоритм расчёта блочного критерия интервальной прогнозируемости динамического показателя на основе коэффициента Тарситано-Ломбардо* 69
- Е.В. Рабинович, П.И. Вайнмастер, Г.С. Шефель. *Устранение избыточности данных сейсмического мониторинга гидроразрыва пласта* 81

Продолжение на следующей странице

Наши публикации



В.Р. Шагиев, А.М. Ахтямов. *Идентификация закрепления трубопровода с использованием минимального количества собственных частот* 95

В.А. Лефевр. *Общая схема современной психологии. Место рефлексивных исследований в системе наук* 108

В.А. Углев. *Базовая модель процессов рефлексии в интеллектуальных автоматизированных обучающих системах* 111

В.А. Филимонов. *Конструктор моделей субъектов рефлексивных игр: версия 1* 122

Компьютерные науки

Е.А. Тюменцев. *О формальном определении абстракции* 131

Е.А. Тюменцев. *Уточнение к статье «О формализации процесса разработки программного обеспечения»* 144

Т.В. Вахний, А.К. Гуц, И.Ю. Пахотин. *Определение оптимального набора средств защиты компьютерной системы методом Монте-Карло* 148

Д.Н. Лавров. *Разработка структурной модели курса подготовки к ЕГЭ по информатике в профориентационной школе ФКН* 159

КРИТЕРИЙ НОРМАЛЬНОЙ СХОДИМОСТИ ДЛЯ СИММЕТРИЧЕСКИХ ФУНКЦИЙ ОТ ЗАВИСИМЫХ ВЕЛИЧИН

А.Г. Гринь

д.ф.-м.н., профессор, e-mail: griniran@gmail.com

Омский государственный университет, Омск, Россия

Аннотация. Получены необходимые и достаточные условия для справедливости центральной предельной теоремы для симметрических функций от случайных величин, в которой масштабная нормировка осуществляется правильно меняющимися последовательностями произвольного положительного порядка. Эти условия включают в себя и так называемые минимальные условия слабой зависимости.

Ключевые слова: симметрические функции от случайных величин, центральная предельная теорема, минимальные условия слабой зависимости.

Пусть $\{\xi_n\}$ — стационарная в узком смысле последовательность. Будем писать $\xi \stackrel{d}{=} \eta$, $\xi_n \xrightarrow{d} \eta$ и $\xi_n \stackrel{d}{\sim} \eta_n$ в случаях, когда, соответственно, распределения ξ и η совпадают, $\{\xi_n\}$ сходится к η по распределению и когда последовательности $\{\xi_n\}$ и $\{\eta_n\}$ слабо эквивалентны (см., например, [1, § 28.1]). Слабая эквивалентность равносильна поточечной сходимости разности характеристических функций величин $\{\xi_n\}$ и $\{\eta_n\}$ к нулю при $n \rightarrow \infty$ [1, с. 393].

Следуя [2], назовём $\{b_n, n = 1, 2, \dots\}$ правильно меняющейся последовательностью порядка ρ , если $b_{[x]}$, $x > 0$ является правильно меняющейся функцией порядка ρ , где $[x]$ — целая часть x . Через $\hat{\xi}_1, \dots, \hat{\xi}_n$ будем обозначать *независимые* случайные величины такие, что $\hat{\xi}_k \stackrel{d}{=} \xi_k$, $k = 1, 2, \dots, n$.

Пусть при каждом $n \in \mathbb{N}$ определена симметрическая вещественнозначная функция f , то есть $f(x_1, x_2, \dots, x_n) = f(x_{i_1}, \dots, x_{i_n})$, для любых $x_1, \dots, x_n \in \mathbb{R}$ для любой перестановки $\{i_1, \dots, i_n\}$ множества $\{1, \dots, n\}$ (на самом деле определена последовательность функций, но чтобы не загромождать рассуждений, мы не будем подчёркивать зависимость f от n какими-либо индексами и называть f последовательностью).

Пусть $X_n = f(\xi_1, \dots, \xi_n)$, $\mathbf{E}X_n^2 < \infty$, $a_n = \mathbf{E}X_n$, $n = 1, 2, \dots$, $b_n^2 = \mathbf{D}X_n \rightarrow \infty$, а $\mathcal{N}(0, 1)$ — случайная величина, имеющая нормальное распределение с параметрами 0 и 1. Если

$$b_n^{-1}(X_n - a_n) \xrightarrow{d} \mathcal{N}(0, 1), \quad n \rightarrow \infty,$$

то будем говорить, что к последовательности $\{X_n\}$ применима центральная предельная теорема.

Скажем, что последовательность $\{\xi_n\}$ удовлетворяет условию (R_f) , если

$$\frac{X_{n+m}}{b_{n+m}} \underset{d}{\sim} \frac{\hat{X}_n}{b_{n+m}} + \frac{\hat{X}_m}{b_{n+m}}, \quad n+m \rightarrow \infty. \quad (R_f)$$

Если b_n является правильно меняющейся последовательностью порядка $1/2$ и $\gamma_n = b_{n+m}^{-1}(a_n + a_m - a_{n+m}) \rightarrow 0$, $n+m \rightarrow \infty$, то будем говорить, что выполнены условия нормировки (N) . В работе [3] получен следующий результат:

Теорема 1. Пусть $\{\xi_n, n = 1, 2, \dots\}$ — стационарная последовательность и пусть $\mathbf{E}X_n^2 < \infty$. Для того чтобы к последовательности $\{X_n\}$ была применима центральная предельная теорема и выполнялись условия нормировки (N) , необходимо и достаточно, чтобы выполнялось условие (R_f) и последовательность $\{b_n^{-2}(X_n - a_n)^2\}$ была равномерно интегрируема.

Замечание 1. В [3] условие (R_f) интерпретировалось как минимальное условие слабой зависимости, при котором справедлива центральная предельная теорема, и как условие, накладывающее ограничения на вид функции f , заключающиеся по сути в том, что распределения функций $f(\xi_1, \dots, \xi_n)$ слабо эквивалентны распределениям сумм некоторых независимых случайных величин.

В теореме 1 $\{b_n\}$ является правильно меняющейся последовательностью порядка $1/2$. Такая масштабная нормировка используется в предельных теоремах для сумм независимых и зависимых случайных величин, в предельных теоремах для максимумов и т.п. Однако существуют предельные теоремы с нормальным предельным распределением, в которых масштабная нормировка осуществляется последовательностями других порядков. Например, в предельных теоремах для так называемых U -статистик:

$$U_n = U_n(\xi_1, \dots, \xi_n) = \sum_{1 \leq i_1 \neq i_2 \neq \dots \neq i_k \leq n} f(x_{i_1}, \dots, x_{i_k}),$$

где $\{\xi_n\}$ — последовательность независимых одинаково распределённых величин, масштабная нормировка может осуществляться последовательностями, имеющими порядок $n^{k/2}$ (см., например, [4]). Одна из первых предельных теорем о сходимости распределений U -статистик к нормальному распределению получена В. Хёфдингом в [5].

В настоящей работе доказывается аналог теоремы 1 в случае, когда $\{b_n\}$ является правильно меняющейся последовательностью порядка $\rho > 0$.

Пусть

$$\alpha_n = \left(\frac{b_n}{b_{n+m}} \right)^{\frac{1}{2\rho} - 1}, \quad \beta_n = \left(\frac{b_m}{b_{n+m}} \right)^{\frac{1}{2\rho} - 1}.$$

Будем говорить, что последовательность $\{\xi_n\}$ удовлетворяет условию (R_f^*) , если

$$\frac{X_{n+m}}{b_{n+m}} \underset{d}{\sim} \frac{\alpha_n \hat{X}_n}{b_{n+m}} + \frac{\beta_n \hat{X}_m}{b_{n+m}}, \quad n+m \rightarrow \infty. \quad (R_f^*)$$

Если $\{b_n\}$ является правильно меняющейся последовательностью порядка $\rho > 0$ и $\gamma_n = b_{n+m}^{-1}(\alpha_n a_n + \beta_n a_m - a_{n+m}) \rightarrow 0$, $n + m \rightarrow \infty$, то будем говорить, что выполнены условия нормировки (N^*) .

Теорема 2. Пусть $\{\xi_n, n = 1, 2, \dots\}$ — стационарная последовательность и пусть $\mathbf{E}X_n^2 < \infty$. Для того чтобы к последовательности $\{X_n\}$ была применима центральная предельная теорема и выполнялись условия нормировки (N^*) , необходимо и достаточно, чтобы выполнялось условие (R_f^*) и последовательность $\{b_n^{-2}(X_n - a_n)^2\}$ была равномерно интегрируема.

Лемма 1. [6] b_n является правильно меняющейся последовательностью порядка ρ , $\rho > 0$ (а $b_n^{1/\rho}$ — правильно меняющейся последовательностью порядка 1) тогда и только тогда, когда

$$b_{n+m}^{1/\rho} \sim b_n^{1/\rho} + b_m^{1/\rho}, \quad n + m \rightarrow \infty.$$

Доказательство теоремы 2.

Необходимость. Пусть к последовательности $\{X_n\}$ применима центральная предельная теорема, то есть при любом $t \in \mathbf{R}$

$$\mathbf{E} \exp\{itb_n^{-1}(X_n - a_n)\} \rightarrow \exp\left\{-\frac{t^2}{2}\right\}, \quad n \rightarrow \infty \quad (1)$$

и выполнены условия нормировки (N^*) . Так как

$$b_n^{-2}(X_n - a_n)^2 \xrightarrow{d} \mathcal{N}^2(0, 1), \quad b_n^{-2}\mathbf{E}(X_n - a_n)^2 = 1 = \mathbf{E}\mathcal{N}^2(0, 1),$$

то равномерная интегрируемость последовательности $\{b_n^{-2}(X_n - a_n)^2\}$ следует, например, из теоремы 5.4 в [7].

Пусть $t \in \mathbf{R}$ и $m = m(n)$. Обозначим

$$\begin{aligned} \Delta(n) &= |\mathbf{E} \exp\{itb_{n+m}^{-1}X_{n+m}\} - \mathbf{E} \exp\{itb_{n+m}^{-1}\alpha_n X_n\} \mathbf{E} \exp\{itb_{n+m}^{-1}\beta_n X_m\}| = \\ &= \left| \mathbf{E} \exp\left\{it \frac{X_{n+m} - a_{n+m}}{b_{n+m}}\right\} - \mathbf{E} \exp\left\{it\alpha_n \frac{X_n - a_n}{b_{n+m}}\right\} \times \right. \\ &\quad \left. \times \mathbf{E} \exp\left\{it\beta_n \frac{X_m - a_m}{b_{n+m}}\right\} \exp\{it\gamma_n\} \right|. \end{aligned}$$

Поскольку b_n — правильно меняющаяся последовательность порядка ρ , то в силу леммы 1

$$\delta(n) = \frac{\alpha_n^2 b_n^2}{b_{n+m}^2} + \frac{\beta_n^2 b_m^2}{b_{n+m}^2} = \frac{b_n^{1/\rho} + b_m^{1/\rho}}{b_{n+m}^{1/\rho}} \rightarrow 1, \quad n + m \rightarrow \infty, \quad (2)$$

так что для любой последовательности натуральных чисел $\{n_1\}$ существуют $0 \leq c \leq 1$ и подпоследовательность $\{n_2\} \subseteq \{n_1\}$ такая, что

$$\frac{\alpha_{n_2}^2 b_{n_2}^2}{b_{n_2+m_2}^2} \rightarrow c, \quad \frac{\beta_{n_2}^2 b_{m_2}^2}{b_{n_2+m_2}^2} \rightarrow 1 - c, \quad n \rightarrow \infty,$$

где $m_2 = m(n_2)$. Если $c = 0$ ($c = 1$), то при $n \rightarrow \infty$

$$\alpha_{n_2} b_{n_2+m_2}^{-1} (X_{n_2} - a_{n_2}) \rightarrow 0 \quad (\beta_{n_2} b_{n_2+m_2}^{-1} (X_{m_2} - a_{m_2}) \rightarrow 0)$$

по вероятности, следовательно, $\Delta(n_2) \rightarrow 0$, $n \rightarrow \infty$. Если же $0 < c < 1$, то в силу (1) и (N*)

$$\begin{aligned} \Delta(n_2) &= \left| \mathbf{E} \exp \left\{ it \frac{X_{n_2+m_2} - a_{n_2+m_2}}{b_{n_2+m_2}} \right\} - \mathbf{E} \exp \left\{ it \alpha_{n_2} \frac{X_{n_2} - a_{n_2}}{b_{n_2+m_2}} \right\} \times \right. \\ &\quad \times \mathbf{E} \exp \left\{ it \beta_{n_2} \frac{X_{m_2} - a_{m_2}}{b_{n_2+m_2}} \right\} \exp \{ it \gamma_{n_2} \} \left. \right| \rightarrow \\ &\rightarrow \left| \exp \left\{ -\frac{t^2}{2} \right\} - \exp \left\{ -\frac{ct^2}{2} \right\} \exp \left\{ -\frac{(1-c)t^2}{2} \right\} \right| = 0. \end{aligned}$$

Таким образом, доказано, что из любой последовательности $\{\Delta(n_1)\}$ можно выделить сходящуюся к нулю подпоследовательность. Это означает, что $\Delta(n) \rightarrow 0$, $n \rightarrow \infty$, то есть выполнено условие (R_f^*) .

Достаточность. Пусть выполнено условие (R_f^*) и последовательность $\{b_n^{-2}(X_n - a_n)^2\}$ равномерно интегрируема. В силу известной теоремы Прохорова (см., например, [7]) последовательность $\{b_n^{-1}(X_n - a_n)\}$ является относительно компактной, так что из любой последовательности натуральных чисел можно выбрать подпоследовательность $\{n_1\}$, $n_1 = n_1(n)$ такую, что при $n \rightarrow \infty$

$$b_{n_1}^{-1}(X_{n_1} - a_{n_1}) \xrightarrow{d} \xi, \quad b_{m_1}^{-1}(X_{m_1} - a_{m_1}) \xrightarrow{d} \eta, \quad b_{n_1+m_1}^{-1}(X_{n_1+m_1} - a_{n_1+m_1}) \xrightarrow{d} \zeta, \quad (3)$$

где $m_1 = m(n_1) \rightarrow \infty$ (случай, когда $\{m_1\}$ – ограниченная последовательность, оговорим позже), а ξ, η и ζ – случайные величины. При этом, поскольку последовательность $\{b_n^{-2}(X_n - a_n)^2\}$ равномерно интегрируема, то

$$\mathbf{E} \xi^2 = \lim_{n \rightarrow \infty} \mathbf{E} b_{n_1}^{-1}(X_{n_1} - a_{n_1})^2 = 1, \quad \mathbf{E} \eta^2 = \mathbf{E} \zeta^2 = 1, \quad \mathbf{E} \xi = \mathbf{E} \eta = \mathbf{E} \zeta = 0. \quad (4)$$

Из ограниченных последовательностей

$$\mu_{n_1} = \frac{\alpha_{n_1} b_{n_1}}{\sqrt{\alpha_{n_1}^2 b_{n_1}^2 + \beta_{n_1}^2 b_{m_1}^2}}, \quad \nu_{n_1} = \frac{\beta_{n_1} b_{m_1}}{\sqrt{\alpha_{n_1}^2 b_{n_1}^2 + \beta_{n_1}^2 b_{m_1}^2}}$$

выберем подпоследовательности $\{\mu_{n_2}\}$ и $\{\nu_{n_2}\}$ такие, что

$$\mu_{n_2} \rightarrow \mu, \quad \nu_{n_2} \rightarrow \nu, \quad n \rightarrow \infty, \quad \mu^2 + \nu^2 = 1. \quad (5)$$

Тогда из (3) и (5) выводим

$$\frac{\alpha_{n_2}(\widehat{X}_{n_2} - a_{n_2}) + \beta_{n_2}(\widehat{X}_{m_2} - a_{m_2})}{\sqrt{b_{n_2}^2 + b_{m_2}^2}} = \mu_{n_2} \frac{\widehat{X}_{n_2} - a_{n_2}}{b_{n_2}} + \nu_{n_2} \frac{\widehat{X}_{m_2} - a_{m_2}}{b_{m_2}} \xrightarrow{d} \mu \widehat{\xi} + \nu \widehat{\eta}. \quad (6)$$

Понятно, что $\alpha \widehat{\xi} + \beta \widehat{\eta}$ имеет невырожденное распределение.

Далее, в силу соотношений (R_f^*) и (3),

$$b_{n_2+m_2}^{-1}(\alpha_{n_2}\hat{X}_{n_2} + \beta_{n_2}\hat{X}_{m_2} - a_{n_2+m_2}) \xrightarrow{d} \zeta, \quad n \rightarrow \infty, \quad (7)$$

где ζ имеет невырожденное распределение. По теореме о сходимости типов [1, с. 216] из (6) и (7) вытекает

$$\frac{\sqrt{b_{n_2}^2 + b_{m_2}^2}}{b_{n_2+m_2}} \rightarrow C_1, \quad \frac{\alpha_{n_2}a_{n_2} + \beta_{n_2}a_{m_2} - a_{n_2+m_2}}{b_{n_2+m_2}} \rightarrow C_2, \quad 0 < C_1, |C_2| < \infty.$$

Отсюда с помощью (6) выводим, что вместе с последовательностями $\{b_{n_2}^{-2}(X_{n_2} - a_{n_2})^2\}$ и $\{b_{m_2}^{-2}(X_{m_2} - a_{m_2})^2\}$ равномерно интегрируемыми являются последовательности

$$\{b_{n_2+m_2}^{-2}(\alpha_{n_2}(\hat{X}_{n_2} - a_{n_2}) + \beta_{n_2}(\hat{X}_{m_2} - a_{m_2}))^2\},$$

$$\{b_{n_2+m_2}^{-2}(\alpha_{n_2}\hat{X}_{n_2} + \beta_{n_2}\hat{X}_{m_2} - a_{n_2+m_2})^2\};$$

из (7) получаем теперь

$$\gamma_{n_2} = b_{n_2+m_2}^{-1}(\alpha_{n_2}a_{n_2} + \beta_{n_2}a_{m_2} - a_{n_2+m_2}) \rightarrow \mathbf{E}\zeta = 0,$$

$$\delta_{n_2} = b_{n_2+m_2}^{-2}(\alpha_{n_2}^2 b_{n_2}^2 + \beta_{n_2}^2 b_{m_2}^2) \rightarrow \mathbf{E}\zeta^2 = 1, \quad n \rightarrow \infty.$$

В силу (2) последнее соотношение означает, что

$$\delta(n_2) = \frac{b_{n_2}^{1/\rho} + b_{m_2}^{1/\rho}}{b_{n_2+m_2}^{1/\rho}} \rightarrow 1, \quad n + m \rightarrow \infty. \quad (8)$$

Таким образом, мы показали (в случае, когда $m(n) \rightarrow \infty$), что для всякой последовательности натуральных чисел найдётся подпоследовательность $\{n_2\}$ такая, что $\gamma_{n_2} \rightarrow 0$, $\delta_{n_2} \rightarrow 1$, $n \rightarrow \infty$. Это означает, что $\gamma_n \rightarrow 0$, $\delta_n \rightarrow 1$, $n \rightarrow \infty$.

Если же $m(n)$ – ограниченная последовательность, то $\{a_m\}$ и $\{b_m\}$ также ограничены, а из (8) нетрудно вывести, что $b_n \rightarrow \infty$, $n \rightarrow \infty$ и, следовательно, $\gamma_n \rightarrow 0$, $\delta_n \rightarrow 1$, $n \rightarrow \infty$.

В силу леммы 1 выполнены условия нормировки (N^*) .

Пусть теперь выполнены условия (N^*) и (R_f^*) . Тогда если $k = k(n) \rightarrow \infty$ достаточно медленно, то $b_{nk} \sim k^\rho b_n$ и из условия (R_f^*) , где

$$\alpha_n = \left(\frac{b_n}{b_{nk}}\right)^{\frac{1}{2\rho} - 1} \sim k^{\rho - \frac{1}{2}}$$

получаем

$$b_{nk}^{-1}X_{nk} \stackrel{d}{\sim} \sum_{j=1}^k k^{\rho - \frac{1}{2}} b_{nk}^{-1}Y_j \stackrel{d}{\sim} \sum_{j=1}^k k^{-\frac{1}{2}} b_n^{-1}Y_j,$$

где $Y_1, \dots, Y_k$, – независимые случайные величины, $Y_j \stackrel{d}{=} f(\xi_{(j-1)n+1}, \dots, \xi_{jn}) \stackrel{d}{=} X_n$, $j = 1, \dots, k$. Отсюда с помощью условия (N^*) получаем

$$b_{nk}^{-1}(X_{nk} - a_{nk}) \stackrel{d}{\sim} \sum_{j=1}^k \frac{Y_j - a_n}{\sqrt{k}b_n}. \quad (9)$$

Для того чтобы к последовательности серий $\left\{ \frac{Y_j - a_n}{\sqrt{k}b_n}, j = 1, \dots, k, n = 1, 2, \dots \right\}$ независимых случайных величин была применима центральная предельная теорема, достаточно, чтобы выполнялось условие Линдеберга: при любом $\varepsilon > 0$

$$L_{nk}(\varepsilon) = \frac{1}{kb_n^2} \sum_{j=1}^k \mathbf{E}\{(Y_j - a_n)^2, |Y_j - a_n| > \varepsilon \sqrt{k}b_n\} \rightarrow 0, \quad n \rightarrow \infty.$$

Используя определение величин Y_j и равномерную интегрируемость последовательности $\{b_n^{-2}(X_n - a_n)^2\}$, получаем

$$L_{nk}(\varepsilon) = \mathbf{E} \left\{ \frac{(X_n - a_n)^2}{b_n^2}, \frac{|X_n - a_n|}{b_n} > \varepsilon \sqrt{k} \right\} \rightarrow 0, \quad n \rightarrow \infty,$$

что вместе с (9) даёт $b_{nk}^{-1}(X_{nk} - a_{nk}) \xrightarrow{d} \mathcal{N}(0, 1)$, $n \rightarrow \infty$.

Представим теперь произвольное натуральное m в виде $m = kn + r$, $0 \leq r < m$, $k = [m/n]$. Из условий (R_f^*) и (N^*) выводим

$$\frac{(X_m - a_m)}{b_m} \stackrel{d}{\sim} \frac{\alpha_n(X_{nk} - a_{nk})}{b_m} + \frac{\beta_n(Z_n - a_r)}{b_m}, \quad (10)$$

где $Z_n \stackrel{d}{=} f(\xi_{km+1}, \dots, \xi_{km+r}) \stackrel{d}{=} X_r$ не зависит от X_{nk} ,

$$\alpha_n = \left(\frac{b_{nk}}{b_m} \right)^{\frac{1}{2\rho} - 1}, \quad \beta_n = \left(\frac{b_r}{b_m} \right)^{\frac{1}{2\rho} - 1}.$$

Правильно меняющаяся функция положительного порядка эквивалентна неубывающей функции [2, с. 26], так что при достаточно больших n $\max_{1 \leq r \leq n} b_r^2 b_n^{-2} < 2$, если $k = k(n) \rightarrow \infty$ растёт достаточно медленно, то что $\sigma_{kn}^2 \sim k\sigma_n^2$, а так как $m/nk \rightarrow 1$, $n \rightarrow \infty$, то $b_n \sim b_{nk}$ и $\alpha_n \rightarrow 1$, $n \rightarrow \infty$. Соотношение (10) принимает тогда вид

$$\frac{(X_m - a_m)}{b_m} \stackrel{d}{\sim} \frac{\hat{X}_{nk} - a_{nk}}{b_{nk}} + \frac{\hat{X}_r - a_r}{b_r} \cdot \frac{\beta_n b_r}{b_m},$$

где

$$\frac{\beta_n b_r}{b_m} \sim \left(\frac{b_r}{b_{nk}} \right)^{\frac{1}{2\rho}} \leq 2^{\frac{1}{2\rho}} k^{-\frac{1}{2}} \rightarrow 0, \quad n \rightarrow \infty,$$

поэтому $\frac{\hat{X}_r - a_r}{b_r} \cdot \frac{\beta_n b_r}{b_m} \rightarrow 0$ по вероятности. Из (10) следует теперь

$$b_m^{-1}(X_m - a_m) \stackrel{d}{\sim} b_{nk}^{-1}(X_{nk} - a_{nk}) \xrightarrow{d} \mathcal{N}(0, 1), \quad m \rightarrow \infty.$$

Теорема доказана.

ЛИТЕРАТУРА

1. Лоэв М. Теория вероятностей. М. : ИЛ, 1962.
2. Сенета Е. Правильно меняющиеся функции. М. : Наука, 1985.
3. Гринь А.Г. О центральной предельной теореме для симметрических функций от зависимых величин // Математические структуры и моделирование. 2017. № 1(41). С. 5–11.
4. Королюк В.С., Боровских Ю.В. Теория U -статистик. Киев : Наукова Думка, 1989.
5. Hoeffding W. A class of statistics with asymptotically normal distribution // Ann. Math. Statist. 1948. V. 19. P. 293–325.
6. Гринь А.Г. О минимальном условии слабой зависимости в предельных теоремах для стационарных последовательностей // Теория вероятн. и ее примен. 2009. Т. 54, № 2. С. 344–354.
7. Биллингсли П. Сходимость вероятностных мер. М : Наука, 1977.

THE CRITERION OF NORMAL CONVERGENCE FOR SYMMETRIC
FUNCTIONS OF THE DEPENDENT VARIABLES

A.G. Grin

Dr. Sc. (Phys.-Math.), Professor, e-mail: griniran@gmail.com

Dostoevsky Omsk State University, Omsk, Russia

Abstract. The necessary and sufficient conditions for fulfillment of the central limit theorem for symmetric functions of random variables, in which the scale normalization is carried out by regularly varying sequences of arbitrary positive order are obtained in this article. These conditions include the so-called minimal conditions of the weak dependence.

Keywords: symmetric functions of random variables, central limit theorem, minimal conditions of the weak dependence.

Дата поступления в редакцию: 30.01.2018

АНАЛИЗ В КОСМИЧЕСКИХ РАССЛОЕНИЯХ НА ОСНОВЕ ГРУПП $U(1,1)$ И $U(2)$: СЛУЧАИ $SU(2,2)$ -ДЕЙСТВИЙ В ИХ 2- И 4-НАКРЫТИЯХ

А.В. Левичев¹

профессор, д.ф.-м.н., с.н.с., e-mail: alevichev@gmail.com

А.Ю. Пальянов^{2,3}

к.ф.-м.н., с.н.с., e-mail: palyanov@iis.nsk.su

¹Институт математики СО РАН им. С.Л. Соболева, Новосибирск, Россия

²Институт систем информатики СО РАН им. А.П. Ершова, Новосибирск, Россия

³Новосибирский государственный университет, Новосибирск, Россия

Аннотация. В данной работе, являющейся непосредственным продолжением предыдущей статьи этих же авторов (А.В. Левичев, А.Ю. Пальянов, Математические структуры и моделирование, 2016, № 4(40), с. 24–38), продолжено получение тех методов, которые необходимы для анализа однородных векторных расслоений на основе параллелизующей группы $U(1,1)$. В частности, изучен вопрос совместимости (введённого Панейтцем и Сигалом) условия согласованности с $SU(2,2)$ -действиями в $U(1,1), U(2)$ и в их 2- и 4-накрытиях. В контексте $U(2)$ -параллелизации исправлены неточности доказательства справедливости (технически важной при работе с индуцированными представлениями) формулы Панейтца-Сигала (см. Corollary 4.1.2 of S.M. Paneitz and I.E. Segal, J. Funct. Anal., **47** (1982), p. 78–142.).

Ключевые слова: параллелизация расслоений над пространством-временем; конформные $SU(2,2)$ -действия в $U(1,1), U(2)$ и в их 2- и 4-накрытиях; условие согласованности и индуцированные представления.

1. Мотивация

Данное исследование продолжает то, которое составило содержание статьи [1] (поэтому мы не будем повторять приведённые на с. 24–27 этой статьи обозначения и мотивацию). Отметим, что за прошедшие (со времени публикации [1]) два года появились дополнительные доводы за продолжение такого рода исследований. Именно, в статье [2] два основных объекта работы [1] (пространства-времени $\mathbf{D}$ и $\mathbf{F}$) были применены для изучения свойств многоуровневой модели кварк-глюонной среды. Эта модель (см. [3]) вводится, исходя из цепочки канонических (т. е., отвечающих главным минорам соответствующих матриц) вложений групп: $U(2)$ в $U(3)$, $U(2)$ в $U(4)$ и т. д. Эти группы названы уровнями (материи): $U(2)$ — нулевым (т. е. нашим обычным), $U(3)$ — первым, $U(4)$ — вторым и т. д. Такая терминология соответствует стандартным

поколениям (кварков). Насколько известно авторам, многоуровневая модель является первой, в которой такие понятия теоретической физики, как аромат и цвет, введены строго математически.

Конечно, лишь время и дальнейшие исследования могут внести ясность в вопрос об адекватности (и степени успеха) этой модели. Так как она обладает высоким уровнем симметрии, то для получения (в её рамках) формул, описывающих сечения реакций, можно применить метод индуцированных представлений групп (см. [4], с. 129–131). В этом методе фундаментальную роль играют операторы переплетения. Как известно, они являются линейными отображениями между пространствами вовлечённых волновых функций (в нашем случае их носителями являются $\mathbf{D}$ и $\mathbf{F}$), т. е. между пространствами представлений. Мы уже отмечали ([1], с. 28) важность (конформного) соответствия W между $\mathbf{D}$ и $\mathbf{F}$. Затем на основе W канонически вводится линейное отображение между двумя пространствами представлений.

Напомним (см. [1], с. 26–28), что каждая из групп Ли G, G_F изоморфна $SU(2, 2)$. Группа G действует на $\mathbf{D}$, а группа G_F — в $\mathbf{F}$. Каждое из действий — дробно-линейное.

Согласно Сигалу (Irving E. Segal, 1918–1998), фундаментальным (при моделировании частиц) является глобальное дробно-линейное действие конформной группы G на группе $U(2)$. На основе этого действия получена, в частности, классификация элементарных частиц спина $1/2$. Их четыре, и они различаются порядком вхождения (своих пространств представлений) в композиционный ряд: $p < v_m < v_e < e$. Отсюда — стабильность протона p (так как ему соответствует инвариантное подпространство). Эти четыре частицы были получены в рамках определённого индуцированного представления конформной группы G . Соответствие со стандартной физикой частиц было осуществлено на основе того, что стационарная подгруппа группы G в её действии на $U(2)$ изоморфна группе Пуанкаре (с дилатациями), а мир Минковского канонически вложен в $U(2)$. Понятно, что при переходе от $\mathbf{D}$ -параллелизации к $\mathbf{F}$ -параллелизации необходимо выбрать подгруппу Пуанкаре в G_F (чему и посвящена следующая секция).

2. О реализации подгруппы Пуанкаре в G_F

Группа G_F была введена как $\{\tilde{g} = WgW : g \in G\}$: см. [1], (1.7). Она действует (дробно-линейно, но имеются сингулярности) в $\mathbf{F} = U(1, 1)$. При рассмотрении действия G_F на $\mathbf{F}$ произведением $g_2 g_1$ двух преобразований называем то дробно-линейное преобразование g , которое соответствует произведению матриц:

$$M(g) = M(g_2)M(g_1).$$

Такое преобразование g имеет, вообще говоря, более обширную область определения, нежели композиция преобразований g_2 и g_1 . Соответствующий пример приведён в Секции 4.

Напомним, что отображение W определено формулой [1], (2.1) и является биекцией пространства-времени $\mathbf{F}$ на свой образ в $\mathbf{D}$. Эта биекция задаётся 4 на

4 матрицей, обозначаемой тем же самым символом W . Обратное отображение обозначаем той же самой буквой, так как для матрицы W выполнено условие $W^2 = 1$. Из формулы для отображения W следует, что образом единичной матрицы из $\mathbf{F}$ является единичная матрица в $\mathbf{D}$. Аналогично для минус единицы (что в данном контексте более важно):

$$W(-1) = -1. \quad (2.1)$$

Из [6], теорема 2.1 следует, что стационарная подгруппа элемента -1 в $\mathbf{D}$ (при дробно-линейном действии [1], (1.9) группы $G = SU(2, 2)$ на $\mathbf{D}$) изоморфна 11-мерной группе Пуанкаре P . Так как в [6] рассматривались всевозможные группы $SU(n, n)$, то (на с. 81) эта группа была обозначена P_2 .

Из (2.1) и [1], (2.4) следует, что любое $\tilde{g}$ из G_F корректно определено в точке -1 . Поэтому стационарная подгруппа (матрицы -1 в $U(1, 1)$) группы G_F задаётся совокупностью матриц $\{\tilde{g} = WgW : g \in P_2\}$.

3. Параллелизация индуцированных расслоений

Структура данной секции аналогична таковой в [6], Секция 4, с. 98–116 (там же можно найти необходимые сведения об индуцированных расслоениях; кроме того, о *параллелизациях* подробно говорится в [1], Секция 1). При этом сейчас мы не только излагаем (новый) материал (относящийся к F -параллелизации), но и приводим те (математически необходимые) дополнения, относящиеся к D -параллелизации, которые отсутствуют в [6].

Сначала рассмотрим вопрос выполнения / невыполнения ключевого условия (фундаментальной) теоремы Панейтца-Сигала (см. [6], с. 99, Теорема 4.1). В этой теореме через ϕ обозначен гомоморфизм (некоторой) группы Ли G в группу C^∞ -гомеоморфизмов (некоторой) её подгруппы. Упомянутое условие будем называть *условием согласованности*. Оно состоит в том, что если x, y являются элементами этой подгруппы, то

$$\phi(x)y = xy, \quad (3.1)$$

т. е. результат $\phi(x)y$ действия x (как элемента группы G) на y (как элемент подгруппы) сводится к произведению xy этих элементов.

В [6] основное внимание было уделено действию $\tilde{G}$ на $\tilde{\mathbf{D}}$; здесь $\tilde{}$ означает универсальную накрывающую. Но интересен и вопрос действий конечных накрытий $G^{(k)}$ группы G на конечных накрытиях $\mathbf{D}^{(m)}$ пространства-времени $\mathbf{D}$. Несколько соответствующих случаев рассматриваются ниже.

Обозначим через N основную параллелизующую группу, применявшуюся в работах школы Сигала. Она будет первым примером (вышеупомянутой) подгруппы в $G = SU(2, 2)$. В [7] было приведено явное описание этой подгруппы: N состоит из всех таких матриц n вида

$$n = \begin{bmatrix} w & 0 \\ 0 & a \end{bmatrix}, \quad (3.2)$$

что $w \in U(2)$ и $a^2(\det w) = 1$. Другими словами, каждая матрица n составлена из 2 на 2 матриц w, a . Здесь a — скалярная матрица (причём задающее её комплексное число обозначено той же самой буквой a).

Рассмотрим N как 2-накрытие пространства-времени $\mathbf{D}$, задаваемое формулой $Z = f(n) = a^2 w$. Отметим, что это «другое» 2-накрытие для $\mathbf{D}$, нежели введённое Сигалом $\mathbf{D}^{(2)}$ (см. [1], с. 25). В частности, $\mathbf{D}^{(2)}$ не является (исходно) подгруппой в $G = SU(2, 2)$. Нетрудно проверить (сопоставив каждой n вида (3.2) пару (a, aw)), что эти два накрытия эквивалентны. Такое сопоставление групп N и $\mathbf{D}^{(2)}$ является изоморфизмом, поэтому используемое ниже обозначение $N = \mathbf{D}^{(2)}$ вполне корректно. Кроме того, из вышеизложенного следует (см. [1], с. 33), что действие ϕ группы G на N задаётся следующей коммутативной диаграммой

$$\begin{array}{ccc} n & \xrightarrow{\quad\quad\quad} & \phi(g)(n) \\ \downarrow & & \downarrow \\ Z = a^2 w & \xrightarrow{\quad\quad\quad} & g(Z) = (AZ + B)(CZ + D)^{-1} \end{array} \quad (3.3)$$

Утверждение 3.1. Если на $N = \mathbf{D}^{(2)}$ задано действие (3.3), то условие согласованности нарушается.

Доказательство. Рассмотрим формулу (3.3) для случая блочно-диагонального (с 2 на 2 блоками A, D) элемента g группы G :

$$\begin{array}{ccc} n & \xrightarrow{\quad\quad\quad} & \tilde{n} \\ \downarrow & & \downarrow \\ Z = a^2 w & \xrightarrow{\quad\quad\quad} & \tilde{Z} = a^2 A w D^{-1} \end{array} \quad (3.4)$$

С необходимостью, здесь

$$\tilde{n} = \begin{bmatrix} \tilde{w} & 0 \\ 0 & a(\det A) \end{bmatrix}, \quad (3.5)$$

где $\tilde{w} = (\det A)^{-2} A w D^{-1}$.

Теперь возьмём такой блочно-диагональный элемент g_1 группы G , что $A = w_1$, а D — скалярная матрица, задаваемая числом a_1 . Тем самым, g_1 принадлежит подгруппе $N = \mathbf{D}^{(2)}$. Напомним (см. (3.2)), что $(a_1)^2(\det w_1) = 1$. Поэтому фигурирующая в (3.5) унитарная матрица $\tilde{w}$ равна $(a_1)^3 w_1 w$. Значит, условие согласованности не выполняется.

Замечание 3.1. В [6], с. 101 подгруппа N рассматривалась как 4-накрытие для $U(2)$ (на этой с. 101 надо положить $n = 2$). Именно, слой над каждой матрицей Z состоит из следующих (четырёх) матриц:

$$n = (\det Z)^{-1/4} \begin{bmatrix} Z & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.6)$$

Один из авторов (А.В. Л.), изучая работы Сигала и его школы, довольно долго не мог понять: почему в них не рассматривается возможность 2-накрытия группы $U(2)$? По-видимому, сейчас ответ на этот вопрос получен: см. наши Утверждения 3.1 и 3.2. Предваряя Утверждение 3.2, опишем (более формально, чем это было сделано в [6]) 4-накрытие подгруппой N группы $U(2)$. Именно, накрытие теперь сопоставляет каждой матрице n вида (3.2) матрицу $Z = a^{-1}w$. Ясно, что для такой (унитарной) матрицы Z выполняется равенство $(\det Z)^4 = a^{-1}$.

Теперь (т. е. когда $N = \mathbf{D}^{(4)}$, а не $N = \mathbf{D}^{(2)}$) действие ϕ группы $G = SU(2, 2)$ на N задаётся следующей коммутативной диаграммой:

$$\begin{array}{ccc} n & \xrightarrow{\quad\quad\quad} & \phi(g)(n) \\ \downarrow & & \downarrow \\ Z = a^{-1}w & \xrightarrow{\quad\quad\quad} & g(Z) = (AZ + B)(CZ + D)^{-1} \end{array} \quad (3.7)$$

Утверждение 3.2. Если на $N = \mathbf{D}^{(4)}$ задано действие (3.7), то условие согласованности выполняется.

Доказательство. Применим диаграмму (3.7) для блочно-диагонального (с 2 на 2 блоками $A = w_1, D = a_1$) элемента g подгруппы N . Очевидно, что $g(Z) = a^{-1}w_1w(a_1)^{-1}$. Отсюда следует, что $\phi(g)(n)$ равно произведению gn , как двух элементов подгруппы N . Утверждение 3.2 доказано.

Замечание 3.2. Действие на N , удовлетворяющее условию типа (3.7), нередко называют *подъёмом действия* (с $\mathbf{D}$ на N).

Обозначим через N_F новую параллелизующую группу. Аналогично (3.2), введём явное описание этой подгруппы: N_F состоит из всех таких матриц n_f вида

$$n_f = \begin{bmatrix} Y & 0 \\ 0 & b \end{bmatrix}, \quad (3.8)$$

что $Y \in U(1, 1)$ и $b^2(\det Y) = 1$. Здесь b — скалярная матрица (причём задающее её комплексное число обозначено той же самой буквой b). Так как для каждой матрицы вида (3.8) выполнено условие (1.6) (см. [1], с. 27), то N_F является подгруппой в G_F . Если зафиксировано накрывающее отображение, переводящее n_f в принадлежащую $U(1, 1)$ матрицу $U = b^2Y$, то N_F обозначаем через $\mathbf{F}^{(2)}$. Отметим, что в [1], с. 25 было приведено альтернативное описание этого же самого 2-накрытия группы $U(1, 1)$.

Действие группы G_F в $\mathbf{F}^{(2)}$ было введено коммутативной диаграммой (A6) (см. [1], с. 34). В контексте следующего утверждения используем ϕ для обозначения этого действия.

Утверждение 3.3. Условие согласованности (3.1) для $\mathbf{F}^{(2)}$ нарушается.

Доказательство. Рассмотрим формулу (A7) из [1], с. 34 для случая блочно-диагонального (с 2 на 2 блоками A, D) элемента g группы G_F :

$$\begin{array}{ccc} n_f & \longrightarrow & \tilde{n}_f \\ \downarrow & & \downarrow \\ U = b^2 Y & \longrightarrow & \tilde{U} = b^2 A Y D^{-1} \end{array} \quad (3.9)$$

С необходимостью, здесь

$$\tilde{n}_f = \begin{bmatrix} \tilde{Y} & 0 \\ 0 & b(\det A) \end{bmatrix}, \quad (3.10)$$

где $\tilde{Y} = (\det A)^{-2} A Y D^{-1}$.

Теперь возьмём такой блочно-диагональный элемент g_1 группы G_F , что $A = Y_1$, а D — скалярная матрица, задаваемая числом b_1 . Тем самым, g_1 принадлежит подгруппе $\mathbf{F}^{(2)}$. Напомним, что $(b_1)^2(\det Y_1) = 1$. Получается, что фигурирующая в (3.10) псевдо-унитарная матрица $\tilde{Y}$ равна $(b_1)^3 Y_1 Y$. Тем самым, условие согласованности не выполняется.

Утверждение 3.4. Если задан подъём (с $\mathbf{F}$ на $N_F = \mathbf{F}^{(4)}$) действия группы G_F , то условие согласованности выполняется.

Доказательство. Теперь (т. е. когда $N_F = \mathbf{F}^{(4)}$, а не $N_F = \mathbf{F}^{(2)}$), для действия ϕ группы G_F в N_F выполняется следующая коммутативная диаграмма:

$$\begin{array}{ccc} n_f & \longrightarrow & \phi(g)(n_f) \\ \downarrow & & \downarrow \\ U = b^{-1} Y & \longrightarrow & g(U) = (AU + B)(CU + D)^{-1} \end{array} \quad (3.11)$$

Применим эту диаграмму для блочно-диагонального (с 2 на 2 блоками $A = Y_1, D = b_1$) элемента g подгруппы N_F . Очевидно, что $g(U) = b^{-1} Y_1 Y (b_1)^{-1}$. Отсюда следует, что $\phi(g)(n_f)$ равно произведению gn_f , как двух элементов подгруппы N_F . Утверждение 3.4 доказано.

Замечание 3.3. Мы вынуждены отметить, что Следствие 4.1.2 публикации [6] сформулировано нечётко (по сравнению с их же Теоремой 4.1). Именно, в Теореме 4.1 «подстилающим пространством» является подгруппа N (задаваемая соотношением (3.2) и четырежды-накрывающая $\mathbf{D}$), в то время как в основной формуле [6] (Corollary 4.1.2, с. 101) аргументами рассматриваемых волновых функций (т. е. сечений индуцированного расслоения) являются матрицы Z из $\mathbf{D} = U(2)$. Несомненно, что затруднение могло быть вызвано лишь (четвёртым) множителем $(\phi(g^{-1})(x))$ формулы (3.13), см. ниже. К сожалению, у самих авторов работы [6] спросить невозможно: Панайтц погиб в 1983 году, а Сигал скончался в 1998.

В наиболее общем варианте вышеупомянутой основной формулы (см. [6], с. 99) через $U(g)\psi$ обозначено сечение индуцированного расслоения, которое получено из произвольно взятого («исходного») сечения ψ применением оператора $U(g)$. Здесь g — произвольный элемент группы G , а через U обозначено индуцированное представление. Как мы отмечали ранее (см. [1], с. 26), аргументами x этих сечений являются элементы подгруппы N . Сама же формула позволяет вычислить значение сечения $U(g)\psi$ в точке x как

$$R(g^*)\psi(\phi(g^{-1})(x)). \quad (3.12)$$

Здесь R — индуцирующее представление, а элемент g^* (он принадлежит «индуцирующей» подгруппе $\mathbf{P}$) является следующим произведением (пяти сомножителей) в G :

$$g^* = x_0 x^{-1} g (\phi(g^{-1})(x)) x_0^{-1}. \quad (3.13)$$

В нашем контексте x_0 — это матрица -1 в N .

Недостаток изложения в этом месте статьи [6] состоит в том, что возможность действия группы G на $N = \mathbf{D}^{(4)}$ не доказана (к примеру, в нескольких случаях в статье говорится о корне четвёртой степени из определителя матрицы Z , но ведь при заданной Z таких корней четыре!). Таким образом, корректность сомножителя $(\phi(g^{-1})(x))$ в (3.13) остаётся под вопросом.

Начнём с обсуждения более общего (чем в нашем Утверждении 3.4) случая. Легко проверяется следующее

Утверждение 3.5. Если элемент

$$k = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix}, \quad (3.14)$$

максимальной компактной подгруппы в G применён к $n = (w, q)$ вида (3.2), то результатом является такой $\tilde{n} = (\tilde{w}, \tilde{q})$, что $\tilde{q} = q\bar{a}$, $a^2 = \det A$.

Замечание 3.4. Получается, что (для задания глобального действия) надо уметь однозначно выбирать квадратный корень из $\det A$. Отсюда следует вывод об отсутствии глобального подъёма действия группы G на $N = \mathbf{D}^{(4)}$. В частности, само употребление (в середине с. 100 статьи [6]) выражения $\phi(g^{-1})(x)$ при доказательстве Следствия 4.1.2 является некорректным.

Исправим положение следующим образом: воспользуемся тем, что (согласно [6], Следствие 4.1.1) можно ограничиться локальным действием группы G на

$N = \mathbf{D}^{(4)}$. Проверим правильность выражения для g^* (см. четвертую из формул на с. 101 статьи [6]; к сожалению, ни одна из (шести) формул на этой с. 101 не снабжена номером). Её достаточно проверить в двух отдельных случаях: для максимальной компактной подгруппы в G (см. выше наше Утверждение 3.5) и для одномерной подгруппы, состоящей из элементов вида

$$k = \begin{bmatrix} C & -S \\ -S & C \end{bmatrix}, \quad (3.15)$$

где $C = \cosh(t/2)$, $S = \sinh(t/2)$, скалярные два-на-два матрицы. Упомянутая достаточность обеспечена тем, что максимальная компактная подалгебра и соответствующая подгруппе (3.15) одномерная подалгебра порождают всю $su(2, 2)$. Некоторые дальнейшие свойства подгруппы (3.15) упомянуты в Секции 4.

Дополним Утверждение 3.5 следующим выбором корня квадратного из $d = \det A$. В силу локальности действия можно считать, что (содержащая 1) дуга единичной окружности между $d, \bar{d}$ меньше, чем π . Тогда в качестве a (фигурирующего в Утверждении 3.5) выбираем тот квадратный корень из d , который лежит на этой дуге. Теперь выражение $\phi(g^{-1})(x)$ корректно определено. Нетрудно (в силу диагональности всех пяти сомножителей, фигурирующих в формуле 3.13) проверить, что «наше» g^* совпадает с таковым в (четвертой) формуле с. 101 статьи [6].

В (более сложном для проверки) случае (3.15) нам надо сначала установить, что (при заданных (w, q) из N и g из G) однозначно определяется такой элемент $(\tilde{w}, \tilde{q})$, что его проекция $\tilde{Z} = \tilde{q}^{-1}\tilde{w}$ на $\mathbf{D}$ совпадает с матрицей $(CZ - S)(C - SZ)^{-1}$. Здесь матрицы (w, q) , $(\tilde{w}, \tilde{q})$ принадлежат $\mathbf{D}^{(4)}$, а $Z = q^{-1}w$. Очевидно, что равенство $(CZ - S)(C - SZ)^{-1} = \tilde{q}^{-1}\tilde{w}$ эквивалентно

$$\tilde{q}(Cw - qS) = \tilde{w}(qC - Sw). \quad (3.16)$$

Обозначим определитель искомой матрицы $\tilde{w}$ через h ; в дальнейшем же будет использовано и соотношение $h(\tilde{q})^2 = 1$. Разложим унитарную матрицу w на произведение числа q^{-1} и матрицы u из группы $SU(2)$, см. [8], формула (A4). Так как q задано, то это разложение однозначно. Введём обозначения C_q, S_q для действительной и мнимой частей числа q^2 . Из (3.16) следует, что

$$\begin{aligned} C^2(C_q - iS_q) - 2SCu_4 + S^2(C_q + iS_q) = \\ = h^2\{S^2(C_q - iS_q) - 2SCu_4 + C^2(C_q + iS_q)\} \end{aligned} \quad (3.17)$$

В связи с этим, выражение для h^2 таково:

$$h^2 = (x + iy)^2 / (x^2 + y^2), \quad (3.18)$$

где $x = C_2C_q - S_2u_4$, $y = -S_q$, а через C_2, S_2 обозначены гиперболические косинус и синус параметра t . Эти обозначения оправданы, так как t равен удвоенному исходному аргументу $t/2$. Через m обозначим положительный квадратный

корень из $(x^2 + y^2)$. Так как действие однопараметрической подгруппы $g(t)$ непрерывно, то нами доказано следующее

Утверждение 3.6. Определитель h матрицы $\tilde{w}$ равен $(x + iy)m^{-1}$.

Теперь найдём число $\tilde{q}$. Из Утверждения 3.6 следует, что $(\tilde{q})^2 = (x - iy)m^{-1}$. Обозначим через d положительный квадратный корень из m^{-1} . Тогда (при вещественном t , достаточно близком к нулю, и при q , достаточно близком к числу 1 — на единичной окружности в комплексной плоскости) «правильный» выбор квадратного корня из $\tilde{q}^2$ таков: $\tilde{q} = d(a + ib)$, где a — это положительный квадратный корень из $(x + m)/2$, а $b = S_q/2a$.

Из вышеизложенного следует

Утверждение 3.7. Результат действия элемента (3.15) на (w, q) равен такому му $(\tilde{w}, \tilde{q})$, что $\tilde{q} = d(a + ib)$, а $\tilde{w} = \tilde{q}(C\bar{q}w + S)(C + S\bar{q}w)^{-1}$.

Теперь все сомножители в (3.13) корректно определены, т. е. «наше» g^* таково:

$$g^* = \begin{bmatrix} w^*C\tilde{w} & S\tilde{q}w^* \\ \bar{q}S\tilde{w} & \bar{q}\tilde{q}C \end{bmatrix}. \quad (3.19)$$

Опять-таки, (3.19) совпадает с g^* , подсчитанным согласно формуле со с. 101 статьи [6].

Теперь сформулируем [6], Corollary 4.1.2 (для интересующего нас случая $n = 2$) аккуратнее. В приводимом ниже соответствующем утверждении через $\mathbf{P}$ обозначена стационарная подгруппа (точки -1 в $U(2) = \mathbf{D}$) группы G в её (дробно-линейном) действии на $\mathbf{D}$, а под N понимается подгруппа, задаваемая соотношением (3.2) и являющаяся 4-кратной накрывающей пространства-времени $\mathbf{D}$. Через ϕ обозначим то действие группы G в N , которое рассматривалось в наших Утверждениях (3.5–3.7).

Утверждение 3.8. Пусть (индуцирующее представление) R является конечномерным представлением подгруппы Пуанкаре $\mathbf{P}$. Тогда элемент g^* из (3.12) корректно определён и совпадает с таковым на с. 101 статьи [6] (см. четвёртую формулу на этой с. 101).

Доказательство следует из (доказанных выше) Утверждений 3.5–3.7.

Приложение. Область определения композиции двух преобразований из группы G_F может не совпадать с таковой для их произведения

Композицию преобразований g_1 и g_2 обозначаем $g_2 \circ g_1$. Под произведением g_2g_1 понимается то дробно-линейное преобразование, матрица которого равна соответствующему произведению двух матриц (этих преобразований). В данном приложении достаточно рассмотреть лишь те преобразования g пространства-времени $\mathbf{F}$, которые задаются матрицами (3.15).

Считаем, что t — не ноль (т. е. g не является тождественным преобразованием). Интересно, что $\tilde{g}$, как матрица, совпадает с g . Поэтому мы опускаем тильду (она была введена в самом начале нашей Секции 2) над g . Этот пример

важен (см. [6], с. 85), так как подалгебра изометрий (она состоит из соответствующих блочно-диагональных матриц) и генератор однопараметрической группы (3.15) порождают всю (15-мерную) алгебру Ли $su(2, 2)$ — как в $\mathbf{D}$ -, так и в $\mathbf{F}$ -случае.

Пусть преобразование g_1 задаётся некоторой матрицей (3.15) с положительным t . Пусть преобразование g_2 совпадает с g_1 . Докажем существование такой точки U (U — это матрица, принадлежащая $\mathbf{F}$), которая сингулярна для $g_2 \circ g_1$, но не является сингулярной для преобразования $g_2 g_1$.

Для любой из 2 на 2 матриц $(A, B, \dots)$, вовлечённых в приводимые ниже подсчёты, их элементы обозначаются так:

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}. \quad (4.1)$$

В [3], Proposition 1 установлено, что множество $Sing\{g_1\}$ всех сингулярностей преобразования g_1 является W -образом множества следующих матриц $Z \in \mathbf{D}$:

$$Z_1 = Z_4 = SC(1 - pq)/(c^2 - pqS^2), Z_2 = p/(C^2 - pqS^2), Z_3 = q/(C^2 - pqS^2). \quad (4.2)$$

Параметры p и q — произвольные комплексные числа модуля 1 (лишь бы их произведение не равнялось 1). Тем самым, $Sing\{g_1\}$ диффеоморфно (посредством отображения W) тору с вырезанной окружностью. Через U обозначим такую $W(Z)$, что $p = 1, q = -1$ в (4.2). Ясно, что U принадлежит множеству $Sing\{g_2 \circ g_1\}$: ведь уже составляющая g_1 не может быть применена к U !

ЛИТЕРАТУРА

1. Левичев А.В., Пальянов А.Ю. Анализ в космических расслоениях на основе группы $U(1, 1)$: основные таблицы инфинитезимального $SU(2, 2)$ -действия // Математические структуры и моделирование. 2016. № 4(40). С. 26–40.
2. Levichev A.V. On key properties of the intertwining operators ornament in the matrix multi-level model of the quark-gluon media // Материалы Всероссийской конференции с международным участием «Знания–Онтологии–Теории» (ЗОНТ-2017). Том 2. Новосибирск : Изд-во Ин-та математики, 2017. С. 41–47.
3. Levichev A.V. Towards a matrix multi-level model of quark-gluon media // Journal of Progressive Research in Mathematics. 2016. Vol. 10, Issue 2. P. 1493–1496. URL: <http://scitecresearch.com/journals/index.php/jprm/article/view/974/0> (дата обращения: 15.01.2018).
4. Менский М.Б. Метод индуцированных представлений: пространство-время и концепция частиц. Москва : Наука, 1976.
5. Levichev A.V. Pseudo-Hermitian realization of the Minkowski world through DLF theory // Physica Scripta. 2011. No. 83(1). P. 1–9. URL: <http://iopscience.iop.org/article/10.1088/0031-8949/83/01/015101> (дата обращения: 15.01.2018).

6. Paneitz S.M., Segal I.E. Analysis in space-time bundles I: General considerations and the scalar bundle // Journal of Functional Analysis. 1982. No. 47. P. 78–142. URL: <http://www.sciencedirect.com/science/article/pii/002212368290101X> (дата обращения: 15.01.2018).
7. Levichev A.V. Certain Chronometric Bundles over Compact Worlds: Triviality of Scalar and Spinor Bundles // Siberian Advances in Mathematics. 2003. No. 4. P. 1–9.
8. Levichev A., Palyanov A. On Separation between Metric Observers in Segal's Compact Cosmos // Journal of Modern Physics. 2015. No. 6. P. 2040–2049. URL: <http://dx.doi.org/10.4236/jmp.2015.614210> (дата обращения: 15.01.2018).

ANALYSIS IN SPACETIME BUNDLES BASED ON GROUPS $U(1,1)$ AND $U(2)$: THE CASE OF $SU(2,2)$ -ACTIONS IN THEIR 2- AND 4-COVERS

A.V. Levichev¹

Dr.Sc. (Phys.-Math.), Professor, Senior Researcher, e-mail: alevichev@gmail.com

A.Yu. Palyanov^{2,3}

Ph.D. (Phys.-Math.), Senior Researcher, e-mail: palyanov@iis.nsk.su

¹Sobolev Institute of Mathematics, Novosibirsk, Russia

²A.P. Ershov Institute of Informatics Systems, Novosibirsk, Russia

³Novosibirsk State University, Novosibirsk, Russia

Abstract. In this paper (which is the direct continuation of A.V.Levichev and A.Yu.Palyanov, Mathematical structures and modeling, 2016, N. 4(40), 24-38) we continue to develop mathematical tools which are necessary in order to perform analysis of homogeneous vector bundles on the basis of the parallelizing group $U(1,1)$. In particular, the compatibility of the (introduced by Paneitz and Segal) coherence condition with the $SU(2,2)$ -action on finite covers of $U(1,1)$ and of $U(2)$ was studied: such a compatibility is violated in case of 2-covers but it holds in case of 4-covers. In terms of the $U(2)$ -approach, we have corrected the proof of one (fundamental for the implementation of any induced representation) Paneitz-Segal formula given in Corollary 4.1.2 of S. M. Paneitz and I. E. Segal, J. Funct. Anal. **47**(1982), 78-142.

Keywords: parallelizations of space-time bundles, conformal $SU(2,2)$ -actions in $U(1,1)$, $U(2)$, 2- and 4-covers, coherence condition and induced representations.

Дата поступления в редакцию: 17.01.2018

«СИНДРОМ САЙМОНА» В ПРОБЛЕМЕ НАДЁЖНОСТИ КОМПЬЮТЕРНЫХ ДОКАЗАТЕЛЬСТВ

В.А. Еровенко

д.ф.-м.н., профессор, e-mail: erovenko@bsu.by

Белорусский государственный университет,
Минск, Республика Беларусь

Аннотация. Статья посвящена философско-математическим аспектам использования компьютера в формализованном математическом доказательстве. Проблема надёжности компьютерных доказательств отражается в востребованности экспериментальной математики в научном познании.

Ключевые слова: философия математических доказательств, кризис переусложненности доказательств, доверие компьютерным доказательствам.

Введение

Проследивая эволюцию становления математических теорий, обратим внимание на интересную работу начала XXI века английского математика Брайана Дэвиса с названием «Куда идёт математика?». В ней выявляются новые философско-методологические аспекты обоснования современной математики, согласно которым к концу прошлого века математика испытала потрясения, благодаря которым могут возникнуть нежелательные последствия. Так, следуя анализу Брайана Дэвиса, уже начиная с 70-х годов прошлого века, в обосновании современной математики произошли ещё два новых «методологических потрясения», которые оказались непредсказуемыми, как и «кризис», вызванный работой Курта Гёделя. «Оба они связаны с проблемой переусложненности: доказательства стали настолько длинными и сложными, что ни один учёный не взял бы на себя смелость однозначно подтвердить или оспорить их правильность» [1, р. 1351]. Второе потрясение относится к математическим доказательствам, проводимым с использованием компьютерных технологий, в правомерности использования которых сомневаются некоторые математики. Соответствующая философская проблема кратко формулируется следующим образом: можно ли считать надёжным такое доказательство, часть которого выполнена на компьютере?

По существу машинные доказательства необозримых фрагментов математических теорем — это, наряду с фундаментальными математическими объектами, открытие «новых миров», заставляющих вновь провести переоценку ценностей таких понятий, как надёжность, обозримость и строгость доказательства. В философии математики надёжность математического доказательства определяется отсутствием в нем контрпримеров, а строгость доказательства связана

с тем, что оно не содержит в себе также неявных предпосылок. Естественно, что ломка традиционных канонов методологии математики приводит к широкому спектру оценок — от новой эпохи в математике до отказа для машинных доказательств от статуса реальных достижений. «Для учёного, занимающегося компьютерными вычислениями, ключевым вопросом является вопрос о том, как быстро растёт время, необходимое для решения любой задачи с ростом её объёма» [2, с. 47]. Но до сих пор не доказано, что алгоритма решения переборных задач за полиномиальное время для классических компьютеров не существует, то есть существуют неразрешимые пока методологические проблемы, которые оказались гораздо более относительными, чем это можно было изначально предположить.

Проблема компьютерных доказательств

Существует много вопросов топологического характера, которые формулируются очень просто, но на которые интуиция не даёт удовлетворительных ответов. Таким примером может служить знаменитая «проблема четырёх красок». Задача состояла в доказательстве того, что для раскраски любой географической карты так, чтобы две страны, имеющие общую границу, как бы они ни были расположены, были окрашены по-разному, достаточно не более чем четырёх красок. Математическая интуиция ничего не говорит о том, верно ли это утверждение или нет, даже противоречащего ему примера приведено не было, и было доказано, что пяти красок достаточно для всех карт. Эта проблема оставалась нерешённой почти 125 лет. Хотя для некоторого конечного числа областей эта гипотеза Фрэнсиса Гатри была доказана, продвижение к бесконечно многим областям было безуспешным, пока эта проблема не была решена с помощью компьютера, перевернув тем самым традиционные представления о математическом доказательстве. В связи с привлечением компьютера для решения задачи о четырёх красках, предложенного американскими математиками Кеннет Аппелем и Вольфгангом Хакеном в 1977 году, возникла современная проблема «компьютерного фактора». Но такое нестандартное решение вызвало возражения некоторых ведущих математиков, считавших компьютерную проверку неадекватной, поскольку она не гарантирует безотказной работы в «недрах» компьютера и соответствующих сбоев в её логике. С одной стороны, возникает вопрос: можно ли признать необозримое логическое доказательство математическим доказательством? С другой стороны, традиционные математические доказательства тоже допускали пробелы в силу их очевидности или обоснованности в других работах.

В общем при желании все неявное и сознательно пропущенное в сложном математическом доказательстве можно было восстановить, однако использование 300 часов машинного времени в математическом доказательстве, как в приведённой выше задаче, такую возможность, вообще говоря, исключает. Другой пример подобного рода связан с решением поставленной четырёхста лет назад проблемы Кеплера о пушечных шарах, состоящей в нахождении плотнейшей упаковки бесконечной системы одинаковых непересекающихся шаров в

евклидовом трёхмерном пространстве. В своей гипотезе Кеплер предположил, что наиболее плотная упаковка шаров одинаковых размеров, то есть когда при заданном количестве шаров объём пространства между шарами минимален, даётся при их пирамидальном упорядочивании по отношению друг к другу. Последовательное изложение американским математиком Томасом Хейлзом доказательства гипотезы Кеплера заняло около 300 страниц текста, который он в 1999 году отправил для публикации. Но только в 2006 году после проверки этой статьи двенадцатью рецензентами её удалось опубликовать в *Annals of Mathematics* на 120 страницах, содержащих часть доказательства, напрямую не связанную с компьютерными вычислениями. Но машинная проверка тоже подтвердила правильность гипотезы Кеплера, поэтому уже можно утверждать, что естественная укладка шаров слоями друг на друга даёт наиболее плотную упаковку.

Однако доказательство проблемы четырёх красок и гипотезы Кеплера реально демонстрирует возможности компьютеров при проверке трудоёмких вычислительных процедур, позволяя математикам концентрироваться на концептуальных сторонах проверки. Одним из современных эффективных средств математического познания является сейчас «компьютерная математика», в которой по сути синтезируются различные направления обоснования математики. Она является новым познавательным инструментом, поэтому есть философские вопросы, анализ которых требует более глубокого проникновения в их сущность. Что философия и методология математики говорит о современном кризисе, связанном с применением компьютера в доказательстве? Философско-математическое осмысление решения с помощью компьютера проблемы четырёх красок было инициировано статьёй американского философа Томаса Тимошко, который по существу стал центральной фигурой, возникшей после этой публикации философской полемики вокруг проблемы «надёжности компьютерных доказательств». Сложившаяся ситуация напоминает притчу, согласно которой вместо длинных доказательств в «марсианской математике» утверждения стали печататься под рубрикой «Так сказал Саймон». Эту метафору можно считать подобной утверждению «доказано с помощью компьютера». Апелляция к Саймону означает, что, получив важные результаты традиционными математическими методами, он со временем перестал представлять доказательства, утверждая, что доказательство слишком длинное, чтобы его приводить, но «я его осуществил». Авторитет математика Саймона был так высок, что если математики не могли воспроизвести доказательство, полученные им результаты принимались как истинные.

Эта притча порождает проблему: можно ли относить к современной математике результаты, истинность которых основана на утверждении, что так сказал Саймон? По мнению Томаса Тимошко, логика утверждений «так сказал Саймон» и «проверено с помощью компьютера» идентичны по смыслу, поскольку компьютеры, используемые в математических доказательствах, — это тоже некие другие авторитеты. В 80-е годы прошлого века эту проблему авторитетный философ науки А.Н. Кочергин образно назвал «синдромом Саймона» [3, с. 73]. Однако есть причины, которые не позволяют считать компьютерное до-

казательство убедительным, несмотря на то, что, хотя компьютерная программа формального доказательства реализуется по законам формальной логики, в неё тоже могут вкрасться ошибки, а вера основывается на надёжности компьютера, в работе которого иногда случаются сбои и который может содержать ошибки в программном обеспечении. В могуществе и бессилии компьютера отчасти отражается «могущество и бессилие человека», поэтому и прорыв в мир суперсложной математики возможен лишь при отказе от некоторых стереотипов научных традиций, поскольку развивающиеся технологии уже стали частью нашей новой реальности. Если философски акцентировать эти проблемы, то по существу они сводятся к проблеме «доверия к компьютерам», ведь для многих существенна проверка, выполненная именно человеком, хотя главное — что в итоге сделано.

В отличие от других доказательств, математическое доказательство обладает «высокой степенью надёжности». Эта надёжность гарантируется не доказательствами отдельных математических результатов, а коллективным пониманием взаимосвязанности математической картины мира и эффективности математического знания в приложениях. Критика сосредоточена на компьютерных вычислениях, рассматриваемых в качестве доказательства. В такой ситуации общезначимым критерием является надёжность полученных результатов, и до тех пор, пока соблюдается это условие, вычисления, произведённые с помощью компьютера, будут столь же убедительны, как и сделанные человеком. Но поскольку мы не имеем перечня условий, определяющих надёжность математического доказательства, то вошедшие в математическую практику компьютерные доказательства ставят вопрос о методологической допустимости компьютерных доказательств, решать который, исходя из рациональных критериев, пока приходится в философском плане. Тем не менее, современные компьютерные доказательства реально способствовали новому пониманию подходов к обоснованию математики, которые соотносятся с вопросами об убедительности, образно говоря, традиционного «ручного» доказательства, сделанного профессиональными математиками.

В трудно обозримом математическом рассуждении, реализованном с использованием компьютера, приходится переводить относительное в абсолютное с помощью конечного и неопределённого числа проверок. Поэтому обоснование правильности компьютерных вычислений попадает под такие же методологические ограничения, что и результаты о неразрешимости некоторых важных математических проблем, причём обосновывать их тем сложнее, чем эффективней соответствующая компьютерная программа. Например, пакеты компьютерной алгебры иногда не учитывают особенности поведения функции или не обращают внимание на её гладкость, в результате чего разные пакеты по-разному дифференцируют одну и ту же кусочно заданную функцию. Однако использование «кремниевой логики» в близкой перспективе меняет практику математического доказательства. Они все же учитывают реальные практические запросы, а компьютерные доказательства обозначили принципиально новый этап осмысления роли компьютера.

Философско-математический анализ даёт возможность выявить причины,

которые не позволяют считать любое компьютерное доказательство убедительным, несмотря на веру в то, что оно соответствует идеалу формального доказательства. Во-первых, эта вера основывается на надёжности работы современного компьютера, в работе которого иногда случаются сбои и который может содержать ошибки в программном обеспечении. Для подтверждения этого заметим, что компьютерные доказательства применяются как для получения новых результатов, так и для численной проверки уже сделанных теоретических доказательств, и, хотя они эвристически полезны, компьютерные доказательства пока чаще характеризуются как «вероятные», а не «абсолютно надёжные». Во-вторых, компьютерная программа доказательства пишется в соответствии с законами формальной логики, поэтому в неё тоже могут вкрасться ошибки. Но творчески мыслящий и философски рефлексирующий математик рассматривает ошибки как закономерный этап на пути решения сложной задачи. Можно заключить, что в чистой математике с развитием компьютерных технологий появилась возможность делать, как говорит Ю.И. Манин, «физические эксперименты в ментальной реальности».

План великого немецкого математика Давида Гильберта по формальному обоснованию математических теорий состоял, во-первых, в предложении аксиоматизировать их в духе разработанной им теории доказательств, а во-вторых, в попытке доказать непротиворечивость соответствующих систем аксиом. Эта труднейшая задача, поставленная Гильбертом, так до сих пор и не решена, несмотря на то, что в период разработки направления формализма господствовало также убеждение, что математические формулы, выражающие математические факты, являются отражением абстрактных и наиболее истинных математических высказываний. Хотя академик Ю.В. Матиясевич в диспуте «Математическое доказательство: вчера, сегодня, завтра» выступал как ярый сторонник полностью формализованных доказательств, востребованных для большей убеждённости в правильности доказательств, которые могут проверить компьютеры. Уже сегодня, считает он, «компьютеры могут проверить формализованные доказательства нетривиальных теорем» [4, с. 17]. Но не каждое «серьёзное» математическое доказательство является полностью логически формализованным. Имеющиеся доказательства не проясняют ситуацию, поскольку критический этап математического доказательства требует применения компьютера, и его «ответ» подменяет собой выявление математически формализованной истинности.

Заключение

Компьютерному моделированию поддаются лишь некоторые частные процессы, поскольку при исследовании модели используются рассуждения, не носящие дедуктивного характера, хотя современная математика в большой степени разделяется на две области исследования с разными методологическими аппаратами и с разными интеллектуальными ценностями — на «численную» и «нечисленную» математику. Не случайно первым знаменитым компьютерным доказательством, которое привлекло усиленное внимание философов матема-

тики, стало нестандартное доказательство теоремы о четырёх красках. Когда математики Нейл Робертсон, Даниель Сандерс, Пол Сеймур и Робин Томас решили проверить доказательство Аппеля–Хакена, то оказалось, что проще написать свою программу, чем разбираться в чужой, а компьютерную часть доказательства они могли распечатать, чтобы её проверить. В 2008 году появилось ещё одно доказательство гипотезы о четырёх красках, когда французскому учёному Жоржу Гонтье [5] удалось в итоге проверить сведение к перебору содержательной части доказательства и ещё формальную корректность алгоритма, осуществлявшего такой перебор, используя для этого «универсальную программу», способную даже проверить доказательство любой теоремы, которая выводится с помощью используемой аксиоматики. Это отчасти снимает критику компьютерных доказательств, связанную с тем, что они не воспроизводятся на бумаге.

Под влиянием компьютерной вычислительной технологии в доказательстве многих теорем «нестрогие» математические доказательства стали встречаться в математической практике гораздо чаще, поэтому их надёжность, вообще говоря, не зависит от гипотетически возможного «чистого», но «строго аксиоматически» не реализованного математического доказательства. Проблема состоит ещё в том, что математический язык человека и язык программного обеспечения компьютера не всегда являются согласованными. В заключение уместно напомнить афористичное высказывание Платона, что «круглое невежество — не самое худшее зло, хуже накопление плохо усвоенных знаний». Если у профессиональных математиков спросить, являются ли выявленные философами математиками «кризисы» показателем нездорового состояния математического знания, то ответ, скорее всего, будет — нет, поскольку кризисы есть проявление вполне закономерной и естественной особенности сложного процесса развития и роста математического знания. А воображение человека может опережать любые технологические достижения, так как будущее математики — только «гипербола настоящего».

ЛИТЕРАТУРА

1. Davies B. Whither mathematics? // Notices of the American Mathematical Society. 2005. Vol. 52, No. 11. P. 1350–1356.
2. Ааронсон С. На что способны квантовые компьютеры? // В мире науки (Scientific American). 2008. № 6. С. 46–53.
3. Кочергин А.Н. Машинное доказательство теорем как нетрадиционная исследовательская программа в математике // Исследовательские программы в современной науке. Новосибирск : Наука, 1987. С. 70–89.
4. Матиясевич Ю.В. Математическое доказательство: вчера, сегодня, завтра // Компьютерные инструменты в образовании. 2012. № 6. С. 13–24.
5. Gonthier G. Formal proof — the four-color theorem // Notices of the American Mathematical Society. 2008. Vol. 55, No. 11. P. 1382–1393.

"SIMON'S SYNDROME" IN PROBLEM RELIABILITY OF COMPUTER PROOFS

V.A. Erovenko

Dr.Sc. (Phys.-Math.), Professor, e-mail: erovenko@bsu.by

Belarusian State University,
Minsk, Republic of Belarus

Abstract. The article is devoted to philosophical and mathematical aspects of use of the computer in the formalized mathematical proof. The problem of computer proofs reliability is reflected in a demand of experimental mathematics in scientific knowledge.

Keywords: philosophy of mathematical proofs, crisis of recomplecity of proofs, trust to computer proofs.

Дата поступления в редакцию: 17.08.2017

МЕТОДОЛОГИЧЕСКИЙ ПРАГМАТИЗМ ПРОБЛЕМНО-ОРИЕНТИРОВАННОГО ОБУЧЕНИЯ МАТЕМАТИКЕ

Н.В. Михайлова

к.ф.н., доцент, e-mail: michailova_mshrc@mail.ru

Белорусский национальный технический университет,
Минск, Республика Беларусь

Аннотация. Статья посвящена проблемно-ориентированному обучению высшей математике на основе проблемного подхода к конкретным задачам и идее методологического прагматизма, которая выражается в критике идеи абсолютного обоснования математики и образовательной аргументации.

Ключевые слова: методологический прагматизм, проблемное обучение, философия математического образования.

Введение

Проблемно-ориентированное обучение математике является по сути философско-математической реакцией на длительный процесс дифференциации в обосновании математических теорий, и его можно определить, как выявление целостных свойств системы обоснования, единство уровней интеграции при дифференциации направлений обоснования математики и явном доминировании тенденций понимания системной целостности. Методологический прагматизм в обосновании современной математики проявляется не только в признании необходимости разработки новых концептуальных подходов к обоснованию, но и во «вторичной концептуализации» работающих направлений обоснования, например, выявлении сильных сторон направлений формализма и интуиционизма. Методологический прагматизм математики имеет специфическую природу, что приводит к философскому переосмыслению таких методологически важных в философии математики понятий, как «объективность», «истинность» и «рациональность», находя в них практически вполне допустимые в гносеологии компромиссные решения. Основная идея методологического прагматизма обоснования современной математики находит своё выражение в критике идеи абсолютного обоснования и прагматической аргументации, состоящей в том, что используемые направления обоснования считаются продуктивными, если они соответствуют эпистемологически оправданному и математически обоснованному критерию.

Методологический прагматизм математики задаёт в совокупности проблемное поле современных исследований по обоснованию математики, а поскольку

проблема обоснования современной математики до сих пор не решена, то попытаемся сузить эту проблему, сосредоточившись на проблемно-ориентированном обучении. С точки зрения математического знания проблемно-ориентированное обучение математике включает в себя несколько аспектов: во-первых, доведение математических теорий до принятого современного уровня строгости; во-вторых, полноценная аргументация существования новых математических объектов; в-третьих, учёт саморазвития математических теорий при решении конкретных задач и избавление их от возможных противоречий. В контексте проблемно-ориентированного обучения заметим, что в истории становления математики надёжными представлялись математические теории, которые соответствовали различным уровням теоретической строгости, формирующимся под влиянием критической познавательной установки, направленной на практическое решение математических задач при системно-целостном анализе математических знаний.

Проблемно-ориентированное обучение математике

Интерес к проблеме взаимодействия философии, математики и образования особенно остро переживается тогда, когда возникают проблемные ситуации. Тогда обращение к философии становится важным как для анализа развития математического знания, так и для понимания связанных с этим проблем математического образования. В истории науки подвергались и продолжают подвергаться изменению ряд существенных сторон научного мышления, что должно найти также отражение и в философии математического образования. Философия современного математического образования — рефлексия над образованием, рассматривающая математическое образование с процессуальной и с содержательной стороны, способствующая развитию критического мышления как одного из проявлений рефлексии, поскольку в указанном контексте её можно рассматривать ещё и как реализацию проблемно-ориентированного обучения. Сущность проблемно-ориентированного обучения традиционно раскрывается переводом с греческого слова «проблема», что означает «задача, вопрос», — это «проблемная ситуация», образовательный эффект реализации которой не имеет однозначного решения.

Известно, что при становлении математического знания стало явно проявляться внутреннее противоречие между его непостижимо эффективной способностью получать конкретные математические результаты и философско-методологическими трудностями их обоснования. Но при преподавании высшей математики в учебную литературу стал постепенно проникать своеобразный профессиональный «пафос обоснования», что по сути отвлекало от подлинного содержания ключевого предмета в общем математическом образовании, затемняя и приглушая свойственные анализу естественнонаучные и практические мотивировки. В таком контексте основная цель проблемного обучения математике состоит не только в пробуждении интереса студентов к занятиям математикой, но и в направлении их на самостоятельные поиски истины при решении проблемно-ориентированных задач, используя для этого собственные

интеллектуальные и эмоциональные ресурсы. Как отмечает специалист в области проблемного обучения высшей математике О.В. Зимина: «В обучении, в отличие от производственной и исследовательской деятельности, проблемные ситуации приходится выявлять, т. е. делать явными и ощутимыми для учащихся, в том числе на уровне эмоций» [1, с. 67]. Дело в том, что излишнее акцентирование на абстрактно-логической форме изложения математического материала при обучении математике в техническом университете приводит к искажению когнитивных практик образования и к нарушению методологического баланса между рациональными и эмоциональными сферами проблемного обучения.

Проблемы обучения и понимания как аспекты познавательной деятельности анализировал Платон, рассматривая её в широком социально значимом контексте. Так, например, процесс познания Платоном рассматривается не как самостоятельная рациональная деятельность, а как часть системы восприятия и понимания, убеждения и обучения, поскольку, несмотря на его теорию анамнеза, для познания математических идей необходима подсказка. Подобная проблема Сократа о «знании о незнании» обсуждалась в философской литературе под названием «парадокс Менона» [2]. В знаменитом диалоге «Менон» Платон устами Сократа утверждает, что человек, никогда не учившийся математике, с помощью хорошо подобранных вопросов может сам открывать математические истины. По существу, проблемный подход к обучению математике берет своё начало со времён Сократа, но в современной педагогике практика проблемного обучения стала разрабатываться во второй половине XX столетия как активная «технология обучения». Педагогический метод Сократа по сути является способом выявления проблемной ситуации. В его основе лежит лекция-беседа, когда лектор предлагает слушателям как верные, так и неверные проблемные идеи, которые принимаются или опровергаются в ходе беседы. Математические результаты, следуя Сократу, с точки зрения проблемного обучения, следует излагать не так, как они появились и были придуманы, а так, чтобы при надлежащей подготовке их мог придумать обучаемый под руководством опытного учителя.

Но чтобы воспользоваться педагогическим методом Сократа, необходимо все же заново переосмыслить закономерности развития математического знания и его основные методологические проблемы. Современные теории математики и так реально очень сложны. В практическом смысле философию современного математического образования можно определить, как самостоятельную область философского знания, предметом которой являются закономерности функционирования и развития в аспектах образования. Нет абсолютного понятия математической строгости, так как это понятие существенно зависит от области математических исследований, например, в чистой или прикладной математике они различны, поскольку в прикладной математике рациональные рассуждения, в отличие от дедуктивных рассуждений чистой математики, имеют немалое значение, что должно учитываться при преподавании математики. С точки зрения проблемного обучения математике «стерильная» последовательность определений, аксиом и теорем может иногда отталкивать.

Проблемно-ориентированное обучение даже при решении стандартных задач содержит различные методологические этапы проблемного обучения, включая в качестве необходимой компоненты выявление смысла математических понятий. Например, комплексное число по существу имеет разные смыслы в представлениях: (a, b) , где пара действительных чисел не содержит идеи мнимой единицы; $a + ib$, но в этой сумме с мнимой единицей не просматривается идея вращения; ae^{ib} , где уже есть идея вращения, хотя все перечисленные смыслы по-разному характеризуют одну и ту же математическую структуру, а именно, поле комплексных чисел.

Хорошим примером синтеза математического знания и методологии является становление философии математического образования. В таком контексте заслуживает внимания проблемно-ориентированная позиция, рассматривающая предмет философии математического образования как синтез общепедагогических и частных научных, прежде всего педагогических, подходов к изучению образования, базирующихся на новых методических подходах к изучению знания, точнее «знания о знании», то есть когнитивных наук, в комплексном построении образовательного процесса. Философская сущность проблемно-ориентированного обучения современной математике, предполагающего общее акцентированное внимание на проблемной задачной ситуации, отражает позицию работающих математиков и может быть методологически прояснена следующим образом: «Фактически приходится говорить о практических задачах и задачах теоретических, поставленных в рамках одной и той же системы. И одна и та же практическая задача заставляет нас строить разные формальные системы. И в одной из формальных реконструкций этой практической задачи можно доказать, что ответ нельзя получить, а в другой формальной реконструкции этой же практической задачи ответ можно получить» [3, с. 52]. Поэтому следует обратить особое внимание на взаимодействие когнитивных практик, для которых в системе математического образования ведущую обучающую роль играет целеполагающий выбор проблемно-ориентированных задач при достижении образовательной цели.

Проблемно-ориентированный подход, например, в обосновании математического анализа предполагает уход от жёсткой дихотомии «внутреннее — внешнее», так как современная математика обосновывается не сама по себе, а в рамках математических теорий. При «онтодидактическом» анализе математических доказательств формальная составляющая не должна превалировать и настаивать на окончательности формальных выводов, хотя математическое вычисление по сути неразрывно связано с конкретно выбранной формализацией математики. Так, например, в общем курсе математического анализа доказывалось, что гармонический ряд формально расходится, поскольку его частичные суммы $S_n = 1 + 1/2 + 1/3 + 1/4 + \dots + 1/n$ неограниченно возрастают, так как сумма 2^k членов гармонического ряда больше чем $1 + k/2$. Но с содержательной точки зрения не обращают внимание на то, что этот ряд расходится очень и очень медленно, так как ещё Леонард Эйлер, изучавший свойства гармонического ряда, нашёл, что $S_{1000} \approx 7,84$, а $S_{1000000} \approx 14,39$. Даже с учётом компьютерных мощностей для того чтобы, например, частичная сумма гармонического ряда

превысила 100, необходимо просуммировать около 10^{43} элементов ряда. Этот пример говорит о методологической роли проблемно-ориентированного обучения в реальной практике математических вычислений. Заметим также, что при использовании компьютера надо иметь в виду, что он не только может помочь выявить, но также и скрыть рассмотренный «вычислительный феномен» гармонического ряда. В частности, речь идет о том, что при компьютерном вычислении частичных сумм гармонического ряда получается не бесконечность, а все же некоторое число, которое, строго говоря, тем больше, чем меньше «машинный ноль».

А в чем собственно состоит проблемно-ориентированная задача в образовательном контексте? Её суть в том, что, если некоторые математические утверждения являются абсолютно очевидными для профессиональных математиков, то для студентов, изучающих курс высшей математики, они уже могут представлять серьёзные трудности, особенно для некоторых студентов с изначально слабой математической подготовкой. Следует также подчеркнуть специфику даже краткого курса высшей математики, состоящего не только из определений понятий и доказательств их свойств, но и заключающегося в понимании взаимосвязей между ними. В современном курсе высшей математики наряду с вопросами исследовательского плана могут быть поставлены проблемно-ориентированные задачи текущих исследований математиков, а также методологические задачи, направленные на выявление культурного потенциала математики [4]. Можно констатировать, что математические курсы вступили сейчас в пору своей методологической зрелости, поскольку стали уже не только предметом математических исследований и методологическим инструментарием многих применений и приложений, но и фундаментальной учебной дисциплиной, по сути указывающей на востребованность проблемно-ориентированного подхода к обучению высшей математике.

Проблемное обучение, явно использующее проблемно-ориентированное обоснование современного математического анализа, показывает, что для того, чтобы образование было полноценным, оно не должно сводиться только к формальному изучению математического аппарата, например, когда сознательно не пытаются обойти порочный круг, вкрадывающийся в традиционный вывод первого замечательного предела; то есть если $x \rightarrow 0$, то отношение $\sin(x)/x$ стремится к 1, традиционное доказательство которого начинается с установления при $0 < |x| < \pi/2$ неравенства $\sin|x| < |x| < \operatorname{tg}|x|$. Стандартное обоснование этих исходных неравенств основывается на интерпретации величины $|x|/2$ как площади сектора единичного круга, опирающегося на дугу длины $|x|$. «Данное обоснование (проводимое в рамках начальной части курса математического анализа) содержит классическую ошибку «порочного круга» ("circularis vitiosus"), которую многие видят, но предпочитают не замечать» [5, с. 34]. Заметим, что, в математическом анализе, в отличие от геометрии, тригонометрические функции рассматриваются как функции числового аргумента, а именно, в виде радианной меры угла, поэтому в выражении $\sin(x)/x$ речь идёт о синусе числа x как синусе угла в x радиан, где радиан, по определению, как единица измерения — это центральный угол, отсекающий дугу равную радиусу

окружности. Проблема состоит в следующем: чтобы посчитать площадь кругового сектора, надо воспользоваться тем же самым предельным соотношением, которое доказывается, а это и есть «порочный круг».

Заключение

Можно ли говорить о возможности синтеза разных методик проблемного обучения и принятой традиционной дидактической системы математического образования? Можно и необходимо, если проблемно-ориентированная система обучения математике сочетается с организацией творческой деятельности на всех видах лекционных и практических занятий. Для этого необходима интеграция или синтез методологических, методических и дидактических принципов проблемного обучения математике. Философская рефлексия над идеями и методами проблемно-ориентированного обучения математике позволяет считать, что такой синтез возможен при выполнении ряда условий. Хотя цели математического образования в философском контексте «математической образованности» подвижны во времени, основной целью математического образования в контексте методологического прагматизма должно стать воспитание умения математически исследовать практические задачи. Проблемная ситуация в проблемно-ориентированном обучении создаётся путём формулирования теоретических утверждений в виде задач, для решения которых необходим методологический прагматизм.

При проблемно-ориентированном обучении необходимо учитывать также то, что «нетривиальное решение» можно найти, чаще всего, «проходом через ошибку», поскольку выполнить все ограничения при реализации концепции обоснования порою возможно, лишь выходя за их пределы, компенсируя этим сделанные ошибки. Вспоминая знаменитые доклады Анри Пуанкаре и Давида Гильберта на первом и втором Международных математических конгрессах, в математике XX века можно, следуя им, выделить следующие две основные линии: «линию Пуанкаре», связанную с развитием математики в естественнонаучной области, и «линию Гильберта», ориентированную на внутреннее развитие математики. Но генезис развития математики позволяет сейчас говорить ещё и о третьей линии, которую можно назвать «проблемно-ориентированной». Поэтому можно заключить, что методологический прагматизм проблемно-ориентированного обучения математике имеет специфическую природу, что приводит к философскому переосмыслению таких методологически важных в философии современной математики понятий, как «объективность», «истинность», «рациональность», находя в них практически вполне допустимые в гносеологии компромиссные решения.

ЛИТЕРАТУРА

1. Зими́на О.В. Проблемное обучение высшей математике в технических вузах // Математика в высшем образовании. 2006. № 4. С. 55–78.

2. Михайлова Н.В. Методологический парадокс Сократа в философии математического образования // Педагогика. 2016. № 4. С. 12–18.
3. Проблемно-ориентированный подход к науке: философия математики как концептуальный прагматизм. Новосибирск : Наука, 2001. 154 с.
4. Михайлова Н.В. Культурный потенциал математики в контексте философии математического образования // Вестник Воронежского государственного университета. Серия: Философия. 2016. № 1. С. 55–58.
5. Шведенко С.В. Две заметки по математическому анализу // Математическое образование. 2011. № 3–4. С. 34–37.

METHODOLOGICAL PRAGMATISM OF THE PROBLEM-ORIENTED TRAINING IN MATHEMATICS

N.V. Michailova

Ph.D. in Philosophy, Associate Professor, e-mail: michailova_mshrc@mail.ru

Belarusian National Technical University,
Minsk, Republic of Belarus

Abstract. Article considers problem-oriented training in the higher mathematics, problematic approach to specific objectives and the idea of methodological pragmatism. The methodological pragmatism is expressed in criticism of the idea of absolute justification of mathematics and the educational argument.

Keywords: methodological pragmatism, problem training, philosophy of mathematical education.

Дата поступления в редакцию: 25.08.2017

КОРРЕЛЯЦИЯ РЕФЛЕКСИИ И МЕЖВРЕМЕННЫХ ПЕРЕХОДОВ

А.К. Гуц

д.ф.-м.н., профессор, зав. каф. кибернетики, e-mail: aguts@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Исследуются этические проблемы, связанные с работой квантовой машины времени.

Ключевые слова: квантовая машина времени, этические проблемы, межвременные переходы.

Введение

Как известно, серьёзный разговор о машине времени, переносящей людей из одной исторической эпохи в другую, из одного времени в другое, стал возможным благодаря открытию Гёделем замкнутых времениподобных кривых в общей теории относительности [1].

Проблемы, связанные с созданием машины времени, являются не только техническими, но и этическими. В статье мы покажем, что этические проблемы — это не только парадокс дедушки, но и те, что связаны с человеческой рефлексией¹ как прошлых исторических эпох, так и будущих.

1. Межвременные переходы между историческими эпохами

В рамках описания Реальности как пространства-времени M^4 мы, в случае наличия глобального времени t , можем говорить о пространственных сечениях M_t^3 пространства-времени, каждое из которых отвечает конкретному значению времени $t = t$. Каждое M_t^3 — это то, что можно назвать исторической эпохой. У историков, правда, историческая эпоха — это непрерывное семейство $\{M_t^3 : t_n < t < t_{n+1}\}$, в котором неизменными остаются общественные и производственные отношения. Но для нас будет сейчас достаточным метить историческую эпоху одним сечением M_t^3 .

Нас интересуют переходы между историческими эпохами, скажем, переход из $M_{t_0}^3$ в $M_{t'}^3$. Такие переходы должны осуществляться с помощью специального технического устройства $\mathcal{A}$, называемого *машиной времени*.

¹Рефлексия в психологии — это обращение субъекта на самого себя, на своё сознание, на продукты собственной активности или какое-либо переосмысление. Рефлексия в философии — форма умственной деятельности человека, направленная на осмысление своих действий, всей человеческой культуры и её основ.

Мы не будем обращаться к теории машины времени, основанной на классической физической теории, восходящей к общей теории относительности [1]. Воспользуемся теорией квантовой машины времени [2, 3], основанной на квантовой геометродинамике Уилера-ДеВитта [4].

1.1. Квантовое описание пространства-времени

Пространство-время Реальности M^4 в квантовой геометродинамике Уилера-ДеВитта появляется как интерференция когерентной квантовой суперпозиции, или волнового пакета:

$$\Psi^{(4)}[\mathcal{G}] = \int_K c_k \Psi_k^{(3)}[\mathcal{G}] dk, \quad c_i \in \mathbb{C}, \quad (1)$$

где $\Psi_k^{(3)}[\mathcal{G}]$ — частная волновая функция, являющаяся функционалом от 3-мерной римановой геометрии ${}^{(3)}\mathcal{G} = (M^3, h_{\alpha\beta})$ и удовлетворяющая функциональному уравнению Уилера-ДеВитта. Здесь M^3 — это 3-многообразие, моделирующее реальное физическое пространство. На нём и рассматриваются всевозможные 3-геометрии ${}^{(3)}\mathcal{G}$.

Мы видим, что то, что считается Реальностью, существующей в *форме* четырёхмерного непрерывного континуума M^4 и называемой пространством-временем, в действительности является квантовой сущностью, т. е. цепью интерференционных «горных пиков» по выражению Halliwell'a [5] в суперпространстве Уилера. Вдоль этой цепи «горных пиков» вводится искусственно расстояние между ними, воспринимаемое людьми как *физическое время* t . Поэтому имеем семейство 3-геометрий ${}^{(3)}\mathcal{G}(t)$, или 3-метрику $h_{\alpha\beta}(x, t)$, удовлетворяющих уравнениям Эйнштейна [2, 4].

Отметим, что наша историческая эпоха $M_t^3 = (M^3, h_{\alpha\beta}(x, t))$ — это всего лишь «пик» с номером t в цепи интерференционных «горных пиков».

Как правило, не обсуждается смысл квантовой системы Ω , описываемой посредством волнового пакета (1), и её состояний Ω_k , $k \in K$, для которых находятся соответствующие волновые функции $\Psi_k^{(3)}[\mathcal{G}]$.

Очевидно, что Ω — это Внешний мир, Квантовая Реальность, а её состояния Ω_k — это формы её существования, которые в соответствии с принципами квантовой механики в процессе, именуемом в квантовой механике *измерением*, материализуются (локализуются).

1.2. Исторические эпохи и квантовая машина времени

Примем, что каждое состояние Ω_k квантовой реальности Ω — это 3-мерный мир, в котором практически ничего не меняется; он вневременен. В этом мире находится *наблюдатель*, способный осуществлять измерения реальности, точнее, её пространственной геометрии. Иначе говоря, Ω_k — это стационарное пространство-время $M_k^4 = M_k^3 \times \mathbb{R}$, в котором осуществляется «замороженное» историческое существование наблюдател(я/ей) в конкретном 3-пространстве

$M_{t(k)}^3$. Историки такое существование называют историческими эпохами. Гёте использовал термин «гештальт».

Формула (1) говорит, что эволюционирующая Реальность появляется как интерференция исторических эпох, как последовательность «горных пиков», в высоту каждого из которых вносит вклад каждая историческая эпоха. Исторические эпохи *существуют одновременно*. Важное понятие квантовой теории — наблюдатели. У нас наблюдатели находятся не вне этой суперпозиции, а внутри — внутри каждой исторической эпохи, для которых собственная историческая эпоха представляется истинной объективной реальностью. Находясь в эпохе Ω_{k_0} , наблюдатель X с помощью аппаратуры $\mathcal{A}$ (= машины времени) проводит наблюдение (измерение) геометрии эпохи $\Omega_{k'}$ ($t(k') < t(k_0)$), и пакет (1) коллапсирует. В результате пространственно ограниченная часть 3-пространства эпохи (состояния) $\Omega_{k'}$ материализуется (локализуется) в эпохе Ω_{k_0} :

$$\int_K c_k \Psi_k^{(3)\mathcal{G}} d\mu(k) \xrightarrow[\substack{\uparrow \\ X \\ \text{измерение}}]{\substack{\text{наблюдатель} \\ X}} \Psi_{k'}^{(3)\mathcal{G}} \subset \Omega_{k_0} \quad (2)$$

с вероятностью $|c_{k'}|^2$. Появляется «вход» в эпоху $\Omega_{k'}$. Симметрично, ограниченная часть пространства эпохи Ω_{k_0} материализуется (локализуется) в 3-пространстве эпохи $\Omega_{k'}$. Иначе говоря, аппаратура $\mathcal{A}$ (машина времени) материализуется в прошлом.

В отличие от классической теории в квантовой теории «до момента взаимодействия квантового объекта с измерительным прибором, как правило, измеряемой величины просто не существует» (Белинский). Измерение материализует (локализует) то, что измеряет.

Определённого рода деятельность (измерения) наблюдателей эпохи Ω_{k_0} относительно величин, типичных только для эпохи $\Omega_{k'}$, разрушит сцепленность последней с остальными эпохами и сцепит её с данной. Это измерение локализует частично каждую из двух эпох в другой. **Возникает переход из одной эпохи в другую.** Это есть не что иное, как *квантовая машина времени* [3].

2. Квантовая машины времени и рефлексия

В случае квантовой машины времени путешественник во времени — наблюдатель в терминах квантовой теории — пытается вернуться к событиям прошлого Реальности. Результат его измерения — это *локализация* (материализация) в эпохе наблюдателя X части пространства прошлой эпохи.

Следовательно, наш наблюдатель должен иметь образ прошлой эпохи, точнее, *иметь о ней корректное представление*, корректный образ прошлой эпохи. Иначе невозможно точно настроить аппаратуру $\mathcal{A}$ на переход в нужную историческую эпоху.

Однако, как показано в [1,6], принципиально невозможно иметь корректный образ прошлой эпохи. Но даже если корректный образ имелся бы, в силу вероятностного характера результата коллапса волнового пакета, трудно ожидать, что ожидания совпадут с тем, что обнаружится.

Рефлексирующие наблюдатели

Наш наблюдатель не есть просто измерительная аппаратура, он — аппаратура, наделённая осознанием. Человека отличает от аппаратуры наличие эмоций. Он является *рефлексирующим субъектом*. Попробуем описать действия и его эмоции в работе с квантовой машиной времени. Воспользуемся теорией рефлексирующих индивидов В.А. Лефевра [7].

Наш наблюдатель — это абстрактный индивид, у него есть внутренние образы своего взаимодействия с другими индивидами, которые способны оценивать себя, других и ситуации, в которые они вовлечены. Индивид обладает рекурсивной структурой: образы индивидов тоже могут иметь образы ситуации, себя и других и тоже способны проводить их оценку.

Будем для удобства отождествлять наблюдателя из эпохи Ω_k с самой эпохой. Тем самым наш наблюдатель олицетворяет общественное сознание, знание, умение своей исторической эпохи.

Начнём присваивать эпохам булевы значения 0 или 1.

Пишем $\Omega_k = 0$, если в этой эпохе наблюдатель не обладает аппаратурой, способной осуществлять коллапс волнового пакета (1), и $\Omega_k = 1$ — если такая аппаратура имеется.

Таким образом, множество эпох $\{\Omega_k\}_{k \in K}$ — это булевы переменные на булевой алгебре с двумя фиксированными элементами 0 и 1, и булевыми операциями $a + b$, $a \bullet b$, $\bar{a}$, $a^b = \bar{b} + a$ [7].

Рефлексирующий наблюдатель Лефевра

$$X = (\Omega_{k_0})^{A*B}, \quad (3)$$

где A — это образ своей эпохи Ω_{k_0} у X ,

B — это образ эпохи $\Omega_{k'}$ у X ,

$*$ — это образ их отношений у X (операция $+$ или $\bullet$),

— это наблюдатель X из эпохи Ω_{k_0} , о котором говорилось в § 1.2.

Основание степени в (3) называется *корнем*.

Поскольку наблюдатель X оперирует не с реальными эпохами, а с их образами A, B , и для него они сама реальность, то он создаёт их образы (модели):

$$A = (\Omega_{k_0})^{\hat{a}\hat{b}}, \quad B = (\Omega_{k'})^{\hat{a}\hat{b}}. \quad (4)$$

Модель A *корректна*, если $\hat{a} = \Omega_{k_0}$, т. е. корень $\hat{a}$ совпадает с корнем A , и *некорректна*, если $\hat{a} = \overline{\Omega_{k_0}}$.

Модель B *корректна*, если $\hat{b} = \Omega_{k'}$, и *некорректна*, если $\hat{b} = \overline{\Omega_{k'}}$.

Иначе говоря, модели своей эпохи — A и другой — B корректны, если образ образа правилен: наблюдатель правильно описывает и свою реальность, и другую реальность, образы которых сам же и создал.

3. Корреляция между рефлексией и межвременными переходами

Итак, наш наблюдатель — это абстрактный индивид Лефевра

$$X = (\Omega_{k_0})^{(\Omega_{k_0})^{\hat{a}} \bullet (\Omega_{k'})^{\hat{b}}}, \quad (5)$$

где $\hat{a} = a = \Omega_{k_0}$ или $\hat{a} = \overline{\Omega_{k_0}}$ и $\hat{b} = b = \Omega_{k'}$ или $\hat{b} = \overline{\Omega_{k'}}$.

Пусть a принимает значение 0 с частотой x , и $b = 0$ — с частотой y . Условимся частоту, с которой булева формула M принимает значение 0, обозначать $\|M\|$.

Теорема 1. (Лефевр, [7, с. 395]).

Для наблюдателя-индивида (5): если модели A, B некорректны, то

$$\|A \bullet B\| = \|A\| + \|B\| - \|A + B\| \text{ и } \|A + B\| > 0. \quad (6)$$

В остальных случаях,

$$\|A \bullet B\| = \|A\| + \|B\|. \quad (7)$$

■

Формула (6) может рассматриваться как аналог формулы, говорящей о квантовой интерференции двух исторических эпох Ω_{k_0} и $\Omega_{k'}$:

$$\begin{aligned} & |\Psi(\Omega_{k_0}) + \Psi(\Omega_{k'})|^2 = \\ & = |\Psi(\Omega_{k_0})|^2 + |\Psi(\Omega_{k'})|^2 + 2|\Psi(\Omega_{k_0})||\Psi(\Omega_{k'})|\cos(\theta_0 - \theta'), \end{aligned}$$

а формула (7) — аналог отсутствия интерференции.

Лефевр предлагает следующие интерпретации [7, с. 392]:

$\|A\|$ — это частота импульсов «чувства вины» у наблюдателя X перед наблюдателем из другой эпохи;

$\|B\|$ — это частота импульсов «чувства осуждения» наблюдателя другой эпохи, имеющееся у наблюдателя X ;

$\|A \bullet B\|$ — частота «импульсов страдания».

Наблюдатели — это сознающие люди с чувствами. Чувство вины у жителя эпохи Ω_{k_0} перед жителями прошлых эпох? С чего бы это? Хотя можно вспомнить русского философа Николая Федорова, который писал о сыновьем долге по воскрешению предков. Вина за невыполняемый долг?

Осуждение? Если $t(k') > t(k_0)$ то, конечно, хотелось, чтобы они, люди из будущего, к нам прилетели с подарками. Не летят. Вот мы их за это и осуждаем. Труднее понять, за что можно осуждать людей прошлого, т. е. в случае $t(k') < t(k_0)$? Впрочем, у историков масса претензий к деятелям прошлого...

Если наблюдатель в эпохе Ω_{k_0} имеет некорректные образы своих образов собственной эпохи и той, куда он хотел бы попасть, то наблюдается интерференция «переживаний».

Если в обеих эпохах (во второй — по мысли наблюдателя первой эпохи) нет аппаратуры для коллапса пакета (1), то имеем интерференцию всех эпох, и наблюдатель пребывает в классической эволюционирующей Вселенной. При этом $a = 0$ и $b = 0$. **Значит, $A = B = 0$, и обе модели некорректны** (Лефевр, [7, с. 398]). Следовательно, имеет место интерференция переживаний.

Космологическая система коррелирует с этической

Построить корректный образ прошлой эпохи (или людей из прошлого) [6] или даже собственной крайне сложно. Поэтому у нас нет квантовой машины времени. Вот и страдаем :-).

ЛИТЕРАТУРА

1. Гуц А.К. Элементы теории времени. 2-е изд., доп. М. : Издательство ЛКИ, 2011. 376 с.
2. Гуц А.К. Временные эффекты коллапса волнового пакета в суперпространстве Уилера // Международный научный семинар «Нелинейные модели в механике, статистике, теории поля и космологии» GRACOS-16. Лекции школы и материалы семинара (5–7 ноября 2016 г., Казань). Казань : Казанский (Приволжский) федеральный университет, 2016. С. 273–280.
3. Гуц А.К. Не-гёделевская машина времени // Математические структуры и моделирование. 2016. № 3(39). С. 48–58.
4. Гуц А.К. Физика реальности. Омск : Изд-во КАН, 2012. 424 с.
5. Halliwell J.J. Introductory lectures on quantum cosmology // Quantum cosmology and baby universes / Eds. Coleman S., Hartle B., Piian T., Weinberg S. World Scientific Publishing Co. Pte. Ltd., 1991. P. 159–244.
6. Гуц А.К. Многовариантная история России. М. : АСТ / СПб. : Полигон, 2000. 381 с.
7. Лефевр В.А. Алгебра совести. М.: Когито-Центр, 2003. 412 с.

CORRELATION OF REFLECTION AND INTERTEMPORAL TRANSITIONS

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: aguts@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. Ethical problems associated with the work of a quantum time machine are investigated.

Keywords: quantum time machine, ethical problems, intertemporal transitions.

Дата поступления в редакцию: 06.02.2018

CAN MASS BE NEGATIVE?

Vladik Kreinovich¹

Ph.D. (Phys.-Math.), Professor, e-mail: vladik@utep.edu

Sergei Soloviev²

Ph.D. (Phys.-Math.), Professor, e-mail: sergei.soloviev@irit.fr

¹University of Texas at El Paso, El Paso, Texas 79968, USA

²Institut de Recherche en Informatique de Toulouse (IRIT), 31062 Toulouse, France

Abstract. Overcoming the force of gravity is an important part of space travel and a significant obstacle preventing many seemingly reasonable space travel schemes to become practical. Science fiction writers like to imagine materials that may help to make space travel easier. Negative mass – supposedly causing anti-gravity – is one of the popular ideas in this regard. But can mass be negative? In this paper, we show that negative masses are not possible – their existence would enable us to create energy out of nothing, which contradicts to the energy conservation law.

Keywords: negative mass, equivalence principle, anti-gravity, energy conservation law.

1. Formulation of the Problem

Overcoming the force of gravity is an important part of space travel and a significant obstacle preventing many seemingly reasonable space travel schemes to become practical. Science fiction writers like to imagine materials that may help to make space travel easier. Negative mass – supposedly causing anti-gravity – is one of the popular ideas in this regard. But can mass be negative?

2. About Our Approach

The considerations below are based on existing fundamental concepts and well known experiences. For example, we assume:

- that the masses add up when we combine two bodies,
- that forces that act on a solid body add up,
- that it is always possible to rigidly bind two bodies together,
- that for every body, one can construct an anti-body with the same mass,

etc. So, strictly speaking, we consider the possibility of a negative mass in the context of these concepts and experiences. Our conclusion is that the negative mass is incompatible with these concepts and experiences – and incompatible on a very basic level.

Comment. It is worth mentioning that while we take Newton's formula for the gravitational force as given, this formula itself can be derived from the fundamental assumptions – like additivity of masses and forces; see, e.g., [3].

3. Reminder: There Are Different Types of Masses

To properly answer the question of whether negative masses are possible, it is important to take into account that there are, in principle, three types of masses:

- inertial mass m_I that describes how an object reacts to a force F : the object's acceleration a is determined by Newton's law $m_I \cdot a = F$; and
- active and passive gravitational masses m_A and m_P : gravitation force exerted by Object 1 with active mass m_{A1} on Object 2 with passive mass m_{P2} is equal to $F = G \cdot \frac{m_{A1} \cdot m_{P2}}{r^2}$, where r is the distance between the two objects; see, e.g., [1, 4].

Since $F = m_I \cdot a$, the formula for the gravitational force can be rewritten as

$$m_{I2} \cdot a_2 = G \cdot \frac{m_{A1} \cdot m_{P2}}{r^2}. \quad (1)$$

Can any of these masses be negative?

4. All Three Masses Are Proportional to Each Other

General idea. To answer the above question, let us recall that, due to energy conservation and the properties of anti-particles, all three masses are proportional to each other; see, e.g., [2]. For completeness – and to make sure that the corresponding arguments are applicable to negative masses as well – let us recall the corresponding arguments.

Active and passive masses are proportional to each other: case of positive masses. Let us first show that the active and passive masses are always proportional to each other, i.e., that

$$\frac{m_{A1}}{m_{P1}} = \frac{m_{A2}}{m_{P2}}$$

for every two objects. We will first show it for bodies of positive mass.

Indeed, suppose that for some pair of bodies, this is not true, i.e.,

$$\frac{m_{A1}}{m_{P1}} \neq \frac{m_{A2}}{m_{P2}}.$$

Multiplying both sides of this inequality by both passive masses, we conclude that $m_{A1} \cdot m_{P2} \neq m_{A2} \cdot m_{P1}$. Thus, the gravitational force exerted by Object 1 on Object 2 is different from the gravitational force exerted by Object 2 on Object 1.

So, if we combine these two objects by a rigid rod, the overall force acting on the resulting 2-object system would be different from 0. Thus, if this system was originally immobile, it will start moving with a constant acceleration. We can

then stop this system, use the gained kinetic energy to perform some work, and thus, get back to the original configuration – with some work done. We can repeat this procedure as many times as we want. This way, without spending anything, we can get as much work done as we want (and/or as much energy stored as we want). This possibility to get energy from nothing, without changing anything, clearly contradicts to energy conservation law, according to which such perpetuum mobile is impossible.

This contradiction shows that for positive masses, active and passive masses should be proportional to each other: $m_A = c \cdot m_P$ for some constant c . If we plug in this expression for the active mass into the formula (1) for gravitational force, we conclude that

$$m_{I2} \cdot a_2 = G \cdot c \cdot \frac{m_{P1} \cdot m_{P2}}{r^2},$$

i.e.,

$$m_{I2} \cdot a_2 = G' \cdot \frac{m_{P1} \cdot m_{P2}}{r^2},$$

where we denoted $G' \stackrel{\text{def}}{=} G \cdot c$. Thus, we get the formula similar to the formula (1), but with the passive and active masses equal to each other.

So, we can conclude that when masses are positive, active and passive masses are equal: $m_A = m_P$.

Active and passive masses are proportional to each other: general case.

What if at least one of the masses – either active or passive – is negative? In this case, the argument about gaining energy does not necessarily apply: e.g., when the product $m_P \cdot m_A$ is negative, the 2-object system does not gain energy, only loses it.

A slight modification of this thought experiment, however, enables us to gain energy. Indeed, let us consider an object with different active and passive masses $m_A \neq m_P$. Instead of considering this object on its own as before, let us attach it to another object with a big positive mass $M_A = M_P > \max(|m_A|, |m_P|)$. This combination has active mass $C_A = m_A + M_A$ and passive mass $C_P = m_P + M_P$. Since $M_A = M_P > \max(|m_A|, |m_P|)$, both these combined masses are positive. Since $M_A = M_P$ and $m_A \neq m_P$, we conclude that $C_A \neq C_P$. So, the active and passive masses of the combined object are positive and different – and we already know that this leads to a contradiction with the energy conservation law. So, for negative masses, active and passive gravitational masses are also always equal.

Since the active gravitational mass is always equal to the passive gravitational mass, in the following text, we will simply talk about gravitational mass m_G .

Gravitational and inertial masses are proportional to each other: case of positive masses. The important property that we will use is that any type of matter, when combined with the corresponding antimatter, can annihilate, i.e., get transformed into photons, and these photons can get transformed into some other types of matter. For example, we can start with iron and anti-iron, annihilate them, and then get gold and anti-gold. We will also take into account that experiments seems to confirm that matter and corresponding anti-matter have the same

inertial and gravitational properties, in particular, the same value of the inertial and gravitational mass; see, e.g., [4].

We want to prove that for all materials, the ratio $\frac{m_G}{m_I}$ of gravitational and inertial masses is the same. Indeed, let us assume that there exist two materials for which this ratio is different, i.e., for which $\frac{m_{G1}}{m_{I1}} \neq \frac{m_{G2}}{m_{I2}}$. Without losing generality, we can assume that the ratio is smaller for the first material: $\frac{m_{G1}}{m_{I1}} < \frac{m_{G2}}{m_{I2}}$. This means that if we select two objects of the same inertial mass $m_{I1} = m_{I2}$ from the first material and from the second material, then the gravitational mass of the first object is smaller: $m_{G1} < m_{G2}$.

We can then get the following scheme for getting energy out of nothing. We place a body and an identical anti-body of the first material at some distance r from the gravitational attractor of some mass M – e.g., from the Earth. We then move both bodies a small distance h away from the Earth. The corresponding force is $F = G \cdot \frac{2m_{G1} \cdot M}{r^2}$, thus the energy that we need to spend for this move is equal to

$$F \cdot h = G \cdot h \cdot \frac{2m_{G1} \cdot M}{r^2}.$$

Once we reached the distance $r + h$, we annihilate both objects, and use the resulting photons to create a pair of a body and anti-body of material 2. Then, we move the new object back to the distance r . This way, the force is equal to $F = G \cdot \frac{2m_{G2} \cdot M}{r^2}$, thus the energy that we gain is equal to

$$F \cdot h = G \cdot h \cdot \frac{2m_{G2} \cdot M}{r^2}.$$

At the distance r , we annihilate both objects, and use the resulting photons to create the original pair of the body and anti-body of Material 1.

Now, we are back to the original state, but, since $m_{G2} > m_{G1}$, we gained more energy that we spent – i.e., as a result, we get energy out of nothing. The impossibility of such a perpetuum mobile shows that, at least for positive masses, gravitational and inertial masses should be proportional to each other: $m_G = \text{const} \cdot m_I$. Thus, if we select the same unit for measuring both gravitational and inertial masses, we can conclude that when masses are positive, gravitational and inertial masses are equal $m_G = m_I$.

Gravitational and inertial masses are proportional to each other: general case. What if at least one of the masses – either gravitational or inertial – is negative? In this case, the above argument does not necessarily apply: e.g., if the inertial mass of some material is negative, we cannot transform it into a material with a positive inertial mass.

A slight modification of this thought experiment, however, enables us to gain energy. Indeed, let us consider an object with different gravitational and inertial masses $m_G \neq m_I$. Instead of considering this object on its own as before, let us

attach it to another object with a big positive mass $M_G = M_I > \max(|m_G|, |m_I|)$. This combination has gravitational mass $C_G = m_G + M_G$ and inertial mass $C_I = m_I + M_I$. Since $M_G = M_I > \max(|m_G|, |m_I|)$, both these combined masses are positive. Since $M_G = M_I$ and $m_G \neq m_I$, we conclude that $C_G \neq C_I$. So, the combined object has positive and different gravitational and inertial masses – and we already know that this leads to a contradiction with the energy conservation law. Since the gravitational mass is equal to the inertial mass, in the following text, we will simply talk about the mass m .

Comment. The fact that all the masses are proportional to each other is one of the formulations of the famous *equivalence principle* – one of the main underlying principles of Einstein's General Relativity; see, e.g., [1, 4]. Note that in our analysis, we did not use this principle – we derived the proportionality of different types of masses from fundamental concepts.

Conclusion: unfortunately, there is no such thing as anti-gravity. An unfortunate conclusion is that for every object, whether its mass m is negative or positive, its acceleration in the gravitational field of a body of mass M is determined by the formula $m \cdot a = G \cdot \frac{m \cdot M}{r^2}$, thus $a = G \cdot \frac{M}{r^2}$. This acceleration does not depend on the mass of the attracted body – so all objects follow the same trajectory, negative masses same as positive ones.

5. So Are Negative Masses Possible?

Finally, we can answer the question of whether negative masses are possible. Suppose that negative masses are possible. Then, by attaching an object with a negative mass $m < 0$ to a regular object with a similar positive mass $|m| = -m$, we get a combined object whose overall mass M is 0 (or at least is close to 0). Since the mass M is close to 0, even a very small force F will lead to a huge acceleration $a = \frac{F}{M}$. Thus, without spending practically any energy, we can accelerate the combined object to as high a velocity as we want. Once the object reaches this velocity, we dis-attach the negative-mass object – let it fly away. As a result, we now have an object of positive mass $|m|$ with a very large kinetic energy – and we can use this energy to perform useful work.

This scheme is not as clear-cut as the previous schemes, since here, we do not exactly go back to the original state – we lose a negative-mass body. However, we can do this for negative-mass body of arbitrarily small size – and still gain a lot of energy. Thus, while we *cannot* gain energy and get back to exactly the same original state, we *can* get back to a state which is as close to the original state as we want – and still gain as much energy as we want. This clearly contradicts to the idea of energy conservation. Thus, negative masses are not possible.

Acknowledgments

This work was supported in part by the National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

REFERENCES

1. Feynman R., Leighton R., Sands M. The Feynman Lectures on Physics. Boston :Addison Wesley, 2005.
2. Kreinovich V. Astronomical Tests of Relativity: Beyond Parameterized Post-Newtonian Formalism (PPN), to Testing Fundamental Principles // Relativity in Fundamental Astronomy / S. Klioner, P.K. Seidelmann, M.H. Soffel (eds.). Proceedings of IAU Symposium. No. 261. Cambridge : Cambridge University Press, 2009. P. 56–61.
3. Kreinovich V., Sriboonchitta S. Quantitative justification for the gravity model in economics // Predictive Econometrics and Big Data / V. Kreinovich, S. Sriboonchitta, N. Chakpitak (eds.). Cham : Springer Verlag, 2018. P. 214–221.
4. Misner C.W., Thorne K.S., Wheeler J.A. Gravitation. San Francisco : Freeman Publ., 1973.

МОЖЕТ ЛИ МАССА БЫТЬ ОТРИЦАТЕЛЬНОЙ?

В. Крейнлович¹

к.ф.-м.н., профессор, e-mail: vladik@utep.edu

С. Соловьев²

к.ф.-м.н., профессор, e-mail: sergei.soloviev@irit.fr

¹Техасский университет в Эль Пасо, США

²ИТ-исследовательский институт Тулузы (IRIT), 31062 Тулуза, Франция

Аннотация. Преодоление силы тяжести является важной частью космических путешествий и значительным препятствием, мешающим многим, казалось бы, разумным схемам космических путешествий стать практичными. Писатели-фантасты любят придумывать материалы, которые могут облегчить космические путешествия. Отрицательная масса — предположительно вызывающая антигравитацию — является одной из популярных идей в этом отношении. Но может ли масса быть отрицательной? В этой работе мы показываем, что отрицательные массы невозможны — их существование позволило бы нам создать энергию из ничего, что противоречит закону сохранения энергии.

Ключевые слова: отрицательная масса, принцип эквивалентности, антигравитация, закон сохранения энергии.

Дата поступления в редакцию: 27.11.2017

THE SUMS OF $M_I \cdot V_I$ AND $M_I \cdot V_I^2$ ARE PRESERVED, WHY NOT SUM OF $M_I \cdot V_I^3$: A PEDAGOGICAL REMARK

John McClure

Ph.D. (Eng.), Professor, e-mail: mcclure@utep.edu

Olga Kosheleva

Ph.D. (Phys.-Math.), Associate Professor, e-mail: olgak@utep.edu

Vladik Kreinovich

Ph.D. (Phys.-Math.), Professor, e-mail: vladik@utep.edu

University of Texas at El Paso, El Paso, Texas 79968, USA

Abstract. Students studying physics sometimes ask a natural question: the momentum – sum of $m_i \cdot v_i$ – is preserved, the energy – one half of the sum of $m_i \cdot v_i^2$ – is preserved, why not sum of $m_i \cdot v_i^3$? In this paper, we give a simple answer to this question.

Keywords: momentum conservation law, energy conservation law, teaching physics.

1. Formulation of the problem

Students studying physics sometimes ask a natural question:

- the momentum – sum of $m_i \cdot v_i$ – is preserved,
- the energy – one half of the sum of $m_i \cdot v_i^2$ – is preserved, so
- why not sum of $m_i \cdot v_i^3$?

In this paper, we give a simple pedagogical answer to this question.

2. Our explanation

To answer the above question, let us consider a simple 2-body 1-D mechanical problem. We have two small (point-like) solid objects with masses $m_1 = 2$ and $m_2 = 1$ on a straight line, the second object is to the left of the first one. Originally:

- the first object is at rest (its velocity is $v_1 = 0$), while
- the second object moves towards the first object with velocity $v_2 = 1$.

What will happen when the second object hits the first one?

The situation is invariant with respect to rotations around the line connecting the two objects, so the resulting trajectories will also be invariant – and thus, both objects will continue to move along the same line. The only remaining question is: what will be the new velocities v'_1 and v'_2 ?

We know that the momentum is preserved, and we know that kinetic energy is preserved. Thus, we can conclude that

$$m_1 \cdot v'_1 + m_2 \cdot v'_2 = m_1 \cdot v_1 + m_2 \cdot v_2 \quad (1)$$

and

$$m_1 \cdot (v'_1)^2 + m_2 \cdot (v'_2)^2 = m_1 \cdot v_1^2 + m_2 \cdot v_2^2. \quad (2)$$

Substituting the known values $m_1 = 2$, $m_2 = 1$, $v_1 = 0$, and $v_2 = 1$ into these formulas, we conclude that

$$2v'_1 + v'_2 = 1 \quad (1a)$$

and

$$2(v'_1)^2 + (v'_2)^2 = 1. \quad (2a)$$

From (1a), we conclude that

$$v'_2 = 1 - 2v'_1. \quad (3)$$

Substituting this expression into the formula (2a), we get

$$2(v'_1)^2 + (1 - 2v'_1)^2 = 1.$$

Opening the parentheses, we get

$$2(v'_1)^2 + 1 - 4v'_1 + 4(v'_1)^2 = 1.$$

Subtracting 1 from both sides and bringing similar terms together, we get

$$6(v'_1)^2 - 4v'_1 = 0,$$

i.e., equivalently,

$$v'_1 \cdot (6v'_1 - 4) = 0.$$

Thus, we have two options:

- either $v'_1 = 0$,
- or $6v'_1 - 4 = 0$, so that $v'_1 = \frac{2}{3}$.

In the first case, when $v'_1 = 0$, from the formula (3), we conclude that $v'_2 = 1$. This means that the the first object remains immobile, and the second object continue with the same speed – so that it somehow passed through the first object. For solid objects (not for ghosts), this is not possible.

Since the case $v'_1 = 0$ is not physically possible, we are left with the second option $v'_1 = \frac{2}{3}$. In this case, the formula (3) implies that

$$v'_2 = 1 - 2v'_1 = 1 - \frac{4}{3} = -\frac{1}{3}.$$

One can check that in this case, the equalities (1) and (2) – describing preservation of the sums of $m_i \cdot v_i$ and $m_i \cdot v_i^2$ – are satisfied, while the sum of $m_i \cdot v_i^3$ are *not* preserved: indeed, in this case,

$$m_1 \cdot v_1^3 + m_2 \cdot v_2^3 = 2 \cdot 0^3 + 1 \cdot 1^3 = 1,$$

while

$$m_1 \cdot (v'_1)^3 + m_2 \cdot (v'_2)^3 = 2 \cdot \left(\frac{2}{3}\right)^3 + \left(-\frac{1}{3}\right)^3 =$$

$$= 2 \cdot \frac{8}{37} - \frac{1}{27} = \frac{16}{27} - \frac{1}{27} = \frac{15}{27} \neq 1.$$

In this example, we use specific weights and velocities, to make computations simpler, but we could get the same conclusion by taking almost any combination of initial masses and velocities.

Conclusion: we cannot require that the sum of the terms $m_i \cdot v_i^3$ is preserved: this way, there would be no way to describe a simple collision.

Acknowledgments

This work was supported in part by the National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

REFERENCES

1. Feynman R., Leighton R., Sands M. The Feynman Lectures on Physics. Boston : Addison Wesley, 2005.

СУММЫ $M_I \cdot V_I$ И $M_I \cdot V_I^2$ СОХРАНЯЮТСЯ, ПОЧЕМУ НЕ СОХРАНЯЕТСЯ СУММА $M_I \cdot V_I^3$: ПЕДАГОГИЧЕСКОЕ ЗАМЕЧАНИЕ

Дж. МакКлюр

к.т.н., профессор, e-mail: mcclure@utep.edu

О. Кошелева

к.ф.-м.н., доцент, e-mail: olgak@utep.edu

В. Крейнович

к.ф.-м.н., профессор, e-mail: vladik@utep.edu

Техасский университет в Эль Пасо, США

Аннотация. Студенты, изучающие физику, иногда задают естественный вопрос: импульс — сумма $m_i \cdot v_i$ — сохраняется, — энергия — половина суммы $m_i \cdot v_i^2$ — сохраняется, почему не сохраняется сумма $m_i \cdot v_i^3$? В этой статье мы даём простой ответ на этот вопрос.

Ключевые слова: закон сохранения импульса, закон сохранения энергии, преподавание физики.

Дата поступления в редакцию: 27.11.2017

МОДЕЛИРОВАНИЕ РАВНОВЕСНОЙ ЭВОЛЮЦИИ ФОРМИРОВАНИЯ ЛЕСНОГО БИОЦЕНОЗА НА СПЛОШНЫХ ВЫРУБКАХ

Л.А. Володченкова

к.б.н., доцент, e-mail: volodchenkova2007@yandex.ru

А.К. Гуц

д.ф.-м.н., профессор, e-mail: guts@omsu.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Показано, что восстановление леса после вырубки может описываться как равновесная по Нэшу эволюция.

Ключевые слова: восстановление леса, равновесная эволюция, равновесие Нэша.

Введение

В восстановлении леса после вырубок в большинстве случаев участвует несколько древесных видов. Выявлено, что их доли в составе формирующихся лесов не остаются постоянными и часто изменяются резко, скачком (проявляя дискретность) при плавном непрерывном изменении абиотических факторов. Такие явления легко описываются моделями теории катастроф Рене Тома.

Используя модель формирования лесной растительности, предложенную Г.П. Быстраем и Н.С. Ивановой [1], покажем, что восстановление лесного биоценоза может представлять эволюцию, равновесную по Нэшу.

1. Модель формирования нового древостоя

Г.П. Быстрой и Н.С. Иванова [1] предложили теоретико-катастрофическую модель формирования лесной растительности на сплошных вырубках. После сплошных рубок возможно несколько альтернативных линий развития растительности. Из всего разнообразия возможных вариантов смен древесных видов авторы рассматривают только взаимоотношения берёзы (*Betula pendula* Roth. и *B. pubescens* Ehrh.) и сосны (*Pinus sylvestris* L.) – наиболее распространённых на Урале и в Зауралье древесных видов – в процессе зарастания вырубок и формирования нового древостоя.

Модель представляет дифференциальное уравнение вида

$$\frac{d\rho}{dt} = -|k_1|\rho + |k_2|T\rho^2 - |k_3|\rho^3 + |k_4|H, \quad (1)$$

где

ρ – характеристика, описывающая интенсивность возобновления древесной растительности: плотность (масса) подроста сосны (*Pinus sylvestris*) и подроста берёзы (*Betula pendula* и *B. pubescens*);

$T = (\rho_0 - \rho_m)/\rho_0$ – безразмерная характеристика интенсивности развития травянистого яруса;

$\rho_0 = (\rho_S - \rho_B)/2$ – средняя суммарная масса (плотность) сосны (*Pinus sylvestris*) и берёзы (*Betula pendula* и *B. pubescens*);

ρ_m – масса трав (плотность);

k_i – некоторые другие параметры экосистемы, подлежащие определению.

Управляющий параметр H – характеристика богатства лесорастительных условий (мощность почвы). Мощность почвы – комплексный фактор, характеризующий запас в почве элементов минерального питания и влаги и широко используемый в лесной типологии. Увеличение H (мощности почв) приводит к угнетению сосны в большей степени, чем берёзы.

Благодаря заменам

$$\eta = \frac{\rho}{\rho_c} - T\rho_0^*,$$

$$a = -3[T^2(\rho_0^*)^2 - 1], \quad b = -\frac{H}{H_c} + 3T\rho_0^* - 2T^3(\rho_0^*)^3,$$

$$\rho_0^* = |k_2|/(3|k_3|\varrho_c),$$

где ρ_c – некоторый масштаб плотности: плотность древесной растительности в критической точке, в которой плотность сосны и берёзы равны (смешанный древостой);

$$|k_1| = 3/t_0, \quad |k_2| = 3/(t_0\rho_c), \quad |k_3| = 1/(t_0\rho_c^2), \quad |k_4| = \rho_c/(H_c t_0)$$

и t_0 характеризует смену временного масштаба: $t = t : t_0$, уравнение (1) приводится к следующему виду:

$$\frac{d\eta}{dt} = -\frac{\partial F}{\partial \eta}, \tag{2}$$

$$F = \frac{1}{4}\eta^4 + \frac{1}{2}a\eta^2 + b\eta.$$

Это соответствует модели катастрофы «сборка».

Модель позволяет исследовать устойчивость текущих стационарных равновесных состояний (т. е. удовлетворяющих уравнению $d\eta/dt = 0$) лесной системы и проследивать динамику формирования растительности (см. [1, 2]).

Однако сам процесс переходов при использовании теории катастроф является скачкообразным, т. е. не представляется как решение дифференциального уравнения.

Мы предлагаем посмотреть на уравнение (2) с точки зрения теории дифференциальных игр, когда эволюция растительности управляется параметрами a и b и они могут динамически меняться в каждый момент времени, и траектория $\eta = \eta(t)$ может быть оптимально равновесной в смысле Нэша [4]. Покажем, что такая равновесная эволюция (a^*, b^*) по Нэшу возможна.

Для отыскания такого равновесия воспользуемся теорией дифференциальных игр [3, 4].

2. Алгоритм нахождения равновесий Нэша

Рассмотрим дифференциальную игру вида

$$\frac{dx}{dt} = f(x, u), \quad u = (u_1, \dots, u_m),$$

где u_i – управляющий параметр i -го игрока.

Если игрок i формирует «свое» управляющее воздействие в виде только функции времени $u_i(t)$ на всю продолжительность игры, то $u_i(t)$ – это *программное управление* игрока. Однако игрок может выбирать своё управление в зависимости от того, в каком положении x в момент времени t находится система. В таком случае игрок конструирует управляющее воздействие в виде функции $u_i(t, x)$, зависящей уже от позиции $\{t, x\}$, и для $u_i(t, x)$ используется термин *позиционное управление* игрока [5]. Часто пишут просто $u_i(x)$.

Мы будем искать позиционное управление, позиционное равновесие Нэша.

Для дифференциальной игры N -игроков

$$\frac{dx}{dt} = f(x) + \sum_{j=1}^N g_j(x) u_j, \quad f(0) = 0,$$

$$x \in \mathbb{R}, \quad u_j \in \mathbb{R},$$

$$J_i(x, u_1, \dots, u_N) = \int_0^{+\infty} [Q_i(x) + \sum_{j=1}^N R_{ij}(u_j)^2] dt, \quad (i = 1, \dots, N),$$

$$Q_i > 0, \quad R_{ii} > 0, \quad R_{ij} \geq 0,$$

существование равновесий Нэша

$$J_i(u_1^*, u_2^*, u_i^*, \dots, u_N^*) \leq J_i(u_1^*, u_2^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*), \quad \forall u_i, \quad i \in N \quad (3)$$

сводится к крайне сложной задаче отыскания положительно определённого решения $V_i(x) > 0$ нелинейного уравнения Гамильтона-Якоби

$$\begin{aligned} (V_i)'_x(x) f(x) + Q_i(x) - \frac{1}{2} (V_i)'_x \sum_{j=1}^N [g_j(x)]^2 (R_{jj})^{-1} (V_j)'_x + \\ + \frac{1}{4} \sum_{j=1}^N R_{ij} [g_j(x)]^2 [(R_{jj})^{-1}]^2 [(V_j)'_x]^2 = 0, \end{aligned} \quad (4)$$

по которому строится равновесие Нэша [4, Theorem 10.4-2]:

$$u_i^*(x) = u_i(V_i(x)) = -\frac{1}{2} R_{ii} g_i(x) (V_i)'_x, \quad i \in N. \quad (5)$$

Мы рассматриваем восстановление растительности как игру (2) с ненулевой суммой и двумя игроками, соответствующими управляющим параметрам a и b .

В нашем случае $N = 2$, игрок 1 – это интенсивность развития травянистого яруса a , игрок 2 – это взаимосвязь травянистого яруса и богатства лесорастительных условий (мощность почвы):

$$f(x) = -\eta^3, \quad g_1(\eta) = -\eta, \quad g_2(\eta) = -1$$

и при $R_{11} = R_{22} = 1, R_{12} = R_{21} = 0$ уравнения Гамильтона-Якоби имеют вид:

$$\begin{aligned} Q_1 + (V_1)'_{\eta} f(\eta) - \frac{1}{4}[g_1(\eta)]^2 [(V_1)'_{\eta}]^2 - \frac{1}{2}[g_2(\eta)]^2 (V_1)'_{\eta} (V_2)'_{\eta} &= 0, \\ Q_2 + (V_2)'_{\eta} f(\eta) - \frac{1}{4}[g_2(\eta)]^2 [(V_2)'_{\eta}]^2 - \frac{1}{2}[g_1(\eta)]^2 (V_1)'_{\eta} (V_2)'_{\eta} &= 0. \end{aligned} \quad (6)$$

Полагая, что

$$V_1(\eta) = V_2(\eta) = \frac{1}{2}\eta^2,$$

получаем уравнения Гамильтона-Якоби в виде

$$\begin{aligned} Q_1 &= \frac{5}{4}x^4 + \frac{1}{2}x^2 > 0, \\ Q_2 &= \frac{3}{2}x^4 + \frac{1}{4}x^2 > 0. \end{aligned} \quad (7)$$

Мы выбираем эти функции Q_1, Q_2 именно такими, поэтому уравнения Гамильтона-Якоби удовлетворяются.

Наши функции Q_1, Q_2 положительно определённые. Поэтому по теореме 10.4-2 из [4] имеем равновесие Нэша

$$a^* = \frac{1}{2}\eta^2, \quad b^* = \frac{1}{2}\eta, \quad (8)$$

найденное по формулам (5).

Выигрышные функции

$$\begin{aligned} J_1(x, p, \tau) &= \int_0^{+\infty} [Q_1(x) + a^2] dt, \\ J_2(x, p, \tau) &= \int_0^{+\infty} [Q_2(x) + b^2] dt. \end{aligned} \quad (9)$$

Равновесная по Нэшу эволюция формирования древостоя в данном случае задаётся уравнением

$$\frac{d\eta}{dt} = -\frac{3}{2}\eta^3 - \frac{1}{2}\eta.$$

Интегрируя это уравнение, находим семейство оптимальных равновесных траекторий

$$\arctg(\eta) = -\frac{\sqrt{3}}{2}t + const.$$

3. Заключение

Мы показали, что восстановление лесного биоценоза может быть описано как равновесная эволюция по Нэшу. Найдена только одна равновесная эволюция, хотя не исключено, что их может быть намного больше, поскольку при восстановлении леса возможно участие различных видов растительности (березняки разнотравно-вейниковые, сосняки брусничниковые и т. д.).

ЛИТЕРАТУРА

1. Быстрой Г.П., Иванова Н.С. Подходы к моделированию динамики лесной растительности на основе теории катастроф // Аграрный вестник Урала. 2010. № 2(68). С. 75–79.
2. Володченкова Л.А., Гуц А.К. Кибернетика катастроф лесных экосистем. Омск : Полиграфический центр КАН, 2012. 220 с.
3. Гуц А.К., Володченкова Л.А. Климатский лес как нэшевское равновесное состояние лесных экосистем // Математические структуры и моделирование. 2017. № 1(41). С. 38–44.
4. Lewis F.L., Vrabie D.L., Syrmos V.L. Optimal Control. John Wiley & Sons, Inc., 2012. URL: <http://www.uta.edu/utari/acs/FL/%20talks/CDC/%20Orlando/%202011-%20online/%20synch/%20PI.pdf> (дата обращения: 01.02.2018).
5. Тынянский Н.Т., Жуковский В.И. Дифференциальные игры с ненулевой суммой (кооперативный вариант) // Итоги науки и техн. Сер. Мат. анализ. 1979. Т. 17. С. 3–112.

MODELING THE EQUILIBRIUM EVOLUTION OF THE FOREST BIOCENOSIS ON SOLID FELLING

L.A. Volodchenkova

Ph.D. (Biology), Associate Professor, e-mail: volodchenkova2007@yandex.ru

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: guts@omsu.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. It is shown that the restoration of the forest after felling can be described as a Nash equilibrium.

Keywords: restoration of the forest, evolution, the Nash equilibrium.

Дата поступления в редакцию: 02.02.2018

ОБ АСИМПТОТИЧЕСКОМ ПОВЕДЕНИИ РЕШЕНИЙ ЗАДАЧИ КОШИ ДЛЯ СИСТЕМЫ ФУНКЦИОНАЛЬНО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ГИПЕРБОЛИЧЕСКОГО ТИПА

Г.Д. Анисимова

аспирант, e-mail: gdanisimova@gmail.com

Омский государственный технический университет, Омск, Россия

Аннотация. Для указанного в названии класса систем доказаны необходимые и достаточные условия экспоненциальной устойчивости и дихотомии решений задачи Коши в терминах нулей (λ, μ) определителя матричного пучка – символа функционально-дифференциального оператора в левой части системы. Приведён иллюстрирующий пример.

Ключевые слова: переход к разностной задаче Коши, характеристизация спектра разрешающего оператора.

Введение

1. Работа является продолжением цикла работ по теории устойчивости для функционально-дифференциальных уравнений (ФДУ), выполненных в последние годы группой сотрудников и аспирантов Омского технического университета [1–8]. Развитый в [5–8] подход к анализу поведения при большом времени решений обыкновенных линейных ФДУ сведением к такой же задаче для разностного уравнения вида $u_n = \Gamma u_{n-1}$ с компактным оператором Γ в пространстве начальных данных распространён на подкласс ФДУ с частными производными. Доказаны необходимые и достаточные условия экспоненциальной устойчивости и дихотомии решений задачи Коши для одномерной стационарной гиперболической системы, возмущённой слагаемыми с запаздывающим аргументом, в терминах нулей квазимногочлена $\det \Delta(\lambda, \mu)$, где $\Delta(\lambda, \mu)$ – символ оператора в левой части системы. Основной результат проиллюстрирован на модельной задаче управления с запаздыванием в управляющем устройстве.

2. Рассмотрим в полуплоскости $\Pi = \mathbb{R} \times [0, \infty)$ задачу Коши

$$\begin{cases} Lu = Du|_{(x,t)} + \int_0^1 [dB(s)] u(x, t-s) = 0, & (x, t) \in \Pi \setminus \Pi_0, \\ u|_{\Pi_0} = \varphi \in E, & \Pi_0 = \mathbb{R} \times [0, 1]. \end{cases} \quad (1)$$

Здесь

$$D = \frac{d}{dt} + A \frac{d}{dx}, \quad A = \text{diag}(a_1 I_1, \dots, a_m I_m), \quad a_1 > \dots > a_m, \quad a_k \neq 0, \quad (2)$$

I_k – единичная матрица порядка N_k , $\sum N_k = N$, $B: [0, 1] \rightarrow \mathbb{C}^{N \times N}$, $V_0^1(B) < \infty$,
 $u = \begin{bmatrix} u^1 \\ \dots \\ u^m \end{bmatrix}$, $\varphi = \begin{bmatrix} \varphi^1 \\ \dots \\ \varphi^m \end{bmatrix}$, u^k , φ^k – столбцы размера N_k , E – банахово пространство непрерывных ограниченных функций $\Pi_0 \rightarrow \mathbb{C}^N$ с нормой $\|\varphi\|_E = \sup |\varphi|$, $|\cdot|$ – эрмитова норма в $\mathbb{C}^N$.

При условиях 2 через каждую точку $(x, t) \in \Pi$ проходят характеристики

$$q_k(x, t) = \{(\sigma, \tau) \in \mathbb{R}^2 \mid \sigma = \sigma_k(x, t, \tau) = x + a_k(\tau - t)\}, k = \overline{1, m}. \quad (3)$$

Оператор D далее понимается в обобщённом смысле:

$$D = \text{diag}(D_1, \dots, D_n), \quad (4)$$

где $D_k u^k$ – производная по t вдоль q_k . Под решением задачи 1 понимается непрерывная функция $u: \Pi \rightarrow \mathbb{C}^N$ с C^1 -гладкими в $\overline{\Pi \setminus \Pi_0}$ вдоль «своих» характеристик q_k компонентами u^k , удовлетворяющая 1. Класс таких функций далее обозначается $\widehat{C}^1$. Из выполняемых в разделе 1 построений, в частности, следует однозначная разрешимость задачи 1 в классе $\widehat{C}^1$.

Далее в разделе 1 построена эквивалентная 1 разностная задача Коши в фазовом пространстве E

$$\begin{cases} u_n = \Gamma u_{n-1}, & n = 1, 2, \dots, \\ u_0 = \varphi \in E. \end{cases} \quad (5)$$

Доказана компактность оператора Γ . В разделе 2 описан спектр Γ в терминах нулей (λ, μ) квазимногочлена, указанного в п.1. В разделе 3 на этой базе доказаны итоговые результаты. В разделе 4 приведён иллюстрирующий пример.

3. Приведём здесь для удобства ссылок используемый далее факт из теории интеграла Стильтьеса.

ПРЕДЛОЖЕНИЕ 1 [9]. Пусть функция $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ непрерывна и функция $g: [c, d] \rightarrow \mathbb{R}$ имеет ограниченную вариацию. Тогда функция

$$h(x) = \int_c^d f(x, s) dg(s)$$

непрерывна и верна формула

$$\int_a^b h(x) dx = \int_c^d \left(\int_a^b f(x, s) dx \right) dg(s).$$

Этот результат без труда переносится на случай f со значениями в $\mathbb{C}^N$ и g со значениями в $\mathbb{C}^{N \times N}$.

1. Переход к разностной задаче Коши

Поставим в соответствие $\varphi \in E$ функции $\Pi_0 \rightarrow \mathbb{C}^N$

$$B_0\varphi|_{(x,t)} = \int_0^t [dB(s)]\varphi(x, t-s), \quad B_1\varphi|_{(x,t)} = \int_t^1 [dB(s)]\varphi(x, 1+t-s), \quad (6)$$

$$S\varphi|_{(x,t)} = \int_0^t \begin{bmatrix} \varphi^1(\sigma_1, \tau) \\ \dots \\ \varphi^m(\sigma_m, \tau) \end{bmatrix} d\tau, \quad P\varphi|_{(x,t)} = \begin{bmatrix} \varphi^1(x - a_1t, 1) \\ \dots \\ \varphi^m(x - a_mt, 1) \end{bmatrix}, \quad (7)$$

$$T_k\varphi = SB_k\varphi, \quad k = 0, 1,$$

где σ_k – функции 3. Нетрудно убедиться, с учётом предложения 1, что формула 7 задаёт линейные ограниченные операторы $E \rightarrow E$.

Лемма 1. Оператор $I + T_0$ имеет ограниченный обратный $E \rightarrow E$.

Доказательство. Из 6, 7 следует неравенство

$$|T_0\varphi| \leq Mt \sup |\varphi|, \quad M = \bigvee_0^1 (B), \quad t \in [0, 1].$$

Индукцией по n получим

$$|T_0^n\varphi| \leq \frac{Mt^n}{n!} \sup |\varphi|, \quad n = 0, 1, \dots$$

Отсюда следует требуемое. ■

Положим

$$\Gamma = (I + T_0)^{-1} (P - T_1). \quad (8)$$

Лемма 2. Формула 8 задаёт компактный оператор $E \rightarrow E$.

Доказательство. Пусть $\mathfrak{M}$ – ограниченное множество в E . Равенство $\psi = \Gamma\varphi$, $\varphi \in \mathfrak{M}$, равносильно соотношению

$$\psi = P\varphi - (T_0\psi + T_1\varphi). \quad (9)$$

Из 9 попутно следует

$$\psi(x, 0) = \varphi(x, 1). \quad (10)$$

Положим при $(x, t) \in \widehat{\Pi}_0 = \mathbb{R} \times [-1, 1]$

$$\alpha(x, t) = \begin{cases} \varphi(x, 1+t), & t \in [-1, 0), \\ \dots & \\ \psi(x, t), & t \in [0, 1]. \end{cases} \quad (11)$$

Функции 11 равномерно по $\varphi \in \mathfrak{M}$ ограничены и, с учётом 10, непрерывны в $\widehat{\Pi}_0$. Применяя к обеим частям 9 оператор 4, после вычислений с учётом равенств $DP\varphi = 0$, $DS\varphi = \varphi$ найдём:

$$D\psi|_{(x,t)} = \int_0^1 [dB(s)] \alpha(x, t-s), \quad \psi \in \Gamma\mathfrak{M}. \quad (12)$$

В силу предложения 1 функция в правой части 12 непрерывна в Π_0 . Представим матрицу B в блочном виде $B = [b_{kj}]$ с блоками размера $N_k \times N_j$. Вытекающие из 12 соотношения для $D_k\psi^k$ после замены

$$f_k(x, t) = \psi^k(x + a_k t, t)$$

принимают вид

$$f'_{kt} = \sum_j \int_0^1 [db_{kj}(s)] \alpha^j(x + a_k t, t-s), \quad k = \overline{1, m}. \quad (13)$$

Из 9, 13 вытекают равномерные по $\psi \in \Gamma\mathfrak{M}$ оценки

$$|f_k| \leq c_1, \quad |f'_{kt}| \leq c_2, \quad c_k = \text{const} > 0, \quad k = \overline{1, m}. \quad (14)$$

Будем рассматривать f_k как функции от t со значениями в банаховом пространстве C_k непрерывных ограниченных функций $\mathbb{R} \rightarrow \mathbb{C}^{N_k}$ с равномерной нормой. В силу второй оценки 14 можно рассматривать f'_{kt} как производную $\frac{d}{dt}f_k(t)$ в топологии C_k . Обозначим

$$M_k = \{f_k(t) \mid \psi \in \Gamma\mathfrak{M}\}, \quad k = \overline{1, m}.$$

Из оценок 14 следует: при каждом k множество M_k – равномерно ограниченное и равностепенно непрерывное семейство функций в банаховом пространстве $C([0, 1], C_k)$ с нормой

$$\|f_k\| = \sup_{t \in [0, 1]} \|f_k\|_{C_k} = \sup_{\Pi_0} |f_k(x, t)|.$$

Таким образом, при каждом k множество M_k – предкомпакт в $C([0, 1], C_k)$ и, с учётом $\sup_{\Pi_0} |f_k(x, t)| = \sup_{\Pi_0} |\psi^k(x, t)|$, множество $\{\psi^k \mid \psi \in \Gamma\mathfrak{M}\}$ – предкомпакт в банаховом пространстве непрерывных ограниченных функций $\Pi_0 \rightarrow \mathbb{C}^{N_k}$. Отсюда следует требуемое. ■

Замечание 1. Из формул 11, 12 для функции $\psi = \Gamma\varphi$, с учётом предложения 1, следует: для любой $\varphi \in E$

$$D\Gamma\varphi \in E, \quad \|D\Gamma\varphi\|_E \leq \text{const} \cdot \|\varphi\|_E. \quad (15)$$

Рассмотрим разностную задачу Коши 5 с оператором 8. Очевидно, задача 5 имеет единственное решение $u_n = \Gamma^n \varphi$.

Лемма 3. Функция $u: \Pi \rightarrow \mathbb{C}^N$ является решением класса $\widehat{C}^1$ задачи Коши 1 точно тогда, когда последовательность

$$u_n(x, t) = u(x, t + n), \quad (x, t) \in \Pi_0, \quad n = 0, 1, \dots \quad (16)$$

является решением задачи 5.

Доказательство. Пусть u – решение класса $\widehat{C}^1$ задачи 1. Записывая систему 1 в точке $(x, t + n)$, указанной в 16, представляя интеграл в 1 в виде суммы интегралов по $[0, t]$, $[t, 1]$ и вводя обозначения 16, получим соотношение

$$(D + B_0)u_n + B_1u_{n-1} = 0, \quad n = 1, 2, \dots, \quad (17)$$

с операторами 4, 6; учтено равенство

$$u(x, t + n - s) = u_{n-1}(x, 1 + t - s), \quad s \in [t, 1]. \quad (18)$$

Применяя к обеим частям 17 оператор S , с учётом равенств

$$SDu_n = u_n(x, t) - \begin{bmatrix} u_n^1(x - a_1t, 0) \\ \dots \\ u_n^m(x - a_mt, 0) \end{bmatrix}, \quad u_n(x, 0) = u_{n-1}(x, 1), \quad n = 1, 2, \dots,$$

получим равносильное 5 уравнение

$$(I + T_0)u_n = (P - T_1)u_{n-1}, \quad n = 1, 2, \dots \quad (19)$$

Обратно, пусть u_n – решение задачи 5. Определим, с учётом вытекающего из 19 равенства $u_n(x, 0) = u_{n-1}(x, 1)$, функцию $u: \Pi \rightarrow \mathbb{C}^N$ равенством 16. Выполнение начального условия и непрерывность функции u очевидны. Применяя к обеим частям 19 оператор D , получим равенство 17, равносильное, с учётом 18, равенству 1 при $t \in [n, n + 1]$, $n \geq 1$, с односторонними производными $D_k u^k$ на прямых $t = n$, $t = n + 1$. Записывая в предположении $n \geq 2$ равенство 17 при $t = 0$ и такое же равенство с заменой n на $n - 1$ при $t = 1$ и сравнивая полученные соотношения, найдём $Du_n|_{t=0} = Du_{n-1}|_{t=1}$.

Лемма доказана. ■

Следствие 1. Задача Коши 1 однозначно разрешима в классе $\widehat{C}^1$.

Обозначим C_p банахово пространство кусочно-непрерывных ограниченных функций $v: \mathbb{R} \rightarrow \mathbb{C}^N$ со скачками в целых точках с нормой $\sup |v|$; C_p^1 – банахово пространство кусочно-гладких ограниченных функций $v: \mathbb{R} \rightarrow \mathbb{C}^N$ со скачками v, v' в целых точках с нормой $\sup |v| + \sup |v'|$. Будем предполагать функции v соответственно непрерывными и гладкими слева.

Лемма 4. Оператор $\mathfrak{L} = \frac{d}{dx} - \Lambda$, $\Lambda \in \mathbb{C}^{N \times N}$, действующий из C_p^1 в C_p , имеет ограниченный обратный $\mathfrak{L}^{-1}: C_p \rightarrow C_p^1$ точно тогда, когда спектр матрицы Λ не пересекается с мнимой осью.

Доказательство. В силу теоремы Банаха об обратном операторе утверждение леммы равносильно следующему: уравнение $\mathfrak{L}v = h$ при любой $h \in C_p$ имеет единственное решение $v \in C_p^1$ точно тогда, когда выполняется указанное условие. Доказательство достаточности проводится повторением с очевидными отличиями аналогичного утверждения в [5, с. 119] для *непрерывного* случая. Если $\mu \in i\mathbb{R}$ – собственное число Λ с собственным вектором f , уравнение $\mathfrak{L}v = 0$ имеет, помимо решения $v = 0$, ещё одно решение из C_p^1 : $v = e^{\mu x} f$; таким образом, условие леммы необходимо. ■

2. Характеризация собственных чисел оператора Γ

В силу структуры спектра компактного оператора в банаховом пространстве спектр оператора Γ состоит из точки $\xi = 0$ и не более чем счётного множества собственных чисел с единственной возможной предельной точкой $\xi = 0$.

Поставим в соответствие оператору 1 матричный пучок

$$\Delta(\lambda, \mu) = \lambda I + \mu A + \int_0^1 e^{-\lambda s} dB(s), \quad \lambda, \mu \in \mathbb{C}. \quad (20)$$

Теорема 1. Число $\xi \neq 0$ является собственным числом оператора Γ точно тогда, когда при некотором $\mu \in i\mathbb{R}$ и некотором λ из множества $M_\xi = \{\lambda = \ln \xi + 2k\pi i, k \in \mathbb{Z}\}$ выполняется равенство

$$\det \Delta(\lambda, \mu) = 0. \quad (21)$$

Доказательство. 1. Пусть пара (λ, μ) , $\lambda \in \mathbb{C}$, $\mu \in i\mathbb{R}$, удовлетворяет 21 и вектор $f \in \mathbb{C}^N$, $|f| = 1$ таков, что $\Delta(\lambda, \mu)f = 0$. Положим

$$\xi = e^\lambda, \quad u = e^{\lambda t + \mu x} f.$$

Подстановка в 1 даёт: $Lu = e^{\lambda t + \mu x} \Delta(\lambda, \mu)f = 0$. В силу леммы 3 функции

$$u_n(x, t) = e^{\lambda(t+n) + \mu x} f, \quad (x, t) \in \Pi_0, \quad n = 0, 1, \dots \quad (22)$$

удовлетворяют уравнению 5, в частности, верно равенство $\Gamma u_0 = u_1$ или, что то же, $\Gamma u_0 = \xi u_0$.

2. Пусть $\xi \neq 0$, $\Gamma\varphi = \xi\varphi$ при некоторой $\varphi \in E$, $\varphi \neq 0$. В силу первого соотношения 15 $D\varphi \in E$. Так как $\{e^{2k\pi it}, k \in \mathbb{Z}\}$ – полная ортогональная система на $[0, 1]$, существует $\lambda \in M_\xi$, такое, что

$$h(x) = \int_0^1 e^{-\lambda t} \varphi(x, t) dt \neq 0. \quad (23)$$

Функция 23 непрерывна и ограничена на оси.

Зафиксируем последовательность $\varepsilon_n \downarrow 0$. Построим последовательность кусочно-гладких, гладких слева функций $\varphi_n: \Pi_0 \rightarrow \mathbb{C}^N$ со скачками φ_n, φ_n' , φ_n' на отрезках $\{x = n \in \mathbb{Z}, 0 \leq t \leq 1\}$ такую, что

$$|\varphi_n - \varphi| + |D\varphi_n - D\varphi| \leq \varepsilon_n, \quad n = 1, 2, \dots \quad (24)$$

Имеет место равенство

$$\Gamma \varphi_n = \xi (\varphi_n + \delta_n), \quad (25)$$

где $\delta_n = \varphi - \varphi_n + \xi^{-1} \Gamma (\varphi_n - \varphi)$. Из 24 и неравенства 15 следует оценка

$$|\delta_n| + |D\delta_n| \leq \text{const } \varepsilon_n, \quad n = 1, 2, \dots \quad (26)$$

Представим 25 в виде

$$\xi (I + T_0) (\varphi_n + \delta_n) + T_1 \varphi_n = P \varphi_n. \quad (27)$$

Применяя к обеим частям оператор D , получим

$$\xi (D + B_0) (\varphi_n + \delta_n) + B_1 \varphi_n = 0$$

(учтены соотношения $DS\varphi = \varphi$, $DP\varphi = 0$). Подстановка

$$\varphi_n = e^{\lambda t} \hat{\varphi}_n, \quad \lambda\text{-число (23),} \quad (28)$$

даёт после вычислений с учётом $e^\lambda = \xi$ равенство

$$\frac{\partial \hat{\varphi}_n}{\partial t} + A \frac{\partial \hat{\varphi}}{\partial x} + b_n(x, t, \lambda) + \xi e^{-\lambda t} (D + B_0) \delta_n = 0, \quad (29)$$

где $b_n = \int_0^1 e^{-\lambda s} [dB(s)] \beta_n(x, t - s)$, $\beta_n(x, t) = \begin{cases} \hat{\varphi}_n(x, 1 + t), & t \in [-1, 0), \\ \hat{\varphi}_n(x, t), & t \in [0, 1]. \end{cases}$

Применим к обеим частям 29 оператор $\int_0^1 \cdot dt$. Имеем

$$\int_0^1 \beta_n(x, t - s) dt = \int_0^s \hat{\varphi}_n(x, 1 + t - s) dt + \int_s^1 \hat{\varphi}_n(x, t - s) dt = h_n(x),$$

$$h_n(x) = \int_0^1 \hat{\varphi}_n(x, t) dt, \quad n = 1, 2, \dots$$

С учётом этого и вытекающего из 27, 28 соотношения $\hat{\varphi}_n(x, 1) - \hat{\varphi}_n(x, 0) = \delta_n(x, 0)$ после преобразований получим равенство

$$\mathfrak{L} h_n = \left(\frac{d}{dx} + \Lambda \right) h_n = \theta_n(x), \quad n = 1, 2, \dots, \quad (30)$$

где

$$\Lambda = A^{-1} \left[\lambda I + \int_0^1 e^{-\lambda s} dB(s) \right], \quad \theta_n = A^{-1} \left[\delta_n(x, 0) + \xi \int_0^1 e^{-\lambda t} (D + B_0) \sigma_n dt \right].$$

Нетрудно убедиться:

$$h_n \in C_p^1, \quad h_n \rightarrow h \text{ в } C_p, \quad \theta_n \in C_p.$$

Так как с учётом $h \neq 0$, $\|h_n\|_{C_p} \geq \text{const} > 0$ ($n \geq n_0$) и ввиду 26, $\theta_n \rightarrow 0$ в C_p , из 30 вытекает: оператор $\mathfrak{L}$ не имеет обратного $C_p \rightarrow C_p^1$. Тогда в силу леммы 4 существует $\mu \in i\mathbb{R}$ такое, что выполняется равенство $\det(\mu I - \Lambda) = 0$, равносильное равенству 21 с матрицей 20. Теорема доказана. ■

3. Признаки устойчивости и дихотомии

Имея ввиду установленное леммой 3 соответствие между решениями класса $\widehat{C}^1$ задачи Коши 1 и решениями разностной задачи Коши 5, примем следующие определения.

Будем говорить, что решение $u = 0$ системы 1 экспоненциально устойчиво, если это имеет место для решения $u_n = 0$ уравнения 5: для решений задачи Коши 5 верна при некоторых $a, \nu > 0$ оценка

$$\|\Gamma^n \varphi\|_E \leq a e^{-\nu n} \|\varphi\|_E. \quad (31)$$

Будем говорить, что имеет место экспоненциальная дихотомия решений класса $\widehat{C}^1$ задачи Коши 1, если фазовое пространство E задачи Коши 5 разлагается в прямую сумму подпространств

$$E = E_1 \dot{+} E_2, \quad E_k \neq \{0\}, \quad (32)$$

так, что для решений задачи Коши 5 верны при некоторых $a, \nu > 0$ оценки

$$\begin{aligned} \varphi \in E_1 &\Rightarrow \|\Gamma^n \varphi\|_E \leq a e^{-\nu n} \|\varphi\|_E, \\ \varphi \in E_2 &\Rightarrow \|\Gamma^n \varphi\|_E \geq a e^{\nu n} \|\varphi\|_E. \end{aligned} \quad (33)$$

Теорема 2. *Решение $u = 0$ системы 1 экспоненциально устойчиво точно тогда, когда при каждом $\mu \in i\mathbb{R}$ уравнение 21 не имеет λ -корней с реальной частью $\operatorname{Re} \lambda \geq 0$.*

Доказательство. 1. Пусть требование теоремы выполняется. Так как спектр компактного оператора состоит из множества (возможно пустого) собственных чисел и точки $\xi = 0$, в этом случае, в силу теоремы 1, спектр оператора Γ лежит целиком в круге $|\xi| < 1$. Ввиду замкнутости спектра существует $r \in (0, 1)$ такое, что спектр Γ лежит в круге $|\xi| < r$. Из формулы Коши-Рисса

$$\Gamma^n \varphi = (2\pi i)^{-1} \int_{|\xi|=r} \xi^n (\xi I - \Gamma)^{-1} \varphi d\xi \quad (34)$$

следует для решений задачи Коши 5 оценка 31 при $a = r \cdot \max_{|\xi|=r} \|(\xi I - \Gamma)^{-1}\|$, $\nu = \ln r^{-1}$.

2. Пусть требование теоремы не выполняется: существует удовлетворяющая уравнению 21 пара (λ, μ) , $\operatorname{Re} \lambda \geq 0$, $\operatorname{Im} \mu = 0$. Тогда, очевидно, не выполняется оценка 31 для решения 22 задачи 5. Теорема доказана. ■

Теорема 3. *Экспоненциальная дихотомия решений задачи Коши 1 имеет место точно тогда, когда при каждом $\mu \in i\mathbb{R}$ уравнение 21 не имеет λ -корней на мнимой оси и имеет хотя бы один λ -корень в полуплоскости $\operatorname{Re} \lambda > 0$.*

Доказательство. 1. Пусть требование теоремы выполняется. С учётом структуры спектра Γ и теоремы 1 это означает: спектр $\sigma(\Gamma)$ оператора Γ распадается на две непустые компоненты σ_1, σ_2 , лежащие соответственно внутри и вне окружности $|\xi| = 1$, и тогда пространство E распадается [10] в прямую сумму 32 инвариантных подпространств, таких, что спектр сужения оператора Γ на подпространство E_k – множество σ_k . Из формулы вида 34 при $\varphi \in E_1$ следует первая оценка 33. Так как спектр оператора $\Gamma^{-1}: E_2 \rightarrow E_2$ – множество $(\sigma_2)^{-1}$, лежащее в круге $|\xi| < 1$, из формулы вида 34 для Γ^{-1} при $\varphi \in E_2$ следует вторая оценка 33 с переменной местами левой и правой частей и переменной знака неравенства на противоположный.

2. Пусть требование теоремы не выполняется. Нетрудно усмотреть, с учётом структуры спектра Γ и теоремы 1, что в этом случае имеет место одно из двух: либо спектр Γ лежит целиком в круге $|\xi| < 1$, либо существует собственное число ξ_0 оператора Γ с модулем 1. В первом случае, в силу теоремы 2, первая оценка 33 выполняется при всех $\varphi \in E$, тем самым не выполняется требование $E_2 \neq \{0\}$ в 32. Во втором случае имеем для некоторой $\varphi \in E$, $\|\varphi\|_E = 1$, $\Gamma\varphi = \xi_0\varphi$, откуда следует равенство $\|\Gamma^n\varphi\|_E = \|\xi_0^n\varphi\|_E \equiv 1$, несовместимое с оценками 33. Теорема доказана. ■

Замечание 2. Утверждения теорем 2, 3 сохраняются, если матрица A в 1 имеет вид $A = Z^{-1}\hat{A}Z$, где $\hat{A}$ – диагональная матрица 2. Это следует из равенства $\det \Delta(\lambda, \mu) = \det \hat{\Delta}(\lambda, \mu)$, где пучок $\hat{\Delta}$ строится по коэффициентам $\hat{A}, \hat{B} = ZB(s)Z^{-1}$ системы, получающейся заменой $\hat{u} = Zu$.

4. Пример

Рассмотрим систему управления теплопереносом в бесконечном стержне с запаздыванием в управляющем устройстве

$$\begin{cases} Du(x, t) + B_0 u(x, t) + b\sigma(x, t-1) = 0, & (x, t) \in \Pi \setminus \Pi_0, \\ \sigma(x, t) = d^* u(x, t), & u|_{\Pi_0} = \varphi \in E. \end{cases} \quad (35)$$

Здесь D – оператор 4,

$$u = \begin{bmatrix} T \\ q \end{bmatrix}, \quad A = \begin{bmatrix} 0 & (c\rho)^{-1} \\ k\varepsilon^{-1} & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon^{-1} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \quad (36)$$

T, q – температура и плотность теплового потока, ρ, c, k – плотность, удельная теплоёмкость и теплопроводность, ε – время релаксации, b, d – векторы управляемости и наблюдаемости, σ – сигнал обратной связи, E – банахово пространство непрерывных ограниченных функций $\varphi: \Pi_0 \rightarrow \mathbb{R}$ с нормой $\sup |\varphi|$. Нетрудно убедиться, что матрица A подобна матрице

$$\hat{A} = \text{diag}(a, -a), \quad a = \sqrt{\frac{k}{\varepsilon c \rho}}.$$

Система 35 при $b = 0$ описывает теплоперенос в стержне в рамках гиперболической теплопроводности [11, 12], в частности, возмущения распространяются вдоль стержня с конечной скоростью a . Система 35 имеет вид 1 при

$$B(s) = \begin{cases} 0, & s = 0, \\ B_0, & 0 < s < 1, \\ B_0 + bd^*, & s = 1. \end{cases}$$

Подстановка в 20 даёт

$$\Delta(\lambda, \mu) = \lambda I + \mu A + B_0 + e^{-\lambda} bd^*. \quad (37)$$

С учётом замечания в конце раздела 4 устойчивость решения $u = 0$ системы 35 зависит от расположения нулей (λ, μ) определителя пучка 37. При $b = 0$ имеем: $\det \Delta(0, 0) = \det B_0$. Тем самым, *при отсутствии управления* не выполняется требование теоремы 2, и решение $u = 0$ системы 35 не является экспоненциально устойчивым. Покажем: при условиях

$$0 < b_1 d_1 < \frac{1}{1 + \varepsilon}, \quad b_2 = d_2 = 0 \quad (38)$$

решение $u = 0$ экспоненциально устойчиво. Вычисления с учётом 36–38 дают

$$\Delta(\lambda, \mu) = A(\mu I - \Lambda), \quad \Lambda = - \begin{bmatrix} 0 & k^{-1}(1 + \varepsilon\lambda) \\ c\rho(\lambda + b_1 d_1 e^{-\lambda}) & 0 \end{bmatrix}.$$

Отсюда следует для μ – корней уравнения 21 равенство

$$\mu_j^2 = \frac{c\rho}{k} (1 + \varepsilon\lambda) (\lambda + b_1 d_1 e^{-\lambda}), \quad j = 1, 2.$$

Подстановка $\lambda = \alpha + i\omega$ приводит это равенство к виду

$$\begin{aligned} \mu_j^2 &= \frac{c\rho}{k} (p + qi), \\ p &= \alpha + \varepsilon(\alpha^2 - \omega^2) + b_1 d_1 [(1 + \varepsilon\alpha) \cos \omega + \varepsilon\omega \sin \omega] e^{-\alpha}, \\ q &= \omega [1 + 2\varepsilon\alpha + b_1 d_1 (\varepsilon \cos \omega - (1 + \varepsilon\alpha) \frac{\sin \omega}{\omega})] e^{-\alpha}. \end{aligned} \quad (39)$$

Пусть выполняется требование 38 и $\alpha \geq 0$. Возможны два случая.

1. $\omega = 0$. В этом случае $p + qi = \alpha + \varepsilon\alpha^2 + b_1 d_1 (1 + \varepsilon\alpha) e^{-\alpha} > 0$.
2. $\omega \neq 0$. Из оценки

$$|q| \geq |\omega| r(\alpha), \quad r(\alpha) = 2\varepsilon\alpha + 1 - b_1 d_1 (\varepsilon\alpha + 1 + \varepsilon) e^{-\alpha},$$

с учётом

$$r(0) = (1 + \varepsilon) \left(\frac{1}{1 + \varepsilon} - b_1 d_1 \right) > 0, \quad r'(\alpha) = 2\varepsilon + b_1 d_1 (1 + \varepsilon\alpha) e^{-\alpha} > 0$$

следует: в этом случае $q \neq 0$. Таким образом при условии $\operatorname{Re} \lambda \geq 0$ правая часть уравнения 39 заведомо не является квадратом мнимого числа, и μ – корни уравнения 21 не лежат на мнимой оси. В силу теоремы 2 отсюда следует требуемое.

Случай любого запаздывания $\tau > 0$ в 35 приводится к случаю $\tau = 1$ масштабным преобразованием.

Благодарности

Автор выражает глубокую благодарность профессору, д.ф.-м.н. Романовскому Рэму Константиновичу (1929–2016 гг.) за оказанную помощь при написании настоящей статьи.

ЛИТЕРАТУРА

1. Алексенко Н.В. Устойчивость решений почти периодических систем функционально-дифференциальных уравнений запаздывающего типа // Изв. вузов. Математика. 2000. Вып. 2. С. 3–6.
2. Алексенко Н.В., Романовский Р.К. Метод функционалов Ляпунова для линейных дифференциально-разностных систем с почти периодическими коэффициентами // Дифференц. уравнения. 2001. Т. 37, № 2. С. 147–153.
3. Троценко Г.А. Об устойчивости решений почти периодической системы функционально-дифференциальных уравнений нейтрального типа // Изв. вузов. Математика. 2003. Вып. 36. С. 77–81.
4. Романовский Р.К., Троценко Г.А. Метод функционалов Ляпунова для линейных дифференциально-разностных систем нейтрального типа с почти периодическими коэффициентами // Сиб. матем. журн. 2003. Т. 44, № 2. С. 444–453.
5. Романовский Р.К., Назарук Е.М. О дихотомии линейных автономных систем функционально-дифференциальных уравнений // Матем. заметки. 2014. Т. 95, № 1. С. 129–135.
6. Романовский Р.К., Назарук Е.М. Прямой метод Ляпунова для линейных систем функционально-дифференциальных уравнений в пространстве Соболева // Сиб. матем. журн. 2014. Т. 55, № 4. С. 846–857.
7. Романовский Р.К., Назарук Е.М. Дихотомия решений функционально-дифференциальных уравнений в пространстве Соболева // Дифференц. уравнения. 2015. Т. 51, № 4. С. 459–471.
8. Романовский Р.К., Назарук Е.М. Дихотомия решений дифференциально-разностных уравнений в пространстве Соболева // Докл. Акад. Наук. 2015. Т. 461, № 4. С. 394–397.
9. Халмош П. Теория меры. М. : ИЛ, 1953. 282 с.
10. Далецкий Ю.Л., Крейн М.Г. Устойчивость решений дифференциальных уравнений в банаховом пространстве. М. : Наука, 1970. 536 с.
11. Лыков А.В. Проблема тепло-массообмена. Минск : Наука и техника, 1976. 312 с.
12. Корнеев С.А. Гиперболические уравнения теплопроводности // Изв. РАН. Сер. Энергетика. 2001. Вып. 4. С. 117–125.

**ON THE ASYMPTOTIC BEHAVIOR OF SOLUTIONS OF THE CAUCHY
PROBLEM FOR SYSTEM OF FUNCTIONAL-DIFFERENTIAL EQUATIONS
OF HYPERBOLIC TYPE**

G.D. Anisimova

Postgraduate Student, e-mail: gdanisimova@gmail.com

Omsk State Technical University, Omsk, Russia

Abstract. For the class of systems specified in the title we prove necessary and sufficient criteria for exponential stability and dichotomy of solutions of the Cauchy problem in terms of the zeros (λ, μ) of determinant of a matrix pencil – symbol of the functional-differential operator in the left part of the system. An illustrative example is considered.

Keywords: transition to the difference Cauchy problem, characterization of the spectrum of resolving operator.

Дата поступления в редакцию: 20.10.2017

АЛГОРИТМ РАСЧЁТА БЛОЧНОГО КРИТЕРИЯ ИНТЕРВАЛЬНОЙ ПРОГНОЗИРУЕМОСТИ ДИНАМИЧЕСКОГО ПОКАЗАТЕЛЯ НА ОСНОВЕ КОЭФФИЦИЕНТА ТАРСИТАНО-ЛОМБАРДО

Ю.М. Краковский¹

д.т.н., профессор, e-mail: kum@stranzit.ru

А.Н. Лузгин²

к.т.н., преподаватель, e-mail: alexln@mail.ru

¹Иркутский государственный университет путей сообщения, Иркутск, Россия

²Иркутский государственный университет, Иркутск, Россия

Аннотация. Предложен и протестирован алгоритм расчёта блочного критерия интервальной прогнозируемости динамического показателя на основе непараметрического коэффициента корреляции Тарситано-Ломбардо. Экспериментально показано, что значение блочного критерия интервальной прогнозируемости, рассчитанное по предыстории динамического показателя, позволяет количественно оценить целесообразность проведения интервального прогнозирования динамического показателя, исходя из его статистических свойств. Интервальное прогнозирование заключается в определении интервала (из двух заранее заданных), в котором будет находиться будущее значение показателя на основе оценок вероятностей этих событий. При этом разделительная граница интервалов задаётся расчётным способом, исходя из статистических свойств динамического показателя. Предложенный алгоритм расчёта блочного критерия интервальной прогнозируемости динамического показателя реализован на языке программирования «R».

Ключевые слова: интервальное прогнозирование, оценка интервальной прогнозируемости, динамический показатель, непараметрическая корреляция, точность интервального прогнозирования.

Введение

В последнее время все большее внимание исследователей уделяется технологиям прогнозирования различных динамических показателей (ДП), как важнейшему фактору, направленному на повышение эффективности принятия управленческих решений в организациях и предприятиях, осуществляющих свою деятельность в условиях неопределённости. Кроме того, наблюдается возрастающий интерес к вероятностным методам прогнозирования [1, 2]. Прежде всего, это объясняется тем, что вероятностные прогнозы позволяют получить

количественную оценку неопределённости самого прогноза (этой оценкой является оценка вероятности будущего события). Как правило, такая задача решается с помощью машинного обучения (machine learning) [3].

Частным и достаточно распространённым случаем вероятностного прогнозирования является вероятностное прогнозирование бинарных исходов, когда в будущий момент времени может произойти только одно из двух возможных событий. Необходимость бинарного прогнозирования часто возникает во многих прикладных областях: метеорологической, энергетической, экономической, социальной, геологической. Так, например, в работе [4] используются данные по суммарному количеству осадков за один день в городе Рединге (Великобритания). В работе строится модель, на основе которой оцениваются вероятности того, будет или нет следующий день дождливым. Из двух оценок вероятностей выбирается максимальная и делается прогноз. Работа [5] посвящена вопросам бинарного прогнозирования цен на электроэнергию на материковой части Испании, а работа [6] — прогнозированию ценовых пиков на рынке электроэнергии Австралии. Применение бинарного прогнозирования в сфере экономики рассматривается в работе [7]. Более подробный обзор прикладных аспектов применения бинарного прогнозирования можно найти в работе [8].

Задачу бинарного прогнозирования можно решать с помощью различных методов прогнозирования. Например, методом логистической регрессии, методом Байеса, методом опорных векторов или случайных лесов [9–11].

Разновидностью бинарного прогнозирования является интервальное прогнозирование (ИП), описанное авторами в работах [12, 13]. Суть ИП заключается в определении интервала из двух заранее заданных интервалов, в котором будет находиться будущее значение ДП, на основе оценок вероятностей этих событий. При этом разделительная граница интервалов задаётся расчётным способом, исходя из статистических свойств ДП.

Задача ИП предполагает построение математических стохастических моделей. В работе [12] применялась вероятностная кластерная модель, а в работе [13] — вероятностная нейронная модель. При построении таких моделей и расчёте их параметров используется доступная предыстория значений ДП, которая непосредственно влияет на качество построенной модели и, как следствие, на точность ИП. Очевидно, что точность прогнозирования является важной характеристикой ИП.

На практике истинная точность ИП будущих значений ДП неизвестна, и для её оценки приходится применять методы кросс-валидации (cross-validation) по ретроспективным (уже известным) значениям ДП [14] (например, такой подход широко практикуется в соревнованиях Kaggle [15]). Если оценка точности ИП приемлема (неприемлема), то метод прогнозирования применять целесообразно (нецелесообразно). Формализация и определение приемлемой точности ИП даны в работе [16]. Главный недостаток здесь состоит в том, что, оценивая точность ИП по ретроспективным значениям ДП методами кросс-валидации, полностью игнорируются статистические характеристики ДП, что может приводить к неадекватным и неприемлемым результатам ИП новых значений этого ДП.

Возникает вопрос: можно ли по предыстории ДП количественно оценить целесообразность осуществления ИП, не осуществляя само ИП на основе соответствующих моделей? Решение данного вопроса подразумевает предварительное изучение (расчёт) некоторых статистических характеристик ДП и разработку некоторого критерия для оценки целесообразности применения ИП на основе соответствующих моделей. Вариант такого критерия под названием «критерий интервальной прогнозируемости» (КИП) был предложен в работе [17] на основе модифицированного рангового коэффициента корреляции Спирмена, с помощью которого оценивалась целесообразность (нецелесообразность) осуществления ИП на основе предложенных моделей [12,13]. При этом авторами была дана рекомендация осуществлять ИП ДП в текущий момент времени только в том случае, если рассчитанное по имеющейся выборке значение КИП не меньше 0,60. Однако в ходе последующих экспериментов [18] с использованием ДП с различными статистическими характеристиками на интервале тестирования были выявлены случаи, когда при значениях КИП значительно меньше 0,60 точность ИП была приемлемой (то есть значения КИП были неадекватны реальности). Для решения данной проблемы и улучшения точности оценок КИП различных ДП авторами был разработан модифицированный вариант, который получил название «блочный критерий интервальной прогнозируемости» (БКИП).

В данной работе описан алгоритм расчёта БКИП и проведено его сравнение с КИП по данным ДП с различными статистическими характеристиками на заданном интервале тестирования. При этом были заранее отобраны такие ДП, где значения КИП дают как адекватную, так и неадекватную характеристику целесообразности ИП. Для проверки точности ИП на заданном интервале тестирования использовалась модель вероятностной нейронной сети.

Учитывая, что алгоритм расчёта БКИП основан на непараметрическом коэффициенте корреляции, предложенном в 2013 году А. Тарситано и Р. Ломбардо (коэффициент корреляции Тарситано-Ломбардо) [19], рассмотрим его подробно.

1. Непараметрический коэффициент корреляции Тарситано-Ломбардо

Любой ДП формализуется как временной ряд:

$$Q = \{q_t; t \in T\}. \quad (1)$$

Здесь q_t — значения ДП, доступные в дискретные моменты времени t ; $t \in T$; $T = \{0, \dots, n-1\}$; n — количество доступных значений.

Выделим из Q два произвольных вектора с равным количеством элементов f и обозначим их так:

$$Q_i^f = \{q_i, \dots, q_{i+f-1}\}, Q_j^f = \{q_j, \dots, q_{j+f-1}\}. \quad (2)$$

Здесь $q_i, \dots, q_{i+f-1}$ — элементы вектора Q_i^f ; $q_j, \dots, q_{j+f-1}$ — элементы вектора Q_j^f ; индекс f — определяет размерность вектора; индексы $i = 0, \dots, n-f$ и

$j = 0, \dots, n - f$ — определяют индексы первого элемента каждого вектора в исходной выборке Q (1). При этом $i \neq j$.

Ранжируем элементы каждого вектора (2) в порядке возрастания и обозначим полученные векторы рангов так:

$$R_i^f = \{r_i, \dots, r_{i+f-1}\}, R_j^f = \{r_j, \dots, r_{j+f-1}\}. \quad (3)$$

Здесь $r_i, \dots, r_{i+f-1}$ — ранги элементов вектора Q_i^f ; $r_j, \dots, r_{j+f-1}$ — ранги элементов вектора Q_j^f .

Далее создадим векторы анти-рангов элементов каждого вектора (2) так:

$$\begin{aligned} A_i^f &= \{a_i = f + 1 - r_i, \dots, a_{i+f-1} = f + 1 - r_{i+f-1}\}, \\ A_j^f &= \{a_j = f + 1 - r_j, \dots, a_{j+f-1} = f + 1 - r_{j+f-1}\}. \end{aligned} \quad (4)$$

Здесь a_{i+f-1} — анти-ранги элементов вектора Q_i^f ; a_{j+f-1} — анти-ранги элементов вектора Q_j^f .

Введем следующую функцию:

$$\mu = \begin{cases} x, & x \geq 1; \\ x^{-1}, & x < 1. \end{cases} \quad (5)$$

Здесь x — положительное вещественное число.

Рассчитаем величины:

$$\begin{aligned} a(R_i^f, A_j^f) &= \frac{1}{f} \sum_{k=0}^{f-1} \mu \left(\frac{r_{i+k}}{a_{j+k}} \right), \quad b(A_i^f, A_j^f) = \frac{1}{f} \sum_{k=0}^{f-1} \mu \left(\frac{a_{i+k}}{a_{j+k}} \right), \\ c(R_i^f, R_j^f) &= \frac{1}{f} \sum_{k=0}^{f-1} \mu \left(\frac{r_{i+k}}{r_{j+k}} \right), \quad d(A_i^f, R_j^f) = \frac{1}{f} \sum_{k=0}^{f-1} \mu \left(\frac{a_{i+k}}{r_{j+k}} \right). \end{aligned} \quad (6)$$

Здесь r_{i+k} — элементы вектора R_i^f ; r_{j+k} — элементы вектора R_j^f ; a_{i+k} — элементы вектора A_i^f ; a_{j+k} — элементы вектора A_j^f .

Дополнительно рассчитаем такую величину:

$$s = \frac{1}{n^2} \cdot \left(\sum_{k=1}^{[f/2]} \frac{n + 1 + k}{k} \right)^2 - 1. \quad (7)$$

Здесь $[\cdot]$ — операция взятия целой части вещественного числа.

Непараметрический коэффициент корреляции Тарситано-Ломбардо между двумя векторами Q_i^f и Q_j^f рассчитывается так:

$$r(Q_i^f, Q_j^f) = \frac{a(R_i^f, A_j^f) \cdot d(A_i^f, R_j^f) - b(A_i^f, A_j^f) \cdot c(R_i^f, R_j^f)}{s}. \quad (8)$$

Здесь $a(R_i^f, A_j^f)$, $b(A_i^f, A_j^f)$, $d(A_i^f, R_j^f)$, $c(R_i^f, R_j^f)$ — величины, рассчитанные по формуле (6); R_i^f — вектор рангов (3) элементов вектора Q_i^f ; R_j^f — вектор

рангов (3) элементов вектора Q_j^f ; A_i^f — вектор анти-рангов (4) элементов вектора Q_i^f ; A_j^f — вектор анти-рангов (4) элементов вектора Q_j^f ; s — величина, рассчитанная по формуле (7).

Данный коэффициент в сравнении с другими подобными непараметрическими коэффициентами корреляции (например, Спирмена или Кенделла) обладает преимуществами, среди которых можно выделить [19]: 1) лучшую чувствительность к коррелированным выборкам различного объёма; 2) лучшую устойчивость к коррелированным выборкам при наличии выбросов. Поэтому он и был выбран авторами при расчётах БКИП.

2. Формализация блочного критерия интервальной прогнозируемости

Преобразуем исходную выборку ДП Q (1) методом конечных разностей первого порядка:

$$s_t = q_t - q_{t-1}, \quad (9)$$

где s_t — преобразованные значения ДП, q_t — исходные значения ДП (1), $t \in [1; n - 1]$. После проведения процедуры преобразования обозначим полученный ДП так:

$$S = \{s_t; t \in \dot{T}\}. \quad (10)$$

Здесь t принимает значения из множества $\dot{T} = 1, \dots, \dot{n}$; $\dot{n} = n - 1$ — объём выборки показателя S (10).

Выделим из S (10) следующий вектор размерностью f и назовём его базовым:

$$S_{\dot{n}-f+1}^f = \{s_{\dot{n}-f+1}, \dots, s_{\dot{n}}\}. \quad (11)$$

Сформируем из S (10) новые выборки (блоки) объёмом m . Объём этих блоков определяется по формуле:

$$m = \left\lceil \frac{\dot{n}}{[\dot{n}/z] + \text{sign}(\dot{n} - z \cdot [\dot{n}/z])} \right\rceil - 1. \quad (12)$$

Здесь z — параметр, ограничивающий количество элементов в блоках так, что $m \leq z$; $\dot{n}$ — объём выборки S (10); $\text{sign}(\cdot)$ — функция знака вещественного числа; $[\cdot]$ — операция взятия целой части вещественного числа.

Нами экспериментально было установлено, что лучшее значение для параметра z находится в диапазоне 450 — 550 значений. В данной работе это значение было задано равным 500 и не менялось на протяжении всех дальнейших экспериментов.

Непосредственно формирование из исходной выборки S (10) блоков с количеством элементов m осуществляется так:

$$S_1 = \{s_1, \dots, s_m\}, S_2 = \{s_{1+g}, \dots, s_{m+g}\}, \dots, S_k = \{s_{1+(k-1) \cdot g}, \dots, s_{m+(k-1) \cdot g}\}. \quad (13)$$

Здесь m — объём выборок, определённый по формуле (12); $g = [m/2]$; $k = 1 + [(\dot{n} - m - 1)/g]$ — число выборок; $[\cdot]$ — операция взятия целой части вещественного числа.

Следует отметить, что при объёме выборки S (10) $\dot{n} \leq z$ будет сформирована единственная выборка $S_1 = s_1, \dots, s_m$, в которой $m = \dot{n} - 1$.

Произвольный вектор размерностью f , взятый из какой-либо выборки i (13), обозначим так:

$$S_{i,j}^f = \{s_{i,j}, \dots, s_{i,j+f-1}\}. \quad (14)$$

Здесь $s_{i,j}, \dots, s_{i,j+f-1}$ — элементы вектора $S_{i,j}^f$; индекс f — определяет размерность вектора; индекс $i = 1, \dots, k$ — номер выборки (13); индекс $j = 1, \dots, m - f + 1$ — номер первого элемента вектора в выборке i .

Применение БКИП основывается на следующей предпосылке: если количественно оценить множественную взаимосвязь между базовым вектором $S_{\dot{n}-f+1}^f$ (11) различной размерности f и векторами $S_{i,j}^f$ (14) блоков (13) на основе предыстории S (10), то можно дать рекомендацию о целесообразности или нецелесообразности ИП показателя Q (1). Отсюда происходит и название критерия.

Взаимосвязь между базовым вектором $S_{\dot{n}-f+1}^f$ (11) и вектором $S_{i,j}^f$ (14) предполагается определять так:

$$\tilde{r}(S_{\dot{n}-f+1}^f, S_{i,j}^f) = \begin{cases} |r(S_{\dot{n}-f+1}^f, S_{i,j}^f)|, & f \neq 1; \\ 1, & f = 1. \end{cases} \quad (15)$$

Здесь $r(S_{\dot{n}-f+1}^f, S_{i,j}^f)$ — непараметрический коэффициент корреляции Тарситано-Ломбардо (8).

БКИП рассчитывается следующим образом.

Фиксируется номер выборки i . Для фиксированного значения f и всех значений $j \in [1; m - f + 1]$ выборки i рассчитывается множество значений $\tilde{r}(S_{\dot{n}-f+1}^f, S_{i,j}^f)$ (15), среди которых находится максимальное:

$$r_{i,max}^f = \max(R_i^f). \quad (16)$$

Здесь $R_i^f = \{\tilde{r}(S_{\dot{n}-f+1}^f, S_{i,1}^f), \dots, \tilde{r}(S_{\dot{n}-f+1}^f, S_{i,m-f+1}^f)\}$ — множество значений $\tilde{r}(S_{\dot{n}-f+1}^f, S_{i,j}^f)$ (15) для фиксированных f и i .

Величины (16) для различных $f = 1, \dots, f_{max}$ образуют множество $R_{i,max} = \{r_{i,max}^1, \dots, r_{i,max}^{f_{max}}\}$, по которому находится медиана:

$$r_{i,med} = med(R_{i,max}). \quad (17)$$

Здесь $med(\cdot)$ — медиана по множеству значений.

Значение f_{max} ограничивает максимальную размерность векторов, используемых при расчётах. Предварительно проведённые эксперименты показали, что для расчёта БКИП по выборкам объёма $m \geq 20$ достаточно использовать векторы с максимальной размерностью f_{max} равной:

$$f_{max} = [\pi \cdot \sqrt{m} + \ln(m)], \quad (18)$$

где m — объём блоков (13), $[\cdot]$ — операция взятия целой части вещественного числа.

Рассчитав множество значений $r_{i,med}$ для всех $i = 1, \dots, k$, БКИП находится так:

$$u_{n-1} = med(R_{med}). \quad (19)$$

Здесь $R_{med} = r_{1,med}, \dots, r_{k,med}$ — множество значений $r_{i,med}$ (17) для заданных значений i .

3. Алгоритм расчёта блочного критерия интервальной прогнозируемости

С учётом выше изложенного алгоритм БКИП в общем виде содержит следующие этапы:

- 1) подготовка исходных данных Q (1);
- 2) преобразование Q (1) методом конечных разностей (9) для получения S (10);
- 3) задание параметра z : $450 < z < 550$;
- 4) вычисление количества элементов в блоках m (12) в зависимости от заданного z ;
- 5) формирование блоков из исходной выборки S (10) с количеством элементов m ;
- 6) вычисление максимальной допустимой размерности векторов f_{max} (18);
- 7) задание номера выборки $i = 1$ и пустого множества $R_{med} = \{\}$;
- 8) задание размерности векторов $f = 1$ и пустого множества $R_{i,max} = \{\}$;
- 9) выделение из S (10) базового вектора S_{n-f+1}^f (11);
- 10) задание $j = 1$ и пустого множества $R_i^f = \{\}$;
- 11) выделение из блока i вектора $S_{i,j}^f$ (14);
- 12) вычисление $\tilde{r}(S_{n-f+1}^f, S_{i,j}^f)$ (15) и добавление его во множество R_i^f ;
- 13) если $j < m - f - 1$, то $j = j + 1$ и возвращаемся на шаг 11; иначе шаг 14;
- 14) вычисление $r_{i,max}^f$ (16) и добавление его во множество $R_{i,max}$;
- 15) если $f < f_{max}$, то $f = f + 1$ и возвращаемся на шаг 9; иначе шаг 16;
- 16) вычисление $r_{i,med}$ (17) и добавление его во множество R_{med} ;
- 17) если $i < k$, то $i = i + 1$ и возвращаемся на шаг 8; иначе шаг 18;
- 18) вычисление значения БКИП u_{n-1} (19).

4. Тестирование блочного критерия интервальной прогнозируемости

Алгоритмы расчёта КИП и БКИП реализованы на свободно распространяемом языке программирования для статистической обработки данных с открытым исходным кодом «R» [20]. Для проверки точности ИП на основе вероятностной нейронной сети был использован программный комплекс «Интервальное прогнозирование нестационарных динамических показателей» [21].

Были выбраны следующие ДП с различными статистическими характеристиками:

1) ежедневное количество новорождённых детей (CN) в городе Квебеке (Канада) [22];

2) ежедневное среднее значение давления над уровнем моря (SP) в городе Мадрасе (Индия)[23];

3) ежедневное среднее значение «возвратного» индекса (IR) акций американской корпорации IBM [24];

4) ежедневное среднее значение температуры (MT) окружающей среды озера Люцерн (Швейцария) [25];

5) значения, полученные с помощью генератора псевдослучайных чисел с нормальным законом распределения вероятностей при математическом ожидании равном 0 и дисперсии равной 1 (NR);

6) последовательность кумулятивных сумм значений показателя NR (CNR).

Исходный объём выборки каждого ДП был различен и равен $n = 500, 1000, 1500, 2000, 2500$ значений.

Точность ИП проверялась для w последних значений каждого ДП, при $w = [0.1 \cdot n]$ (интервал тестирования). Здесь $[\cdot]$ — операция взятия целой части вещественного числа. То есть для тестирования использовалось 10% значений от всего объёма взятой выборки. Для этих же самых значений рассчитывались значения КИП и БКИП. При этом предыдущие значения ДП в объёме равном $n - j, j = w, \dots, 1$ использовались для построения и обучения вероятностной нейронной сети, в расчётах КИП и БКИП.

Для характеристики точности ИП использовался критерий:

$$pl = l / (l + e). \quad (20)$$

Здесь pl — доля оправдавшихся прогнозов; l — число оправдавшихся прогнозов; e — число ошибочных прогнозов; $0 \leq pl \leq 1$. pl — это мера точности ИП. Чем больше это значение — тем лучше. Мы будем считать [16], что точность ИП приемлема, если $pl > 0,6$ (21).

Для тестирования адекватности оценок КИП и БКИП использовались их средние значения на интервале тестирования:

$$\bar{k} = \left(\sum_{i=1}^w k_{n-w-1+i} \right) / w, \quad \bar{u} = \left(\sum_{i=1}^w u_{n-w-1+i} \right) / w. \quad (21)$$

Здесь $k_{n-w-1+i}$ — значения КИП на интервале тестирования; $\bar{k}$ — среднее значение на интервале тестирования; $u_{n-w-1+i}$ — значения БКИП на интервале тестирования; $\bar{u}$ — среднее значение БКИП на интервале тестирования.

Напомним, что при осуществлении ИП на основе вероятностной нейронной сети необходимо задать следующие параметры: f — размерность обучающих векторов вероятностной нейронной сети; p — время упреждения; α — параметр, влияющий на границу интервалов при ИП. Эти параметры были заданы такими: $f = 5$, $p = 1$, $\alpha = 0.0$. Для алгоритма расчёта БКИП был установлен параметр $z = 500$.

В таблице 1 приведены средние значения КИП ($\bar{k}$), БКИП ($\bar{u}$) и значения pl (21).

Таблица 1. Результаты тестирования точности ИП, КИП и БКИП

ДП	n	500	1000	1500	2000	2500
CN	pl	0.75	0.75	0.77	0.75	0.75
	$\bar{k}$	0.62	0.64	0.63	0.62	0.61
	$\bar{u}$	0.84	0.84	0.83	0.81	0.81
SP	pl	0.74	0.78	0.79	0.80	0.80
	$\bar{k}$	0.81	0.79	0.79	0.78	0.78
	$\bar{u}$	0.87	0.87	0.87	0.86	0.85
IR	pl	0.60	0.62	0.66	0.67	0.69
	$\bar{k}$	0.23	0.15	0.15	0.13	0.12
	$\bar{u}$	0.65	0.63	0.66	0.66	0.66
MT	pl	0.62	0.72	0.71	0.68	0.66
	$\bar{k}$	0.37	0.32	0.31	0.31	0.31
	$\bar{u}$	0.64	0.65	0.64	0.63	0.64
NR	pl	0.66	0.68	0.67	0.66	0.66
	$\bar{k}$	0.19	0.17	0.14	0.12	0.10
	$\bar{u}$	0.65	0.62	0.63	0.60	0.60
CNR	pl	0.40	0.44	0.45	0.45	0.48
	$\bar{k}$	0.21	0.14	0.12	0.10	0.10
	$\bar{u}$	0.56	0.57	0.55	0.57	0.57

Как видно из приведённой таблицы для показателей IR, MT и NR точность прогнозирования приемлема, однако значение КИП характеризует эти показатели как неприемлемые (значения выделены жирным шрифтом) с точки зрения осуществления ИП. Напротив, во всех случаях БКИП даёт адекватные оценки целесообразности прогнозирования ДП: там, где точность ИП менее 0,6, значение БКИП также менее 0,6 и наоборот. Также по данным таблицы 1 видно, что с увеличением среднего значения БКИП точность ИП повышается. Это позволяет предложить БКИП вместо КИП для проверки целесообразности или нецелесообразности ИП ДП как более точный и адекватный критерий. При этом как и для КИП рекомендация остаётся прежней и заключается в том, что значение БКИП должно быть не менее 0,6. В противном случае проводить ИП нецелесообразно.

Заключение

В работе был предложен блочный критерий интервальной прогнозируемости динамического показателя на основе непараметрического коэффициента корреляции Тарситано-Ломбардо и соответствующий алгоритм его расчёта. Экспериментально показано, что значение блочного критерия интервальной прогнозируемости, рассчитанное по предыстории динамического показателя, позволяет количественно оценить, целесообразно или нет проводить интервальное прогнозирование динамического показателя, исходя из его статистических свойств. Рекомендовано осуществлять интервальное прогнозирование динамического показателя в текущий момент времени только в том случае, если рассчитанное по имеющейся выборке значение блочного критерия интервальной прогнозируемости не меньше 0,6. Рекомендовано на практике вместо обычного критерия интервальной прогнозируемости применять блочный критерий интервальной прогнозируемости как более точный и адекватный в своих оценках для различных динамических показателей.

ЛИТЕРАТУРА

1. Elliot G., Lieli R.P. Predicting binary outcomes // J Econom. 2013. No. 174. P. 15–26.
2. Lahiri K., Yang L. Forecasting binary outcomes // Handb Econ Forecast. 2013. No. 2. P. 1025–1106.
3. James G., Witten D., Hastie T., Tibshirani T. An Introduction to Statistical Learning with Applications in R. Springer Texts in Statistics, 2013. 426 p.
4. Little M.A. Generalized Linear Models for Site-Specific Density Forecasting of UK Daily Rainfall // Monthly Weather Review. 2009. No. 37(3). P. 1029–1045.
5. Anbazhagan S., Kumarappan N. Binary classification of day-ahead deregulated electricity market prices using neural networks // Power India Conference. 2013.
6. Eichler M., Grothe O., Manner H., Turk D. Models for short-term forecasting of spike occurrences in Australian electricity markets: a comparative study // Journal of Energy market. 2014. V. 7, No. 1. P. 55–81.

7. Nyberg H. Studies on binary time series models with applications to empirical macroeconomics and finance. 2010. URL: <https://helda.helsinki.fi/bitstream/handle/10138/23519/studieso.pdf?sequence=1> (дата обращения: 17.01.2017).
8. Краковский Ю.М., Лузгин А.Н. Прикладные аспекты применения интервального прогнозирования в системном анализе // Современные технологии. Системный анализ. Моделирование. 2017. N. 2(54). С. 115-121.
9. Crawford M. M., Ham J., Chen Y., Ghosh, J. Random forests of binary hierarchical classifiers for analysis of hyperspectral data // IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data. 2003. P. 337-345.
10. Ng A.Y., Jordan M. I. On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes // Advances In Neural Information Processing Systems. 2002. URL: <https://ai.stanford.edu/~ang/papers/nips01-discriminativegenerative.pdf> (дата обращения: 06.05.2017).
11. Platt C.J. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. 1999. URL: citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.41.1639&rep=rep1&type=pdf (дата обращения: 12.03.2017).
12. Краковский Ю.М., Лузгин А.Н. Алгоритм интервального прогнозирования динамических показателей на основе робастной вероятностной кластерной модели // Наука и образование. 2016. No. 11. С. 113-126. URL: <http://technomag.neicon.ru/doc/849839.html> (дата обращения: 20.08.2017).
13. Краковский Ю.М., Лузгин А.Н. Алгоритм интервального прогнозирования динамических показателей на основе вероятностной нейросетевой модели // Современные технологии. Системный анализ. Моделирование. 2016. № 4(50). С. 126-132.
14. Arlot S., Celisse A. A survey of cross-validation procedures for model selection // Statistics Surveys. 2010. No. 4. P. 40-79.
15. Kaggle: Your Home for Data Science. URL: <https://www.kaggle.com/competitions> (дата обращения: 11.09.2017)..
16. Краковский Ю.М., Лузгин А.Н. Стохастический критерий оценки приемлемой точности вероятностного бинарного прогнозирования динамических показателей // Вестник ВГУ, серия: Системный анализ и информационные технологии. 2017. № 2. С. 98-104. URL: <http://www.vestnik.vsu.ru/pdf/analiz/2017/02/2017-02-15.pdf> (дата обращения: 17.09.2017).
17. Краковский Ю.М., Лузгин А.Н. Оценка прогнозируемости динамических показателей на основе коэффициента ранговой корреляции // Наука и образование. 2016. № 9. С. 60-73. URL: <http://technomag.neicon.ru/doc/845015.html> (дата обращения: 22.04.2017).
18. Лузгин А.Н. Предпосылки модификации коэффициента интервальной прогнозируемости // Actualscience. 2017. Т. 3, № 3. С. 162-163.
19. Tarsitano A., Lombardo R. A Coefficient of Correlation Based on Ratios of Ranks and Anti-ranks // J Econ and Statis. 2013. V. 233, No. 2. P. 206-224.
20. The R project of statistical computing. URL: <http://www.r-project.org> (дата обращения: 05.05.2017).
21. Лузгин А.Н. Программный комплекс «Интервальное прогнозирование нестационарных динамических показателей» // Свидетельство о государственной регистрации программы для ЭВМ №2015617751 от 22.07.2015.

22. DataMarket. URL: <https://datamarket.com/data/> (дата обращения: 15.04.2017).
23. Madras Monthly Sea Level, CRU. URL: http://www.comp-engine.org/timeseries/time-series_data/data-11114/ (дата обращения: 15.04.2017).
24. Center for Research in Security Prices. URL: <http://www.crsp.com> (дата обращения: 15.04.2017).
25. Air Quality Data of Switzerland. URL: <https://cran.r-project.org/web/packages/SwissAir/> (дата обращения: 15.04.2017).

CALCULATION ALGORITHM OF A BLOCK CRITERION OF THE DYNAMIC INDICATOR INTERVAL FORECASTABILITY BASED ON TARSIANO-LOMBARDO'S COEFFICIENT

Y.M. Krakovsky¹

Dr.Sc. (Eng.), Professor, e-mail: kum@stranzit.ru

A.N. Luzgin²

Ph.D. (Eng.), Instructor, e-mail: alexln@mail.ru

¹Irkutsk State University of Railway Transport, Irkutsk, Russia

²Irkutsk State University, Irkutsk, Russia

Abstract. Calculation algorithm of a block criterion of the interval forecastability for dynamic indicator based on Tarsitano-Lombardo's nonparametric correlation coefficient was proposed and tested. It was experimentally shown that the value of the block criterion of the interval forecastability, which was calculated from prehistory of dynamic indicator values, makes it possible to quantitatively estimate the appropriateness of performing interval forecasting of the dynamic indicator based on its statistical properties. The interval forecasting is to determine the interval from two predetermined intervals in which the future value of the indicator will be located. The forecasting is based on probability estimates of these events. In this case, the separation division of the intervals is set by the calculation method based on statistical characteristics of the dynamic indicator. The proposed calculation algorithm of the block criterion of the dynamic indicator interval forecastability was implemented in the programming language "R".

Keywords: interval forecasting, interval predictability estimating, dynamic indicator, nonparametric correlation, interval forecasting accuracy.

Дата поступления в редакцию: 05.10.2017

УСТРАНЕНИЕ ИЗБЫТОЧНОСТИ ДАННЫХ СЕЙСМИЧЕСКОГО МОНИТОРИНГА ГИДРОРАЗРЫВА ПЛАСТА

Е.В. Рабинович

профессор, д.т.н., e-mail: erabinovich1952@gmail.com

П.И. Вайнмастер

аспирант, e-mail: wmaster91@mail.ru

Г.С. Шефель

доцент, к.ф.-м.н., e-mail: ligr1x@yahoo.com

Новосибирский государственный технический университет, Новосибирск, Россия

Аннотация. Важным этапом мониторинга гидроразрыва пластов является построение графической модели трещиноватой зоны. Однако реальные данные, регистрируемые в процессе гидроразрыва, характеризуются большой интенсивностью сейсмической активности в области мониторинга, что затрудняет корректное построение такой модели. Поэтому актуальной является задача устранения избыточности сейсмических данных. В статье рассматривается оригинальный алгоритм, позволяющий устранить избыточность за счёт выделения ограниченного набора локальных гипоцентров сейсмической активности. В основе алгоритма лежат методы кластерного и факторного анализа. Факторный анализ использован для выявления схожих характеристик и сокращения количества переменных, используемых для описания сейсмического источника. Кластерный анализ применён для обнаружения «облаков» сейсмической эмиссии и выделения их локальных гипоцентров. Полученные результаты использованы для построения трёхмерной графической модели трещины.

Ключевые слова: гидравлический разрыв пласта, трещина, кластерный и факторный анализ.

Введение

Гидравлический разрыв пласта (ГРП) — эффективный метод повышения продуктивности нефтяных скважин. Метод заключается в создании искусственной трещины в целевом пласте, которая служит каналом для притока добываемого флюида из продуктивной части пласта в призабойную зону скважины.

Создание трещины осуществляется путём закачки в пласт специальной жидкости с расклинивающим агентом под давлением, значительно превышающим горное давление. Благодаря такому воздействию в пласте происходит образование трещины, которая в процессе проведения ГРП увеличивается в

направлениях пониженной прочности породы. После снятия избыточного давления и прекращения закачки жидкости разрыва трещина остаётся в закреплённом состоянии благодаря расклинивающему агенту (проппанту), создавая высокопроводящий канал для поступления в скважину дополнительного флюида.

Проведение ГРП является дорогостоящей операцией, сопряжённой с риском получения негативных последствий. Так, при превышении критического давления разрыва глинистых экранов возможен гидродинамический контакт добывающей скважины с этими пластами или с нагнетательными скважинами. Такой контакт приводит к резкому обводнению добываемого флюида, вследствие чего скважина становится непригодной к работе. Чтобы избежать риска обводнения, специалистам по работе со скважиной необходимо знать форму и пространственное расположение трещины гидроразрыва.

Существующие математические расчёты не позволяют с достаточной точностью и достоверностью определить параметры раскрывающихся трещин, а бурение исследовательских скважин с отбором кернов является весьма дорогостоящим. Поэтому возникла актуальная необходимость разработки новых средств для определения формы, размеров и динамики образования трещин ГРП.

Одним из таких средств является технология наземного сейсмического мониторинга ГРП, предложенная и описанная в работах [1, 2]. Технология базируется на локации сейсмических сигналов наземной нерегулярной системой приёма и дальнейшей их обработке с целью определения параметров трещины. В данной работе рассматривается один из важных аспектов данной технологии — устранение избыточности микросейсмических данных и определение гипоцентров сейсмической эмиссии.

1. Постановка задачи

Острые трещины, разрываемое при ГРП, имеет размер много меньше расстояния до поверхности приёма. Разрыв породы с высокой степенью достоверности можно считать точечным источником сейсмических импульсов. Это подтверждается регистрацией сейсмоприёмниками возбуждаемых при разрывах пласта волн в виде последовательности сейсмических импульсов.

Регистрация сейсмических импульсов осуществляется наземной нерегулярной системой приёма и представляет собой антенну, состоящую из n одинаковых сейсмоприёмников (регистрирующих датчиков), расположенных в пространстве так, чтобы шумы на их входах были взаимно независимыми. Антенна «сканирует» зону обнаружения, зарегистрированные в ходе гидроразрыва колебания записываются сеймостанцией.

Для определения местоположения сейсмических источников на уровне продуктивного пласта формируется виртуальная трёхмерная сетка с шагом 1 м и с центром, совпадающим с координатами забоя. От каждого узла сетки вычисляются времена распространения сейсмической волны до каждого сейсмоприёмника по формуле:

$$t_k = \frac{\sqrt{h^2 + L_k^2}}{v}, k = 1, \dots, n,$$

где h — расстояние от узла до поверхности, L_k — расстояние между проекцией источника на поверхность земли и k -м сейсмоприёмником, v — средняя скорость распространения сейсмической волны в среде.

С целью фокусировки антенны на каждом узле виртуальной сетки и повышения отношения сигнал / шум осуществляется накопление энергии сейсмотрасс, записанных всеми приёмниками, с учётом рассчитанных временных сдвигов (задержек). Каждая суммарная сейсмограмма характеризует сейсмический сигнал, полученный от одного узла виртуальной трёхмерной сетки за некоторый временной отрезок проведения ГРП. На основе полученных данных осуществляется обнаружение искомого сейсмического импульса с помощью квадратурного корреляционного [3] или статистического [4] алгоритмов и, если импульс найден, сопоставление данного узла сетки с точечным сейсмическим источником. Каждый такой источник характеризуется координатами X , Y в горизонтальной плоскости пласта, глубиной расположения источника H , максимальной амплитудой колебания A и моментом времени T , в который была зарегистрирована максимальная амплитуда.

Однако реальные данные, регистрируемые в ходе гидроразрыва, характеризуются большой интенсивностью сейсмической активности во всём объёме области мониторинга. Такая избыточность может быть вызвана резонансными свойствами среды и повторным возникновением сейсмических сигналов в тех областях, где уже зарегистрирована трещина при увеличении её толщины. Избыточность сейсмических данных затрудняет корректное построение графической модели трещины. Поэтому актуальной является задача устранения избыточности микросейсмических данных, в частности — выделение гипоцентров сейсмической активности среди набора импульсных источников.

2. Теория

В процессе распространения трещины, как правило, образуются локальные «облака» сейсмической эмиссии, включающие в себя набор близко расположенных источников со сходными характеристиками. Построение графической модели трещины может обеспечить знание координат гипоцентров таких облаков сейсмичности.

В данной работе предложен оригинальный алгоритм определения координат гипоцентров, основанный на применении методов факторного и кластерного анализа.

Кластерный анализ является способом разбиения многомерных объектов на однородные группы. С его помощью осуществляется выделение «облаков» импульсных сейсмических источников. Центры «масс» кластеров используются в качестве гипоцентров, набор которых является достаточным для дальнейшей обработки. Таким образом, кластерный анализ устраняет избыточность наблюдений, мешающую построению адекватной графической модели.

Факторный анализ, в свою очередь, является средством сокращения количества переменных, используемых для описания объекта. С его помощью осуществляется предварительная подготовка данных к кластеризации, заключающаяся в устранении влияния схожих характеристик, искажающих результат кластерного анализа, и ускорении процедуры обработки данных.

Как показано в [5], схожие характеристики при кластеризации получают больший вес, т. к. расстояния между кластерами вычисляются из разностей по каждой переменной. Вследствие этого результат кластеризации искажается: преимущественно учитываются схожие параметры с большим весом. В исследуемой задаче положение сейсмического источника задаётся координатами (X, Y, H) , влияние которых может подавить другие параметры объекта (A, T) и исказить итоговый результат. Для устранения такого негативного воздействия применён факторный анализ.

Ускорение обработки данных особенно актуально при работе с сейсмическими сигналами, т. к. количество источников сейсмических импульсов составляет десятки тысяч, а их обработка должна осуществляться в режиме «реального времени». Применён критерий «отсева» факторов, который позволяет, с одной стороны, оставить только набор значимых факторов (содержащих полезную информацию), а с другой — отсеять факторы, принципиально не влияющие на результаты исследования.

В рамках данной работы также исследован другой вариант применения факторного анализа — для устранения выбросов на завершающем этапе алгоритма (после кластеризации).

Для того чтобы осуществить предварительную обработку сейсмических данных в соответствии с описанными требованиями, был использован вариант факторного анализа — метод главных компонент (МГК). Данный метод используется во многих прикладных задачах и зачастую применяется в «тандеме» с кластерным анализом [5].

Суть метода заключается в следующем. Совокупность исследуемых объектов представляется в виде точек в n -мерном пространстве, где n — число исследуемых признаков. Осуществляется построение прямой линии, «пронизывающей» облако точек так, что проекция точек на неё имеет максимальное среднеквадратичное отклонение (СКО). Так образуется первая главная компонента. На следующем шаге определяется новая прямая, ортогональная первой и также максимизирующая СКО проекции точек — вторая главная компонента. Процесс продолжается до тех пор, пока количество прямых не достигнет размерности пространства n . Таким образом осуществляется переход к новой системе координат, где переменными являются факторы — линейные комбинации исходных признаков объекта.

Для оценки качества результатов факторного анализа мы руководствовались двумя основными характеристиками. Первая характеристика — это «объяснённая» фактором дисперсия [6] — доля от общей дисперсии данных, которую составляет дисперсия фактора. Объяснённая дисперсия характеризует разброс данных вдоль координатной оси, заданной фактором: чем она выше, тем больше информации об исследуемом объекте несёт фактор. Вторая важная характери-

стика — матрица факторных нагрузок, содержащая корреляции между факторами и исходными переменными. На её основе делается вывод о связи между новыми переменными (факторами) и исходными параметрами объекта.

Как уже упоминалось, каждый сейсмический источник характеризуется пятью параметрами: X , Y , H , A , T . В результате применения МГК были предварительно определены пять новых факторов, среди которых проводилась процедура отсева. Для этого был использован критерий Кайзера: если величина дисперсии фактора меньше дисперсии исходной переменной, то фактор «отсеивается» как помеховая составляющая. Важно отметить, что предварительно данные были стандартизованы так, что среднее значение каждой переменной стало равным 0, а дисперсия — 1. Поэтому, согласно критерию Кайзера, отсеивались факторы с дисперсией меньшей единицы.

Из пяти факторов были отсеяны два, поскольку их дисперсия оказалась 0,74 и 0,75. Три оставшихся фактора, дисперсия которых превысила единицу, были сохранены. Дисперсия этих факторов оказалась выше дисперсии исходных переменных, благодаря тому, что соответствующие оси лучше описывают разброс данных. Иными словами, СКО проекции данных на оси факторов оказалось больше, чем СКО проекции данных на оси исходных переменных.

В результате, был осуществлён переход от множества переменных (X , Y , H , A , T) к множеству факторов (F_1 , F_2 , F_3).

Матрица факторных нагрузок для факторов F_1 , F_2 , F_3 приведена в таблице 1. Жирным шрифтом выделены корреляции (нагрузки), превышающие по модулю величину «порога достоверности» — минимального значения корреляции, при котором корреляция считается значимой (0,7).

Таблица 1. Факторные нагрузки

Переменная/фактор	F_1	F_2	F_3
X	-0,776800	0,164289	0,026227
Y	-0,751429	-0,150475	0,025240
H	0,001522	0,003128	0,999430
A	0,253022	-0,773072	0,004291
T	-0,306739	-0,715682	0,014660
Дисперсия фактора	1,26174	1,159484	1,050418
Доля объяснённой дисперсии (%)	26,5235	23,1897	21,0084

Отрицательные корреляции говорят о том, что ось соответствующей главной компоненты (фактора) и исходная ось направлены в разные стороны. Подобный эффект проявляется при ортогональном вращении факторов в рамках МГК, и никаких выводов на основании знака корреляции делать не следует.

На рис. 1 приведён трёхмерный график факторных нагрузок, построенный на основе таблицы 1. Он отражает расположение исходных переменных в новой системе координат, заданной факторами. Координаты по осям $Factor_1$, $Factor_2$

и $Factor_3$ представляют собой величину корреляции между переменными и соответствующими факторами.

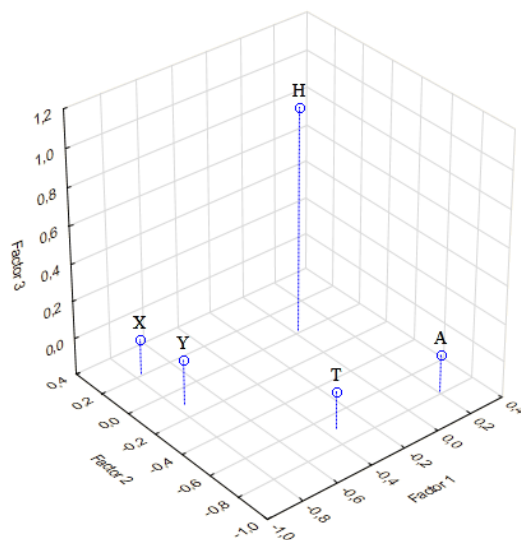


Рис. 1. График факторных нагрузок

Для оценки результатов факторного анализа были проанализированы дисперсии, объяснённые факторами, и факторные нагрузки (табл. 1).

Фактор F_1 имеет наибольшее значение дисперсии — он объясняет 26,5 % от общей дисперсии выборки, т. е. сильнее всего влияет на разброс данных. При этом высокие нагрузки у этого фактора (0,78 и 0,75 соответственно) обнаружены с переменными X и Y , определяющими координаты точечного источника внутри трёхмерной сетки пласта. Таким образом, первый фактор характеризует пространственное положение точечного сейсмического источника.

Фактор F_2 в свою очередь наиболее коррелирован с переменными A и T (соответствующие коэффициенты корреляции равны 0,77 и 0,72). Можно сделать вывод, что данный фактор характеризует параметры динамики импульса, возбуждаемого источником (момент достижения максимума колебания и соответствующую амплитуду). Эти параметры действительно взаимосвязаны: области трещиноватой зоны, сформированные в близкие моменты времени, как правило, имеют схожие значения амплитуды колебаний, что и объясняет высокие коэффициенты корреляции между характеристиками. Доля объяснённой фактором F_2 дисперсии составила 23,19 % от общей дисперсии данных, т. е. он является вторым по значимости влияния на разброс данных (после фактора F_1).

Наконец, по фактору F_3 значимая корреляция была обнаружена только с переменной H , определяющей глубину, на которой располагается рассматриваемая плоскость внутри пласта (0,99). Т. к. величина корреляции оказалась близка к 1, можно сделать вывод, что выделенный фактор почти полностью повторяет исходную переменную и не подвергся изменениям при обработке. Это подтверждается и тем, что доля дисперсии фактора F_3 от общей дисперсии

данных составила 21 %, или почти $1/5$, — т. е. доля сопоставима с вкладом одной переменной из пяти, присутствующих до обработки.

Суммарная доля дисперсии, объяснённой тремя выделенными факторами F_1 , F_2 и F_3 , достигла 71 % от общей дисперсии исходного набора данных. Применение критерия Кайзера позволило отсеять два «незначимых» фактора, которые объясняли оставшиеся 29 % от общей дисперсии. В рамках МГК такие факторы интерпретируются как помеховая (шумовая) составляющая, не содержащая полезной информации об исследуемом объекте. При этом размерность данных сократилась на 40 % (с пяти признаков до трёх).

Таким образом, в результате применения факторного анализа было достигнуто предварительное устранение избыточности, позволяющее ускорить следующий далее процесс кластеризации.

Учитывая вышесказанное, факторы F_1 , F_2 , F_3 были выбраны в качестве замены исходных переменных при дальнейшей обработке.

Для выделения «облаков» сейсмичности были применены средства кластерного анализа. Очевидным вариантом представлялось выделение кластеров среди всей области локации, однако такое решение имеет серьёзные недостатки. В ходе экспериментов было обнаружено, что при таком подходе центры масс кластеров оказываются сконцентрированными близко к забою скважины, а удалённые области не получают представления. Альтернативным решением является разбиение области локации на множество вспомогательных областей меньшего размера и выделение естественных кластеров внутри каждой из них. Так как вспомогательные области распределены по всему объёму исследуемой области (и, особенно, на границах трещиноватой зоны), упомянутый недостаток удалось устранить.

Вспомогательные области были представлены в виде параллелепипедов, имеющих управляемые параметры — длину, ширину и высоту, и покрывающих всю область локации. Чтобы избежать потери информации, которая может находиться на границе параллелепипедов, параллелепипеды частично перекрывают друг друга. Пример такого разбиения области локации представлен на рис. 2. Чёрные точки внутри представляют точечные источники сейсмических импульсов, а трубка в центре — скважину.

В качестве метода кластеризации был использован агломеративный алгоритм, заключающийся в последовательном объединении точек в кластеры и построении кластерного дерева. Важным этапом является выбор метрики для расчёта парных расстояний между объектами, а также стратегии кластеризации, использование которых окажется наиболее эффективным в рамках рассматриваемой предметной области.

Были исследованы два метода расчёта расстояния между объектами: метод наиболее удалённых соседей (Complete-link) и метод попарного среднего (Average-link).

Метод удалённых соседей определяет межклассовое расстояние как наибольшую дистанцию между объектами двух кластеров. Формула для расчёта дистанции $d(r, s)$ между двумя кластерами r и s имеет следующий вид:

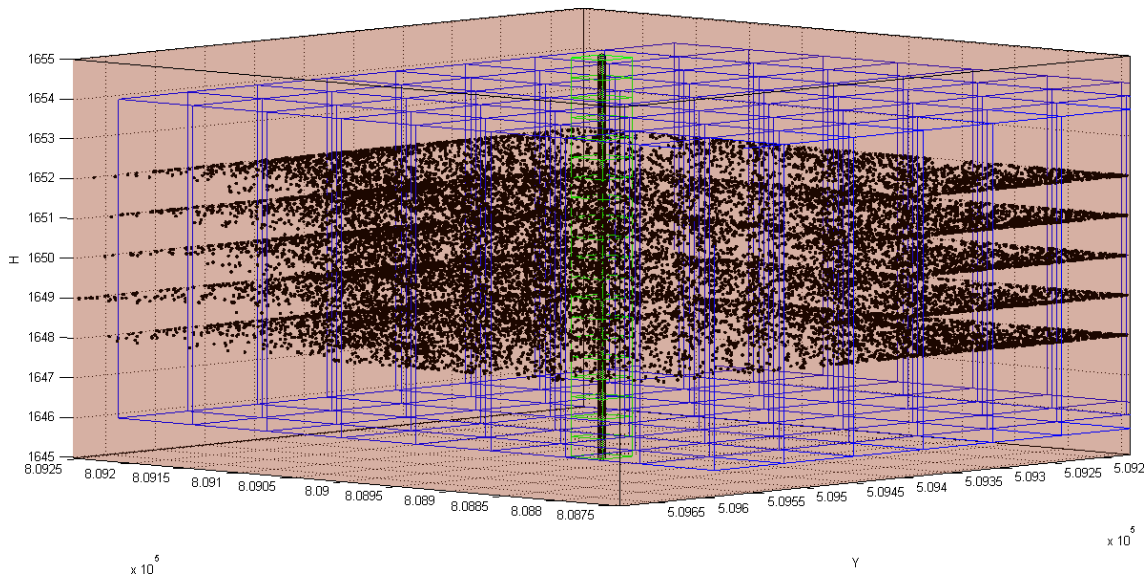


Рис. 2. Разбиение области локации на вспомогательные области (параллелепипеды)

$$d(r, s) = \max[\text{dist}(X_{ri}, X_{sj})],$$

где $i = 1, \dots, n_r$, $j = 1, \dots, n_s$.

Таким образом, при построении кластерного дерева дистанция между двумя кластерами рассчитывается по максимальному расстоянию между их объектами. Такой подход положительно сказывается на охвате всей области локации потому, что способствует формированию кластеров на границах трещиноватой зоны, т. к. расстояние между ними в этом случае максимально.

Метод невзвешенного попарного среднего, в свою очередь, является наиболее универсальным и определяет расстояние $d(r, s)$ между двумя кластерами как среднее расстояние между всеми парами объектов в них:

$$d(r, s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \text{dist}(x_{ri}, x_{sj}).$$

В отличие от метода удалённых соседей, он не способствует выделению кластеров в приграничных зонах области локации, но, тем не менее, распределяет кластеры наиболее равномерно.

Для сравнения описанных методов была проведена серия экспериментов. Смоделирована тестовая выборка источников сейсмической активности, предварительно разделённых на несколько групп по различным параметрам: в первом опыте — по пространственным характеристикам, представленным факторами F_1 и F_3 , во втором — по характеристикам возбуждаемого сейсмического импульса, представленным фактором F_2 , а в третьем — по совокупности всех параметров. Для каждого из опытов были рассмотрены два случая: когда количество классов равно трём и данные в них расположены «компактно» (так,

что классы хорошо различимы между собой) и когда количество групп равно десяти, а данные в некотором смысле «размыты» между группами.

В качестве меры для расчёта расстояния между точками были использованы Евклидова мера и мера Чебышёва. Мера Евклида представляет расстояние между двумя точками x и y в виде геометрического расстояния в многомерном пространстве:

$$d(X_j, X_i) = \sqrt{\sum_{k=1}^N (x_{ki} - x_{kj})^2}.$$

Мера Чебышёва, в свою очередь, определяет расстояние как максимальную разность между двумя аналогичными параметрами объектов:

$$d(X_j, X_i) = \sum_{k=1}^N |x_{ki} - x_{kj}|.$$

Результаты сравнения сведены в таблицу 2.

Таблица 2. Результаты сравнения методов кластеризации

Метод	Мера расстояния	Среднее кол-во кластеров при «компактном» разбиении	Среднее кол-во кластеров при «размытом» разбиении	Средний % объектов, верно отнесённых к кластерам
Complete-Link	Euclidean	4	16	68%
Complete-Link	Chebyshev	3	12	92%
Average-Link	Euclidean	3	11	94%
Average-Link	Chebyshev	4	14	81%

Из таблицы видно, что результаты, наиболее близкие к исходному разбиению (3 группы при «компактном» разбиении и 10 групп при «размытом»), были получены во втором и третьем опытах. Здесь также наиболее высоким оказался средний процент объектов, верно отнесённых к кластерам — 92 % и 94 % соответственно. Также можно сделать вывод, что метод Completelink лучше работает с мерой Чебышёва, а метод Averagelink — с Евклидовой мерой.

Учитывая, что наилучшие результаты кластеризации были получены в третьем эксперименте, в качестве стратегии кластеризации был выбран метод Averagelink, а в качестве меры расстояния между объектами — мера Евклида.

Стоит заметить, что в некоторых случаях были обнаружены «лишние» (по сравнению с исходным разбиением объектов) кластеры. Чтобы избежать их появления, был введён критерий оптимального разбиения, описываемый формулой:

$$d_{\min}(C_i, C_j) > \max\{h(C_i), h(C_j)\},$$

где $d_{\min}(C_i, C_j)$ — минимальное расстояние между парой объектов из двух кластеров C_i и C_j ; $\max\{h(C_i), h(C_j)\}$ — максимальное из расстояний среди объектов одного кластера. Оптимальным считается то количество кластеров, при котором все кластеры удовлетворяют описанному условию. Подробное рассмотрение критериев оптимального кластерного разбиения является темой отдельной статьи и описано в [7].

В результате кластеризации был получен набор локальных «облаков» сейсмичности. Для каждого из них был найден гипоцентр — центр масс соответствующего кластера (в пространстве исходных характеристик объектов — X , Y , H , A , T). В результате удалось превратить исходное количество узлов виртуальной сетки локации, представляющих источники сейсмических импульсов (приблизительно 20000), в набор гипоцентров (приблизительно 700), используемых для графического построения модели зоны трещиноватости. При этом каждый узел характеризуется пятью параметрами, а каждый гипоцентр — тремя, что в совокупности приводит к сокращению объёма обрабатываемых данных почти в 50 раз.

На завершающем этапе был вновь применён факторный анализ — теперь к полученному набору гипоцентров, чтобы проанализировать изменение факторных нагрузок после кластеризации, а также выявить и устранить выбросы в наборе гипоцентров.

Таблица 3. Факторные нагрузки после обнаружения гипоцентров

Переменная/фактор	F_1	F_2	F_3
X	0,781574	0,186051	0,206770
Y	0,750792	-0,185946	-0,003375
H	0,052999	0,016990	-0,973803
A	0,205652	-0,797268	0,161902
T	0,224919	-0,737341	-0,044297
Дисперсия фактора	1,273637	1,240787	1,019232
Доля объяснённой дисперсии (%)	25,4727	24,8157	20,3846

Сравнение табл. 3 с табл. 1 показывает, что количество факторов не изменилось. Не изменились и корреляции факторов с исходными переменными. Так, у фактора F_1 высокие нагрузки обнаружены с переменными X и Y (пространственные характеристики); у фактора F_2 — с переменными A и T (параметры импульса), у фактора F_3 — с H , близкая к единице (глубина трещины). При этом нагрузки между фактором F_2 и переменными A и T незначительно выросли (на 0,02), а нагрузка между фактором F_3 и переменной H уменьшилась (также на 0,02). Процент дисперсии, объяснённой новыми факторами, остался тем же — 71 %.

После обработки, заменившей избыточные данные, представляющие «облака» сейсмичности нечётких очертаний, набором ограниченного числа гипоцентров, сохранились все существенные зависимости между переменными, характеризующими источники сейсмических импульсов.

Для устранения выбросов по каждому из факторов был применён критерий Шовене. Всего было найдено 6 выбросов, связанных с фактором F_2 , что составляет 0,01 % от общего количества данных. Проанализировав характеристики соответствующих гипоцентров, было обнаружено, что выбросы связаны с высокими значениями амплитуды в этих узлах сетки локации.

3. Результаты экспериментов

Трёхмерная модель трещины, построенная с помощью триангуляции Делоне на исходных данных, полученных в результате применения описанного алгоритма, приведена на рис. 3. Область локации представляет параллелепипед с квадратным основанием со сторонами по 512 м и высотой 2 м, трещина ориентирована в нем наклонно.

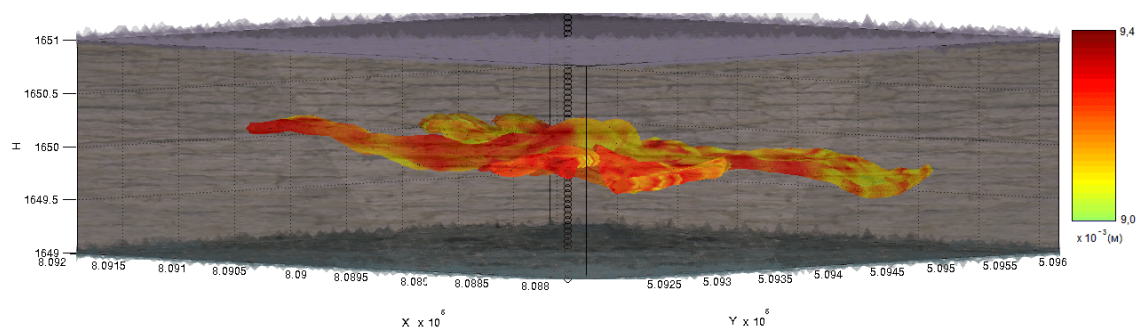


Рис. 3. Трёхмерная модель зоны трещиноватости целевого пласта

Фрагменты трещины имеют раскраску, отображающую амплитуду сейсмических импульсов, которая даёт косвенную оценку толщины закреплённой трещины в целевом пласте.

На рисунках 4 и 5 приведены примеры результатов кластеризации в рамках одной из «вспомогательных» зон (параллелепипедов). Рисунок 4 демонстрирует разбиение на кластеры до применения факторного анализа (т. е. кластеризацию по переменным X , Y , H , A и T , предварительно нормированным), а рисунок 5 — после применения факторного анализа (кластеризацию по факторам F_1 , F_2 и F_3). Оба графика приведены в системе координат (X, Y, H) для визуального сравнения результатов кластеризации.

Из рисунка 4 видно, что в первом случае получились «размытые» кластеры, точки которых рассредоточены по всей области параллелепипеда. Этот результат может быть объяснён тем, что из-за ограниченной разрешающей способности сейсмической антенны, кластеризации подвергались гипоцентры с неточно определёнными характеристиками. Кроме того, имеет место влияние случайных шумов и помех на параметры исходных точечных источников.

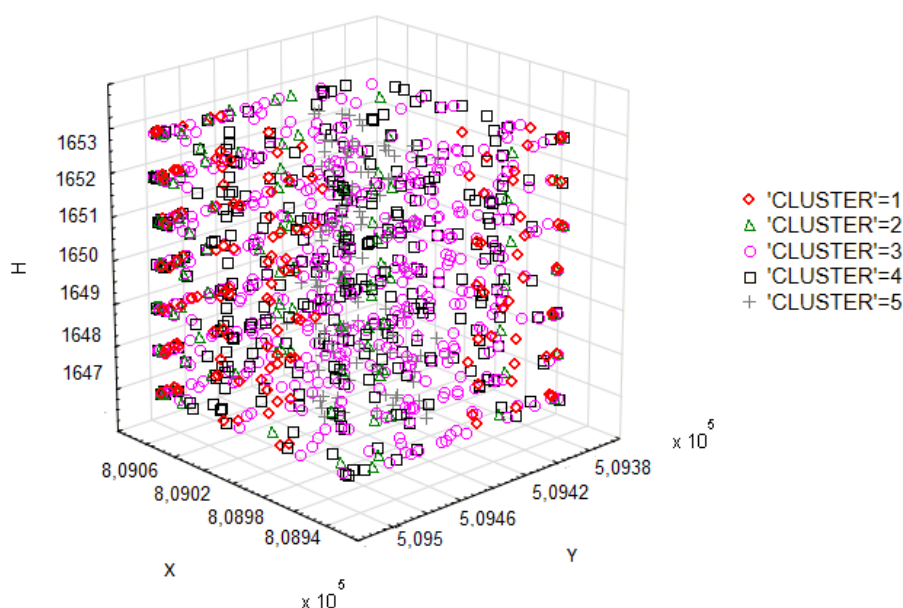


Рис. 4. Результат кластеризации по исходным признакам X , Y , H , A , T

После предварительного выделения факторов и кластеризации на их основе (рис. 5) результат стал более компактным и имеет лучшее разрешение. Можно сделать вывод, что два фактора, «отсеянные» по критерию Кайзера на этапе факторного анализа, в большой мере описывали шумовые характеристики, которые влияли на результат кластеризации в предыдущем эксперименте. Устранив их, удалось получить кластеры, гипоцентры которых могут быть более чётко интерпретированы как источники сейсмических импульсов.

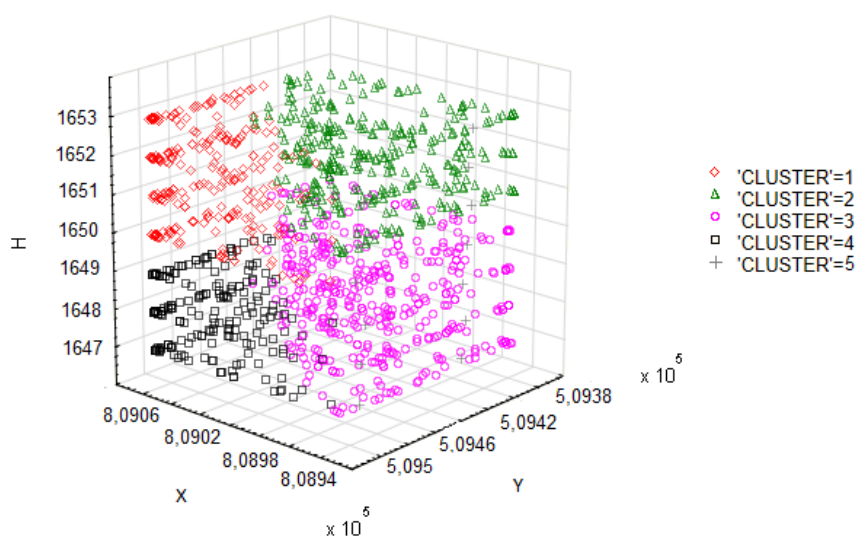
Стоит также отметить, что предварительная фильтрация данных с помощью факторного анализа позволила сократить время кластеризации на 25–30 % (за счёт сокращения количества признаков точечных источников, использованных в обработке).

4. Заключение

В работе предложен оригинальный алгоритм устранения избыточности сейсмических данных при мониторинге ГРП, обеспечивающий определение координат гипоцентров сейсмической активности. В рамках алгоритма применены, исследованы и адаптированы к решаемой задаче методы факторного и кластерного анализа.

Использование факторного анализа (в качестве предварительной обработки данных перед кластеризацией) позволило сократить количество анализируемых признаков объекта с пяти до трёх. В результате были устранены помеховые составляющие, искажающие результат кластеризации, и достигнуто значительное ускорение процесса выделения кластеров (на 25–30 %).

Применение кластерного анализа, в свою очередь, позволило сократить из-

Рис. 5. Результат кластеризации по факторам F_1 , F_2 , F_3

быточность исходных данных путём выделения компактных «облаков» сейсмичности и определения их центров «масс» в качестве единичных гипоцентров микроземлетрясений. Применение кластерного анализа позволило сократить избыточность представления данных почти в 50 раз.

Применение факторного анализа на завершающем этапе показало, что в результате обработки структурные характеристики сейсмических данных не изменились, и все значимые корреляции между переменными были сохранены.

Наконец, использование критерия Шовене позволило устранить выбросы среди гипоцентров, связанные с аномальным значением амплитуды импульса (кол-во выбросов составило 0,01 % выборки).

На основе выделенных гипоцентров с помощью метода Делоне была построена трёхмерная графическая модель трещиноватой зоны пласта.

ЛИТЕРАТУРА

1. Новаковский А.Ю., Рабинович Е.В. Сейсмическая измерительная система для локализации трещин при гидроразрыве нефтяного пласта // Идентификация, измерение характеристик и имитация случайных сигналов (состояние, перспективы развития): сборник материалов конференции. 2009. С. 171–174.
2. Рабинович Е.В., Туркин А.С. Наземная локация микросейсмических сигналов для мониторинга гидравлического разрыва пласта // Доклады ТУСУР. 2012. № 1(25), Ч. 1. С. 104–112.
3. Correlation detection of seismic pulses for hydraulic fracturing monitoring / E.V. Rabinovich [and others] // 16 International conference of young specialists on micro/nanotechnologies and electron devices (EDM). Altai, Erlagol, 2015. P. 162–165.
4. Statistical detection of seismic pulses for hydraulic fracturing monitoring / E.V. Rabinovich [and others] // 11 International forum on strategic technology (IFOST 2016). Novosibirsk, 2016. Pt. 1. P. 375–378.

5. Юдин Г.Б. Кластеры на факторах: proetcontra // Социологические методы в современной исследовательской практике. М. : ГУ-ВШЭ, 2007. С. 296–302.
6. Иберла К. Факторный анализ. М. : Статистика, 1980. 398 с.
7. Пространственная фильтрация сейсмических событий, возникающих при гидравлическом разрыве пласта / Е.В. Рабинович [и др.] // Автоматика и программная инженерия. 2013. № 1. С. 53–60. URL: <http://www.jurnal.nips.ru/sites/default/files/Paper-2013-1-7.pdf> (дата обращения: 28.05.2017).

ELIMINATION OF DATA REDUNDANCY IN SEISMIC MONITORING OF HYDRAULIC FRACTURING

E.V. Rabinovich

Dr.Sc. (Eng.), Professor, e-mail: erabinovich1952@gmail.com

P.I. Vaynmaster

Postgraduate Student, e-mail: wmaster91@mail.ru

G.S. Shefel

Ph.D. (Phys.-Math.), Associate Professor, e-mail: ligr1x@yahoo.com

Novosibirsk State Technical University, Novosibirsk, Russia

Abstract. The important stage of hydraulic fracturing monitoring is constructing a graphical model of fracturing area. However, the real data recorded in the process of hydraulic fracturing are characterized by high intensity of seismic activity in the monitoring area, which hinders correct constructing of such model. Therefore, the problem of eliminating the redundancy of seismic data is important. The original algorithm is considered in this article, which allows eliminating redundancy by detection of local hypocenters of seismic activity. The algorithm is based on the methods of cluster and factor analysis. Factor analysis is used to identify similar characteristics and reduce the number of variables used to describe a seismic source. Cluster analysis is used to detect clouds of seismic emission and find their local hypocenters. The results obtained are used to construct a three-dimensional graphic fracture model.

Keywords: hydraulic fracturing, fracture, factor and cluster analysis.

Дата поступления в редакцию: 21.06.2017

ИДЕНТИФИКАЦИЯ ЗАКРЕПЛЕНИЯ ТРУБОПРОВОДА С ИСПОЛЬЗОВАНИЕМ МИНИМАЛЬНОГО КОЛИЧЕСТВА СОБСТВЕННЫХ ЧАСТОТ

В.Р. Шагиев¹

аспирант, e-mail: shagiev-vadim@mail.ru

А.М. Ахтямов^{1,2}

профессор, д.ф.-м.н., e-mail: akhtyamovam@mail.ru

¹Башкирский государственный университет, Уфа, Республика Башкортостан, Россия

²Институт механики им. Р.Р. Мавлютова Уфимского научного центра РАН, Уфа,
Республика Башкортостан, Россия

Аннотация. Рассматриваются колебания трубопровода с жидкостью. Ранее было показано, что если жидкость не течёт по трубопроводу, то по всем собственным частотам изгибных колебаний трубопровода вид закрепления трубопровода определяется однозначно с точностью до перестановок закреплений на его концах. Задача идентификации краевых условий решалась также и по девяти собственным частотам. В настоящей статье количество собственных значений, с помощью которых можно однозначно с точностью до перестановок закреплений на его концах восстановить краевые условия, уменьшено до пяти. Количество спектральных данных удалось уменьшить за счёт того, что если ранее решалась линейная система 9-ти уравнений, то в настоящей статье решается система пяти нелинейных уравнений относительно четырёх неизвестных коэффициентов из краевых условий, приведённых к канонической форме. Представлен пример решения этой обратной задачи. Приведены также два контрпримера, в которых показано, что меньшего числа собственных значений для идентификации вообще говоря недостаточно. В первом контрпримере показано, что четырёх ненулевых собственных частот ещё недостаточно для идентификации вида закрепления трубопровода. Во втором контрпримере показано, что в отдельных случаях необходима информация о том, является ли нуль собственным значением.

Ключевые слова: краевые условия, собственные частоты, собственные значения, заделка, свободное опирание, плавающая заделка, свободный конец, трубопровод.

Введение

Колебаниям трубопроводов, содержащим жидкость, а также определению собственных частот колебаний стержней и трубопроводов посвящены работы

[1]–[8]. В [9]–[16] решались задачи определения параметров краевых условий и условий сопряжения стержней, трубопроводов и других распределённых механических систем. В частности, в [15] по двум собственным значениям определялись неупругие закрепления трубопроводов, в [16] по 14-ти собственным частотам определялись закрепления трубопроводов с протекающей жидкостью. В [13, 14] показано, что по всем собственным частотам (в том числе и нулевому собственному значению) вид закрепления трубопровода (в случае, когда жидкость не течёт по трубопроводу) определяется однозначно с точностью до перестановок местами закреплений на его концах. Более того, там же было показано, что для такой идентификации достаточно и девяти собственных частот. В статье [15] показано, что для определения одного из 16-ти неупругих видов закреплений трубопровода ((заделка)–(заделка), (заделка)–(свободное опирание), (заделка)–(плавающая заделка), (заделка)–(свободный конец), (свободное опирание)–(заделка), (свободное опирание)–(свободное опирание), (свободное опирание)–(плавающая заделка), (свободное опирание)–(свободный конец), (плавающая заделка)–(заделка), (плавающая заделка)–(свободное опирание), (плавающая заделка)–(плавающая заделка), (плавающая заделка)–(свободный конец), (свободный конец)–(заделка), (свободный конец)–(свободное опирание), (свободный конец)–(плавающая заделка), (свободный конец)–(свободный конец)) с точностью до перестановок закреплений на его концах достаточно одной собственной частоты и информации о том, является ли нулевое значение собственным. В настоящей статье показано, что для идентификации закреплений трубопровода достаточно пяти собственных частот, а четырёх собственных частот для этого недостаточно.

Прямая задача

Необходимо определить собственные частоты изгибных колебаний трубопровода с жидкостью (жидкость не движется), зная вид закреплений на его концах (краевые условия) и физические параметры механической системы (коэффициент жёсткости пружин, закрепляющих трубопровод; масса, плотность, длина, внешний и внутренний радиусы трубы; масса и плотность жидкости, которая находится внутри и полностью заполняет весь объём трубопровода в рассматриваемом участке).

Уравнение малых свободных колебаний трубопровода с протекающей по нему жидкостью (с учётом несжимаемости жидкости) имеет следующий вид [1, с. 163]:

$$EI \frac{\partial^4 W}{\partial X^4} + (m + \tilde{m}) \frac{\partial^2 W}{\partial t^2} + 2\tilde{m}V_0 \frac{\partial^2 W}{\partial X \partial t} + \tilde{m} \left(\frac{p_0}{\rho_0} + V_0^2 \right) \frac{\partial^2 W}{\partial X^2} = 0. \quad (1)$$

Здесь

$$I = \frac{\pi}{4}(r^4 - r_1^4), \quad p_0 = E \left(\frac{\pi r_1}{2L} \right)^2 \cdot \left(\left(\frac{r}{r_1} \right)^4 - 1 \right),$$

$$m = \pi(r^2 - r_1^2)\rho, \quad \tilde{m} = \pi r_1^2 \rho_0,$$

где I — момент инерции трубчатого сечения, EI — жёсткость трубы, p_0 — критическое внутреннее давление, m и $\tilde{m}$ — массы трубы и жидкости, приходящиеся на единицу длины трубы, r и r_1 — радиусы внешнего и внутреннего поперечного сечения, V_0 — скорость движения жидкости, ρ — плотность материала трубы, ρ_0 — плотность жидкости, L — длина трубы.

Вводя обозначения $x = X/L$, $w = W/L$, запишем уравнение (1) следующим образом:

$$\frac{\partial^4 w}{\partial x^4} + \frac{(m + \tilde{m})L^4}{EI} \frac{\partial^2 w}{\partial t^2} + \frac{2\tilde{m}V_0L^3}{EI} \frac{\partial^2 w}{\partial x \partial t} + \frac{\tilde{m}L^2}{EI} \left(\frac{p_0}{\rho_0} + V_0^2 \right) \frac{\partial^2 w}{\partial x^2} = 0.$$

Подстановка $w(x, t) = X(x)e^{iwt}$ приводит к уравнению:

$$X^{(4)} + aX'' + 2biwX' - cw^2X = 0, \quad (2)$$

где

$$a = \frac{\tilde{m}L^2}{EI} \left(\frac{p_0}{\rho_0} + V_0^2 \right), \quad b = \frac{\tilde{m}V_0L^3}{EI}, \quad c = \frac{(m + \tilde{m})L^4}{EI}.$$

Введение обозначений x и w позволило нам обезразмерить переменные

$$a : \frac{\text{кг}}{\text{м}} \cdot \frac{\text{м}^2}{1} \cdot \frac{\text{м} \cdot \text{с}^2}{\text{кг}} \cdot \frac{1}{\text{м}^4} \cdot \frac{\text{м}^2}{\text{с}^2} = 1, \quad w \cdot b : \frac{1}{\text{с}} \cdot \left(\frac{\text{кг}}{\text{м}} \cdot \frac{\text{м}}{\text{с}} \cdot \frac{\text{м}^3}{1} \cdot \frac{\text{м} \cdot \text{с}^2}{\text{кг}} \cdot \frac{1}{\text{м}^4} \right) = 1,$$

$$w^2 \cdot c : \frac{1}{\text{с}^2} \cdot \left(\frac{\text{кг}}{\text{м}} \cdot \frac{\text{м}^4}{1} \cdot \frac{\text{м} \cdot \text{с}^2}{\text{кг}} \cdot \frac{1}{\text{м}^4} \right) = 1.$$

Так как мы рассматриваем случай, когда жидкость не течёт по трубопроводу ($V_0 = 0$), уравнение (2) упрощается и выглядит так:

$$X^{(4)} + aX'' - cw^2X = 0. \quad (3)$$

Краевые условия в общем виде (условия типа Штурма [17, с. 70]):

$$\begin{aligned} U_1(X) &= -a_1X(0) + a_4X'''(0) = 0, & U_3(X) &= b_1X(1) + b_4X'''(1) = 0, \\ U_2(X) &= -a_2X'(0) + a_3X''(0) = 0, & U_4(X) &= b_2X'(1) + b_3X''(1) = 0, \end{aligned} \quad (4)$$

где

$$|a_1| + |a_4| \neq 0, \quad |a_2| + |a_3| \neq 0, \quad |b_1| + |b_4| \neq 0, \quad |b_2| + |b_3| \neq 0.$$

Матрицу, составленную из коэффициентов a_j форм $U_1(X_k)$ и $U_2(X_k)$, обозначим буквой A , а матрицу, составленную из коэффициентов b_j форм $U_3(X_k)$ и $U_4(X_k)$, — буквой B .

$$A = \begin{bmatrix} -a_1 & 0 & 0 & a_4 \\ 0 & -a_2 & a_3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_2 & b_3 & 0 \end{bmatrix}. \quad (5)$$

Заметим, что в рассмотренных нами примерах спектр состоит из простых собственных значений. Для первых собственных значений это проверено с помощью неравенства $\frac{\partial \Delta(\lambda)}{\partial \lambda} \neq 0$ при $\lambda = \lambda_k$, а для последующих собственных

частот это вытекает из теоремы 2 [17, с. 74-75]), которая утверждает, что спектр краевой задачи с условиями Штурма-Лиувилля, начиная с некоторого значения, состоит из простых собственных значений.

Покажем решение прямой задачи на примере случая (заделка)–(свободный конец), коэффициенты в краевых условиях $a_1 = 1, a_2 = 1, a_3 = 0, a_4 = 0, b_1 = 0, b_2 = 0, b_3 = 1, b_4 = 1$:

$$\begin{aligned} U_1(X) &= X(0) = 0, & U_3(X) &= X'''(1) = 0, \\ U_2(X) &= X'(0) = 0, & U_4(X) &= X''(1) = 0. \end{aligned}$$

Характеристическое уравнение для этой задачи в случае, когда жидкость не течёт по трубопроводу ($V_0 = 0$), предполагая коэффициенты a и c равными 1, имеет следующий вид:

$$X^{(4)} + X'' - w^2 X = 0, \quad (6)$$

$\lambda_1, \lambda_2, \lambda_3$ и λ_4 – корни уравнения (6):

$$\lambda_1 = -\lambda_3 = \sqrt{\frac{-1 + \sqrt{1 + 4w^2}}{2}}, \quad \lambda_2 = -\lambda_4 = \sqrt{\frac{-1 - \sqrt{1 + 4w^2}}{2}}. \quad (7)$$

Функции

$$\begin{aligned} X_1(x) &= -\frac{\lambda_2^2}{2(\lambda_1^2 - \lambda_2^2)} e^{\lambda_1 x} + \frac{\lambda_1^2}{2(\lambda_1^2 - \lambda_2^2)} e^{\lambda_2 x} - \frac{\lambda_2^2}{2(\lambda_1^2 - \lambda_2^2)} e^{-\lambda_1 x} + \frac{\lambda_1^2}{2(\lambda_1^2 - \lambda_2^2)} e^{-\lambda_2 x}, \\ X_2(x) &= -\frac{\lambda_2^2}{2(\lambda_1^2 - \lambda_2^2)\lambda_1} e^{\lambda_1 x} + \frac{\lambda_1^2}{2(\lambda_1^2 - \lambda_2^2)\lambda_2} e^{\lambda_2 x} + \frac{\lambda_2^2}{2(\lambda_1^2 - \lambda_2^2)\lambda_1} e^{-\lambda_1 x} - \frac{\lambda_1^2}{2(\lambda_1^2 - \lambda_2^2)\lambda_2} e^{-\lambda_2 x}, \\ X_3(x) &= \frac{1}{2(\lambda_1^2 - \lambda_2^2)} e^{\lambda_1 x} - \frac{1}{2(\lambda_1^2 - \lambda_2^2)} e^{\lambda_2 x} + \frac{1}{2(\lambda_1^2 - \lambda_2^2)} e^{-\lambda_1 x} - \frac{1}{2(\lambda_1^2 - \lambda_2^2)} e^{-\lambda_2 x}, \\ X_4(x) &= \frac{1}{2(\lambda_1^2 - \lambda_2^2)\lambda_1} e^{\lambda_1 x} - \frac{1}{2(\lambda_1^2 - \lambda_2^2)\lambda_2} e^{\lambda_2 x} - \frac{1}{2(\lambda_1^2 - \lambda_2^2)\lambda_1} e^{-\lambda_1 x} + \frac{1}{2(\lambda_1^2 - \lambda_2^2)\lambda_2} e^{-\lambda_2 x} \end{aligned}$$

являются линейно независимыми решениями уравнения (6), которые удовлетворяют условию

$$X_j^{(r-1)}(0) = \begin{cases} 0, & j \neq r, \\ 1, & j = r, \end{cases} \quad j, r = 1, 2, 3, 4.$$

Общее решение уравнения (6) представляется в следующем виде

$$X(x) = C_1 X_1(x) + C_2 X_2(x) + C_3 X_3(x) + C_4 X_4(x).$$

Уравнение частот получают из условия равенства нулю характеристического определителя

$$\Delta(w) = \begin{vmatrix} U_1(X_1) & U_1(X_2) & U_1(X_3) & U_1(X_4) \\ U_2(X_1) & U_2(X_2) & U_2(X_3) & U_2(X_4) \\ U_3(X_1) & U_3(X_2) & U_3(X_3) & U_3(X_4) \\ U_4(X_1) & U_4(X_2) & U_4(X_3) & U_4(X_4) \end{vmatrix} = \begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ U_3(X_1) & U_3(X_2) & U_3(X_3) & U_3(X_4) \\ U_4(X_1) & U_4(X_2) & U_4(X_3) & U_4(X_4) \end{vmatrix}. \quad (8)$$

Вычисления на компьютере приводят к следующим результатам: $w_1 = 3.641309$, $w_2 = 21.72958$, $w_3 = 61.32403$, $w_4 = 120.4921$, $w_5 = 199.4298$, $w_6 = 298.1131$, $w_7 = 416.5395$, $w_8 = 554.7075$, $w_9 = 712.6162$.

Заметим, что для случая, когда жидкость в трубопроводе не движется, собственные частоты не имеют комплексную (мнимую) часть.

Обратная задача

Известны собственные частоты w_k . Ранги матриц A и B (5) равны двум. Требуется восстановить коэффициенты краевых условий a_j и b_j (4), то есть определить вид закрепления.

Ранее [13, 14] было показано, что краевые условия находятся однозначно с точностью до перестановок закреплений местами, и для этого использовалось девять собственных частот. В данной статье мы покажем, что, используя пять собственных частот, можно однозначно определить вид закрепления трубопровода, когда жидкость не движется.

Матрицу, составленную из матриц A , B и двух нулевых матриц, обозначим через C :

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} -a_1 & 0 & 0 & a_4 & 0 & 0 & 0 & 0 \\ 0 & -a_2 & a_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_1 & 0 & 0 & b_4 \\ 0 & 0 & 0 & 0 & 0 & b_2 & b_3 & 0 \end{bmatrix}.$$

Элементы матрицы C обозначим через c_{ij} , а миноры матрицы C , образованные из столбцов с номерами k_1, k_2, k_3, k_4 – через

$$M_{k_1 k_2 k_3 k_4} = \begin{vmatrix} c_{1 k_1} & c_{1 k_2} & c_{1 k_3} & c_{1 k_4} \\ c_{2 k_1} & c_{2 k_2} & c_{2 k_3} & c_{2 k_4} \\ c_{3 k_1} & c_{3 k_2} & c_{3 k_3} & c_{3 k_4} \\ c_{4 k_1} & c_{4 k_2} & c_{4 k_3} & c_{4 k_4} \end{vmatrix}, \quad \text{например, } M_{1256} = (-a_1) \cdot (-a_2) \cdot b_1 \cdot b_2.$$

В терминах матрицы C задача идентификации краевых условий (4) сводится к идентификации матрицы C с точностью до линейных преобразований её строк.

Характеристический определитель (8) можно представить в другом виде:

$$\Delta(w) = \det(C \cdot D), \quad \text{где}$$

$$D = \begin{vmatrix} X_1(0) & X_2(0) & X_3(0) & X_4(0) \\ X'_1(0) & X'_2(0) & X'_3(0) & X'_4(0) \\ X''_1(0) & X''_2(0) & X''_3(0) & X''_4(0) \\ X'''_1(0) & X'''_2(0) & X'''_3(0) & X'''_4(0) \\ X_1(1) & X_2(1) & X_3(1) & X_4(1) \\ X'_1(1) & X'_2(1) & X'_3(1) & X'_4(1) \\ X''_1(1) & X''_2(1) & X''_3(1) & X''_4(1) \\ X'''_1(1) & X'''_2(1) & X'''_3(1) & X'''_4(1) \end{vmatrix}.$$

Так как нам известны собственные частоты w_k , мы можем найти значения λ_1 , λ_2 , λ_3 и λ_4 , что позволяет вычислить функции $X_1(x)$, $X_2(x)$, $X_3(x)$, $X_4(x)$ и все их производные. Поэтому матрица D нам полностью известна.

Используя формулу Бине-Коши, получаем

$$\Delta(w_k) = \sum_{\substack{1 \leq k_1 < k_2 \leq 4 \\ 5 \leq k_3 < k_4 \leq 8}} M_{k_1 k_2 k_3 k_4} f_{k_1 k_2 k_3 k_4} = 0, \quad (9)$$

где $f_{k_1 k_2 k_3 k_4}$ – миноры четвёртого порядка матрицы D , составленные из строк с номерами k_1 , k_2 , k_3 , k_4 .

Уравнение (9) раскрывается следующим образом

$$\Delta(w_k) = x_1 f_{1257}(w_k) + x_2 f_{1268}(w_k) + x_3 f_{1368}(w_k) + x_4 f_{1278}(w_k) + x_5 f_{1378}(w_k) + \\ + x_6 f_{2478}(w_k) + x_7 f_{1357}(w_k) + x_8 f_{2468}(w_k) + x_9 f_{1256}(w_k) + x_{10} f_{3478}(w_k) = 0, \quad (10)$$

где

$$\begin{aligned} x_1 &= M_{1257} - M_{1356} \quad (f_{1257} = -f_{1356}), & x_5 &= M_{1378} - M_{3457} \quad (f_{1378} = -f_{3457}), \\ x_2 &= M_{1268} - M_{2456} \quad (f_{1268} = -f_{2456}), & x_6 &= M_{2478} - M_{3468} \quad (f_{2478} = -f_{3468}), \\ x_3 &= M_{1368} + M_{2457} \quad (f_{1368} = f_{2457}), & x_4 &= M_{1278} + M_{3456} \quad (f_{1278} = f_{3456}), \\ x_7 &= M_{1357}, & x_8 &= M_{2468}, & x_9 &= M_{1256}, & x_{10} &= M_{3478}. \end{aligned} \quad (11)$$

$$cw^2 f_{1356} = -f_{1378}, \quad cw^2 f_{2456} = -f_{2478}, \quad cw^2 f_{1357} = -f_{2468}, \quad cw^4 f_{1256} = f_{3478}. \quad (12)$$

Из (12) видно, что собственные частоты могут совпадать и для случаев, когда закрепления не получаются одно из другого перестановкой их на концах. Здесь c — это коэффициент из (2). Почему это важно, мы покажем ниже.

Подставив в (10) пять собственных частот $w_1, \dots, w_5$, получим систему пяти уравнений относительно неизвестных $x_1, \dots, x_{10}$. Матрицей этой системы явля-

ется следующая матрица:

$$\begin{pmatrix} f_{1257}(w_1) & f_{1268}(w_1) & \dots & f_{1256}(w_1) & f_{3478}(w_1) \\ f_{1257}(w_2) & f_{1268}(w_2) & \dots & f_{1256}(w_2) & f_{3478}(w_2) \\ f_{1257}(w_3) & f_{1268}(w_3) & \dots & f_{1256}(w_3) & f_{3478}(w_3) \\ f_{1257}(w_4) & f_{1268}(w_4) & \dots & f_{1256}(w_4) & f_{3478}(w_4) \\ f_{1257}(w_5) & f_{1268}(w_5) & \dots & f_{1256}(w_5) & f_{3478}(w_5) \end{pmatrix}. \quad (13)$$

Выразим переменные $x_1, \dots, x_{10}$ через $a_1, \dots, a_4$ и $b_1, \dots, b_4$

$$x_1 = a_1 \cdot a_2 \cdot b_1 \cdot b_3 + a_1 \cdot a_3 \cdot b_1 \cdot b_2; \quad x_2 = -a_1 \cdot a_2 \cdot b_2 \cdot b_4 - a_2 \cdot a_4 \cdot b_1 \cdot b_2;$$

$$x_3 = a_1 \cdot a_3 \cdot b_2 \cdot b_4 + a_2 \cdot a_4 \cdot b_1 \cdot b_3; \quad x_4 = -a_1 \cdot a_2 \cdot b_3 \cdot b_4 - a_3 \cdot a_4 \cdot b_1 \cdot b_2;$$

$$x_5 = a_1 \cdot a_3 \cdot b_3 \cdot b_4 + a_3 \cdot a_4 \cdot b_1 \cdot b_3; \quad x_6 = -a_2 \cdot a_4 \cdot b_3 \cdot b_4 - a_3 \cdot a_4 \cdot b_2 \cdot b_4;$$

$$x_7 = -a_1 \cdot a_3 \cdot b_1 \cdot b_3; \quad x_8 = -a_2 \cdot a_4 \cdot b_2 \cdot b_4;$$

$$x_9 = a_1 \cdot a_2 \cdot b_1 \cdot b_2; \quad x_{10} = a_3 \cdot a_4 \cdot b_3 \cdot b_4.$$

В результате получим систему пяти нелинейных уравнений от восьми неизвестных.

Эта система, с помощью сведения матрицы C к канонической форме, сводится к решению нелинейной системы пяти уравнений от четырёх неизвестных коэффициентов краевых условий. Четырёх собственных частот для идентификации вида закрепления трубопровода недостаточно. Соответствующий контрпример будет приведён ниже.

Ранее задача идентификации закреплений трубопровода решалась с помощью системы 9-ти линейных уравнений (10) от 10-ти неизвестных x_i ($i = 1, 2, \dots, 10$). Поэтому использовалось девять, а не пять собственных частот.

Пример

Рассмотрим случай, когда нам известны частоты $w_1 = 29.44435$, $w_2 = 72.98521$, $w_3 = 134.6173$, $w_4 = 215.1980$, $w_5 = 315.0639$. Найдём λ_1 , λ_2 , λ_3 и λ_4 (7) для каждой из пяти частот:

$$\lambda_1 = -\lambda_3 = 5.380390_{(w_1)}, 8.513925_{(w_2)}, 11.58094_{(w_3)}, 14.65260_{(w_4)}, 17.73596_{(w_5)}$$

$$\lambda_2 = -\lambda_4 = 5.472531i_{(w_1)}, 8.572452i_{(w_2)}, 11.62404i_{(w_3)}, 14.68668i_{(w_4)}, 17.76413i_{(w_5)}$$

Матрица D для w_1 выглядит следующим образом

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 55.53811 & 10.19385 & 1.831441 & 0.344800 \\ 298.9313 & 55.53811 & 9.849050 & 1.831441 \\ 1587.804 & 298.9313 & 53.70667 & 9.849050 \\ 8538.828 & 1587.804 & 289.0823 & 53.70667 \end{pmatrix}.$$

Раскроем эту матрицу (10) и получим первое уравнение, соответствующее частоте w_1 :

$$-0.480117 x_1 - 0.477539 x_2 - 74.79756 x_3 + 37.22037 x_4 - 416.2470 x_5 - \\ -414.0117 x_6 + 2.671859 x_7 - 2316.421 x_8 - 0.041778 x_9 - 31401.96 x_{10} = 0.$$

Аналогично получим остальные четыре уравнения:

$$2.818855 x_1 + 14.49809 x_2 + 1640.216 x_3 - 813.1440 x_4 + 15015.59 x_5 + \\ + 77229.04 x_6 - 25.70239 x_7 + 136912.5 x_8 + 0.152838 x_9 + 4336816 x_{10} = 0,$$

$$-23.92452 x_1 - 515.7571 x_2 - 31464.06 x_3 + 15651.91 x_4 - 433555.9 x_5 - \\ -9346459 x_6 + 321.6035 x_7 - 5828042 x_8 - 0.863650 x_9 - 2.836231 \cdot 10^8 x_{10} = 0,$$

$$251.4325 x_1 + 13071.78 x_2 + 603128.2 x_3 - 300417.9 x_4 + 1.164389 \cdot 10^7 x_5 + \\ + 6.053564 \cdot 10^8 x_6 - 4576.252 x_7 + 2.119271 \cdot 10^8 x_8 + 6.487102 x_9 + \\ + 1.391246 \cdot 10^{10} x_{10} = 0,$$

$$-3044.613 x_1 - 297594.5 x_2 - 1.176207 \cdot 10^7 x_3 + 5863328 x_4 - 3.022243 \cdot 10^8 x_5 - \\ -2.954079 \cdot 10^{10} x_6 + 70778.30 x_7 - 7.025825 \cdot 10^9 x_8 - 59.06727 x_9 - \\ -5.820246 \cdot 10^{11} x_{10} = 0.$$

Обозначим эту систему через S . Определителем этой системы является следующая матрица

$$\begin{pmatrix} -0.480117 & -0.477539 & \dots & -0.041778 & -31401.96 \\ 2.818855 & 14.49809 & \dots & 0.152838 & 4336816 \\ -23.92452 & -515.7571 & \dots & -0.863650 & -2.836231 \cdot 10^8 \\ 251.4325 & 13071.78 & \dots & 6.487102 & 1.391246 \cdot 10^{10} \\ -3044.613 & -297594.5 & \dots & -59.06727 & -5.820246 \cdot 10^{11} \end{pmatrix}.$$

Нетрудно проверить, что наибольшим по модулю определителем четвёртого порядка этой матрицы является определитель, составленный из элементов 4-8 столбцов и равный $-6.275948 \cdot 10^{24}$. Поэтому удобнее всего выразить переменные $x_4, \dots, x_8$ через остальные. Оставим переменные $x_4, \dots, x_8$ в левой части уравнений системы S , а переменные $x_1, x_2, x_3, x_9, x_{10}$ перенесём в правую часть (со знаком минус). Получим систему уравнений, решив которую методом Крамера, получим выражение переменных $x_4, \dots, x_8$ через остальные:

$$x_4 = -0.009002x_1 - 0.103224x_2 + 2.143385x_3 - 0.001673x_9 - 950.8871x_{10},$$

$$x_5 = -0.000323x_1 - 0.008424x_2 + 0.008858x_3 - 0.000076x_9 + 33.38482x_{10},$$

$$x_6 = -6.67 \cdot 10^{-7}x_1 - 0.000021x_2 + 0.000028x_3 - 1.83 \cdot 10^{-7}x_9 - 12.03323x_{10},$$

$$x_7 = 0.264565x_1 + 0.583920x_2 - 0.818512x_3 + 0.029648x_9 - 1567.256x_{10},$$

$$x_8 = 0.000011x_1 + 0.000326x_2 - 0.000391x_3 + 0.000003x_9 - 34.49125x_{10}.$$

Выразим переменные $x_1, \dots, x_{10}$ через $a_1, \dots, a_4$ и $b_1, \dots, b_4$, как мы уже сделали ранее, и предположим какие-нибудь четыре коэффициента равными единице, например, $a_3 = a_4 = b_3 = b_4 = 1$.

$$\begin{aligned} -a_1a_2 - b_1b_2 = & -0.009002(a_1a_2b_1 + a_1b_1b_2) + 0.103224(a_1a_2b_2 + a_2b_1b_2) + \\ & + 2.143385(a_1b_2 + a_2b_1) - 0.001673a_1a_2b_1b_2 - 950.8871, \end{aligned}$$

$$\begin{aligned} a_1 + b_1 = & -0.000323(a_1a_2b_1 + a_1b_1b_2) + 0.008424(a_1a_2b_2 + a_2b_1b_2) + \\ & + 0.008858(a_1b_2 + a_2b_1) - 0.000076a_1a_2b_1b_2 + 33.38482, \end{aligned}$$

$$\begin{aligned} -a_2 - b_2 = & -6.67 \cdot 10^{-7}(a_1a_2b_1 + a_1b_1b_2) + 0.000021(a_1a_2b_2 + a_2b_1b_2) + \\ & + 0.000028(a_1b_2 + a_2b_1) - 1.83 \cdot 10^{-7}a_1a_2b_1b_2 - 12.03323, \end{aligned}$$

$$\begin{aligned} -a_1b_1 = & 0.264565(a_1a_2b_1 + a_1b_1b_2) - 0.583920(a_1a_2b_2 + a_2b_1b_2) - \\ & - 0.818512(a_1b_2 + a_2b_1) + 0.029648a_1a_2b_1b_2 - 1567.256, \end{aligned}$$

$$\begin{aligned} -a_2b_2 = & 0.000011(a_1a_2b_1 + a_1b_1b_2) - 0.000326(a_1a_2b_2 + a_2b_1b_2) - \\ & - 0.000391(a_1b_2 + a_2b_1) + 0.000003a_1a_2b_1b_2 - 34.49125. \end{aligned}$$

Решение первых четырёх уравнений даёт такие ответы (отбрасываем отрицательные и комплексные значения)

$$a_1 = 26.00974, a_2 = 6.549311, b_1 = 20.24870, b_2 = 5.450090,$$

$$a_1 = 25, a_2 = 5, b_1 = 21, b_2 = 7.$$

Решение первых трёх и пятого уравнений даёт такие ответы (отбрасываем отрицательные и комплексные значения)

$$a_1 = 33.63423, a_2 = 6.985825, b_1 = 12.94287, b_2 = 5.013029,$$

$$a_1 = 25, a_2 = 5, b_1 = 21, b_2 = 7.$$

Таким образом, частотам $w_1 = 29.44435$, $w_2 = 72.98521$, $w_3 = 134.6173$, $w_4 = 215.1980$, $w_5 = 315.0639$ соответствуют краевые условия $a_1 = 25a_4$, $a_2 = 5a_3$, $b_1 = 21b_4$, $b_2 = 7b_3$ либо $a_1 = 21a_4$, $a_2 = 7a_3$, $b_1 = 25b_4$, $b_2 = 5b_3$ (так как жидкость не течёт по трубопроводу).

Анализ погрешности

В таблице 1 представлены результаты расчёта относительной погрешности при возмущении собственной частоты w_1 .

Таблица 1. Относительные погрешности при возмущении w_1

Возмущение w_1	δ для a_1	δ для a_2	δ для b_1	δ для b_2
10^{-10}	1159.156	154.0616	1119.691	154.2552
10^{-9}	1159.155	154.0617	1119.691	154.2553
10^{-8}	1159.154	154.0622	1119.689	154.2558
10^{-7}	1159.136	154.0675	1119.674	154.2611
10^{-6}	1158.956	154.1204	1119.525	154.3139
10^{-5}	1157.162	154.6501	1118.041	154.8429
10^{-4}	1140.016	159.9984	1103.960	160.1841

Частота $w_1 = 29.44435$ — это частота, вычисленная с точностью в 40 значащих цифр, соответствующая крайним условиям, рассмотренным в примере. Возмущая частоту w_1 на определённое значение из левого столбца, мы рассчитывали относительную погрешность по формуле:

$$\delta = \frac{|a_{1 \text{ возмущённое}} - a_1|}{\text{возмущение}}.$$

За a_1 бралось значение при невозмущённой w_1 . Эти результаты говорят нам о том, что желая получить погрешность результата не более 10^{-3} , погрешность собственной частоты w_1 не должна быть выше 10^{-7} .

В таблице 2 представлены результаты расчёта относительной погрешности при возмущении по отдельности каждой частоты на 10^{-15} .

Контрпримеры

Контрпример 1 (Восстановление по четырём собственным частотам). Четырёх собственных частот для идентификации вида закрепления трубопровода недостаточно. Рассмотрим закрепление трубопровода с такими крайними условиями:

$$a_1 = 20.10253, a_2 = 5.015512, a_3 = 1, a_4 = 1,$$

Таблица 2. Относительные погрешности при возмущении всех частот

	δ для a_1	δ для a_2	δ для b_1	δ для b_2
w_1	1159.156	154.0616	1119.691	154.2552
w_2	7902.846	4351.306	8156.099	8156.099
w_3	17547.45	17012.10	18840.45	17027.47
w_4	16513.40	21978.03	18239.51	22008.76
w_5	5649.435	9231.953	6375.944	9250.024
Сумма	48772.29	52727.45	52731.70	56596.61

$$b_1 = 26.27080, b_2 = 6.982448, b_3 = 1, b_4 = 1.$$

Этому закреплению соответствуют собственные частоты $w_1 = 29.44435$, $w_2 = 72.98521$, $w_3 = 134.6173$, $w_4 = 215.1980$, $w_5 = 315.06382$.

Значения собственных частот для упругого закрепления из примера $w_1 = 29.44435$, $w_2 = 72.98521$, $w_3 = 134.6173$, $w_4 = 215.1980$, $w_5 = 315.06388$.

Все наши вычисления выполнены с точностью в 40 значащих цифр, различия между частотами w_1 , w_2 , w_3 , w_4 проявляются на 35–38 знаке, что считается несущественным. А различие у w_5 проявилось на пятом знаке после запятой, что для нас является существенным. Получается, два различных вида закрепления имеют четыре одинаковые ненулевые собственные частоты.

Контрпример 2 (Нулевое собственное значение). Рассмотрим случай (заделка)–(заделка) и (свободный конец)–(свободный конец).

Из равенств (12) видно, что

$$cw^4 f_{1256} = f_{3478}.$$

Минор f_{1256} соответствует закреплению (заделка)–(заделка), а минор f_{3478} соответствует закреплению (свободный конец)–(свободный конец). Равенство показывает нам, что спектры частот этих двух закреплений совпадают полностью, различие лишь в том, что значение $w = 0$ в первом случае не является собственным, а во втором является. Поэтому какие бы четыре ненулевые частоты мы не взяли для решения обратной задачи ($w_1 = 22.09646$, $w_2 = 61.29826$, $w_3 = 120.4937$, $w_4 = 199.4297$, $w_5 = 298.1131$ и т. д.), они все будут соответствовать и закреплению (заделка)–(заделка), и закреплению (свободный конец)–(свободный конец).

Таким образом, лишь информация о том, является ли нуль собственным значением, позволит нам однозначно идентифицировать вид закрепления трубопровода. Таких особенных случаев всего лишь четыре (12).

Заключение

В данной статье показано, что вид и параметры закреплений трубопровода с неподвижной жидкостью однозначно с точностью до перестановок закреплений местами идентифицируются по пяти собственным частотам изгибных колебаний. Приведены контрпримеры, показывающие, что четырёх собственных частот для такой идентификации недостаточно.

Работа поддержана РФФИ грантами 15-01-01095_а, 17-41-020230-р_Башкортостан_а.

ЛИТЕРАТУРА

1. Ильгамов М.А. Колебания упругих оболочек, содержащих жидкость и газ. М. : Наука, 1969. 182 с.
2. Томпсон Дж. М.Т. Неустойчивости и катастрофы в науке и технике. М. : Мир, 1985. 254 с.
3. Вибрации в технике: справочник. Т. 1. Колебания линейных систем / Под ред. В.В. Болотина. М. : Машиностроение, 1978. 352 с.
4. Chu M.T., Golub G.H. Inverse Eigenvalue Problems: Theory, Algorithms, and Applications. Oxford : University Press, 2005.
5. Коллатц Л. Задачи на собственные значения (с техническими приложениями). М. : Наука, 1968. 503 с.
6. Гладвелл Г.М.Л. Обратные задачи теории колебаний. М. – Ижевск : НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2008. 608 с.
7. Хакимов А. Г. Определение плотности жидкости и внутреннего давления в трубопроводе по собственным частотам изгибных колебаний // Проблемы сбора, подготовки и транспорта нефти и нефтепродуктов. 2014. № 1(95). С. 37–43.
8. Абзалимов Р.Р., Саяхова Е.В. Разностно-аналитический метод вычисления собственных значений для уравнений четвертого порядка с разделёнными краевыми условиями // Известия высших учебных заведений. Математика. 2008. № 11. С. 3–11.
9. Гнуни В.Ц., Оганесян З.Б. Определение граничных условий круглой кольцевой пластинки по заданным частотам собственных колебаний // Известия Академии наук Армении. Механика (проблемы механики сплошной среды и конструкций). 1991. Т. 44, № 5. С. 9–15.
10. Ватульян А.О., Осипов А.В. Поперечные колебания балки с локализованными неоднородностями // Вестник ДГТУ. 2012. № 8(69). С. 34–40.
11. Ватульян А.О., Солуянов Н.О. Об определении местоположения и размера полости в упругом стержне // Дефектоскопия. 2005. № 9. С. 44–56.
12. Ильгамов М.А., Хакимов А.Г. Диагностика повреждений консольной балки с надрезом // Дефектоскопия. 2009. Т. 45, № 6. С. 83–89.
13. Ахтямов А.М. Теория идентификации краевых условий и её приложения. М. : Физматлит, 2009. 272 с.
14. Ахтямов А.М., Сафина Г.Ф. Определение виброзащитного закрепления трубопровода // Прикладная математика и техническая физика. 2008. Т. 49, № 1. С. 139–147.

15. Ахтямов А.М., Шагиев В.Р. Идентификация неупругих видов закреплений трубопроводов // Вестник Башкирского университета. 2016. Т. 21, № 1. С. 21–26.
16. Ахтямов А.М., Сафина Г.Ф. О единственности решения и корректности задачи определения параметров закрепления трубы с текущей в ней жидкостью // Прикладная механика и техническая физика. 2016. Т. 57, № 2(336). С. 32–45.
17. Наймарк М.А. Линейные дифференциальные операторы. М. : Наука, 1969. 526 с.

IDENTIFICATION OF PIPE FASTENING USING THE MINIMUM NUMBER OF NATURAL FREQUENCIES

V.R. Shagiev¹

Postgraduate Student, e-mail: shagiev-vadim@mail.ru

A.M. Akhtyamov^{1,2}

Dr.Sc. (Phys.-Math.), Professor, e-mail: akhtyamovam@mail.ru

¹Faculty of Mathematics and Information Technologies, Bashkir State University,
Ufa, Republic of Bashkortostan, Russia

²Mavlyutov Institute of Mechanics, Ufa, Republic of Bashkortostan, Russia

Abstract. The oscillations of a pipeline with a liquid are considered. Previously, it was shown that if the liquid does not flow through the pipeline, then by all its frequencies of flexural vibrations of the pipeline, the type of fixing of the pipeline is uniquely determined up to permutations of fastenings at its ends. The problem of identifying boundary conditions was also solved for nine natural frequencies. In this paper, the number of eigenvalues with which one can uniquely reduce the boundary conditions, up to permutations of the fastenings at its ends, is reduced to five. The number of spectral data was reduced due to the fact that if a linear system of 9 equations was previously solved, then in the present paper a system of five nonlinear equations is solved with respect to four unknown coefficients from the boundary conditions reduced to the canonical form. An example of solving this inverse problem is presented. Two counterexamples are also given in which it is shown that fewer eigenvalues for identification are generally insufficient. In the first counterexample, it is shown that four nonzero eigenfrequencies are still not enough to identify the type of anchorage of the pipeline. The second counterexample shows that in some cases information is needed about whether zero is an eigenvalue.

Keywords: boundary conditions, natural frequencies, eigenvalues, clamping, free support, floating fixing, free end, pipeline.

Дата поступления в редакцию: 17.10.2017

ОБЩАЯ СХЕМА СОВРЕМЕННОЙ ПСИХОЛОГИИ. МЕСТО РЕФЛЕКСИВНЫХ ИССЛЕДОВАНИЙ В СИСТЕМЕ НАУК

В.А. Лефевр

профессор, e-mail: valefebv@uci.edu

Калифорнийский университет, Ирвин, США

Аннотация. Стенограмма дистанционного выступления на конференции «Рефлексивный театр ситуационного центра» 29.11.2016 (Ирвайн, США — Омск, Россия). Отредактировано автором.

Ключевые слова: субъективная психология, психофизика, бихевиоризм, когнитивная психология, рефлексивное направление.

Я хочу обсудить вопрос: «Каково место рефлексивных исследований в психологии?» Для того чтобы описать психологию, мне необходимы четыре понятия: психика, мозг, поведение, среда. Сами исходные понятия я не буду определять, потому что они фундаментальны. У каждого есть семантическая сеть, связанная с этими словами.

Психология возникла во второй половине XIX-го века. Было два направления. Одно из них — субъективная психология, которая была направлена на психику человека. Тогда ещё не умели как следует экспериментировать и думали, что люди, анализируя свои мысли и чувства, могут описать свою психику. В основном этим занимались в Германии. Но оказалось, что этот путь бесплоден. Не удалось найти ничего общего в отчётах различных субъектов.

Второе направление, которое возникло во второй половине XIX-го века, — это психофизика. Вы, наверное, все слышали о законе Фехнера. Этот закон связывает перцепцию интенсивности физического стимула с физической мерой этого стимула. Фехнер обнаружил, что эта зависимость выражается логарифмом, т. е. надо взять логарифм от физической интенсивности звука или интенсивности любого агента, и получится то, что воспринимают люди. Его современники считали, что это такое же важное открытие, как закон Ньютона, что наконец установлена связь между внутренним миром человека и средой. Эта линия продолжается до сих пор. Я об этом скажу несколько дальше.

В начале XX-го века было понято, что это не единственные пути. Сначала в Америке, а потом в других странах возник бихевиоризм. Вы, наверное, слышали это слово. Его схема: поведение — среда. Получена очень важная информация о поведении людей и животных, особенно в работах Скиннера и его последователей. Они проводили эксперименты на животных в специальных камерах. Сейчас существуют десятки, если не сотни, исследователей с этой камерой. Они исследуют, в основном, голубей и крыс. Мне удалось с помощью рефлексивной модели вывести один закон, который они не могли

объяснить. Они отнеслись к этому очень серьёзно. Я получил предложение написать статью вместе со специалистом в этой области, и статья была написана и опубликована в журнале «Behavioral Processes», её перевод есть в «Вопросах психологии». Эта же статья помещена в моей книге «Что такое одушевлённость?» (М. : Когито-Центр, 2017.)

В психофизике был выдвинут другой закон, который стал конкурировать с законом Фехнера. Фехнер считал, что его закон никогда не будет опровергнут, потому что психологи не договорятся, как его опровергать. Тем не менее, Стивенс выдвинул новый закон — степенной. Он заключается в том, что интенсивность восприятия пропорциональна степени для каждого стимула специальной. Таким образом, возникла конкуренция между законом Фехнера и законом Стивенса. Стивенс сравнительно недавно, лет 40 назад, написал статью о том, что надо отдать должное Фехнеру и сдать его закон в утиль.

В 1950-е годы возникла когнитивная психология. Её схема: мозг — поведение. Тогда была использована компьютерная метафора для мозга, т. е. мозг рассматривался в терминах Computer Science. Там была память, оперативная область, кодирование, передача информации, логический вывод рассуждений и т. д. Наиболее важным было открытие Хомского. Он показал, что человек обладает врождённым алгоритмом грамматики.

Теперь я перейду к рефлексивному направлению. Оно не рассматривает работу мозга и строит математическую модель психики. Раньше, в XIX-м веке и в XX-м веке, моделей психики, особенно математических, не было: делались только модели мозга. В рефлексивном подходе была сделана модель психики, и она была связана с поведением и средой. Треугольник (на схеме рис.1) «психика — среда — поведение» — это то, что лежит в основе рефлексивного подхода. Я поясню это на примере рефлексивных игр. Модель субъекта — это модель его психики. Диагональная форма с алгоритмом вычисления — это психика; результат вычислений — это его выбор. А среда — это группа, к которой субъект принадлежит и которая воздействует на его решение.

Эта модель позволила объяснить целый ряд психофизических экспериментов. И не только: она позволила построить теорию выбора, которая вам, наверное, известна немножко. Но самое интересное, оказалось, когда мы перешли к главной проблеме современной психологии: связи психики и мозга. Психологи делятся на два класса по своим убеждениям. К первому, маленькому, классу относятся психологи, которые верят в самостоятельное существование психики. Большинство психологов, особенно американских, на самом деле не верят в существование психики. Они считают, что психика — побочный продукт функционирования мозга. И главная проблема, которая возникла, это выяснить, какова связь психики и мозга. Этим занимается

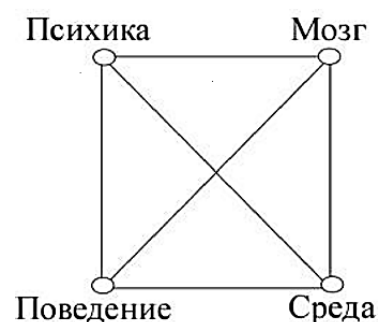


Рис. 1. Основная схема

множество психологов. Великое множество психологов и философов. Некоторые приводят такой пример, что если мозг — это физический объект, то почему, когда мы живём, воспринимаем некий свет своей души? Ведь у физического объекта должно быть физическое функционирование, и там не место психической деятельности.

Эта проблема принимается далеко не всеми, но есть люди, которые много и плодотворно работают в этой области. Было показано, что рефлексивная модель может быть сведена к законам термодинамики: закону сохранения энергии и закону неубывания энтропии. Модель устроена следующим образом. Рассмотрим цепочку абстрактных тепловых двигателей. Тепловой двигатель может произвести максимум работы, когда он обратим. Обычный тепловой двигатель необратим. Любой физический двигатель необратим: он производит меньшую работу, чем максимально теоретически возможная работа. Следующий за ним двигатель производит работу, равную потерянной доступной работе первого двигателя. Третий двигатель совершает работу равную потере доступной работы второго двигателя и т. д. Первый двигатель интерпретируется как субъект, второй — как образ себя у субъекта, третий — как образ себя у образа себя. Из такой цепочки тепловых машин можно вывести многие уравнения рефлексивной теории (она подробно изложена в книге «Что такое одушевлённость?»).

На этом пути удалось вывести набор музыкальных интервалов. Эта проблема существует уже более тысячи лет. Известен набор из 18-ти дробей: $1/2$, $2/3$, $3/4$, $4/7$ и другие. Физик Р. Фейнман написал в своих лекциях, что наше понимание, почему именно эти интервалы являются выделенными, продвинуто не больше, чем во времена Пифагора. С помощью рефлексивной модели можно вывести набор этих интервалов. Причём этот набор верен не только для европейской музыки, основанной на квинте, но и на иных музыкальных системах, основанных на кварте. Оказалось, что там точно такие же выделенные интервалы. Поскольку мы их вывели из первых двух законов термодинамики, это говорит о том, что, возможно, набор этих интервалов фундаментален, т. е. является законом природы. Спасибо за внимание.

GENERAL SCHEME OF MODERN PSYCHOLOGY. PLACE OF REFLEXIVE RESEARCH IN THE SYSTEM OF SCIENCES

V.A. Lefebvre

Ph.D., Professor, e-mail: valefebv@uci.edu

University of California, Irvine, USA

Abstract. Transcript of the remote presentation at the conference "Reflective Theater of the Situation Center" 29.11.2016. It has been edited by the author.

Keywords: subjective psychology, psychophysics, behaviorism, cognitive psychology, reflexive approach.

Дата поступления в редакцию: 29.01.2018

БАЗОВАЯ МОДЕЛЬ ПРОЦЕССОВ РЕФЛЕКСИИ В ИНТЕЛЛЕКТУАЛЬНЫХ АВТОМАТИЗИРОВАННЫХ ОБУЧАЮЩИХ СИСТЕМАХ

В.А. Углев

к.т.н., доцент, e-mail: uglev-v@yandex.ru

ФГАОУ ВО «Сибирский федеральный университет», Железногорск, Россия

Аннотация. Рассматриваются вопросы создания интеллектуальных автоматизированных обучающих систем с функцией рефлексии. При использовании теории рефлексивных игр В.А. Лефевра в структуру планировщика обучающей системы вводится интеллектуальный рефлексивный агент (ИРА), обеспечивающий поиск компромиссных решений при взаимодействии моделей ученика, учителя, предметного тьютора и методиста. Приводятся базовые компоненты модели рефлексивного восприятия ИРА ученика и учебной ситуации с учётом специфики принятия решений для различных этапов обучения.

Ключевые слова: моделирование, индивидуализация обучения, автоматизированное обучение, рефлексивное управление, интеллектуальный рефлексивный агент.

1. Место рефлексивных процессов в автоматизированном обучении

Педагогика как наука о взаимодействии ученика и учителя в век информатизации начинает активно использовать новые возможности [1], связанные, в первую очередь, с развитием глобальных вычислительных сетей и систем искусственного интеллекта. Здесь и далее будет обсуждаться процесс обучения в высших учебных заведениях, хотя отдельные моменты будут справедливы и для обучения в школе. Также будет сделан упор на индивидуализированную подготовку. Роль ученика остаётся неизменной, а вот учитель всё чаще заменяется обучающей системой и лишь опосредованно влияет на учебный процесс. В системах дистанционной и электронной форм обучения учитель всё чаще интерпретируется как предметный тьютор, лишь сопровождающий учебный процесс, а в специализированных обучающих системах с элементами искусственного интеллекта его влияние на процесс обучения зачастую заканчивается подготовкой дидактического материала. Может ли интеллектуальная автоматизированная обучающая система (ИАОС) эффективно заменить педагога-человека? Если мы говорим о педагоге профессионале, то, скорее всего, не заменит никогда. Но профессионалов в любой области мало, да и «массовость» обучения

снижает качество передачи знаний (индивидуализированный подход заменяется «усредняющим»). Огромный потенциал индивидуальной работы ученика с ИАОС во многом нивелируется тем обстоятельством, что машина не способна полноценно «увидеть» ученика [2], фиксируя его действия через достаточно примитивные устройства обратной связи (клавиатура и компьютерная мышь, применение прочих источников информации об ученике в массовом порядке в современных условиях объективно не происходит и тенденции на их использование пока не намечается. В первую очередь это замечание касается web-камер и очков дополнительной и виртуальной реальности). Необходимость формировать представление об ученике в памяти обучающей системы очевидна, и существует достаточно много соответствующих моделей [3]. Их наличие позволяет несколько повысить эффективность функционирования ИАОС и ограничивается возможностями интеллектуальной подсистемы принятия решений и управления (планировщика). Одной из наиболее перспективных функций, которую должен выполнять планировщик, на наш взгляд, является имитация рефлексии [4]. Рассмотрим её подробнее, опираясь на теорию рефлексивных игр В.А. Лефевра [5]. Прежде всего требуется определиться с понятием «рефлексия» применительно к педагогической деятельности. Согласно определению А.А. Безяевой, это «сложный психологический феномен, проявляющийся в способности учителя входить в активную исследовательскую позицию по отношению к своей деятельности и к себе как её субъекту с целью критического анализа, осмысления и оценки её эффективности для развития личности ученика» [6]. Компьютерное моделирование и имитация рефлексии в педагогике должна опираться на формализованную модель, позволяющую оценивать перспективы применения тех или иных педагогических воздействий, которые ИАОС способна реализовать. Они должны быть обеспечены как на логическом уровне (обоснование выбора решений и оценка их результативности), так и на операциональном (ситуационный выбор формы управляющего воздействия). Таким образом, планировщик должен принимать решения, одновременно реализуя ретроспективную, ситуативную и перспективную формы рефлексии, обеспечивая системность функционирования обучающей системы. Когнитивные процессы всегда сопровождаются рефлексией как со стороны ученика, так и со стороны учителя. Если мы заменяем учителя машиной, то ей необходимо ответить не только на базовые вопросы дидактики (чему учить и как учить), но и учесть специфику человеко-машинного взаимодействия [2]. Человек-учитель способен видеть в ученике индивида, обладающего собственными психофизиологическими характеристиками, а также способного к целеполаганию. По этой причине рефлексивное взаимодействие участников учебного процесса, имеющего одним из звеньев автоматизированную обучающую систему, должно быть исследовано системно, учитывая активную позицию всех взаимодействующих сторон. Собственно сам процесс обучения как объект рефлексии в работах В.А. Лефевра напрямую не рассматривается. Первая и, на наш взгляд, недостаточно удачная попытка сделать обобщение идей рефлексии для автоматизированных обучающих систем была предпринята в работе [7], но дальнейшего полноценного исследования не последовало. Поэтому мы выделим базовые вопросы, ответы

на которые должны быть получены при организации ИАОС, имитирующих рефлексивные процессы:

- каковы цели и задачи, которые должен реализовывать планировщик, обладающий функцией рефлексии?
- каков должен быть типовой модельный состав при формализации конфликтующих субъектов?
- какова должна быть базовая модель рефлексивного восприятия различных этапов обучения, отражающая конфликт интересов его участников?
- как сформировать образ всех конфликтующих субъектов в автоматизированном режиме?
- каким образом промоделировать частные механизмы процесса рефлексии в разрезе этапов обучения и решаемых задач?
- каким образом согласовать педагогический, физиологический и психологический уровни организации учебного процесса, управляемого рефлексивным планировщиком?
- как обосновать выбор типа управляющего воздействия планировщиком?
- как привести ученика к выбору наиболее рационального решения, опираясь на выбранный метод рефлексивного управления?
- каким образом обеспечить механизмы целеполагания и принятия решений на логическом уровне, опираясь на технические возможности рефлексивной обучающей системы?
- каким образом обосновать легитимность решений, принимаемых обучающей системой, опираясь на методологические требования и специфику современных образовательных стандартов?
- каким образом планировщику отрефлексировать результаты своих действий?
- как эффективно обобщать и визуализировать различные факторы, способствуя поиску компромиссных решений?

Рассмотрим в рамках данной работы первые три вопроса, раскрыв их основные положения.

2. Формирование образа ученика в традиционном и дистанционном обучении

В процессе непосредственного аудиторного / индивидуального обучения рефлексия имеет место быть как со стороны ученика, так и со стороны учителя: оба оценивают свои возможности, оба имеют различные цели, оба вынуждены принимать решения. Так как задача управления учебным процессом традиционно возлагается на учителя, то будет необходимо рассмотреть то, как у него формируется образ ученика. Будем полагать, что он складывается из трёх компонент: текущего впечатления, ранее имевшегося опыта взаимодействия и целевого образа. Для формализации взаимодействия ученик – учитель (индивидуализированное обучение) будем полагать, что ученик и учитель имеют собственные активные позиции по отношению к дидактическому материалу, на котором и происходит обучение. При этом отметим, что учитель не является

автором учебного материала и вынужден опираться на требования методолога, задающего общую «рамку» для процесса обучения в виде базовой (не индивидуализированной) рабочей программы дисциплины и системы норм / требований (трудоемкость дисциплины, виды контроля и пр.). В ситуации обучения, когда индивидуализация обеспечивается на официальном уровне, в процесс обучения вводится предметный тьютор, сопровождающий обучение, в первую очередь, через консультирования ученика. При этом между ними, согласно [5], может формироваться как отношение взаимопонимания (союза, сплошная линия), так и отношение противостояния (конфликта, пунктирная линия). Это объясняется наличием у всех субъектов собственных целей, а также возможностью влиять на процесс обучения, проявляя целенаправленную активность. Рассмотрим взаимодействие учителя (y), ученика (x) и методиста (m) в процессе индивидуализированного аудиторного обучения с позиции учителя как ведущего субъекта в задаче управления учебным процессом. Учитель и методист (хотя бы формально) будут в отношении союза, т. к. от их согласованного взаимодействия зависит корректность выполнения требований образовательных стандартов, придающих легитимность учебному процессу. Ученик будет вынужден работать с учителем, но отрицательно относиться к нормативам и требованиям, соответствие которым формально контролирует методист. Их взаимодействие показано на рис. 1, и в виде стратифицируемого графа, который можно представить как $y(m+x)$, а также полиномом в виде (1). Заметим, что здесь и далее представлены графы, отражающие наиболее часто встречающиеся типы взаимоотношений действующих лиц процесса индивидуализированного обучения. Их не следует воспринимать в качестве однозначно возможных.

$$[y(m+x)]^{[y][m+x][m]+[x]} \quad (1)$$

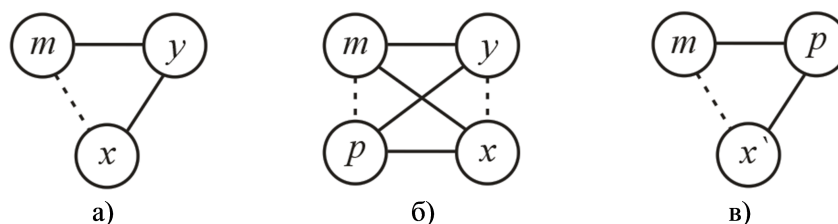


Рис. 1. Сущности, взаимодействующие в индивидуальном учебном процессе

Эффективность обучения во многом обеспечивается за счёт ведущей роли учителя. У него формируется все три компоненты образа ученика через непосредственное взаимодействие с ним, что позволяет выбирать соответствующие учебной ситуации методы управления. Погрешность в образ вносит несовершенство механизмов оценки результатов обучения (оцениваются латентные показатели) и качество контрольно-измерительных материалов. Также имеется некоторое снижение потенциально возможной активности и мотивации к обучению у ученика: учитель, опираясь на нормативы, вынужден применять автори-

тарный стиль управления процессом обучения, манипулируя преимущественно посредством влияний (первый тип управления по [5]). Если очное обучение предполагает введение предметного тьютора (p), отстаивающего права ученика на самореализацию, то граф взаимоотношений будет выглядеть так, как показано на рис. 1(б). Тьютор (также как и учитель) вынужден соблюдать рамки учебного процесса, заданные методистом, но у него в приоритете стоят цели ученика, что приводит к «напряжённости» в процессе согласования содержания учебного процесса с учителем. В результате мотивация к обучению у ученика возрастёт, как и степень индивидуализации. Образ ученика у учителя теперь формируется наиболее эффективно, т. к. были специально выявлены и учтены личностные приоритеты, что можно описать в виде полинома $(m + x)(y + p)$, имеющего диагональную форму (2). Таким образом, тьютор в своей деятельности осуществляет воздействие как на ученика, так и на учителя, управляя через изменение отношений и через изменение порядка значимости компонентов дидактического материала (второй и третий типы управления по [5]).

$$[(m + x)(y + p)]^{[m+x][m]+[x][y+p][y]+[p]} \quad (2)$$

На рис. 1(в) показан стратифицируемый граф для дистанционной формы обучения, когда предметный тьютор сопровождает учебный процесс, дидактический материал для которого когда-то был подготовлен учителем для абстрактного ученика (диагональная форма полинома для этого случая будет практически совпадать с (1)). Качество обучения существенно падает, т. к. предметный тьютор непосредственно не контактирует с учеником (используются современные каналы связи), что заставляет его оценивать успехи обучения на основании впечатления от удалённого общения. В результате формируется упрощённый образ x' , сильно уступающий по качеству при непосредственном взаимодействии с учеником. Первые два компонента образа ученика носят фрагментарный характер, а третий — практически сведён к соблюдению рамок, заданных методистом. Ещё одним недостатком такого взаимодействия является формальное (и не всегда достаточно квалифицированное) выполнение предметным тьютором роли учителя. Анализируя существующую практику по применению данной формы обучения при базовой подготовке специалистов и магистров [8], следует констатировать её низкую эффективность, как раз за счёт слабости связи между сущностями p и x . Как следствие, вновь применение управления за счёт манипулирования влияниями, т. к. для иных типов управления знаний становится недостаточно.

3. Цель, задачи и состав рефлексивного планировщика в ИАОС

Рассмотрим обучение на базе ИАОС: теперь ситуация принципиальным образом меняется. Действующих лиц фактически становится двое — это интеллектуальный планировщик ИАОС и ученик. Не используя рефлексивный подход, модель ученика в рамках автоматизированных обучающих систем пред-

ставляется в виде совокупности различных его характеристик (см., например, модели из [3]). Но так как человеко-машинное взаимодействие имеет меньшие возможности для формирования x' в памяти компьютера, то эффективность вырабатываемых такими обучающими системами воздействий на ученика может быть повышена за счёт введения в ИАОС функций рефлексии. Далее будем полагать, что рассматривается современная ИАОС, в которой наличие функции рефлексии оправдывается необходимостью повышения результативности принятия решений. Более примитивные планировщики, присущие системам первого-третьего поколений обучающих систем [4], обеспечить эффективное выполнение подобной задачи принципиально не в состоянии.

Обучение при помощи ИАОС имеет достаточно чёткую схему реализации, унаследованную от компьютеризированной формы программированного обучения. Отметим базовые этапы, формирующие множество U : u_1 — инициализация индивидуализированного процесса обучения на начальной фазе или при пересмотре учеником целей; u_2 — изучение фрагмента теоретического учебного материала; u_3 — закрепление новых знаний на практике; u_4 — промежуточный контроль знаний; u_5 — итоговый контроль знаний. Как правило, первый и последний этапы происходят однократно, а остальные — итерационно. Переход между фрагментами дидактического материала изначально осуществляется в соответствии с рабочей программой, но для каждого ученика траектория будет различна. В конкретный момент времени t взаимодействие ученика и ИАОС будет определяться активностью в текущем этапе из U в состоянии принятия / выполнения решения a_q , принадлежащего множеству альтернатив A_q решаемой «проблемы» q , в свою очередь принадлежащей общему множеству задач принятия решений Q . Минимальная структура рефлексивного планировщика обучающей системы включает в явном или не в явном виде описание моделей учителя u^* и методиста m^* , обладающих собственными фрагментами базы знаний. Исходя из потребностей индивидуализации и ограниченности возможностей человеко-машинного взаимодействия, планировщику не обойтись и без модели (образа) ученика x' , формируемой в памяти ИАОС. Далее, для обозначения той части планировщика, которая отвечает за имитацию процессов рефлексии, будем использовать словосочетание *интеллектуальный рефлексивный агент* (ИРА) [9]. Каково предназначение ИРА в процессе автоматизированного обучения на базе ИАОС? Оно, по нашему мнению, должно заключаться в повышении качества взаимодействия между обучающей системой и учеником, участвуя в решении задачи выработки планировщиком индивидуализированных педагогических воздействий (принятия решений) на всех этапах учебного процесса за счёт имитации функций рефлексии. Отметим основные укрупнённые задачи, при решении которых в ИАОС должен участвовать ИРА:

- 1) формирование и уточнение образа ученика с опорой на его индивидуальные психофизиологические характеристики и целевые установки;
- 2) комплексный анализ учебной ситуации, а также выдвижение и проверка сопутствующих гипотез;
- 3) формирование набора альтернатив действий, на основании которых осуществляется принятие решений по вопросам организации и сопровожде-

ния учебного процесса;

- 4) объяснение действий ИАОС (аргументированные ответы на вопросы учащегося «почему?», «зачем?», «каким образом?») через естественно-языковой диалог;
- 5) поиск компромиссов между точкой зрения ученика и ИРА;
- 6) формирование ситуационных и комплексных рекомендаций / подсказок (включая предоставление ссылок на элементы тезауруса, справочного материала и пр.).

На рис. 2 (левая часть схемы) представлен граф отношений между указанными моделями из ИРА: видно, что в памяти машины частично имитируются отношения из традиционного взаимодействия с рис. 1(а). Отношение между x^* и y^* в ИРА имеет некоторую двойственность: с одной стороны, ученик вынужден работать с ИАОС для достижения целей обучения, а с другой — авторитет и возможности виртуального учителя значительно уступают аналогичным характеристикам человека-учителя (для обозначения этой двойственности используем тип линии «точка-пунктир»). И рефлектирующая ИАОС, и человек-ученик в процессе обучения будут активны в смысле теории В.А. Лефевра. Направлениями рефлексивного анализа, который должна обеспечить ИАОС, будут следующие группы гипотез: гипотезы этапа инициализации процесса обучения (выбор начальных условий при ответе на вопрос «чему учить?»); гипотезы процесса сопровождения прочих этапов учебного процесса (ситуационный выбор и применение педагогических воздействий при ответе на вопрос «как учить?»). Очевидно, что если ограничиться имеющимся составом моделей при решении задачи управления обучением со стороны системы, то она стремится к авторитарному стилю взаимодействия (модели учителя и методиста формируют группу). Это, в свою очередь, приводит к своеобразному конфликту интересов, при котором ученик, в силу логики взаимодействия, имеет мало шансов аргументировано доказать необходимость учитывать его собственные цели и предпочтения. Таким образом, состав ИРА, опирающийся только на модели ученика, учителя и методиста, не позволяет полноценно решить поставленные перед ним задачи (особенно пятую) и нуждается в дополнении.

Спецификой взаимодействия ученика с автоматизированными обучающими системами является то, что инициатива всегда находится у ученика: это касается как работы с дидактическим материалом, так и ведения диалога с планировщиком системы. В этих условиях игнорировать цели ученика не продуктивно, т. к. он может сократить объёмы работы с электронным курсом в ИАОС, осознанно или не осознанно считая её логику работы не конструктивной. Поэтому введём в состав ИРА модель предметного тьютора p^* , которая будет обеспечивать выполнение четвертой и пятой задач агента, играя роль примирителя (правая часть схемы отношений на рис. 2).

4. Базовая модель рефлексивного восприятия ИРА

Модель рефлексивного восприятия ИРА ученика и учебной ситуации определяется следующими компонентами:

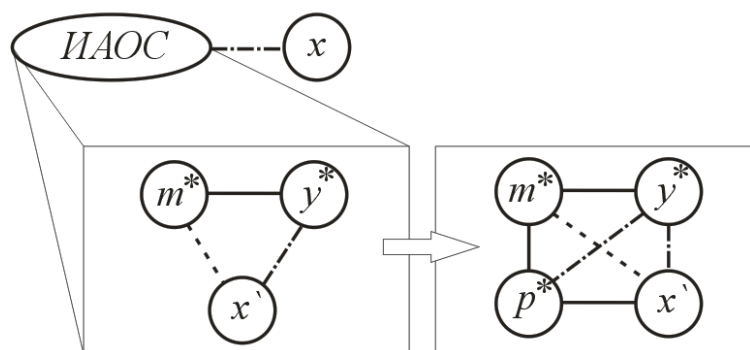


Рис. 2. Детализация структуры ИРА

- состоянием группы образов субъектов на момент принятия решений в памяти ИАОС;
- графом отношений между субъектами для текущего этапа обучения;
- набором действий ИАОС для сложившейся ситуации, представленного в виде универсального множества;
- унарными отношениями реализуемости действий для каждого субъекта в группе;
- матрицей влияний, определяющей желаемое поведение субъекта прочими участниками конфликтной ситуации.

Для рассматриваемой задачи все субъекты (модели из ИРА) являются целенаправленными. Модель x' итерационно уточняется планировщиком, исходя из реального взаимодействия ИАОС с учеником-человеком. Кроме того, цели ученика могут изменяться в процессе обучения, что может привести к пересмотру условий организации учебного процесса. Ведущую роль в принятии системой тех или иных решений будут занимать различные модели, обрабатываемые планировщиком (см. таблицу). На основании этого, между предметным тьютором и учителем, а также учителем и учеником вид взаимодействия союз/конфликт будет зависеть от этапа обучения. Приведём, в качестве примера, стратифицируемый граф отношений для этапа u_2 (рис. 3(а)), касающийся выработки управляющих воздействий (например, корректирование траектории работы с дидактическим материалом электронного курса). Опираясь на образ ученика, у ИРА появляется возможность осуществлять манипулирование путём изменения отношений и манипулирование порядком значимости, реализуемое через диалог в естественно-языковой форме. Диагональная форма полинома, позволяющего описать такую ситуацию, представлена в (3).

$$[y * p * (m * + x)]^{[y*][p*][m*+x]^{[m*]+[x]}} \quad (3)$$

Более сложный пример приведён на правом фрагменте рисунка 3, где рассматривается этап u_1 . Видно, что при принятии базовых решений (выбор ин-

Таблица 1. Ведущая роль моделей в ИРА на различных этапах процесса обучения

Этап	Номера задач	Ведущая модель в ИРА
u_1	1, 4, 5	Тьютор
	3	Учитель и Тьютор
u_2, u_3, u_4	1, 4	Тьютор
	2, 3	Учитель и Тьютор
	6	Учитель
u_5	2	Учитель
	3	Учитель и Методист

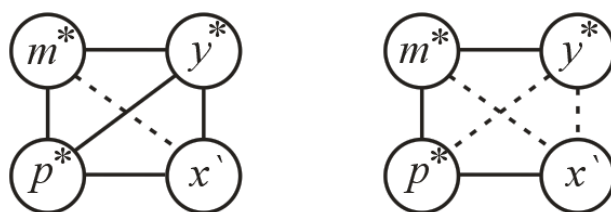


Рис. 3. Примеры графов отношений в процессе рефлексии ИРА

дидуального состава курса, определение базовой траектории обучения, выработка норм контроля, определение методов и стилей обучения), задающих индивидуальную рабочую программу, имеется конфликт отношений, делающий граф не декомпозируемым. С одной стороны, модель предметного тьютора пытается «отстоять» позицию модели ученика, в то время как модели методиста и учителя пытаются достичь своих целей при определении содержания будущего учебного процесса. С другой стороны, p^* способствует поиску компромиссных решений, привлекая дополнительные механизмы логического вывода решателя, опирающиеся на технологию анализа когнитивных карт диагностики знаний [10]. Это уже выходит за рамки традиционной модели В.А. Лефевра, в которой предпочтение отдаётся стратегии последовательного исключения субъектов из конфликта до приведения графа к декомпозируемому виду [5]. Играя роль примирителя, модель предметного тьютора снижает степень напряжённости конфликта между остальными моделями, используя в качестве основного инструмента манипулирование путём изменения отношений и манипулирования порядком значимости. Это особенно ценно для повышения эффективности электронного обучения при работе со слабо замотивированными учащимися, проявляющими так называемый эффект «когнитивного иммунитета» [11].

5. Заключение

Завершая рассмотрение вопроса о рефлексивных функциях в интеллектуальных автоматизированных обучающих системах, следует сделать ряд обобщений:

- 1) рефлексивное восприятие и принятие решений при сопровождении индивидуализированного учебного процесса, затрагивающего все этапы работы учащегося с ИАОС, базируется на копировании модели поведения человека и основывается на поиске компромиссов;
- 2) при реализации рефлексивного управления в автоматизированном обучении наиболее перспективно применять методы, обеспечивающие манипулирование путём изменения отношений и манипулирование порядком значимости по В.А. Лефевру;
- 3) для ряда ситуаций, когда граф отношений становится не декомпозируемым, а принимать решения необходимо, рационально дополнять оценки учебной ситуации, полученные от рефлексивного агента, дополнительными выводами интеллектуального планировщика, опирающегося на собственную базу знаний. Все эти функции в комплексе ориентированы на повышение адекватности работы обучающих систем четвёртого поколения [4]. Механизмы выработки педагогических воздействий, управляющих индивидуализированным процессом обучения, постепенно приближаются к аналогичному процессу у человека-учителя. При этом все решения принимаются ситуационно и аргументировано, т. к. сопровождаются актами рефлексии.

ЛИТЕРАТУРА

1. Беспалько В.П. Обучение и образование с помощью компьютеров. Воронеж: МОДЭК, 2002. 352 с.
2. Углев В.А. О специфике индивидуализации обучения в автоматизированных обучающих системах // Философия образования. 2010. № 2. С. 68–74.
3. Карпенко А.П., Добряков А.А. Модельное обеспечение автоматизированных обучающих систем. Обзор // Наука и образование: научное издание МГТУ им. Н.Э. Баумана [Электронный журнал]. 2011. № 7. URL: (<http://technomag.bmstu.ru/doc/193116> (дата обращения 10.01.2012)).
4. Углев В.А. Автоматизированные обучающие системы нового поколения: основные черты и принципы организации // Перспективные методы и средства интеллектуальных систем: материалы Всероссийского научно-практического семинара и школы молодых ученых. Новосибирск: Изд-во НГТУ. 2015. С. 37–40.
5. Лефевр В.А. Лекции по теории рефлексивных игр. М.: Когито-Центр, 2009. 218 с.
6. Бизяева А.А. Рефлексивные процессы в сознании и деятельности учителя: дис. ... канд. псих. наук. СПб., 1993. 238 с.
7. Артюшина Л.А. Рефлексия как необходимый компонент автоматизированной обучающей системы: постановка проблемы // Алгоритмы, методы и системы обработки данных: сборник научных статей. 2009. Вып. 14. С. 12–19.

8. Углев В.А. Автоматизированные обучающие системы и компьютерное тестирование: системный кризис и перспективы развития // Модернизация российского образования: тренды и перспективы. Ч. 2. Краснодар: Премьер, 2011. С. 3-18.
9. Углев В.А. Интеллектуальный рефлексивный агент в структуре автоматизированных обучающих систем // Нейроинформатика, её приложения и анализ данных: XXIII Всероссийский семинар. Красноярск, 2015. С. 130–135.
10. Uglev V.A. Implementation of Decision-making Methods in Intelligent Automated Educational System Focused on Complete Individualization in Learning // AASRI Procedia, 2014. V. 6. P. 66–72. DOI 10.1016/ j.aasri.2014.05.010.
11. Филимонов В.А. Когнитивный иммунитет как проблема и когнитивная инфраструктура как ресурс // Робототехника и искусственный интеллект. Сб. статей VII Всероссийской научно-технической конференции с международным участием (Железногорск, 11 декабря 2015 г.). Красноярск: Изд-во Сибирского федерального ун-та, 2016. С. 181–186.

THE BASE MODEL OF REFLEXIVE PROCESS IN THE INTELLECTUAL AUTOMATION EDUCATIONAL SYSTEMS

V.A. Uglev

Ph.D.(Eng.), Associate Professor, e-mail: uglev-v@yandex.ru

Siberian Federal University, Zheleznogorsk, Russia

Abstract. The issues of creating intelligent automated training system with the function of reflection are considered. Using the reflexive games theory of V. A. Lefebvre, the structure of the scheduler of training system introduces a reflexive intelligent agent (RIA), which provides trade-offs in the interaction models of a student, teacher, subject tutor and practitioner. The basic components of the reflexive perception model of RIA of student and learning situation are described, taking into account the specifics of decision-making for different stages of learning.

Keywords: modeling, individualization of learning, automated learning, reflexive governance, reflexive intelligent agent.

Дата поступления в редакцию: 28.01.2018

КОНСТРУКТОР МОДЕЛЕЙ СУБЪЕКТОВ РЕФЛЕКСИВНЫХ ИГР: ВЕРСИЯ 1

В.А. Филимонов

д.т.н., с.н.с., профессор, e-mail: filimonov-v-a@yandex.ru

Институт математики им. С.Л. Соболева СО РАН, Омск, Россия

Аннотация. Конструктор субъектов рефлексивных игр рассматривается как компонент поддержки рефлексивного управления в проекте «Ген-Гуру». Проект реализуется с использованием кросс-технологий ситуационного центра. Анализируются ключевые идеи системы моделей В.А. Лефевра. Перечисляются варианты детализации моделей для повышения точности моделирования ситуаций. Рассматривается специфика моделирования таких субъектов, как индивиды и роботы.

Ключевые слова: рефлексивное управление, модели В.А. Лефевра, детализация моделей, субъекты.

Введение

Настоящая статья посвящена описанию фрагмента проекта «Ген-Гуру» [1] в части создания системы поддержки рефлексивного управления. Этот фрагмент относится к созданию конструктора моделей субъектов, потенциально являющихся предметом исследования и рефлексивного управления. Здесь можно выделить две крайние позиции относительно инструментов и технологий. Первую позицию можно назвать «мягкой»; к ней относятся люди, использующие методы, связанные с психологией, эзотерикой и т. п. В качестве примера укажем на У. Черчилля, который использовал астролога Луи де Воля в качестве модели Карла Крафта — астролога А. Гитлера. Вторую позицию назовём «жёсткой» и отнесём к ней модели, использующие достаточно изощённый математический аппарат. В качестве примера укажем на монографию [2]. Наш подход в рамках упомянутого проекта представляет собой комбинацию обеих позиций, поскольку ориентирован на возможность практического применения. Примером такого подхода является «калькулятор стратагем» В.Н. Крючкова [3]. В качестве основы мы используем теоретический аппарат, разработанный В.А. Лефевром [4–8]. В первой части статьи рассмотрены модели субъектов и их взаимодействий в рамках данного аппарата. Далее рассмотрены возможности детализации этих моделей.

1. От формулы человека до эйдос-навигатора

1.1. Модель субъекта

Модели В.А. Лефевра — уникальное явление в науке. Они относительно просты по форме и исключительно глубоки по содержанию. Это позволило с их помощью описать фундаментальные законы поведения и мышления. С этих моделей мы начинаем создание нашего конструктора.

На первом этапе (1967 г.) был предложен комплекс понятий и гипотез, связанных с рефлексивным анализом и управлением, а также соответствующая схематизация [5]. В 1991 г. появилась «формула человека», которая была кибернетической моделью, позволяющей вычислять предполагаемое поведение субъекта в целой системе ситуаций. Было введено биполярное представление субъектом картин мира и его реакции, которое использовано во всех дальнейших моделях. Были введены понятия «негативный полюс» и «позитивный полюс», которые интерпретировались соответственно как «зло» и «добро», и кодировались как «0» и «1». В модели предполагалось, что реакция X_1 субъекта, иначе, объективная готовность выбрать позитивный полюс, определяется тремя параметрами. Этими параметрами являются: воздействие x_1 (давление среды в сторону позитивного полюса), восприятие x_2 этого давления (что определялось опытом субъекта), а также его намерение x_3 выбрать позитивный полюс: $X = G(x_1, x_2, x_3)$. На основе аксиом свободы выбора $G(0, 0, x_3) = x_3$, незлонамеренности $G(1, x_2, x_3) = 1$, вреда абсолютной доверчивости $G(0, 1, x_3) = 0$, и в предположении линейной зависимости от переменных «формула человека» записывается в виде (1):

$$X = G(x_1, x_2, x_3) = x_1 + x_3(1 - x_1)(1 - x_2). \quad (1)$$

В книге «Космический субъект» [5] (1996 г.) рассмотрено много интерпретаций этой формулы, а также дано её обоснование на основе законов термодинамики. В [9] рассмотрена другая система аксиом, приводящая к «формуле человека». В [8] (2013-2017 гг.) в поддержку тезиса Платона о реальном существовании идеальных объектов сформулирована гипотеза об эйдос-навигаторе, определяющем поведение субъектов в точках формирования различных траекторий их существования. В модели субъекта вместо параметра «намерение» x_3 использована его детализация X_2 через «образ себя». Ситуация поведения в случае биполярного выбора описывается парой уравнений (2):

$$X_1 = x_1/(x_1 + x_2 - x_1x_2); X_2 = x_2/(x_1 + x_2 - x_1x_2). \quad (2)$$

В обоих вариантах формулы субъекта его образ себя подобен ему самому:

$$G(.) = F(x_1, F(x_2, x_3)), F(a, b) = 1 - a + ab; X_1/x_1 = X_2/x_2. \quad (3)$$

Такое представление субъекта позволило, в частности, описать феномен нелинейности категориальной экспертной оценки, появление «золотого сечения» в ситуации выбора из одинаковых альтернатив, дать корректные интерпретации законов Вебера-Фехнера и Стивенса, а также закона соответствия (Matching Law).

1.2. Модель взаимодействия двух субъектов

В [6] (1982–2003 гг.) рассматривается взаимодействие субъекта с самим собой (взаимодействие образов себя) и с другим субъектом. Это взаимодействие интерпретируется как автомат, реагирующий на воздействия внешнего мира. Виды взаимодействия трактуются как «конфликт» и «союз», образы себя и других могут иметь квалификацию «корректный» либо «некорректный». Постулируется, что реальный (истинный) образ субъекта доступен только внешнему наблюдателю.

Используется расширенный вариант булевой алгебры (гамма-алгебра) двух элементов: 0 и 1. Автором предложен экспоненциальный формат модели, который позволяет наглядно представить рефлексивную структуру субъекта. Эта структура является одновременно и когнитивной вычислительной схемой. Кратко опишем данный формат.

Логическая импликация $b \rightarrow a = \bar{b} + a$ записывается в виде a^b . Любые цепочки импликаций могут быть записаны в экспоненциальной форме, а сами такие формулы могут быть преобразованы с использованием дополнительных соотношений. Для понимания дальнейшего достаточно знать следующие правила: $a^b = a + (1-b)$ и $a^{b^c} = a + (1-b) \bullet c$. Сами формулы могут быть интерпретированы как субъекты либо ситуации, либо не иметь ни одну из этих интерпретаций. Поясним способ интерпретации на примере формулы (4), приведённой ниже.

$$A + B = a^{\bar{a}} + \bar{b}^a + b^{\bar{b}} \bullet \bar{a}^b. \quad (4)$$

Данная формула интерпретируется следующим образом. Субъекты A и B с позиции внешнего наблюдателя находятся в конфликте (символ «+»), причём у A корректный образ этого отношения, а B ошибочно полагает, что они находятся в союзе (символ «•»). У обоих субъектов некорректные образы себя и партнёра. A видит себя как $\bar{a}$, при этом он принимает корректный образ себя a у образа партнёра за ошибочный. У B ситуация аналогичная.

На рис. 1 представлена общая схема ситуации взаимодействия двух субъектов.

Образы себя, другого субъекта и ситуации в целом позволяют дать содержательную интерпретацию значений характеристик упомянутого автомата для различных ситуаций. Эти интерпретации приведены в табл. 1.

Есть основание предположить, что в обществе существуют субъекты, которые стремятся максимизировать различные характеристики, в нашем случае это гордость, комфорт и т. п. Подчеркнём, что указанные характеристики субъект оптимизирует не у себя (себя он наблюдать не может в принципе), а у образа себя. Чтобы делать это осознанно, субъекты должны иметь достаточно сложную структуру, в частности, иметь представление о себе, партнёре и своих отношениях. В наших обозначениях это означает, что структура таких субъектов должна быть не менее сложной, чем $a^{ar_2b}ar_1b$. Поскольку каждое из

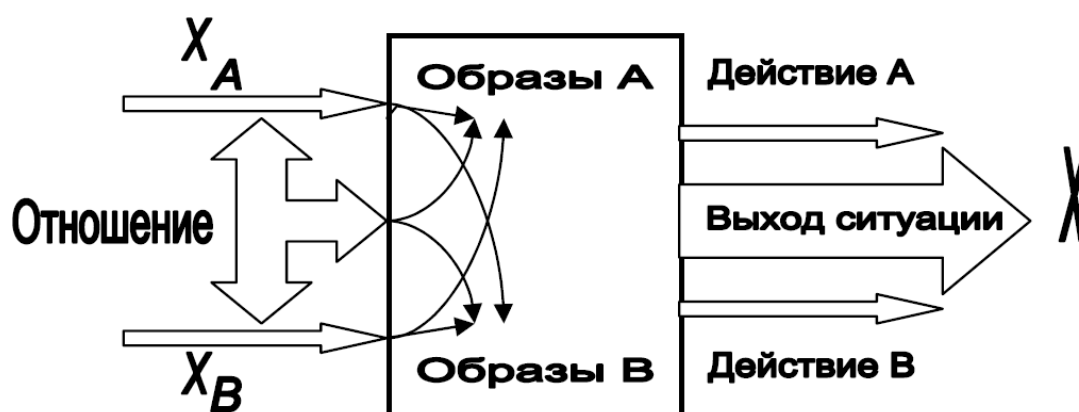


Рис. 1. Схема автомата для описания ситуаций взаимодействия

Таблица 1. Интерпретация выходов автомата

Компонент (предмет оценки)	Оценка компонента	
	«1» (Позитивная)	«0» (Негативная)
Собственное поведение	Гордость	Вина
Поведение партнёра	Одобрение	Обвинение
Ситуация в целом	Комфорт	Страдание

двух отношений r_1 и r_2 в этой формуле может принимать одно из двух значений «+» и «•», получаем всего четыре возможных типа субъектов, названных нормативными. Их определения приведены в табл. 2.

Названия субъектов обусловлены вычислением их характеристик как автоматов по формулам, указанным в табл. 2. Общее предположение состоит в том, что на входы автомата подаются последовательности нулей и единиц, что интерпретируется как побуждение субъектов к совершению поступков: хороших (1) либо плохих (0). Частное предположение заключается в равновероятном появлении на входах нулей и единиц. В этом случае для получения характеристик требуется вычислить среднее значение выходов автоматов для 4-х комбинаций входов: 0, 0; 0, 1; 1, 0; 1, 1. Результаты вычислений для многих вариантов представлены в [6].

Объём и задача статьи обусловили только краткое упоминание одного из важнейших результатов этой работы: обнаружение, теоретическое и экспериментальное, двух этических систем. Фундаментальное различие этих систем заключается в различной оценке результата взаимодействия добра и зла; в частности, ситуация борьбы добра и зла в одной системе рассматривается как добро, а в другой как зло.

Таблица 2. Нормативные субъекты

Нормативные субъекты			
Праведник	Герой	Обыватель	Лицемер
$a^{a+b} \bullet b$	$a^{a \bullet b} \bullet b$	$a^{a+b} + b$	$a^{a \bullet b} + b$
максимум вины и страдания	минимум вины	минимум страдания	максимум гордости и комфорта

1.3. Группы субъектов: рефлексивные игры

В монографии [7] (2009 г.) рассмотрены модели групп субъектов, которых должно быть более двух, чтобы появилась возможность рассчитывать отношения между отношениями. Снова предполагается, что отношения между субъектами могут быть либо конфликтом, либо союзом. Примеры представления групп в виде графов даны на рис. 2, где субъекты обозначены буквами, а отношения — линиями, сплошными для союза ($\bullet$) и пунктирными для конфликта ($+$). Дополнительно для каждого субъекта задаются наборы альтернатив (его потенциально возможных действий). Кроме того, задаётся также матрица влияний, где для каждого субъекта представлены наборы альтернатив, к выбору которых его побуждают другие члены группы, а также его собственные намерения. Для альтернатив не вводятся какие-либо дополнительные параметры типа ценности по различным шкалам. Методика расчёта состоит в последовательной стратификации графа на подграфы (подгруппы, находящиеся только в одном из двух указанных отношений). Основная идея расчёта состоит в том, что для подгруппы, находящейся в союзе, суммарное влияние представляется пересечением множеств альтернатив членов подгруппы, а для конфликта — объединением этих множеств. Обнаружено, что существует недекомпозируемая таким образом структура — граф S_4 , представленный на рис. 2 (а), однако, и для этого случая предложена методика расчёта.

В указанной монографии доказана также теорема о том, что существует группа субъектов с декомпозируемым графом отношений такая, что в ней есть субъект, который, в зависимости от влияний на него других субъектов, может находиться в любом из четырёх возможных состояний: активном, пассивном, свободы выбора и фрустрации (невозможности совершить выбор). Граф соответствующей группы приведён на рис. 2 (б). Полином этого графа представлен уравнением (5), по которому строится диагональная форма (6), и окончательно получается уравнение (7), которое служит для расчёта поведения субъекта a .

$$(a + b(c + d))(e + f), \quad (5)$$

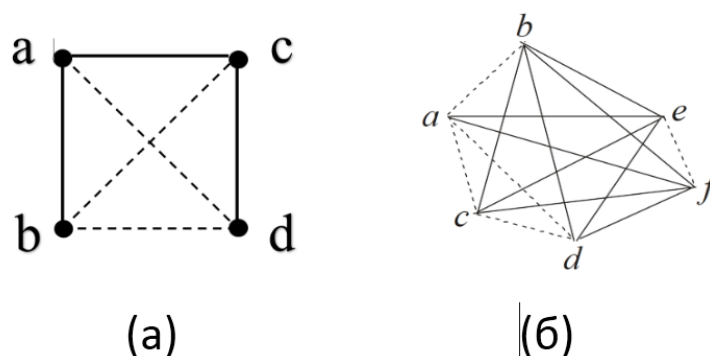


Рис. 2. Недекомпозируемый граф (а); граф к теореме 2 о разнообразии (б)

$$[a+b(c+d)(e+f)][a+b(c+d)][a] + [b(c+d)][b][c+d][c] + [d] + [e+f][e] + [f], \quad (6)$$

$$a = (a + b(c + d)(e + f) + \bar{a}\bar{b}). \quad (7)$$

Эти модели позволили, в частности, объяснить парадокс миротворца и обосновать структуру потенциально совершенной судебной системы. Интересные варианты исследования этих моделей в динамике представлены в работах С. Тарасенко [17].

2. Возможности детализации моделей

2.1. Общие соображения

Мы используем термин «субъект», а не «индивид», поскольку последний относится обычно к биологическим структурам. Развитие робототехники требует рассмотрения в качестве субъектов, помимо индивидов, роботов, а также системы искусственного интеллекта. Отметим здесь инициативу Европарламента по признанию роботов «электронными личностями», а также патент Google 2015 г. на воссоздание личностей живущих и умерших людей в роботах и распределённых вычислительных системах. Уже в настоящее время достаточно распространены чат-боты, которые наделены определёнными способностями к выбору. Упомянем также методы сенсорной подстановки, позволяющие трансформировать одни типы ощущений в другие.

Экспериментальная база для построения моделей субъектов рефлексивного управления нуждается в пополнении. В частности, в [2] отмечено, что применение предложенных в указанной монографии моделей — дело достаточно отдалённого будущего из-за отсутствия экспериментальных данных. Что касается моделей В.А. Лефевра, то здесь можно указать на эксперименты по

сортировке фасоли, проведённые Викторией Лефевр, а также на эксперименты по определению этических систем, описанные в [6] и независимо проведённые В.В. Шаталиной в Великобритании. Упомянем также эксперимент «Лабиринт» [5], проведённый в 1969 г., который продемонстрировал возможность использования информационного противодействия для оптимизации функционирования.

Безусловно, любые варианты моделей и их детализации нуждаются в экспериментальной проверке. Нами запланированы эксперименты такого рода [12]. Отметим, что наиболее сложным компонентом таких экспериментов является формирование конфликтных отношений. Конфликты должны быть достаточно выражены, однако участники при выходе из эксперимента не должны уносить их следы. Конфликты в компьютерных играх не удовлетворяют первому требованию. Конфликты в некоторых психологических тренингах не удовлетворяют второму требованию, даже при проведении дебрифинга.

2.2. Варианты детализации

Различные варианты детализации были рассмотрены в наших работах [10–16]. Одним из достаточно очевидных вариантов является использование моделей личности. Ещё в 1994 г. В.Ю. Крыловым было предложено для обобщения модели субъекта В.А. Лефевра использовать соционику и транзакционную модель Э. Бёрна [9]. Нами был предложен вариант соционики для робототехники, который также может быть использован.

Фундаментальным свойством моделей В.А. Лефевра является двухтактная («ощущение» — «осознание ощущения») циклически повторяемая процедура рефлексивной возгонки, что напрямую не следует из термодинамической модели [8]. Ввиду важности такого представления запишем его формализованно в виде (8) по аналогии с (4). Мы используем следующие обозначения: P — восприятие (от *perception*), A — осознание (от *awareness*), P of Z — ощущение Z , A of Z — осознание Z .

$${}_a[P][AofP][Pof(AofP)][Aof(Pof(AofP))]\dots \quad (8)$$

В публикациях [13, 16] была рассмотрена потенциальная возможность совместного использования моделей гомеостатики Ю.М. Горского и моделей рефлексивного управления В.А. Лефевра для моделирования популяций с рефлексией. В модели Ю.М. Горского принципиальным моментом является наличие противоречия между двумя Исполнителями, которое управляется Шефом. В модели В.А. Лефевра принципиальным является наличие противоречия в форме позитивного и негативного полюсов биполярного выбора субъекта. Именно возможность констатации таких полюсов, как предпосылка биполярного выбора, является фундаментальным свойством одушевлённых систем. В этой модели нет прямого аналога Шефа, но есть возможность подбора группы с

отношениями союза и конфликта, которая будет формировать заданное поведение Исполнителей. В модели В.А. Лефевра есть также внешний навигатор, определяющий поведение одушевлённого существа в точках бифуркации [8].

Ещё одним направлением является вложенное применение моделей В.А. Лефевра: модели субъекта в модели взаимодействий пары субъектов, а затем в модели группы.

3. Заключение

Разработка конструктора субъектов рефлексивных игр является исследовательским проектом, находящимся в стадии реализации в Институте математики им. С.Л. Соболева СО РАН (Омский филиал). Автор приглашает к участию своих коллег как математиков, так и представителей гуманитарных направлений. Кроме того, автор приглашает студентов любых специальностей и уровня подготовки для участия в экспериментах (рефлексивных играх и экспертизах), направленных на пополнение эмпирической базы рефлексивного управления.

Благодарности

Автор выражает благодарность Владимиру Александровичу Лефевру за его поддержку исследований в области рефлексивного управления, в особенности за взаимодействие с участниками ежегодных конференций «Рефлексивный театр ситуационного центра» в Омске, начиная с первой конференции 2007 г. и по настоящее время.

ЛИТЕРАТУРА

1. Филимонов В.А. Исследовательский комплекс «Ген-Гуру» // Матер. Всеросс. конф. с междун. участием «Знания–Онтологии–Теории». Т. 1. Новосибирск : Ин-т математики им. С.Л. Соболева СО РАН, 2007. С. 24–31.
2. Новиков Д.А., Чхартисвили А.Г. Рефлексивные игры. М.: СИНТЕГ, 2003. 149 с.
3. Крючков В.Н. Китайские стратагемы: формулы стратегии. М. : Т/О «Неформат», 2014. 56 с.
4. Лефевр В.А. Общая схема современной психологии. Место рефлексивных исследований в системе наук // Математические структуры и моделирование. 2018. № 1(45). С. 108–110.
5. Лефевр В.А. Рефлексия. М. : Когито-центр, 2003. 496 с.
6. Лефевр В.А. Алгебра совести. М. : Изд-во Когито-центр, 2003. 426 с.
7. Лефевр В.А. Лекции по теории рефлексивных игр. М.: Когито-центр, 2009. 218 с.
8. Лефевр В.А. Что такое одушевлённость. М. : Когито-Центр, 2017. 123 с.
9. Крылов В.Ю., Казанцев А.Ю. Модель рефлексивного поведения В.А. Лефевра: частные случаи, варианты аксиоматики, возможные обобщения // Методологические и теоретические проблемы математической психологии / В.Ю. Крылов. М. : Янус-К, 2000, С. 255–270.

10. Филимонов В.А. Алгебра логики и совести / Вступительное слово В.А. Лефевра. Учеб. пособие по математике для старших классов общеобразовательных и профильных школ. Омск : ОГИС, 2006. 72 с.
11. Филимонов В.А. Вариации на темы В.А. Лефевра // Рефлексивные процессы и управление. 2016. Т. 16, № 1–2. С. 95–99.
12. Филимонов В.А. Экспериментальное исследование моделей теории рефлексивных игр: постановка задачи // Математическое и компьютерное моделирование : сб. материалов IV Международной научной конференции. Омск : Изд-во Омского гос. ун-та, 2016. С. 62–65.
13. Филимонов В.А., Филимонова Т.А. Гомеостатические системы рефлексивного управления популяциями: подход к анализу и проектированию // Математические структуры и моделирование. 2014. № 3(31). С. 99–108.
14. Филимонов В.А. Когнитивная инфраструктура обучения людей и роботов (текст пленарного доклада) // Робототехника и искусственный интеллект — 2014 : материалы VI междун. науч.-практ. конф. Красноярск : Центр информации, 2014. С. 205–209.
15. Филимонов В.А. Соционика для роботов: прототип модели // Робототехника и искусственный интеллект : сб. статей VIII Всероссийской научно-технической конференции с международным участием. Красноярск : Изд-во Сибирского федерального ун-та, 2016. С. 215–220.
16. Филимонов В.А. Концепция альтернативной нейронной сети на основе гомеостатики и рефлексивного управления // Математическое и компьютерное моделирование : сб. материалов V Международной научной конференции. Омск : Изд-во Ом. гос. ун-та, 2017. С. 79–81.
17. Tarasenko S. Applications of the Reflexive Game Theory: Advanced Topics. Lulu, 2017. 134 p.

CONSTRUCTOR OF MODELS OF SUBJECTS OF REFLEXIVE GAMES: VERSION 1

Filimonov V.A.

Dr.Sc. (Eng.), Senior Researcher, Professor, e-mail: filimonov-v-a@yandex.ru

Sobolev Institute of Mathematics, Omsk, Russia

Abstract. The constructor of subjects of reflexive games is considered as a component of support for reflexive control in the project "Gen-Guru". The project is implemented with the use of cross-technologies of situational center. Key ideas of V. A. Lefebvre's model system are analyzed. List for the options of detail models to improve the accuracy of modeling situations is suggested. We consider the specifics of modeling such subjects as individuals, and robots.

Keywords: reflexive control, models of V. A. Lefebvre, details of the models, subjects.

Дата поступления в редакцию: 18.02.2018

О ФОРМАЛЬНОМ ОПРЕДЕЛЕНИИ АБСТРАКЦИИ

Е.А. Тюменцев

ген. директор, e-mail: etyumentcev@hwdtech.ru

LLC Hello World! Technologies, Омск, Россия

Аннотация. В статье предлагается формальное определение абстрагирования как отображения множества сущностей во множество подслов всех слов некоторого алфавита, иерархии абстракций как отношения на множестве подслов языка и интерпретации. Также получен ряд ограничений, которым обладает механизм абстрагирования.

Ключевые слова: абстракция, барьер абстракции, наследование, Гради Буч, Барбара Лисков, Роберт Мартин.

Введение

Абстракция — одно из важнейших понятий в программировании. По словам Ахо и Ульмана [1]: «*Computer Science is a science of abstraction — creating the right model for thinking about a problem and devising the appropriate mechanizable techniques to solve it*». При этом, как отмечено, например, в [2], его понимание и изучение, и, как следствие, применение на практике, вызывает серьёзные трудности. На наш взгляд, возможная причина связана с неформальным характером широко используемых определений. Например, вот одно из наиболее популярных, предложенное Гради Бучем [3]:

«Абстракция выделяет существенные характеристики некоторого объекта, отличающие его от всех других объектов».

Его часто используют в университетских курсах по программированию, цитируют или перефразируют в учебных пособиях и учебниках, например: [4, с. 172], [5, с. 375].

В своей книге Буч отмечает, что похожие формулировки можно встретить у Даль, Дейкстры и Хоара [6], Шоу [7], Берниса, Грея, Наумана [8] и др.

Отметим одну из особенностей этих определений: они объясняют назначение абстракции, но при этом никак не поясняют, что именно является абстракцией.

Ещё одно определение даёт Барбара Лисков в книге [9]:

«The process of abstraction can be seen as application of many-to-one mapping».

Однако она не поясняет ни область значений, ни область определения такого отображения, ни саму природу отображения «многие-к-одному».

В настоящей статье мы предлагаем формальное определение абстракции, иллюстрируем его применение на нескольких примерах и формулируем несколько ограничений, связанных с абстракцией, в виде утверждений. Публикуемые результаты были представлены на конференции [10].

1. Некоторые базовые определения

Нам потребуются базовые понятия теории формальных языков. Для удобства читателя приведём их в этой статье.

Начнём с символа, или буквы. Как и водится с исходными понятиями, мы не можем дать точное определение символа, а лишь попытаемся пояснить его смысл. Итак, *символ* — это некий знак, который рассматривается как неделимое единое целое. Чаще всего мы будем обозначать символы латинскими буквами a , b , c .

Все, что мы можем сказать про два произвольных символа, — это одинаковы они или нет. Причём одинаковость символов может трактоваться по-разному, в зависимости от ситуации. Например, следует ли считать одинаковыми буквы a и A латинского алфавита — зависит от того, хотите ли вы различать написание слов в верхнем и нижнем регистрах. Или одинаковые при написании буквы a кириллицы и буквы a латинского алфавита все же являются различными символами.

Тогда *алфавит* — это произвольное непустое конечное множество символов. Мы будем обозначать алфавит буквой Σ .

Словом в алфавите Σ называется произвольная конечная последовательность букв $a_1 a_2 \dots a_n$, где $a_i \in \Sigma, i \in (1, 2, \dots, n)$.

Подсловом слова $w = a_1 a_2 \dots a_n$, где $a_i \in \Sigma, i \in (1, 2, \dots, n)$ в алфавите Σ называется любая подпоследовательность $a_i a_{i+1} \dots a_j$, где $i \geq 1, i \leq j, j \leq n$.

На множестве слов определена операция конкатенации слов.

Формальным языком над алфавитом Σ называется произвольное подмножество конечных слов в алфавите Σ .

Пример 1. Рассмотрим язык C++. Процедура

```
void f()
{
    for(int i = 0; i < 10; ++i)
    {
        std::cout << i;
    }
}
```

является словом языка C++, а лексемы `void` `{ }` `( )` `::` `=` `;` `for` и т. д. — символами алфавита. Среди всех подслов данного слова будут

```

void f()
void f
std::cout << i;
for(int i = 0; i < 10; ++i)
{
    std::cout << i;
}
int i = 0;
i = 0
= 0
< 10; ++

```

Пусть L — формальный язык над некоторым алфавитом Σ . Будем обозначать $\check{L}$ — множество всех подслов каждого слова языка L .

Предложение 1. Для любого алфавита Σ множество Σ^* всех конечных слов в этом алфавите счётно.

Доказательство. Это хорошо известный факт из стандартного курса математической логики. ■

2. Определение

Определение 1 (Формальное определение абстракции). Пусть Σ — произвольный алфавит символов, $A \subset \Sigma^*$ — произвольное подмножество множества всех слов в алфавите Σ , X — некоторое множество. Тогда сюръективное отображение $F : X \rightarrow A$ называется *абстрагированием множества X над алфавитом Σ* , или, кратко, *абстрагированием множества X* , если это не вызывает разночтений. При этом элементы множества X будем называть *сущностями*, а элементы множества A — *абстракциями*.

Предлагаемое определение не привязано ни к какой методологии разработки программного обеспечения, и лишь отражает необходимость переноса сущностей предметной области на конструкции используемого языка программирования.

Пример 2 (Десятичная запись целого числа). Рассмотрим алфавит

$$\Sigma = \{-0123456789\}$$

и подмножество DR всех цепочек Σ^* вида:

$$d_1 d_2 \dots d_n,$$

где $n \geq 1, d_1 \in \{123456789\}, d_i \in \{0123456789\}, 2 \leq i \leq n$, или вида:

$$-d_1 d_2 \dots d_m,$$

где $m \geq 1, d_1 \in \{123456789\}, d_j \in \{0123456789\}, 2 \leq j \leq m$, либо цепочку 0.

Тогда отображение $IntD : \mathbb{Z} \rightarrow DR$, которое сопоставляет каждому целому числу его запись в десятичной системе счисления является абстрагированием.

Пример 3 (Двоичная запись целого числа в Java). Рассмотрим алфавит $\Sigma = \{-b01\}$ и подмножество BR всех цепочек Σ^* вида:

$$0bd_1d_2 \dots d_n,$$

где $n \geq 1, d_i \in \{01\}, 1 \leq i \leq n$, или вида:

$$-0bd_1d_2 \dots d_m,$$

где $m \geq 1, d_j \in \{01\}, 1 \leq j \leq m$, либо

$$0b0.$$

Тогда отображение $IntB : \mathbb{Z} \rightarrow BR$, которое сопоставляет каждому положительному целому числу строку символов вида $0b$ [запись числа в двоичной системе счисления], отрицательному целому числу строку символов вида $-0b$ [запись модуля числа в двоичной системе счисления], $0 - 0b0$, является абстрагированием.

Определение 2. Пусть $L \subset \Sigma^*$ – формальный язык над алфавитом Σ , $F : X \rightarrow Abs$ – некоторое абстрагирование множества X над алфавитом Σ . Тогда абстрагирование F применимо к языку L тогда и только тогда, когда $\forall x \in F(X) : x \in \check{L}$.

Пример 4. Абстрагирование $IntD$ из примера 2 применимо к языкам C++, Java, PHP, Python и многим другим, а $IntB$ из примера 3 применимо к языку Java, начиная с версии 7.

Как видно из последних примеров, образы одной и той же сущности предметной области могут отличаться относительно разных абстрагирований. Но может быть и обратная ситуация, когда образы нескольких сущностей относительно разных абстрагирований совпадают. В тексте программы представлены только абстракции. Значит, для понимания программного кода и предметной области необходимо по абстракциям восстановить набор использованных абстрагирований. В некоторых случаях, например, магические константы (в программе используются числовые константы без пояснения смысла), это сделать довольно сложно. Именованные константы, говорящие названия переменных и методов — это встраивание информации для простого и быстрого восстановления информации об абстрагированиях. Сравните:

2 и `int maxPlayers = 2;`

`int f(int a, int b);` и `int sum(int arg1, int arg2);`

3. Иерархия абстракций

Хорошо известный факт (классификация Хомского), что каждый формальный язык можно описать с помощью формальной грамматики. Формальная грамматика – это частный случай исчисления. Следовательно, используя правила вывода соответствующего исчисления – правила формальной грамматики – можно одни подслова и слова выводить из других.

Определение 3. Пусть L – произвольный формальный язык над некоторым алфавитом Σ , $Gr(L)$ – исчисление, в котором выводимы все слова языка L . Определим на множестве всех его подслов $\check{L}$ отношение Hr следующим образом:

$$\forall u, w \in \check{L} : (u, v) \in Hr \iff \exists w_1, w_2, \dots, w_n,$$

где $n \geq 1$, – вывод в исчислении $Gr(L)$, такой что $w_n = w$, $\exists j \in (1, 2, \dots, n) : w_j = u$, причём u встречается только один раз среди w_j , последовательность $w_1, \dots, w_{j-1}, w_{j+1}, \dots, w_n$ выводом не является. Отношение Hr называется *отношением иерархии на множестве всех подслов $\check{L}$, заданного с помощью исчисления $Gr(L)$* , или, кратко, если это не вызывает разночтений, *отношением иерархии*.

Предложение 2. Пусть L – произвольный формальный язык над некоторым алфавитом Σ , $Gr(L)$ – исчисление, в котором выводимы все слова языка L . Отношение иерархии Hr на множестве всех подслов $\check{L}$, заданного с помощью исчисления $Gr(L)$, является:

- рефлексивным,
- транзитивным.

Доказательство. Следует непосредственно из определения отношения иерархии и понятия выводимости в исчислении. ■

Определение 4. Пусть L – произвольный формальный язык над некоторым алфавитом Σ , $Gr(L)$ – исчисление, в котором выводимы все слова языка L , Hr – отношение иерархии на множестве всех подслов $\check{L}$, заданного с помощью исчисления $Gr(L)$. Предположим, что u, w – две абстракции, которые определены с помощью некоторых абстрагирований, применимых к языку L . Тогда u называется низкоуровневой абстракцией по отношению к w , а w – высокоуровневой абстракцией по отношению к u тогда и только тогда, когда $(u, w) \in Hr$.

4. Барьер абстракции и обобщение

С абстракцией связывают понятие барьера абстракции, когда в силу ограниченного числа выделенных характеристик одна и та же абстракция может сразу соответствовать нескольким сущностям предметной области, другими словами, эти сущности становятся неразличимыми. С точки зрения данного

нами определения существование барьера — это утверждение о том, что не всякая абстракция F обязательно является инъективным отображением.

Пример 5. Рассмотрим абстракцию

```
class Person
{
    std::string lastname;
    std::string firstname;
    ...
};
```

Эта абстракция не позволяет различать людей, которые являются тезками.

Предложение 3. Пусть L — произвольный формальный язык над некоторым алфавитом Σ . Существует такой набор сущностей X , что для любого абстрагирования F существует сущность $x \in X$, для которой

1. либо $F(x)$ не определена,
2. либо $\exists y \in X : F(x) = F(y)$.

Доказательство. Согласно утверждению 1 множество Σ^* счётно. Рассмотрим множество действительных чисел $\mathbb{R}$ и произвольное абстрагирование $F : \mathbb{R} \rightarrow A$, где $A \subset \Sigma^*$.

Предположим, что $\forall x_1, x_2 \in \mathbb{R} : F(x_1) \neq F(x_2)$, то есть F — инъективно. Тогда F определено на всем множестве и получаем взаимнооднозначное отображение $\mathbb{R}$ и A , но в таком случае $\mathbb{R}$ не более чем счётно, либо F определено не на всем множестве $\mathbb{R}$.

Предположим, что F определено на всем множестве $\mathbb{R}$. Тогда F не может быть инъективным, иначе получим опять взаимнооднозначное отображение $\mathbb{R}$ и A . ■

Определение 5. Пусть Σ — произвольный алфавит символов. Рассмотрим $F : X \rightarrow A$, где $A \subset \Sigma^*$, — некоторое абстрагирование множества X над алфавитом Σ . Будем говорить, что F является абстрагированием с обобщением, если и только если

$$\exists x, y \in X : F(x) = F(y).$$

Определение 6. Пусть Σ — произвольный алфавит символов. Рассмотрим $F : X \rightarrow A$, где $A \subset \Sigma^*$, — некоторое абстрагирование множества X над алфавитом Σ . Будем говорить, что абстракция a , $a \in A$, обобщает сущности $x, y \in X$, если и только если

$$F(x) = F(y) = a.$$

Утверждение 3 показывает, что есть такие наборы сущностей, для которых абстрагирование будет либо определённым частично, либо абстрагированием с обобщением. Возникает естественный вопрос: как часто встречаются такие наборы сущностей?

Предложение 4. Пусть Σ – произвольный алфавит символов, w – некоторое слово в этом алфавите, $F : X \rightarrow A$, где $A \subset \Sigma^*$, – некоторое абстрагирование множества X над алфавитом Σ . Предположим, что слово w содержит все абстракции множества A в качестве подслов. Тогда:

- либо F является абстрагированием с обобщением,
- либо X – конечное множество,
- либо F определено частично на множестве X .

Доказательство. Конечность множества A вытекает из предположения, что слово w содержит все абстракции множества A в качестве подслов, а любое слово конечно. По определению абстрагирования F – сюръективное отображение, а, следовательно, $|X| \geq |A|$.

Если F является инъективным, то F – взаимнооднозначное отображение, и тогда $|X| = |A|$.

Если X не является конечным множеством, то либо F определено частично, в этом случае оно может быть инъективным на своей области определения, либо будет абстрагированием с обобщением, то есть не будет инъективным.

Если F определено на всем множестве X , то либо F инъективно, а значит F взаимнооднозначно и $|X| = |A|$, либо F – абстрагирование с обобщением. ■

Практический смысл данного утверждения заключается в том, что программисту приходится постоянно иметь дело либо с обобщением, либо соглашаться с тем, что его программа определена только для неполного набора сущностей. На конечность множества сущностей рассчитывать не приходится, так как разработка программы – процесс, развивающийся во времени. А с течением времени часто наше понимание об исходных сущностях либо меняется само по себе, либо под воздействием внешних факторов. Например, множество способов оплаты в магазине постоянно меняется, потому что магазины придумывают разные способы привлечения клиента: бонусы, подарочные сертификаты и т. д., появляются новые технологии, например, NFC. Другими словами, имея на данный момент конечный набор сущностей, нельзя быть уверенным, что он не поменяется в будущем.

Предложение 5. Пусть Σ – произвольный алфавит символов, $F : X \rightarrow A$, где $A \subset \Sigma^*$, – некоторое абстрагирование с обобщением множества X над алфавитом Σ , причём $|A| > 1$. Предположим, что Y – множество такое, что $|Y| > 1$, $id : Y \rightarrow AY$ – взаимнооднозначное абстрагирование множества Y над алфавитом Σ . Тогда существует такое отображение $G : X \rightarrow Y$, для которого не существует $g : A \rightarrow AY$, чтобы следующая диаграмма была коммутативной:

$$\begin{array}{ccc} X & \xrightarrow{F} & A \\ G \downarrow & & \downarrow g \\ Y & \xrightarrow{id} & AY \end{array}$$

Доказательство. Проведём построением примера. Поскольку абстрагирование с обобщением, причём $|A| > 1$, то существуют такие различные

$$x_1, x_2, x_3 \in X : F(x_1) = F(x_2) \neq F(x_3).$$

Поскольку $|Y| > 1$, то найдутся два различных $y_1, y_2 \in Y$. Определим действие $G : X \rightarrow Y$ на x_1 как $G(x_1) = y_1$, а $G(x_2) = y_2$. По условию $id(y_1) \neq id(y_2)$. Поскольку $F(x_1) = F(x_2)$, то $\forall g : A \rightarrow AY : g(F(x_1)) = g(F(x_2))$, но, по построению функции G , имеем $id(G(x_1)) \neq id(G(x_2))$. ■

Практическое значение этого утверждения заключается в том, что какое бы множество абстракций размерности больше одного программист не выбрал, найдётся функция, отличная от константной, которая не может быть реализована с помощью этого набора абстракций. То есть не существует «универсальных» абстракций, подходящих на все случаи жизни. Есть конструкции языков программирования, например, if-else, оператор множественного выбора switch требуют явного перечисления всех вариантов, а значит поощряют создание множеств абстракций размерности не менее, чем два. То есть реализация новых бизнес-требований может потребовать правок программного кода (соответствующих абстракций), а это, в свою очередь, как показано в [11], может приводить к снижению скорости разработки программного продукта.

5. Некоторые свойства и ограничения абстрагирований

Широко известный факт из университетского курса математической логики – образ любого отображения задаёт отношение эквивалентности на области определения. Следовательно,

Замечание 1. Пусть Σ – некоторый алфавит символов, $F : X \rightarrow A$, где $A \subset \Sigma^*$, – произвольное абстрагирование множества X над алфавитом Σ . Множество абстракций A определяет отношение эквивалентности на множестве сущностей X .

Более того, в условиях предыдущего замечания,

Замечание 2. Если множество X является носителем какой-либо алгебраической системы, то абстрагирование F может оказаться конгруэнцией.

Замечание 3. Если множества X и A являются носителями каких-либо алгебраических систем, то абстрагирование F может оказаться морфизмом.

Предложение 6. Пусть Σ – произвольный алфавит символов. Множество всех возможных абстрагирований над алфавитом Σ несчётно.

Доказательство. Согласно утверждению 1 множество Σ^* счётно. Рассмотрим все различные множества $X_\alpha = \{\alpha\}$, где $\alpha \in \mathbb{R}$, состоящие из одного действительного числа. Таких множеств несчётно. Тогда и абстрагирований вида $F_\alpha : X_\alpha \rightarrow A$, где $A = \{a\}, a \in \Sigma^*$, — множество абстракций A состоит из одного слова. Множество все абстрагирований F_α , где $\alpha \in \mathbb{R}$, будет несчётным. ■

Предложение 7. Пусть Σ — произвольный алфавит символов. Задача построения произвольного абстрагирования алгоритмически не разрешима.

Доказательство. Следует из факта существования универсальной машины Тьюринга и утверждения 6. ■

6. Интерпретация абстракций

Сами по себе слова формального языка — это лишь конечные цепочки символов. Тем не менее, они несут смысловую нагрузку, которую можно определить с помощью понятия интерпретации.

Определение 7. Пусть Σ — некоторый алфавит, L — формальный язык над алфавитом Σ . Тогда отображение $I : \check{L} \rightarrow Y$, где Y — некоторое множество, будем называть интерпретацией языка L на множестве Y .

Замечание 4. Скорее всего, интерпретация будет определена частично на множестве $\check{L}$.

Например, для пустого оператора языка C++

;

который значит «ничего не делать», может не существовать интерпретации, или, для выражения

5+5+

Замечание 5. Для определения интерпретации можно использовать отношение иерархии Hr (см. определение 3) на множестве всех подслов $\check{L}$, заданного с помощью исчисления $Gr(L)$.

Пример 6. Рассмотрим формальный язык Fs , который описывает множество арифметических формул над кольцом целых чисел $\mathbb{Z}$. Определим формальную грамматику для данного языка. Множество терминальных символов

$$T = \{0123456789 + -()\}.$$

Начальный нетерминальный символ — это «Formula». Грамматика будет учитывать приоритет операций.

$Formula ::= Formula + Add \mid Formula - Add \mid Add$

$Add ::= -Number \mid Number \mid (Formula)$

$Number ::= Number\ 0 \mid Number\ 1 \mid Number\ 2 \mid Number\ 3 \mid Number\ 4 \mid$
 $\mid Number\ 5 \mid Number\ 6 \mid Number\ 7 \mid Number\ 8 \mid Number\ 9$

$Number ::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

Множество нетерминальных символов составляет

$$N = \{Formula\ Add\ Number\}.$$

Определим теперь интерпретацию I для языка Fs в кольцо целых чисел $\mathbb{Z}$. Для того чтобы отличать записи слов языка Fs и элементы в $\mathbb{Z}$, будем к числам из $\mathbb{Z}$ добавлять индекс z .

Рассмотрим произвольное слово $w \in Fs$ и восходящий вывод этого слова с помощью определённой выше грамматики:

$$w_1, w_2, \dots, w_n,$$

где $w_1 = w$, $w_n = Formula$, $n \in \mathbb{N}$.

Определим теперь серию функций

$$V_i(\mathbb{N}, N) \rightarrow \mathbb{Z},$$

где $i \in (1, 2, \dots, n)$, индукцией по номеру i .

Пусть $i = 1$. Тогда функция $V_1(\mathbb{N}, N) \equiv 0$.

Предположим, что все функции V_i с номерами $i < M$, $M > 1$, $M \in \mathbb{N}$, уже определены. Построим определение функции V_M .

Если $M > n$, тогда $V_1(\mathbb{N}, N) \equiv 0$. Иначе, рассмотрим слово w_M . Запишем слово w_M как uRv , где R – нетерминальный символ, который получился в результате применения одного из правил. Обозначим как r позицию нетерминального символа R в слове w_M , а количество символов в правой части правила, в результате которого получился R , как c . Рассмотрим произвольный нетерминальный символ A , входящий в запись слова w_M . Пусть p – его позиция в записи этого слова.

Если $p < r$, то определим $V_M(p, A) = V_{M-1}(p, A)$.

Если $p > r$, тогда $V_M(p, A) = V_{M-1}(p - c + 1, A)$.

Смысл последних двух выражений – значение функции не меняется для тех нетерминальных символов, которые не были затронуты применением последнего правила. Рассмотрим случай, когда $p = r$. То есть случай, когда $A = R$. Для того чтобы определить значение $V_M(R)$, необходимо рассмотреть все варианты правил.

- Правило $Number ::= 0$, тогда

$$V_M(r, Number) = 0.$$

Здесь 0 – это целое число. Аналогично поступаем в случае правил

$$Number ::= 1, \dots, Number ::= 9.$$

- Правило $Number ::= Number\ 0$, тогда

$$V_M(r, number) = V_{M-1}(r, Number) * 10 + 0.$$

Аналогично поступаем в случае правил

$$Number ::= Number\ 1, \dots, Number ::= Number\ 9.$$

- Правило $Add ::= Number$, тогда

$$V_M(r, Add) = V_{M-1}(r, Number) \bmod 2^{32}.$$

Сравнение по модулю 2^{32} связано с тем, что в вычислительных машинах имеется ограниченное число разрядов для представления данных. В данном примере предполагаем, что строим интерпретацию для 32-разрядной машины.

- Правило $Add ::= -Number$, тогда

$$V_M(r, Add) = V_{M-1}(r + 1, Number) \bmod 2^{32}.$$

- Правило $Add ::= (Formula)$, тогда

$$V_M(r, Add) = V_{M-1}(r + 1, Formula).$$

- Правило $Formula ::= Formula + Add$, тогда

$$V_M(r, Formula) = V_{M-1}(r, Formula) + V_{M-1}(r + 2, Add) \bmod 2^{32}.$$

- Правило $Formula ::= Formula - Add$, тогда

$$V_M(r, Formula) = V_{M-1}(r, Formula) - V_{M-1}(r + 2, Add) \bmod 2^{32}.$$

Во всех остальных случаях отображение V_M равно 0.

Тогда интерпретацией слова w будем считать значение $V_n(1, Formula)$.

Одно и тоже слово может иметь несколько выводов, а значит и интерпретация такого слова может быть различной в зависимости от того, какой вывод был использован для её вычисления. Здесь возможно несколько вариантов.

- Доказать, что значение интерпретации не зависит от выбранного вывода. Пример 6 — как раз такой случай. В статье это не сделано, так как пример нужен только, чтобы показать, как интерпретация строится, и обозначить проблему разных выводов.
- Указать какой-либо способ выбора применяемых правил, чтобы ограничиться только одним возможным выводом. Например, левосторонний вывод в $LL(1)$ -грамматиках.
- Исключить из языка слова, которые допускают неоднозначность интерпретации. Например, следующая запись считается ошибочной в C++:
`f(++i, ++i);`

Заметим, что в примере 6 абстрагирование $IntD : \mathbb{Z} \rightarrow DR$ из примера 2 продолжается до гомоморфизма аддитивной группы $\mathbb{Z}$ и аддитивной группы $\mathbb{Z}_{2^{32}} : I \circ IntD : \mathbb{Z} \rightarrow \mathbb{Z}_{2^{32}}$. Этот факт в программировании известен как переполнение при арифметических операциях над целыми числами. Если же вместо абстрагирования $IntD : \mathbb{Z} \rightarrow DR$ взять построенное по аналогии $IntD_{32} : \mathbb{Z}_{2^{32}} \rightarrow DR$, то его можно продолжить до автоморфизма аддитивной группы $\mathbb{Z}_{2^{32}} : I \circ IntD_{32} : \mathbb{Z}_{2^{32}} \rightarrow \mathbb{Z}_{2^{32}}$.

Этот факт позволяет использовать следующие эквивалентности

$$4294967295 \equiv -1(mod\ 4294967296)$$

$$4294967294 \equiv -2(mod\ 4294967296)$$

...

$$2147483648 \equiv -2147483648(mod\ 4294967296)$$

для представления отрицательных чисел в виде дополнительного кода и сэконо-
мить на транзисторах для операции вычитания при разработке процессора —
вместо неё используется сумматор.

Если попробовать построить по аналогии абстрагирование для веществен-
ных чисел на множество чисел с плавающей точкой, то такое абстрагирова-
ние не будет морфизмом вообще. То есть если взять два вещественных числа,
например, 5.2 и 3, и перемножить, то получим 15.6, тогда как для чисел с
плавающей точкой равенство $5.2 * 3 = 15.6$ окажется неверным. Это связано
с тем, что любой интервал вещественных чисел содержит несчётное число
чисел, а чисел с плавающей точкой — только лишь конечное число, поэтому
при абстрагировании вещественных чисел числами с плавающей точкой возни-
кает проблема округления — представление вещественного числа ближайшим
к нему числом с плавающей точкой. Из-за этого вместо операции равенства
используется сравнение модуля разности чисел с некоторым маленьким числом
 ϵ , а в математике возникло целое направление — вычислительные методы, кото-
рое занимается изучением вычислений с плавающей точкой, чтобы их результат
был максимально близок к вычислениям в вещественных числах.

Ещё один момент, который необходимо иметь ввиду при построении абстра-
гирований: интерпретация может меняться, например, из-за изменения архи-
тектуры процессора.

Пример 7. Выражение $65535 + 1 == 0$ будет истинным на 16-разрядной
архитектуре, но ложным на 32-разрядной.

Определение 8. Пусть Σ — некоторый алфавит, L — формальный язык над
алфавитом Σ . Тогда абстрагирование $F : X \rightarrow A$, где $A \subset L$, называется устой-
чивым относительно интерпретаций $I_1 : \check{L} \rightarrow Y$ и $I_2 : \check{L} \rightarrow Y$, где Y — некоторое
множество, если и только если

$$\forall x \in X : I_1 \circ F(x) = I_2 \circ F(x).$$

Замечание 6. Устойчивость абстракции транзитивна.

ЛИТЕРАТУРА

1. Aho A., Ulman J. Foundations of Computer Science: C Edition (Principles of Computer Science Series). URL: <http://i.stanford.edu/~ullman/focs.html> (дата обращения: 01.10.17).
2. Hazzan O. Reflections On Teaching Abstraction and other Soft Ideas ACM SIGCSE Bulletin. 2008. Vol. 40(2). P. 40–43.
3. Буч Г., Максимчук Р., Энгл М., Янг Б., Коннален Дж., Хьюстон К. Объектно-ориентированный анализ и проектирование с примерами приложений на C++. М. : «И.Д. Вильямс», 2008. 720 с.
4. С.А. Орлов. Технологии разработки программного обеспечения: учебник для вузов. СПб. : Питер, 2004. 527 с.
5. Рапаков Г.Г., Ржеуцкая С.Ю. Программирование на языке Pascal. СПб. : БХВ-Петербург, 2004. 480 с.
6. Dahl O., Dijkstra E., Hoare C. Structured Programming. London : Academic Press. 1972. P. 83.
7. Shaw M. Abstraction Techniques in Modern Programming Languages // IEEE Software. 1984. Vol. 1(4). P. 10.
8. Berzins V, Gray M., Naumann D. Abstraction-Based Software Development // Communications of the ACM. 1984. Vol. 29(5). P. 403.
9. Liskov B., Guttag J. Program Development in Java: Abstraction, Specification, and Object-Oriented Design. Pearson Education, 2000. 464 p.
10. Тюменцев Е.А. Формальное определение абстракции. С. 36-38.
11. Тюменцев Е.А. О формализации процесса разработки программного обеспечения // Математические структуры и моделирование. 2017. № 3(43). С. 96–107.
12. The Liskov Substitution Principle. URL: <http://www.objectmentor.com/resources/articles/lsp.pdf> (дата обращения: 01.12.12).

ABOUT FORMAL DEFINITION OF AN ABSTRACTION

Е.А. Tyumentsev

CEO, e-mail: etyumentsev@hwdtech.ru

LLC Hello World! Technologies, Omsk, Russia

Abstract. The article proposes a formal definition of abstraction as a mapping of a set of entities into a set of subwords of all words of some alphabet, a hierarchy of abstractions as relations on the set of subwords of a language and interpretation. Also received a number of limitations, which has an abstraction mechanism.

Keywords: abstraction, abstraction barrier, inheritance, Grady Booch, Barbara Liskov, Robert Martin.

Дата поступления в редакцию: 31.01.2018

УТОЧНЕНИЕ К СТАТЬЕ «О ФОРМАЛИЗАЦИИ ПРОЦЕССА РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ»

Е.А. Тюменцев

генеральный директор, e-mail: etyumentcev@hwdtech.ru

LLC Hello World! Technologies, Омск, Россия

Аннотация. В статье «О формализации процесса разработки программного обеспечения» даны определения процесса разработки как процесса редактирования текста программы, скорости данного процесса и сформулировано достаточное условие линейной скорости. В определении скорости процесса разработки необходимо существование предела. Но требуемый предел может и не существовать. В этой статье приводится определение скорости процесса разработки, не требующее существования предела, и переформулируется достаточное условие линейной скорости в соответствии с новым определением.

Ключевые слова: формализация, процесс разработки, программное обеспечение, производительность труда, Брукс.

В номере 43 журнала «Математические структуры и моделирование» за 2017 год была опубликована статья [1], в которой процесс разработки представлен как последовательность Pr :

$$P_0 \xrightarrow{c_1} P_1 \xrightarrow{c_2} P_2 \xrightarrow{c_3} \dots$$

в терминах и обозначениях статьи [1], где c_i – одна из операций по редактированию текста программы:

- добавление нового слова,
- добавление одного символа в существующее слово,
- удаление одного символа из существующего слова,
- удаление слова.

При этом скорость процесса разработки определяется как

Определение 8. Пусть $F : \mathbb{N} \rightarrow \mathbb{R}$ – некоторая функция из множества $\mathbb{N}$ в $\mathbb{R}$. Говорят, что процесс разработки Pr обладает скоростью F тогда и только тогда, когда

$$\lim_{x \rightarrow \infty} \frac{|P_n|}{F(n)} = C,$$

где $C > 0$.

Это определение имеет ограничение по применению, так как предел

$$\lim_{n \rightarrow \infty} \frac{|P_n|}{F(n)}$$

может и не существовать. Для определения скорости разработки важен не сам предел, а то, что последовательность $\frac{|P_i|}{F(i)}$, где $i \in (1, 2, \dots)$, можно ограничить снизу и сверху некоторыми константами. Если предел существует, то из этого следует ограниченность, но обратное утверждение не всегда выполняется. В силу этого замечания Предложение 7 верно не для всех процессов разработки (здесь и далее все предложения – это предложения из статьи [1]).

Кроме того, если следовать понятию скорости в общепринятом смысле, то скоростью будет не сама F , а $\frac{F}{n}$, так как n в нашей модели играет роль времени.

На основании данных соображений дадим новые определения.

Определение 1. Пусть $F : \mathbb{N} \rightarrow \mathbb{R}^+$ – некоторая функция из множества $\mathbb{N}$ в $\mathbb{R}^+$, где $\mathbb{R}^+ = \{x \in \mathbb{R}, x \geq 0\}$. Говорят, что процесс разработки Pr обладает скоростью не лучше, чем F с точностью до константы тогда и только тогда, когда

$$\exists k \in \mathbb{N} \forall n > k |P_n| \leq F(n) \cdot n.$$

Замечание 1. В силу предложения 4 скорость любого процесса разработки не может быть лучше, чем 1.

Определение 2. Пусть $F : \mathbb{N} \rightarrow \mathbb{R}^+$ – некоторая функция из множества $\mathbb{N}$ в $\mathbb{R}^+$. Говорят, что процесс разработки Pr обладает скоростью не хуже, чем F с точностью до константы тогда и только тогда, когда

$$\exists k \in \mathbb{N} \forall n > k F(n) \cdot n \leq |P_n|.$$

Определение 3. Пусть $F : \mathbb{N} \rightarrow \mathbb{R}^+$ – некоторая функция из $\mathbb{N}$ в $\mathbb{R}^+$. Говорят, что процесс разработки Pr обладает скоростью F с точностью до константы тогда и только тогда, когда

$$\exists C_1 > 0, C_2 > 0 \exists k \in \mathbb{N} \forall n > k C_1 \cdot F(n) \cdot n \leq |P_n| \leq C_2 \cdot F(n) \cdot n.$$

Замечание 2. Пусть $F : \mathbb{N} \rightarrow \mathbb{R}^+$ – некоторая функция из $\mathbb{N}$ в $\mathbb{R}^+$, Pr – некоторый процесс разработки. Предположим, что $\lim_{n \rightarrow \infty} \frac{|P_n|}{F(n) \cdot n} = C$, где $C \geq 0$. Тогда процесс разработки обладает скоростью F с точностью до константы.

Доказательство. Из существования предела $\lim_{n \rightarrow \infty} \frac{|P_n|}{F(n) \cdot n} = C$, где $C \geq 0$, следует ограниченность последовательности: $\frac{|P_n|}{F(n) \cdot n}$, начиная с некоторого $k \in \mathbb{N}$. То есть

$$\forall C_1 > 0 \exists k \in \mathbb{N} \forall n > k C - C_1 \leq \frac{|P_n|}{F(n) \cdot n} \leq C + C_1.$$



Основной результат статьи [1] — достаточное условие линейной скорости разработки теперь имеет следующую формулировку.

Теорема 1 (Достаточное условие константной скорости процесса разработки). *В условиях предложения 6 предположим, что*

$$\exists C_1 > 0 \exists k_1 \in \mathbb{N} \forall n \geq k_1 |D_n| \leq C_1 \cdot |P_n|.$$

Тогда процесс разработки Pr обладает константной скоростью.

Доказательство. В силу предложения 6

$$\frac{n}{c} \leq |P_n \cup D_n| \leq n.$$

По предложению 1

$$|P_n \cup D_n| = |P_n| + |D_n|,$$

следовательно, по предположению теоремы

$$|P_n| + |D_n| \leq |P_n| + C_1 \cdot |P_n| = (1 + C_1)|P_n|.$$

Получаем, что

$$\frac{n}{c} \leq (1 + C_1)|P_n|,$$

то есть

$$\frac{n}{C(1 + C_1)} \leq |P_n|.$$

С другой стороны, $|P_n| \leq n$. Получаем

$$\frac{n}{C(1 + C_1)} \leq |P_n| \leq n.$$

В силу определения 3 скорость F процесса разработки Pr равна $F(n) = 1$, что и требовалось доказать. ■

Замечание 3. Вообще говоря, константная скорость в нашем случае не значит, что скорость процесса разработки не меняется с течением времени. Важно то, что она не может упасть ниже определённой константы по мере роста размера проекта. Этим свойством обладают не все процессы разработки. В [1] был построен пример процесса разработки, скорость которого асимптотически приближалась к 0 по мере роста размера проекта.

ЛИТЕРАТУРА

1. Тюменцев Е.А. О формализации процесса разработки программного обеспечения // Математические структуры и моделирование. 2017. № 3(43). С. 96–107.

CLARIFICATION OF THE ARTICLE "ABOUT THE FORMALIZATION OF THE SOFTWARE DEVELOPMENT PROCESS"

E.A. Tyumentcev

CEO, e-mail: etyumentcev@hwdtech.ru

LLC Hello World! Technologies, Omsk, Russia

Abstract. In the article "About The Formalization Of The Software Development Process", the definitions of the software development process as a process of editing the text of the program, the speed of the process, and a sufficient condition for linear speed are formulated. In determining the speed of the development process, the existence of a limit is required. However, the required limit may not exist. In this paper, we define the speed of software development process, which does not require the existence of a limit, and reformulates a sufficient condition for linear velocity in accordance with the new definition.

Keywords: formalization, development process, software, productivity.

Дата поступления в редакцию: 05.02.2018

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО НАБОРА СРЕДСТВ ЗАЩИТЫ КОМПЬЮТЕРНОЙ СИСТЕМЫ МЕТОДОМ МОНТЕ-КАРЛО

Т.В. Вахний

к.ф.-м.н., доцент, e-mail: vahniytv@mail.ru

А.К. Гуц

д.ф.-м.н., профессор, e-mail: guts@omsu.ru

И.Ю. Пахотин

студент, e-mail: prakt1k@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. В статье представлено программное приложение, позволяющее на основе теории игр находить наиболее оптимальный набор средств защиты компьютерной системы методом Монте-Карло.

Ключевые слова: защита информации, теория игр, хакерские атаки, оптимальная стратегия, программный продукт, метод Монте-Карло.

Введение

Использование теоретико-игрового подхода позволяет обеспечить оптимизацию выбора программных продуктов для защиты компьютерной системы [1–6]. Однако количество возможных стратегий хакера и возможных стратегий администратора безопасности экспоненциально увеличиваются с ростом возможных вариантов атак и средств защиты соответственно. Это приводит к тому, что увеличивается размер платёжной матрицы игры между администратором безопасности и хакером. Кроме того, если решать позиционную игру для всех возможных стратегий хакера и администратора безопасности, то это потребует больших затрат на вычислительные ресурсы, и к тому же на вычисление оптимальной стратегии администратора безопасности потребуется продолжительное время. Поэтому актуально нахождение решения указанной игры методами, которые менее затратны на вычислительные ресурсы.

В данной работе для нахождения наиболее оптимальных вариантов защиты компьютерной системы предлагается построить двухходовую позиционную игру администратора безопасности со злоумышленником и решить её методом имитационного моделирования. На основе описанного подхода было создано программное приложение, которое позволяет рассчитать оптимальный набор средств защиты компьютерной информации методом Монте-Карло.

1. Постановка задачи и игровой подход

Для поиска наиболее оптимального набора средств защиты компьютерной системы можно провести математическую игру двух сторон, одной из которых является система защиты компьютерной информации (I игрок – администратор), а с другой – возможные атаки хакеров (II игрок – злоумышленник). Один из подходов, моделирующий игру хакера и администратора безопасности, основан на теории позиционных игр [2].

Позиционная игра — это последовательная многоходовая игра, которая состоит в последовательном переходе из одного состояния (позиции) в другое состояние (позицию) путём выбора игроками одного из возможных действий в соответствии с правилами игры. Позиционные игры описываются с помощью дерева игры. На рис. 1 представлено дерево двухходовой позиционной игры. Каждая вершина дерева — это позиции игры, ход из одной позиции игры в другую позицию задаётся дугой. Начальной позиции игры соответствует корень дерева R , конечным позициям — листья, из них ходы игроками уже не совершаются. Каждому листу c_{ij} ($i = 1, \dots, M$; $j = 1, \dots, K$) приписывается выигрыш.

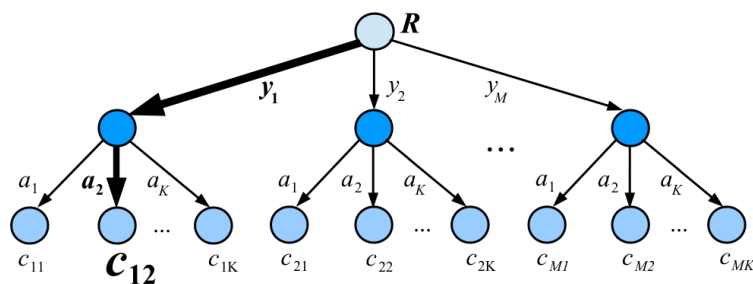


Рис. 1. Дерево двухходовой игры хакера и администратора безопасности

Пусть злоумышленник (игрок II) обладает K стратегиями атаки на компьютерную систему $a_1, a_2, \dots, a_K$, каждая такая стратегия a_j ($j = 1, \dots, K$) — это j -я угроза, способная нарушить работу компьютерной системы. Каждому способу атаки a_j ($j = 1, \dots, K$) соответствует ущерб l_j , который терпит администратор в случае успешного осуществления этой j -ой атаки. Администратор безопасности (игрок I) пусть располагает S средствами защиты $d_1, d_2, \dots, d_S$, каждое из которых обладает соответствующей ценой $g_1, g_2, \dots, g_S$ и эффективностью защиты $e_i = (e_{i1}, e_{i2}, \dots, e_{iK})$ ($i = 1, \dots, S$), где e_{ij} — эффективность i -го средства защиты по нейтрализации j -ой угрозы.

Стратегии администратора — это кортежи

$$y_i = (z_{i1}, z_{i2}, \dots, z_{iS}),$$

где $z_{ij} = 1$, если задействован способ защиты d_j и $z_{ij} = 0$ — в противном случае.

Имея S средств защиты d_k ($k = 1, \dots, S$), администратор может составить $M = 2^S - 1$ стратегий y_i ($i = 1, \dots, M$), из которых мы исключили кортеж $(0, 0, \dots, 0)$, поскольку он представляет полное бездействие администратора.

Ходом администратора является выборка из способов обеспечения защиты информационной системы (разных наборов средств защиты), т. е. выбор стратегии y_i , а ходом злоумышленника — применение способа атаки a_j на компьютерную систему.

Составим платёжную матрицу C , элементами c_{ij} которой (табл. 1) являются потери администратора в том случае, когда он использовал стратегию y_i ($i = 1, \dots, M$), а хакер осуществил атаку a_j ($j = 1, \dots, K$). На дереве игры элементы платёжной матрицы представлены листьями (см. рис. 1).

Элементы платёжной матрицы вычисляются с помощью формулы:

$$c_{ij} = R(y_i, a_j) + G_i,$$

где

$G_i = \sum_{j=1}^S z_{ij} g_j$, — затраты администратора на использование средств защиты i -той стратегии y_i ;

$R(y_i, a_j) = l_j \sum_{l=1}^S (1 - z_{il} e_{lj})$ — остаточный риск атаки a_j при использовании стратегии y_i , т. е. возможный ущерб помноженный на вероятность того, что используемые средства защиты окажутся неэффективны.

Путь от корня R к листу c_{ij} является *партией игры* с исходом игры $F(y_i, a_j) = c_{ij}$. На рис. 1 изображён пример партии, в которой администратор избрал стратегию y_1 , а злоумышленник — атаку a_2 .

Процесс игры состоит в том, что администратор выбирает стратегию защиты y_i , злоумышленник выбирает способ атаки a_j , после чего вычисляется исход партии, заключающийся в том, что администратор терпит ущерб равный c_{ij} .

Нанесение хакером ущерба обычно является скорее следствием его действий, а не самой целью. В действительности при атаке он может преследовать какие-то свои цели, порой известные лишь ему.

Цель администратора — выбор такой стратегии, которая сводит потери от атак к минимуму, авторы на его стороне, и поэтому цели атакующих хакеров в расчёт не принимаются. В силу этого можно считать, что хакер увлечён желанием нанести как можно больший ущерб атакуемой компьютерной системе. При таком предположении выигрыш хакера будет равен проигрышу администратора безопасности c_{ij} .

Таблица 1. Платёжная матрица игры

	a_1	a_2	...	a_K
y_1	c_{11}	c_{12}	...	c_{1K}
y_2	c_{21}	c_{22}	...	c_{2K}
...	...	...	...	...
y_M	c_{M1}	c_{M2}	...	c_{MK}

В качестве стратегий администратора безопасности будем понимать строки y_i ($i = 1, \dots, M$) платёжной матрицы, а в качестве стратегий хакера — её столбцы

a_j ($j = 1, \dots, K$) (табл. 1). Перед своим ходом игроки не знают о ходах друг друга. Для проведения на компьютере игры C надо также знать результаты игры при каждой паре стратегий y_i и a_j .

Администратор безопасности стремится выбрать такую стратегию, которая позволит ему свести к минимуму наносимый компьютерной системе ущерб от реализации тех или иных угроз. Поставим в соответствие каждой i -ой стратегии администратора число $W_i(C)$, вычисляемое с помощью платёжной матрицы C . Критерий выбора оптимальной стратегии для администратора состоит в том, чтобы взять $W = \min_i W_i(C)$. Тогда оптимальной будет такая стратегия y_{i_0} , для которой $W_{i_0}(C) = W$.

Будем считать, что администратор не имеет никакой информации о том, какую стратегию (способы атаки) выберет злоумышленник, тогда вероятности атак можно считать одинаковыми и равными $1/K$. В таком случае можно воспользоваться критерием недостаточного основания Лапласа [5, 7], для которого берём:

$$W_i(C) = \frac{1}{K} \sum_{j=1}^K c_{ij}. \quad (1)$$

Использование критерия недостаточного основания Лапласа оправдано, если минимизация риска проигрыша представляется менее существенным фактором принятия решения, чем максимизация среднего выигрыша. Администратор при необходимости может применить и другие критерии [5, 7] к подбору оптимального набора средств защиты.

2. Нахождение оптимальной стратегии администратора безопасности методом Монте-Карло

С увеличением числа S средств защиты растёт и количество $M = 2^S - 1$ вариантов возможных стратегий защиты компьютерной системы. В результате общее количество возможных партий игры администратора безопасности со злоумышленником можно вычислить по формуле:

$$A = KM = K(2^S - 1). \quad (2)$$

Например, если возьмём число возможных атак хакера $K = 10000$, число средств защиты администратора $S = 20$, то число возможных стратегий администратора $M = 1048575$ и количество возможных партий игры будет равно $A = 10485750000$.

Для нахождения оптимальной стратегии администратора нужно вычислить по формуле (1) все $W_i(C)$ для $i = 1, \dots, M$. Для этого потребуется провести минимум A сложений, а в приведённом выше примере это более 10 миллиардов сложений. Данный способ расчёта требует больших затрат на вычислительные ресурсы и уже при данных здесь значениях K и S оптимальная стратегия администратора может вычисляться продолжительное время. Если ещё в несколько

раз увеличить значения K и S , то невозможно будет данным способом найти W за разумное время даже с использованием вычислительных машин.

Исходя из этого, актуально нахождение решения рассматриваемой игры методами, которые менее затратны на вычислительные ресурсы. При исследовании сложных систем, подверженных случайным воздействиям, можно использовать имитационное моделирование, которое представляет собой численный метод проведения на ЭВМ вычислительных экспериментов с математической моделью. Результаты поведения рассматриваемой системы, полученные при воспроизведении на имитационной модели, являются случайными реализациями. Для нахождения объективного и устойчивого решения требуется многократное нахождение различных решений с последующей статистической обработкой полученных данных. В настоящее время при имитационном моделировании случайных процессов очень широко используется метод статистических испытаний Монте-Карло [8].

Для приблизительного вычисления значений $W_i(C)$, определяемых в (1), можно для каждой стратегии администратора y_i ($i = 1, \dots, M$) провести по N_{y_i} симуляций (с выбранными случайным образом атаками), после чего для полученных результатов расчёта убытков $(x_{y_i1}, x_{y_i2}, \dots, x_{y_in}, \dots, x_{y_iN_{y_i}})$ рассчитать их средние значения для каждой стратегии администратора:

$$\overline{x_{y_i}} = \frac{1}{N_{y_i}} \sum_{j=1}^{N_{y_i}} x_{y_ij}. \quad (3)$$

С ростом значения N_{y_i} математическое ожидание $\overline{x_{y_i}}$ будет стремиться к $W_i(C)$:

$$M\overline{x_{y_i}} \rightarrow W_i(C) \text{ при } N_{y_i} \rightarrow \infty.$$

Критерием выбора оптимальной стратегии администратора может служить выбор стратегии y_{i_0} , для которой

$$\overline{x_{y_{i_0}}} = \min_i \overline{x_{y_i}},$$

т. е. стратегии, для которой среднее значение результатов симуляций убытков (ущерба от атак и затрат на средства защиты) минимально.

Пусть $N = \sum_{i=1}^M N_{y_i}$ – общее количество симуляций. Тогда общее количество операций сложения, которые нужно выполнить для расчёта $\overline{x_{y_i}}$ равно N_{y_i} , а общее количество операций сложения для приблизительного нахождения W равно N . Если $N < A$, т. е. меньше количества возможных исходов игры, то мы получаем выигрыш по времени.

3. Оптимизация симуляций

Если количество симуляций для всех стратегий администратора y_i установить одинаковым, т. е. все значения N_{y_i} ($i = 1, \dots, M$) равными, то данный способ будет неоптимальным.

Например, установив общее количество симуляций N , на шаге n гораздо лучше динамически принимать решение, для какой стратегии проводить очередную симуляцию. Нужно каким-то образом отдавать предпочтения тем стратегиям y_i , для которых значение $\overline{x}_{y_i}$ на шаге n меньше. В таком случае минимальные значения $\overline{x}_{y_i}$ будут быстрее сходиться к своему математическому ожиданию, и мы будем экономить на вычислениях. Мы укажем ниже более эффективный алгоритм выбора стратегий y_i .

4. Реализуемый алгоритм USB1

Для минимизации потерь администратора применим алгоритм UCB1 (Upper-Confidence-Bound) [9] к нашей задаче:

Пусть N – требуемое количество симуляций и $N > M$.

1. Шаг $n = 1$. Проведём M симуляций, по одной симуляции для каждой стратегии администратора y_i .

2. Шаг n , пока $n < N$.

На шаге n проводим симуляцию x_{y_i} для стратегии y_i , для которой минимальна величина

$$\overline{x}_{y_i} - b\sqrt{\frac{2 \ln n}{n_{y_i}}}, \quad (4)$$

где $\overline{x}_{y_i}$ – средние убытки компьютерной системы при использовании администратором стратегии y_i (см. (3));

n_{y_i} – количество уже проведённых симуляций для стратегии y_i ;

n – номер текущей симуляции;

b – константа, используемая для установки нужного баланса между шириной и глубиной поиска. Чем она больше, тем чаще будут рассматриваться варианты, для которых средние убытки компьютерной системы не являются минимальными на текущий момент.

На рис. 2 изображены графики радикала $\sqrt{\frac{2 \ln n}{n_{y_i}}}$ при различных значениях n_{y_i} . Алгоритмический смысл данного радикала состоит в том, что без него на шаге n мы бы каждый раз просто выбирали стратегию с минимальным значением $\overline{x}_{y_i}$. Но так как этот радикал растёт у стратегий, для которых мы не проводим симуляций, то время от времени мы будем проводить симуляции у стратегий, не показывающих минимальный результат на данный момент. Это позволяет сгладить ситуации, когда нам не повезло и для «хороших» стратегий были проведены симуляции с «плохим» результатом.

5. Описание программного продукта и анализ результатов расчёта

В данной статье мы представляем программный продукт [10], который по введённым значениям стоимости средств защиты и ущерба от применения атак

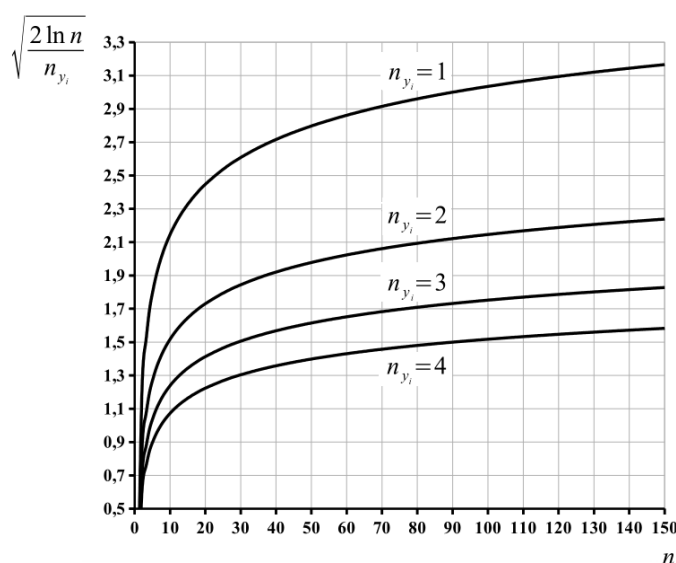


Рис. 2. Зависимость величины радикала от номера текущей симуляции n для различных значений количества проведенных симуляций i -ой стратегии администратора

позволяет рассчитать оптимальный набор средств защиты компьютерной системы методом Монте-Карло с использованием алгоритма UCB1.

Приложение создавалось в среде разработки IntelliJ IDEA с использованием языка программирования Java. В качестве Фреймворка для автоматической сборки был использован Maven, для разработки графического интерфейса на Java использована библиотека Swing.

На рис. 3 показано главное окно программного приложения, предоставляющее доступ к другим окнам.

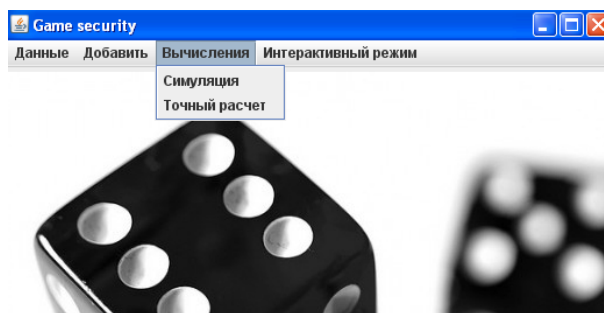


Рис. 3. Интерфейс меню программного приложения

Созданное программное приложение на основе набора введенных средств защиты $(d_1, d_2, \dots, d_5)$ генерирует список из всех возможных стратегий администратора y_i ($i = 1, \dots, M$). Пользователь имеет возможность добавлять новые способы атаки на компьютерную систему и новые средства защиты. В программном приложении можно получить как точное решение задачи, так и решение методом Монте-Карло. Общее количество симуляций является суммой

значения поля «Количество симуляций» (см. рис. 4) и количества стратегий администратора.

После проведения необходимых расчётов программное приложение сортирует стратегии администратора в порядке увеличения средних значений затрат на средства защиты и ущерба от атак.

При выборе вида вычисления «Симуляция» (см. рис. 3) программным приложением производятся расчёты для нахождения значений $\overline{x_{y_i}}$, которые с увеличением количества симуляций приближаются к значениям математического ожидания убытков администратора при выборе соответствующих стратегий. Оптимизация симуляций производилась введением коэффициента смены стратегий (КСС), который является значением b из формулы (4). Этот коэффициент отвечает за то, насколько часто алгоритм будет проводить симуляции для стратегий, у которых на текущий момент убытки не принимают минимального значения [9].

На рис. 4 показаны результаты работы программного приложения при количестве симуляций 50000 и $KCC = 5000$. В таблицу с лучшим результатом выводится три стратегии с наименьшими средними убытками и с наибольшим количеством симуляций (N_{y_i}).

Программное приложение было запущено на случайно сгенерированных данных с абстрактными способами атаки и методами защиты (см. рис. 5). На данных малого размера расчёт оптимальных стратегий администратора по методу Монте-Карло существенно проигрывает точному расчёту. Но интерес представляют ситуации, когда количество методов защиты и способов атаки являются большими числами, причём количество способов атак много больше возможных стратегий администратора. На таких данных алгоритм может иметь большое преимущество перед точным расчётом.

Было произведено моделирование на основе случайных данных, когда количество методов защиты равно 7, а способов атак 10 000. Программное приложение было запущено по несколько раз на количестве симуляций 100, 500, 1000, 5000 и 50 000. При малом числе симуляций полученный методом Монте-Карло результат отличался от точных расчётов. На 5000 симуляциях почти во всех случаях программное приложение точно находило одну из двух лучших стратегий, и она была на первом месте в таблице «Лучший результат». На 50000 симуляциях программное приложение всегда точно определяло две лучшие стратегии в правильном порядке (см. рис. 4 и рис. 6).

На 50000 симуляций потребовалось провести $50000 + 2^7 = 50127$, а не $10000(2^7 - 1) = 1270000$ сложений чисел, что составляет 0,04 от исходного количества. Таким образом, удалось добиться существенного выигрыша по времени вычисления оптимального набора средств защиты компьютерной информации.

В рамках данной работы критерием выбора оптимальной стратегии администратора безопасности был использован критерий недостаточного основания Лапласа, предполагающий, что все атаки равновероятны. В случае, когда атаки неравновероятны и администратору известны их вероятности, можно использовать критерий Байеса [5–7].

Более подробно программный продукт описан в [10].

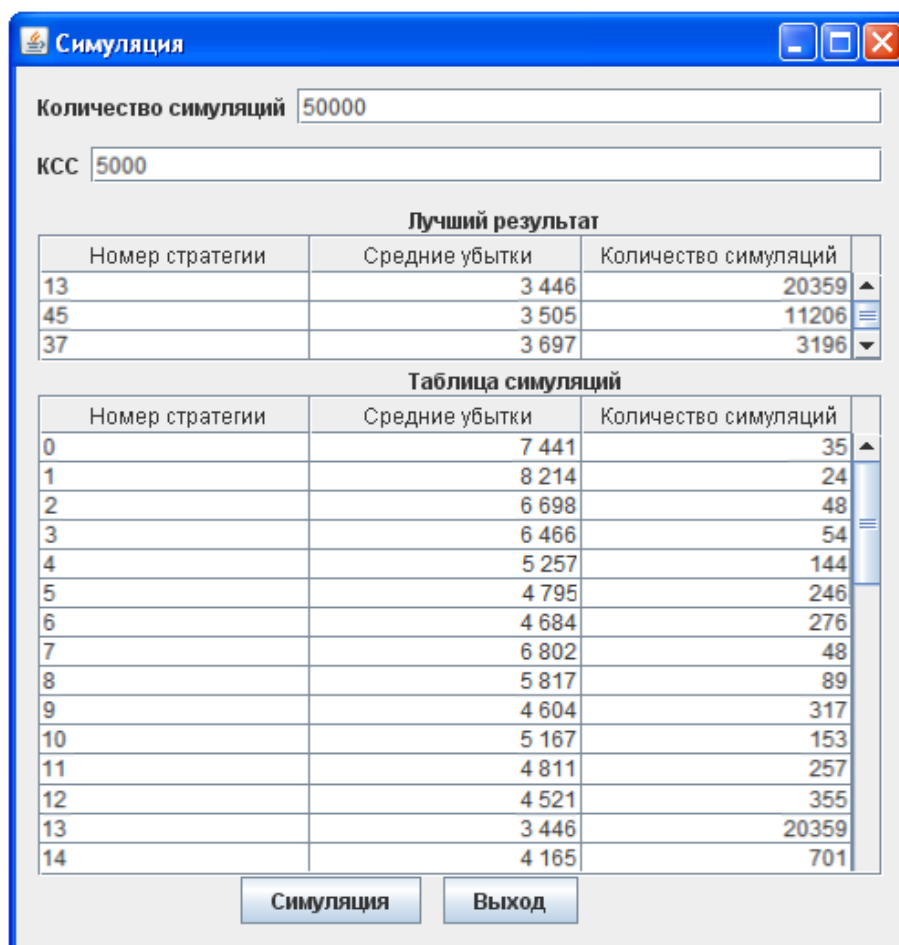


Рис. 4. Результат нахождения оптимальных стратегий администратора методом Монте-Карло при количестве симуляций 50 000

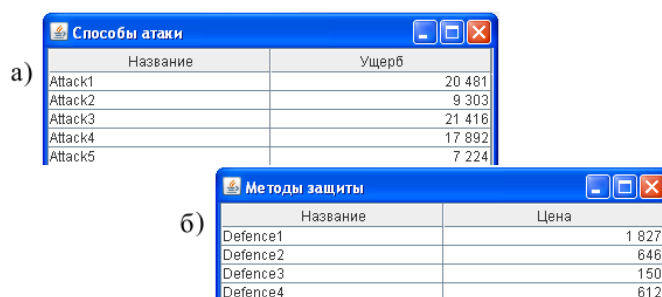


Рис. 5. Добавление в игру абстрактных способов атаки (а) и методов защиты (б)

6. Заключение

Разработанное программное приложение находит оптимальный набор средств защиты компьютерной системы, решая двухходовую позиционную игру администратора безопасности со злоумышленником методом статистических испытаний Монте-Карло. Сравнение полученных результатов с точными расчё-

№	Номер стратегии	Средние убытки
1	13	3 437
2	45	3 459
3	43	3 741
4	37	3 774
5	41	4 259
6	14	4 325
7	11	4 454
8	5	4 521
9	29	4 563
10	12	4 588
11	6	4 623
12	44	4 628
13	38	4 649
14	77	4 662
15	46	4 765

Рис. 6. Результаты точных расчётов убытков (ущерба от атак и затрат на средства защиты) для разных стратегий администратора

тами показало, что при огромном количестве возможных атак хакеров и выборе оптимальной стратегии администратора из большого количества средств защиты данный подход позволяет добиться существенного выигрыша по времени нахождения решения.

ЛИТЕРАТУРА

1. Матричные игры / Под. ред. Н.Н. Воробьёва. М. : ФМ, 1961. 280 с.
2. Гуц А.К., Вахний Т.В. Теория игр и защита компьютерных систем: учебное пособие. Омск : Изд-во ОмГУ, 2013. 160 с.
3. Вахний Т.В., Гуц А.К. Теоретико-игровой подход к выбору оптимальных стратегий защиты информационных ресурсов // Математические структуры и моделирование. 2009. № 19. С. 104–107.
4. Вахний Т.В., Гуц А.К., Константинов В.В. Программное приложение для выбора оптимального набора средств защиты компьютерной информации на основе теории игр // Вестник Омского университета. 2013. № 4(70). С. 201–206.
5. Вахний Т.В., Гуц А.К., Новиков Н.Ю. Матрично-игровая программа с выбором критерия для определения оптимального набора средств защиты компьютерной системы // Математические структуры и моделирование. 2016. № 2(38). С. 103–115.
6. Вахний Т.В., Гуц А.К., Бондарь С.С. Учёт вероятностей хакерских атак в игровом подходе к подбору программных средств защиты компьютерной информации // Математические структуры и моделирование. 2015. № 3(35). С. 91–105.
7. Шевченко Д.В. Методы принятия управленческих решений: задания и методические указания для выполнения расчётно-графической работы. Казань : Познание, 2014. 69 с.
8. Ермаков С.М., Михайлов Г.А. Статистическое моделирование. М. : Физматлит, 1982. 296 с.

9. Finite-time Analysis of the Multiarmed Bandit Problem. URL: <https://homes.di.unimi.it/~cesabian/Pubblicazioni/ml-02.pdf> (дата обращения: 07.01.2017).
10. Пахотин Ю.И. Применение стохастических методов к защите информации: выпускная квалификационная работа. Омск: ОмГУ, 2017. 51 с.

DETERMINING THE OPTIMAL SET OF TOOLS FOR COMPUTER SYSTEM SECURITY BY MONTE-CARLO METHOD

T.V. Vahniy

Ph.D. (Phys.-Math.), Associate Professor, e-mail: vahniytv@mail.ru

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: guts@omsu.ru

I.Yu. Pahotin

Student, e-mail: prakt1k@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. The article presents a software application that allows based on game theory to find the most optimal set of tools for computer information protection by Monte-Carlo method.

Keywords: information security, theory of games, hacker attacks, optimal strategy, software product, Monte-Carlo method.

Дата поступления в редакцию: 10.02.2018

РАЗРАБОТКА СТРУКТУРНОЙ МОДЕЛИ КУРСА ПОДГОТОВКИ К ЕГЭ ПО ИНФОРМАТИКЕ В ПРОФОРИЕНТАЦИОННОЙ ШКОЛЕ ФКН

Д.Н. Лавров

к.т.н., доцент, зав. каф. компьютерных технологий и сетей, e-mail: lavrov@omsu.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Статья посвящена анализу тем единого государственного экзамена по информатике и построению на его основе структурной модели курса для профориентационной школы факультета компьютерных наук. Проведён анализ информационных ресурсов, являющихся источниками задач для подготовки. Структурная модель дополнена темами, обеспечивающими логичность следования и связность тем.

Ключевые слова: структурная модель, информатика, ИКТ, ЕГЭ, профориентация, дополнительное образование.

Введение

Перед профориентационной школой факультета компьютерных наук (ПШФКН) стоит ряд задач:

1. Привлечение абитуриентов на факультет.
2. Знакомство с профессиональной деятельностью.
3. Помощь в подготовке к сдаче вступительных испытаний в форме ЕГЭ.

Все эти задачи неразрывно связаны. В выпускных классах фокус направлен на подготовку к вступительным испытаниям. Эта задача важна как для самих школьников, так и для факультета.

1. Мотивация

К сожалению, в целом уровень подготовки по информатике в школе остаётся невысоким. Хотя и есть отдельные приятные исключения. Уровень подготовки мы видим как по школьникам, обучающимся в профориентационной школе, так и по студентам младших курсов.

Знания по информатике часто носят разрозненный и несистемный характер. Причины заключаются не только в кадровом голоде: высококвалифицированных учителей информатики не хватает, но и в малом количестве часов на данный предмет в школе.

Будущие абитуриенты, как правило, определяются с выбором профессии к выпускному классу и начинают усиленно готовиться к вступительным испытаниям. Для профшколы в 11 классе важнейшей целью является подготовка к ЕГЭ по информатике и ИКТ.

Самый простой и эффективный подход для достижения этой цели — наreshивание задач из тестов прошлых лет. Как правило он сводится к натаскиванию на определённый тип задач, решение по шаблону. Недостатком такого подхода является невосприятие новых типов задач и неспособность решить их самостоятельно. Вырастает поколение «репетиторских» детей. Такие дети не способны к самообучению, слабо самоорганизованны, испытывают сильную потребность в наставнике (репетиторе). На взгляд автора данной статьи, это тупиковый путь. И хотя этот путь, как правило, позволяет решить проблему поступления в вуз, но в вузовской системе с её системой сессионных экзаменов и более слабым контролем такие студенты часто доучиваются только до первой сессии.

Ещё одним, напрямую связанным с предыдущей проблемой, негативным аспектом является наличие разрозненности знаний, их несистемность: «знаем и запоминаем только то, что нужно для решения конкретной задачи», а ещё хуже: «запоминаем лишь шаблон решения», не особо вдаваясь, откуда он взялся и как был получен. Выражается это сигнальной фразой школьника, абитуриента, студента: «Мы таких задач раньше не решали!» Обучающийся должен решать любую задачу нового типа, даже если видит её первый раз!

И в этом ему может помочь системность полученных знаний.

Что такое системность знаний и из чего она складывается? На взгляд автора данной работы складывается системность из:

- твёрдого знания базовых понятий;
- понимания связей между базовыми понятиями;
- понимания связей между различными темами.

То есть системность знаний — это в первую очередь их связность, встроенность в уже существующую систему понятий. В таком виде знания предстают перед нами как сильносвязный граф. И встаёт проблема размотать этот «клубок» знаний, превратив в линейный список тем занятий, логически связанных и вытекающих одна из другой.

Обычно будущие абитуриенты приходят в ПШФКН в начале одиннадцатого класса. Необходимо за один год систематизировать их знания в области информатики и подготовить к вступительным испытаниям, научить самостоятельно думать и быть готовым к любым новым видам задач.

2. Ресурсы для подготовки

Наиболее популярными ресурсами (источниками задач) являются:

1. Банк открытых заданий ФИПИ [1]. По-видимому, самая большая коллекция тестовых задач. Состоит из трёх разделов:

- Информация и информационные процессы.
- Информационная деятельность человека (задания отсутствуют).
- Средства ИКТ.

Теоретического материала, примеров решений на сайте не имеется.

2. Сборники тестовых материалов в формате ЕГЭ [3–5]. Имеются ответы и примеры решения задач. Сами задачи слабо систематизированы, но располагаются так, как обычно располагаются в тесте. В некоторых сборниках имеется теоретический материал.
3. Сайт Константина Юрьевича Полякова [2] — ресурс наиболее востребованный школьными учителями, репетиторами, преподавателями профшкол и курсов подготовки. Задачи систематизированы, имеются ответы, есть теоретический материал.

Константин Юрьевич каждый год проделывает огромную работу по систематизации задач, обновлению теоретического материала и примеров решения задач. Задачи на его сайте распределены по следующим разделам (в скобках указаны номера задачи теста ЕГЭ):

1. Информация
 - 1.1. Вычисление количества информации (13, 9)
 - 1.2. Комбинаторика (10)
 - 1.3. Кодирование (5, 9)
2. Системы счисления
 - 2.1. Двоичное кодирование (1)
 - 2.2. Позиционные системы счисления (16)
3. Логика
 - 3.1. Составление таблиц истинности (2)
 - 3.2. Сложные запросы поисковых систем (17)
 - 3.3. Проверка истинности логических выражений (18)
 - 3.4. Логические уравнения (23)
4. Пользовательский курс
 - 4.1. Анализ информационных моделей (3)
 - 4.2. Сортировка и поиск в базах данных (4)
 - 4.3. Файловая система (4)
 - 4.4. Электронные таблицы (7)
 - 4.5. Адресация в Интернет (12)
 - 4.6. Поиск путей в графе (15)
5. Алгоритмизация и программирование
 - 5.1. Выполнение и анализ простых алгоритмов (6)
 - 5.2. Алгоритмы для исполнителей (6, 14)
 - 5.3. Анализ алгоритмов с циклами (8)
 - 5.4. Рекурсивные алгоритмы (11)

- 5.5. Обработка массивов и матриц (19)
- 5.6. Анализ алгоритмов с циклом и ветвлением (20)
- 5.7. Анализ алгоритмов с циклом и подпрограммой (21)
- 5.8. Перебор вариантов. Динамическое программирование (22)
- 6. Задачи с развёрнутым ответом
 - 6.1. Поиск ошибок в программе (24)
 - 6.2. Обработка массивов (25)
 - 6.3. Анализ стратегии (26)
 - 6.4. Обработка массивов, символьных строк и последовательностей (27).

Как мы видим, систематизировать задачи достаточно сложно. По мнению автора данной статьи, в систематизации К.Ю. Полякова есть небольшое количество неточностей. Так, например, тема «Поиск путей» явно относится к разделу комбинаторики, хотя по формулировкам задач немного выбивается из общего списка задач этого раздела. «Алгоритмы для исполнителей» и «Перебор вариантов. Динамическое программирование» — можно было бы объединить и назвать «Алгоритмы для исполнителей: построение и анализ». Нет связывающих тем. Предполагается, что это сделано в школьном курсе.

В следующем разделе предлагается вариант построения структуры курса, дополненный темами, не входящими в ЕГЭ, но необходимыми для обеспечения связности и логичности излагаемого слушателям профориентационной школы материала. Часть тем объясняет, зачем необходимы знания предыдущих разделов, а часть — дополняет пробелы в общей системе знаний по информатике.

3. Структурная модель

Итак, в результате проведённого анализа и исходя из поставленных целей появилась следующая структурная модель курса для профориентационной школы (в скобках указаны номера задач теста ЕГЭ; в фигурных скобках — ссылки на ранее изученный материал):

- 1. Математические основы информатики
 - 1.1. Теория множеств (17)
 - 1.1.1. Определение
 - 1.1.2. Основные операции
 - 1.1.3. Круги Эйлера
 - 1.1.4. Законы основных операций
 - 1.1.5. Мощность множества
 - 1.1.6. Формула включений-исключений
 - 1.2. Комбинаторика {1.1.6}
 - 1.2.1. Правило произведения (10)
 - 1.2.2. Перестановки
 - 1.2.3. Размещения
 - 1.2.4. Сочетания

- 1.2.5. Подсчёт путей в графе (15)
- 1.3. Теория вероятностей {1.2}
 - 1.3.1. Классическое определение
 - 1.3.2. Независимые события
 - 1.3.3. Несовместные события
 - 1.3.4. Произведение событий
 - 1.3.5. Сумма событий
 - 1.3.6. Условная вероятность
- 1.4. Информация (5, 9, 13)
 - 1.4.1. Событийный подход {1.3}
 - 1.4.1.1. Понятие информации и энтропии
 - 1.4.1.2. Формула Шеннона
 - 1.4.2. Единицы измерения
 - 1.4.3. Алфавитный подход
 - 1.4.3.1. Равномерное кодирование
 - 1.4.3.2. Формула Хартли {1.4.1.2}
 - 1.4.3.3. Неравномерные коды
 - 1.4.3.4. Условие Фано
- 1.5. Системы счисления (1, 10, 16)
 - 1.5.1. Позиционная запись числа
 - 1.5.2. Алгоритм перевода $X_p \rightarrow X_{10}$
 - 1.5.3. Алгоритм перевода целого числа $X_{10} \rightarrow X_p$
 - 1.5.4. Алгоритм перевода дробного числа $X_{10} \rightarrow X_p$
 - 1.5.5. Перевод для оснований $p = 2^q$ {1.4.3}
 - 1.5.6. Выполнение арифметических операций в позиционной системе счисления (ПСС)
- 1.6. Представление чисел в ЭВМ {1.5.5, 1.5.6}
 - 1.6.1. Представление целых
 - 1.6.1.1. Прямые коды
 - 1.6.1.2. Обратные коды
 - 1.6.1.3. Дополнительные коды
 - 1.6.2. Представление действительных
 - 1.6.2.1. Представление с фиксированной точкой
 - 1.6.2.2. Нормализованная форма
 - 1.6.2.3. Представление чисел с плавающей точкой в Pascal
- 1.7. Логика
 - 1.7.1. Элементарные высказывания и понятие логической переменной
 - 1.7.2. Элементарные логические функции и их связь с естественным языком
 - 1.7.3. Таблицы истинности (2)
 - 1.7.4. Алгебра логики (законы логики) {1.1.4}
 - 1.7.5. Нормальные формы (2)
 - 1.7.6. Эквивалентность функций

- 1.7.7. Решение логических уравнений (18) {1.1, 1.5.6, 1.1.6, 1.7.4}
- 1.7.8. Число решений систем логических уравнений (23) {1.2, 1.7.3 }
- 1.7.9. Логические и переключательные схемы
- 1.7.10. Логические задачи
- 1.8. Приложения
 - 1.8.1. Электронные таблицы (7)
 - 1.8.2. Информационные модели (3,4) {1.1, 1.7}
 - 1.8.3. Файловая система (4)
 - 1.8.4. Адресация в Интернет (12)
 - 1.8.4.1. Структура URL
 - 1.8.4.2. Адресация IP v4 {1.5, 1.7}
- 2. Алгоритмизация и программирование
 - 2.1. Исполнители
 - 2.1.1. Чертёжник
 - 2.1.2. Умный мячик
 - 2.1.3. Кузнечик
 - 2.1.4. Путник
 - 2.1.5. Поразрядный вычислитель
 - 2.1.6. Калькулятор
 - 2.1.7. Анализ алгоритмов исполнителей (6, 14, 22)
 - 2.2. Конструкции языков программирования
 - 2.2.1. Конструкции языка Pascal
 - 2.2.2. Конструкции языка Python
 - 2.2.3. Арифметические выражения
 - 2.2.4. Ветвления
 - 2.2.5. Циклы (8, 19, 20)
 - 2.2.6. Подпрограммы (21)
 - 2.2.7. Рекурсия (11) {2.1.7}
 - 2.2.8. Разные задачи (19) {1.1, 1.5, 1.6, 1.7}
 - 2.3. Задачи с развёрнутым ответом
 - 2.3.1. Поиск ошибок в программе (24) {2.2}
 - 2.3.2. Обработка массивов (25) {2.2}
 - 2.3.3. Анализ стратегии (26)
 - 2.3.4. Обработка последовательностей (27) {2.2}

Данная структурная модель представлена на рис. 1.

Заключение

В соответствии с данной структурной моделью (рис. 1) подготовлено учебное пособие для ПШФКН. В течение следующего учебного года оно будет апробировано на слушателях профориентационной школы ФКН и при необходимости будут сделаны корректировки и уточнения.

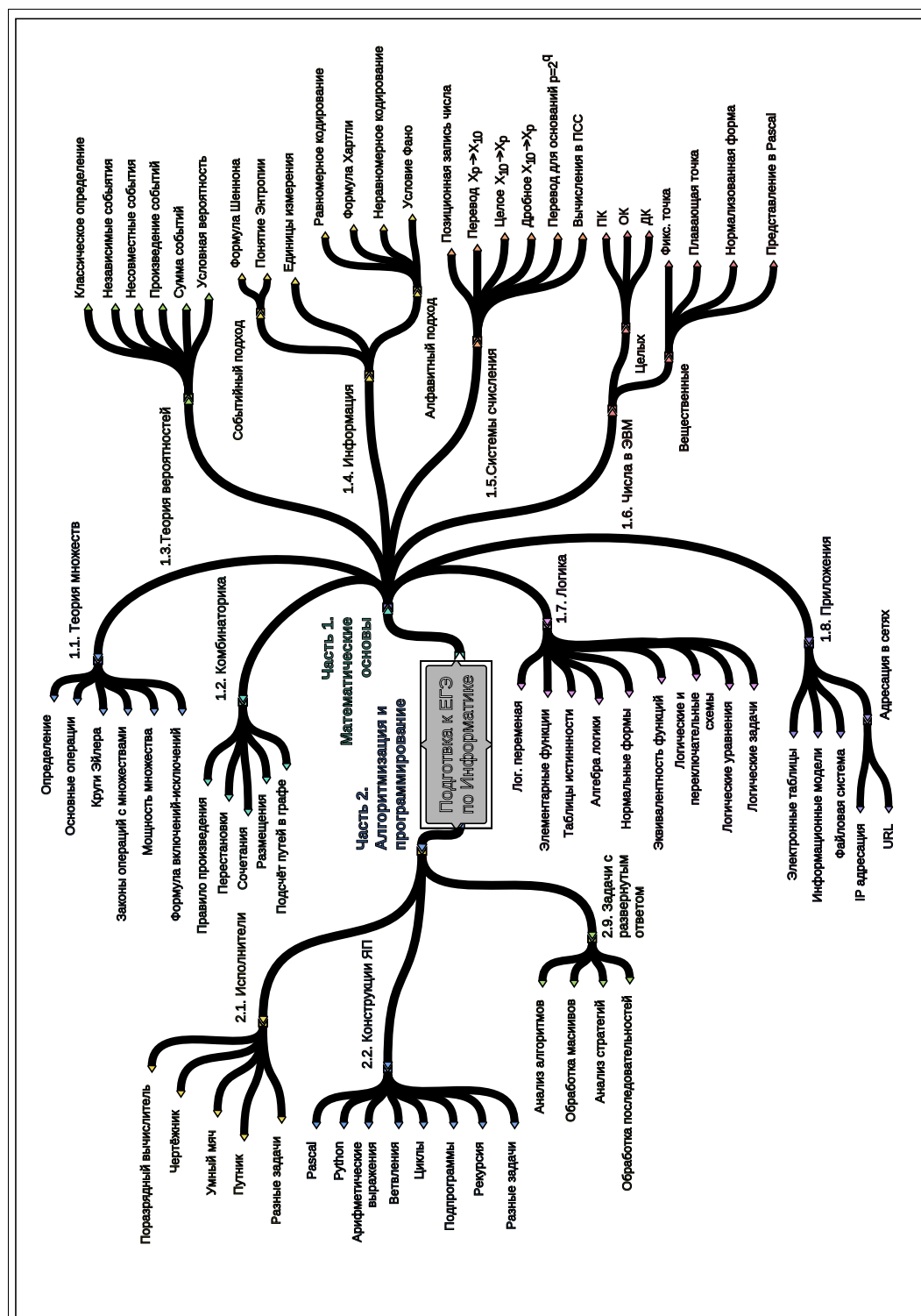


Рис. 1. Структурная модель курса

Полученные школьниками знания по данным предметам будут базой для ряда дисциплин, изучаемых на первом курсе ФКН: информатика, дискретная математика, программирование, ассемблер и др.

Благодарности

Выражаю глубокую признательность Константину Юрьевичу Полякову за разрешение использовать ряд задач с его сайта в пособии для подготовки к ЕГЭ по информатике для профориентационной школы факультета компьютерных наук.

Благодарю Надежду Фёдоровну Богаченко за обсуждение содержательной части данной статьи и конструктивную критику.

ЛИТЕРАТУРА

1. Открытый банк заданий ЕГЭ / Информатика // Федеральный институт педагогических измерений. URL: <http://85.142.162.119/os11/xmodules/qprint/index.php?proj=B9ACA5BVB2E19E434CD6BEC25284C67F> (дата обращения: 15.01.2018).
2. Поляков К.Ю. ЕГЭ по информатике (2018) // kpolyakov.spb.ru Преподавание, наука и жизнь. URL: <http://kpolyakov.spb.ru/school/ege.htm> (дата обращения: 15.01.2018).
3. Лещинер В.Р. ЕГЭ 2018. Информатика. 14 вариантов. Типовые тестовые задания от разработчиков ЕГЭ. М. : Издательство «Экзамен», 2018. 279 с.
4. Крылов С.С., Ушаков Д.М. ЕГЭ 2018. Тренажёр. Информатика. М. : Издательство «Экзамен», 2018. 271 с.
5. Ушаков Д.М. Информатика : большой сборник тематических заданий для подготовки к единому государственному экзамену. М. : Издательство «АСТ», 2018. 321 с.

**DEVELOPMENT OF A STRUCTURAL MODEL OF THE TRAINING COURSE
FOR THE UNIFIED STATE EXAMINATION IN INFORMATICS
IN THE VOCATIONAL GUIDANCE SCHOOL OF THE FCS**

D.N. Lavrov

Ph.D. (Eng.), Associate Professor, e-mail: lavrov@omsu.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. The article is devoted to the analysis of the themes of the unified state exam in informatics and the construction on its basis of the structural model of the course for the vocational guidance school of the Faculty of Computer Science. The analysis of information resources, which are sources of tasks for preparation, has been carried out. The structural model is supplemented with topics that ensure consistency and coherence of topics.

Keywords: structural model, computer science, information and telecommunication technologies, unified state examination, vocational guidance, additional education.

Дата поступления в редакцию: 16.02.2018

Авторам

Предоставляемые данные и документы

Автор предоставляет в редакцию:

- рукопись статьи в формате $\text{\LaTeX}$ (см. требования к оформлению);
- список из трёх экспертов по тематике статьи, давших согласие написать рецензию на представленную работу¹;
- экспертное заключение о возможности открытого опубликования.

Лицензирование

Согласно ГК РФ ст. 1286 лицензионный договор с автором для публикации в периодических изданиях может быть заключён в устной форме. Сам факт получения рукописи статьи редколлегией журнала «Математические структуры и моделирование» является акцептом (принятием) лицензионного договора.

Все статьи в журнале «Математические структуры и моделирование» публикуются под лицензией Creative Commons Attribution 4.0 International (CC-BY). Текст лицензии находится по адресу <https://creativecommons.org/licenses/by/4.0/legalcode>.

Требования к оформлению рукописи

К публикации принимаются рукописи объёмом не более 16 страниц.

Авторам необходимо предоставить следующую информацию на русском и английском языках:

- название статьи;
- список авторов с указанием
 - фамилии, имени и отчества,
 - учёного звания,
 - учёной степени,
 - должности,
 - места работы или учёбы,
 - действующего адреса электронной почты;
- аннотация (абстракт) объёмом от 100 до 250 слов;
- список ключевых слов.

Автор также указывает УДК (универсальный десятичный код) статьи. Его можно подобрать по тематике статьи в справочнике <http://msm.univer.omsk.su/udc/>.

Библиографические ссылки оформляются согласно ГОСТ 7.0.5–2008.

Рукопись статьи представляется в редакцию по электронной почте в двух форматах pdf и tex. Статья должна быть набрана с использованием макропакета $\text{\LaTeX}$ и стиля msmb.cls, предоставляемого редакцией <http://msm.univer.omsk.su/files/msmb.zip>. Рекомендуется установить компилятор $\text{\LaTeX}$, так как именно им пользуются в редакции.

Отклонения в оформлении рукописи от приведённых правил позволяют редколлегии принять решение о снятии статьи с публикации. Статья может быть отклонена по причинам несоответствия тематике журнала или в связи с низким уровнем качества научного исследования.

В статье запрещается переопределять стандартные команды и окружения.

Нумеруемые формулы необходимо выделять в отдельную строку.

Нумерация только арабскими цифрами в порядке возрастания с единицы. Нумеровать следует только те формулы, на которые в тексте имеются ссылки.

¹Необходимы полные данные экспертов (место работы, учёная степень, должность), с указанием способа связи с ними (e-mail, телефон). Редколлегия может обратиться к одному из экспертов из предложенного списка с просьбой написать рецензию или может назначить рецензента из собственного списка.

Запрещается использовать в формулах буквы русского алфавита. Если без них никак не обойтись, то следует использовать команду `\mbox{...}`.

Все рисунки и таблицы должны иметь подпись, оформленную с помощью команды `\caption{...}`.

Файлы с рисунками необходимо представить в формате PDF или EPS (использовать редакторы векторной графики типа InkScape, Adobe Illustrator или Corel Draw).

Используйте стандартные команды переключения на готический, каллиграфический и ажурный шрифты: `\mathfrak`, `\mathcal` и `\mathbb`.

Не допускается заканчивать статью рисунком или таблицей.

В списке литературы обязательно указание следующих данных: для книг — фамилии и инициалы авторов, название книги, место издания, издательство, год издания, количество страниц; для статей — фамилии и инициалы авторов, название статьи, название журнала, год издания, том, номер (выпуск), страницы начала и конца статьи (для депонированных статей обязательно указать номер регистрации).

Кавычки в русском тексте («абвгд») должны быть угловыми, в английском — прямыми верхними кавычками ("abcdeг" или "abcdeг").

Обязательна расшифровка сокращений при первом вхождении термина. Например: ... искусственный интеллект (ИИ)...

Порядок рецензирования

Первичная экспертиза проводится главным редактором (заместителем главного редактора). При первичной экспертизе оценивается соответствие статьи тематике журнала, правилам оформления и требованиям, установленным редакцией журнала к научным публикациям.

Все статьи, поступившие в редакцию научного журнала «Математические структуры и моделирование», проходят через институт рецензирования.

Рецензент выбирается главным редактором журнала из числа членов редколлегии или ведущих специалистов по профилю данной работы.

Рецензенты уведомляются о том, что присланные им рукописи являются частной собственностью авторов и относятся к сведениям, не подлежащим разглашению. Рецензентам не разрешается делать копии статей для своих нужд.

Срок для написания рецензии устанавливается по согласованию с рецензентом.

Рецензия должна раскрывать актуальность представленного материала, степень научной новизны исследования, определять соответствие предлагаемого к публикации текста общему профилю издания и стиль изложения.

Рецензент выносит заключение о возможности опубликования статьи: «рекомендуется», «рекомендуется с учётом исправления замечаний, отмеченных рецензентом» или «не рекомендуется». В случае отрицательной рецензии редакция направляет автору мотивированный отказ, заверенный главным редактором или его заместителем.

В случае несогласия с мнением рецензента автор статьи имеет право предоставить аргументированный ответ в редакцию журнала. Статья может быть направлена на повторное рецензирование, либо на согласование в редакционную коллегию.

При наличии в рецензии рекомендаций по исправлению и доработке статьи автору направляется текст рецензии с предложением учесть их при подготовке нового варианта статьи или аргументированно (частично или полностью) их опровергнуть. Доработанная (переработанная) автором статья повторно направляется на рецензирование и рассматривается в общем порядке. В этом случае датой поступления в редакцию считается дата возвращения доработанной статьи.

После принятия редколлегией решения о допуске статьи к публикации автор информируется об этом и указываются сроки публикации.

Оригиналы рецензий хранятся в редакции в течение пяти лет.

Авторская этика

Авторы публикаций должны гарантировать, что в список авторов включены только лица, соответствующие критериям авторства (лица, внёсшие значительный вклад в работу), и что заслуживающие авторства исследователи не исключены из списка авторов.

Должны работать вместе с редакторами или издателями для скорейшего исправления своих работ в случае обнаружения в них ошибок или упущений после публикации.

Обязаны незамедлительно уведомлять редакцию в случае обнаружения ошибки в любой поданной ими на публикацию, принятой для публикации или уже опубликованной работе.

Не вправе копировать из других публикаций ссылки на работы, с которыми они сами не ознакомились; цитаты и ссылки на другие работы должны быть точными и оформленными в соответствии с предъявляемыми требованиями.

Должны ссылаться максимально правильно и точно на имеющие отношение к публикации предыдущие работы как других исследователей, так и самих авторов, обращаясь, прежде всего к первоисточнику; дословное воспроизведение собственных работ и их перефразирование неприемлемы, они могут быть использованы лишь в качестве основы для новых выводов.

Необходимо указывать авторство данных, текста, рисунков и идей, которые автор получил из других источников — они не должны представляться, как принадлежащие автору публикации; прямые цитаты из работ других исследователей должны выделяться кавычками и соответствующей ссылкой.

Должны соблюдать нормы законодательства о защите авторских прав; материалы, защищённые авторским правом (например, таблицы, цифры или крупные цитаты), могут воспроизводиться только с разрешения их владельцев.

Памятка для перевода должностей, учёных степеней и званий на английский язык

Профессор = Professor

Доцент = Associate Professor

Старший преподаватель = Assistant Professor

Преподаватель = Instructor

Ассистент = Instructor

Аспирант = Postgraduate Student или Ph.D. Student

Соискатель = Ph.D. Doctoral Candidate

Магистрант = Master's Degree Student

Студент = Student

д.ф.-м.н. = Dr.Sc. (Phys.-Math.)

к.ф.-м.н. = Ph.D. (Phys.-Math.)

д.т.н. = Dr.Sc. (Eng.)

к.т.н. = Ph.D. (Eng.)

Инженер-программист = Software Engineer

Старший/младший научный сотрудник = Senior/Junior Scientist Researcher

Электронная почта для отправки статей

lavrov@omsu.ru — зам. главного редактора (ответственный за выпуск) Д.Н. Лавров.

Научный журнал

Математические структуры и моделирование

№1(45)

Главный редактор

А.К. Гуц

Зам. глав. ред., выпускающий редактор

Д.Н. Лавров

Зам. глав. ред., технический редактор

Н.Ф. Богаченко

Корректор:

И.Н. Баловнева

Проверка корректности перевода:

А.Н. Кабанов

Адрес научной редакции

644077, Омская обл., г. Омск, пр-т Мира, д. 55а,
Омский государственный университет

E-mail: guts@omsu.ru, lavrov@omsu.ru

Электронная версия журнала:

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

<http://msm.omsu.ru>



Подписано в печать 12.03.2018. Формат 60 × 84 1/8.

Усл. печ. л. 19,88. Тираж 100 экз. Заказ №45.

Отпечатано на полиграфической базе издательства ОмГУ им. Ф.М. Достоевского
644077, Омская обл., г. Омск, пр-т Мира, д. 55а

ISSN 2222-8772



9 772222 877005



18045 >